Job Tasks, Worker Skills, and Productivity

G. Jacob Blackwood, Cindy Cunningham, Matthew Dey, Lucia Foster, Cheryl Grim, John Haltiwanger, Rachel Nesbit, Sabrina Wulff Pabilonia, Jay Stewart, Cody Tuttle, and Zoltan Wolf*

July 7, 2025

Abstract

We present new empirical evidence supporting the hypothesis that explaining productivity dispersion across businesses requires accounting for differences in how tasks, skills, and occupations are organized. This aligns with growing theoretical attention to the task content of production. We link establishment-level data from the Bureau of Labor Statistics Occupational Employment and Wage Statistics (OEWS) survey with productivity data from the Census Bureau's manufacturing surveys. Our analysis reveals strong relationships between establishment-level productivity and task, skill, and occupation measures within industries. However, these relationships are highly nonlinear and vary by industry. Using these nonlinear, industry-specific patterns helps account for a substantial share of within-industry productivity dispersion across establishments.

JEL codes: D24, J24

Keywords: productivity dispersion, tasks, skills, occupations

*Blackwood: Amherst College; Cunningham, Dey, and Pabilonia: Bureau of Labor Statistics; Foster and Grim: Center for Economic Studies, U.S. Census Bureau; Haltiwanger: University of Maryland; Nesbit: RAND Corporation; Stewart: Upjohn Institute for Employment Research; Tuttle: University of Texas at Austin; Wolf: New Light Technologies. Blackwood, Haltiwanger, and Tuttle were also part-time employees of the U.S. Census Bureau at the time of the writing of this paper. Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau or the U.S. Bureau of Labor Statistics. The Census Bureau and the Bureau of Labor Statistics have ensured appropriate access and use of confidential data and have reviewed these results for disclosure avoidance protection (Project 7526914: CBDRB-FY24-CES020-006, CBDRB-FY25-CES007-10). We thank Enghin Atalay, Thomas Niebel, and John Samuels for helpful comments. Corresponding author: JStewart@Upjohn.org.

I. Introduction

We combine establishment-level productivity and occupational data to provide new insights into how skills, tasks, and occupations contribute to productivity differences across establishments. It is well known that measured productivity varies substantially across establishments, even within narrowly defined industries. For example, publicly available statistics from the Dispersion Statistics on Productivity (DiSP) [1] show that, on average, an establishment at the 90th percentile of the measured total factor productivity (TFP) distribution is nearly 2.9 times as productive as one at the 10th percentile within four-digit NAICS manufacturing industries (Cunningham et al., 2023), a pattern also observed in other studies.[2]

Syverson (2011) reviews potential sources of productivity dispersion, including hard-to-measure factors such as managerial ability and input quality. Cunningham et al. (2023) show that common firm characteristics studied in the firm dynamics literature—such as state, age class, and size class—explain little of the observed dispersion, suggesting a need to examine alternative explanations. We address this by focusing on a critical but underexplored source of heterogeneity: establishment-level differences in the organization and nature of tasks and occupations.

Standard productivity measurement typically aggregates labor input as total hours worked, as in the DiSP data. However, accounting for variation in worker skills

[1] DiSP was developed jointly by the Bureau of Labor Statistics (BLS) and the Census Bureau. See Cunningham et al. (2023) for a detailed description of the development of DiSP. DiSP is available at: https://www.bls.gov/productivity/articles-and-research/dispersion-statistics-on-productivity/ and https://www.census.gov/disp. A restricted-access dataset is available for use by qualified researchers on approved projects in the Federal Statistical Research Data Centers (https://www.census.gov/fsrdc).
[2] See Syverson (2004), Syverson (2011), and Blackwood et al. (2021).

and the types of tasks performed may be essential for accurately measuring both productivity levels and dispersion.[3] Differences in observed productivity may partly reflect the occupational mix and task content of the workforce (see, e.g., Acemoglu and Restrepo, 2019b). To explore this, we integrate establishment-level data from the Bureau of Labor Statistics (BLS) Occupational Employment and Wage Statistics (OEWS) survey with productivity data from the Census Bureau's manufacturing surveys.

This paper builds on a precursor study (Blackwood et al., 2025), conducted before the relevant microdata could be linked. That study examined, at the industry level, the relationship between dispersion in productivity and dispersion in measures of tasks, skills, and occupations within four-digit NAICS manufacturing industries. We adopt most of the task/skill/occupation measures developed in that study, which are described in more detail later. Briefly, these include two composite measures that are constructed using data from the OEWS survey and Occupational Information Network (O*NET); the five aggregate task measures derived from O*NET; and a measure of STEM intensity. In this paper, we also include additional broad occupation groups including the production worker share and the management share.

The composite measures summarize establishment-level variation in occupational and task/skill distributions. One captures variation in the occupational mix and is related to—but distinct from—the BLS skill-adjusted labor input used in official

[3] A few empirical studies allow workers' skill levels to vary. See Iranzo, Schivardi, and Tosetti (2008) and Stoyanov and Zuanov (2022).

TFP estimates.[4] The other captures variation in the use of five aggregate tasks derived from O*NET's work activity and work-context-importance scales, as described in Acemoglu and Autor (2011): nonroutine analytical, nonroutine interpersonal, nonroutine manual physical, routine manual, and routine cognitive tasks.

Blackwood et al. (2025) find that both labor productivity (LP) and TFP dispersion within industries are positively associated with dispersion in the task/skill/occupation measures. These patterns differ quantitatively across different groupings of manufacturing industries but they are especially strong in the high-tech industries. The remarkably high within-industry dispersion of both productivity and task/skill intensities across establishments in high-tech industries implies there is considerable heterogeneity in both the outcomes and the ways of doing business, especially among the most innovative sectors of the economy.

This prior analysis was limited to industry-level variation and could not assess establishment-level relationships. In this paper, we overcome that limitation by linking occupational data from the OEWS survey to establishment-level productivity data from the Collaborative Micro-productivity Project (CMP), resulting in a matched CMP-OEWS dataset. This allows for novel establishment-level analyses. An initial establishment-level dispersion analysis shows that adjusting total hours using a simple scalar measure of task/skill/occupation content has little impact on measured productivity dispersion. While this might seem surprising, this result assumes a (log) linear and uniform relationship across establishments within the same industry—a restrictive assumption.

[4] See https://www.bls.gov/productivity/technical-notes/changes-in-composition-of-labor-total-factor-productivity-2014.pdf for a description of the official labor composition measure. For a more detailed discussion of the theory and measurement issues behind the labor composition index, see Zoghi (2007).

We find instead that the relationship between these task and occupational measures and productivity is highly nonlinear. The strongest relationships between TFP and these measures occur at the extremes—among establishments in the top and bottom quintiles of the task and occupation measures. These nonlinear patterns are even more pronounced among larger establishments.

Motivated by these findings, we re-examine within-industry variation in the OEWS skill, task, and occupation measures. Two key findings emerge from our analysis of the OEWS survey data. First, most of the variation in broad occupation shares occurs within, not between, industries. Second, within-industry variation in occupation shares is especially large in the top and bottom quintiles of the task/skill distribution.

Returning to the matched CMP-OEWS microdata, we conduct a descriptive analysis of variance. While not causal, this analysis shows that a substantial portion of establishment-level TFP dispersion within industries can be accounted for by differences in occupational mix, tasks, and skills. These relationships are nonlinear, vary by industry, and are particularly strong among larger establishments and those in high-tech sectors.

The rest of the paper is organized as follows. Section II describes how we construct the occupation, task, and skill measures. Section III discusses our data sources and the matching procedure. Section IV presents our main results, highlighting the complex relationship between TFP and the occupational/task measures. Section V concludes and outlines directions for future research.

II. Measuring occupations, tasks, and skills

We start by defining our concepts.[5] *Tasks* are activities that when combined with capital and intermediate goods create a good or service and are the true factors of production we would like to measure. However, because we do not observe time spent on different tasks, we use occupations as proxies. An *occupation* can be thought of as a bundle of tasks.[6]

Skill refers to a worker's ability to perform various tasks. It is commonly measured as a function of education and experience; however, due to data constraints, we proxy skills by wages. Complex tasks generally require greater skills, although the relationship between skills and tasks can vary over time and across businesses, presenting a challenge for productivity measurement and highlighting a need for detailed data on tasks and skills.

We now turn to defining our two composite measures, five task measures, and three occupation groups (see also Blackwood et al. (2025)).

A. Bundled Task/Skill Intensity Index (TSB): Counterfactual Wages

Our first measure of task/skill intensity is a counterfactual wage equal to the average wage paid by the establishment if the establishment paid the national average occupational wage for all workers in each occupation it employs. Thus, it accounts for differences in the occupational mix across establishments by attaching a different price to each occupation. By using the national average wage for each occupation, the price

[5] These descriptions are based on the nomenclature from the Revised Handbook of Analyzing Jobs (Employment and Training Administration (1991)) and Acemoglu and Autor (2011).

[6] From Employment and Training Administration, 1991, p. 9, an occupation is "a common set of tasks are performed or are related in terms of similar objectives, methodologies, materials, products, worker actions, or workers characteristics."

of each occupation is the same across establishments. We refer to this as a "bundled" task/skill intensity index (TSB) because tasks are bundled into occupations.

More formally, let $\overline{w}_{ej}$ and L_{ej} denote the mean log wage and the number of workers in occupation *j* at establishment *e*. Suppressing time subscripts for simplicity, the national mean log wage for occupation *j* is given by:

$$\overline{w}_{nj} = \frac{1}{\sum_{e \in E_n} L_{ej}} \sum_{e \in E_n} (\overline{w}_{ej} \times L_{ej}) \tag{2}$$

where E_n is the set of all establishments. The counterfactual mean log wage for establishment *e*, $\widetilde{w}_e$, can then be written as:

$$\widetilde{w}_e = \frac{1}{L_e} \sum_{j \in J_e} (\overline{w}_{nj} \times L_{ej}) \tag{3}$$

where J_e is the set of occupations employed by establishment *e* and L_e is total employment in establishment *e*.

TSB is a simple measure that summarizes the types of tasks employed by the establishment using wages, which proxy for skills, to price those tasks. Given that TSB is based on occupation-specific national average wages, the cross-establishment differences in this measure reflect variation in the occupation mix. Although this is a useful measure, it does not distinguish between different occupations (with different task sets) paying the same wage. Thus, two establishments might have the same task/skill intensity but very different mixes of occupations.[7]

[7] In Blackwood et al. (2023), we illustrate this point by plotting the TSB measure against a dissimilarity index that quantifies how the occupational mix of the establishment differs from the average occupation mix of its four-digit industry. The dissimilarity index that we use is the absolute value of the sum over all occupations (two-digit Standard Occupational Classification (SOC)) of the distances between the establishment's payroll share for that occupation and the industry-wide payroll share for that occupation. It takes on values between zero and one, with higher values indicating an establishment has a much different occupational distribution than the typical establishment in the industry.

B. Unbundled Task/Skill Intensity Index (TSU): Task-Adjusted Counterfactual Wages

Our second measure focuses on tasks and builds on Acemoglu and Autor (2011), who use O*NET data to operationalize the Autor, Levy, and Murnane (2003) taxonomy of tasks. Autor, Levy, and Murnane (2003) develop a two-dimensional categorization of tasks based on whether they are (1) routine or non-routine and (2) cognitive or manual. They further break down non-routine cognitive tasks into analytic and interpersonal tasks. This yields five aggregate tasks: non-routine cognitive (analytical), non-routine (interpersonal), routine cognitive, routine manual, and non-routine manual physical.[8]

We use this methodology to create the same five task indexes for each of the O*NET years where the index variables are available for most occupations (2007, 2008, 2014, and 2017).[9] We merge these five task indexes to the OEWS by occupation and estimate the following regression of the national occupational mean log wage for each year on these five task indexes:

$$\bar{w}_{nj} = \alpha + \sum_{k=1}^{5} \beta_k \tau_{jk} + \varepsilon \qquad (4)$$

where τ_{jk} is the O*NET measure of task *k* for occupation *j*, and $\bar{w}_{nj}$ is defined as in equation (2).[10] The coefficients on the task indexes, β_k, are akin to prices in a hedonic regression. We then calculate the counterfactual average establishment wage as:

$$\hat{w}_e = \frac{1}{L_e} \sum_{k=1}^{5} \hat{\beta}_k \left[\sum_{j \in J_e} \left(L_{ej} \times \tau_{jk} \right) \right] \qquad (5)$$

[8] Acemoglu and Autor (2011) include a sixth category, offshorability, which we do not include here because it is not a task.

[9] We match two prior years of OEWS data to a given O*NET year to obtain the employment weights. When an occupation is covered in both OEWS years, we average the two years; otherwise, we take the value for the one OEWS year with coverage for that occupation. Thus, the 2007 O*NET is matched to 2005 and 2006 OEWS; 2008 O*NET to 2006 and 2007 OEWS; 2014 O*NET to 2012 and 2013 OES; and 2017 O*NET to 2015 and 2016 OEWS.

[10] We first aggregate occupations to a time-consistent SOC classification.

where the summation in square brackets is the total amount of task *k* employed by the establishment and $\hat{\beta}_k$ is the estimated "price" of task *k* estimated from equation (4). That is, the TSU measure can be thought of as the average price of the tasks performed by employees in the establishment.

We refer to this second measure as an "unbundled" task/skill intensity index (TSU) because tasks (weighted by prices) are aggregated without accounting for how they are bundled into occupations. In contrast, TSB captures the occupational mix of an establishment (and the prices of those occupations), so it implicitly takes into account that individual occupations reflect a bundle of tasks (and that the bundle of tasks is not determined randomly). Like the TSB index, there are many combinations of tasks that can result in the same value of TSU.

Both composite measures reflect task/skill differences across establishments and account for the prices of those tasks in the labor market, where prices reflect the skills required to accomplish those tasks (among other things that determine wages). The major difference between these two measures is the first reflects how the tasks are organized into occupations, indirectly accounting for complementarities between tasks that make up an occupation and the benefit of having them performed by the same person, while the second prices the tasks individually and ignores any complementarities between tasks within occupations.

C. Individual Average Task Indexes

In addition to the two task/skill/occupation intensity measures based on counterfactual wages, $\widetilde{w}_e$ and $\widehat{w}_e$, we also construct five task measures based on the average values of the individual O*NET task indexes. For each of the five task indexes described above, we measure an employment-weighted establishment-level average for task index *k* as follows:

$$\bar{\tau}_{ek} = \frac{1}{L_e} \sum_{j \in J_e} \tau_{jk} \times L_{ej} \quad (6)$$

where *k* = 1, …, 5. Thus, $\bar{\tau}_{ek}$ is the average task *k* content of all jobs in establishment *e*. We construct these measures for each establishment for each year in our sample.

D. Occupation Groups

In addition to the measures described above, we also perform analysis with three major occupation groups: Production Workers, STEM Occupations, and Management, as defined by the SOC codes in the OEWS data.[11] Our definition of production workers focuses on workers who do actual production (including material moving) and is therefore narrower than the ASM definition, which includes occupations that are not directly involved with production.[12] The broader task-based definition of production work in the ASM likely means workers in STEM occupations sometimes fall under the production worker definition (e.g., under product development), while most workers in STEM occupations fall into the nonproduction category. Management occupations are included in ASM total employment, but not in the production worker count.

[11] Production Workers include Production Occupations (51-0000) and Material Moving Workers (53-7000). STEM Occupations include Computer and Mathematical Occupations (15-0000), Architecture and Engineering Occupations (17-0000), and Life, Physical, and Social Science Occupations (43-5000). Management includes Management Occupations (11-0000) and Business and Financial Operations Occupations (13-0000).

[12] The ASM definition includes workers engaged in fabricating, processing, assembling, inspecting, receiving, packing, warehousing, shipping (but not delivering), maintenance, repair, janitorial, guard services, product development, auxiliary production for plant's own use (e.g., power plant), recordkeeping, and other closely associated services (including truck drivers delivering ready-mixed concrete). It also includes first-line supervisors.

III. Data and matching

In this section we describe the two datasets we use, the Collaborative Micro-productivity Project (CMP) data and the OEWS survey occupation data, and how we link them.

A. CMP Data

As part of the CMP, BLS and the Census Bureau created an establishment-level productivity database for the manufacturing sector.[13] Data on inputs and output are from the Annual Survey of Manufactures (ASM) and the Census of Manufactures (CM), and are longitudinally linked using information from the Longitudinal Business Database (LBD), which is based on the Census Bureau's Business Register (see Chow et al. (2021)). The LBD provides high-quality longitudinal links and information on the universe of active non-farm private sector employer establishments. The ASM is a five-year panel of manufacturing establishments updated by births in each year and is collected annually.[14] The CM collects data from all manufacturing establishments, except those that are very small, every five years.[15] The CMP microdata combine information from the ASM, CM, and LBD to create measures of inputs, output, and productivity for each establishment (Cunningham et al., 2023).

In preparation for matching CMP data to the OEWS survey data, we address some disagreements between the CM/ASM data and the LBD. Because production functions are calculated industry-by-industry, the most relevant are disagreements in

---

[13] Each year, the CMP team releases a new version of this database; the version used for this paper is Version 7, which covers the 1972–2020 period.
[14] ASM panels start in years ending in "4" and "9."
[15] The CM is collected in years ending in "2" and "7."

industry codes, which can arise for several reasons. For example, ASM and CM industry codes are based on the actual survey responses, whereas the LBD codes are updated with a lag. Thus, an establishment that changes its industry might show up in different industries in the LBD, the ASM, and the CM. As a result, there are three potential “raw” industry codes in the CMP data: the LBD industry code, the CM industry code, and the ASM industry code. A separate but related issue results from changes to the NAICS industry classification system over time, and the differential timing of those changes in the various datasets. For example, the transition from 2007 to 2012 NAICS codes resulted in a major reduction in the number of manufacturing industries, from 473 to 364 six-digit industries. The LBD provides an additional longitudinally consistent industry code—the vintage consistent (VC) industry code (Chow et al. (2021)). This code aims to pick one vintage of NAICS (in this case, the 2017 vintage) and extend that vintage backward so that, during the sample period, industry codes are consistent with the 2017 classification system. At times, the VC process involves consolidation or even imputation of codes based on other characteristics of the establishment.

Because an establishment’s industry code is an integral part of the matching procedure, it is important to have the best possible chance of matching an establishment’s NAICS code to the code that the BLS would assign. Accordingly, we use four different NAICS codes for our CMP dataset: the LBD code, the ASM code (only available for manufacturing observations), the VC code, and a “combined” code created by combining information from the CM, ASM, and LBD. Details of this procedure are in Appendix A.

While this paper focuses on manufacturing establishments, we apply the matching procedure to non-manufacturing observations as well. Therefore, we have a total of approximately 999,000 establishment-year observations in our augmented CMP dataset, with the goal of assigning occupation information to all those establishments using the OEWS survey data.

B. Occupational Employment and Wage Statistics (OEWS) Survey Data

Our occupation data come from the OEWS survey, which is a semi-annual survey of approximately 200,000 establishments in May and November of each year.[16] This survey covers both full-time and part-time workers in private, non-agricultural industries. Employer Identification Numbers (EINs) and NAICS codes come from the BLS's Quarterly Census of Employment and Wages (QCEW), which is the sample frame for the OEWS survey.

The survey instrument asks establishments to provide what is essentially a complete payroll record for the pay period that includes the 12th of the sample month. For each occupation, respondents report the number of employees in each of 12 wage intervals.[17] The OEWS survey uses the Office of Management and Budget's occupational classification system, the SOC, to categorize workers into over 800

---

[16] From 1999 to 2001, the program surveyed approximately 400,000 establishments in November of each year. Starting in November 2002, the program switched to semi-annual sampling with 200,000 establishments sampled each May and November. To keep sample sizes roughly consistent across the various years, we combine November and May panels to create a pseudo-annual sample and assign it the May year value. For this reason, we do not have data for 2002.

[17] Wages in the OEWS survey represent straight-time, gross pay, exclusive of premium pay. Base rate, cost-of-living allowances, guaranteed pay, hazardous-duty pay, incentive pay including commissions and production bonuses, tips, and on-call pay are included, while back pay, jury duty pay, overtime pay, severance pay, shift differentials, non-production bonuses, employer cost for supplementary benefits, and tuition reimbursements are excluded from the reported wage. For a description of the wage intervals, see https://www.bls.gov/oes/mb3-methods.pdf.

detailed occupations, which is much more detailed than the Census occupation codes used in household surveys.

The sample contains both certainty and non-certainty units. The former are generally sampled every three years, while the latter are selected randomly and tend to be smaller establishments. Given this sample design, six consecutive panels can be used to create a representative sample that corresponds to any three-year period.[18] We use this aspect of the sampling scheme in our matching procedure, described in detail below.

We make the same time-consistent adjustment to the OEWS survey industry codes as we make to the LBD and ASM industry codes. This results in two versions of the OEWS survey NAICS codes, one the original version and the other the time-consistent version in which some six-digit industry codes have been aggregated into quasi-five-digit codes.

C. Linking the OEWS Survey Data and the CMP Data

Linking OEWS survey data and CMP data is not straightforward because the establishment identifiers are not the same in the two datasets. However, both datasets have information about the taxpayer ID (EIN) and the industry (the NAICS code) attached to each establishment. We note that the EIN does not necessarily correspond to the Census definition of an enterprise, which depends on operational control. Thus, an enterprise may comprise many EINs. For each establishment in our augmented

[18] Official estimates are typically published for May of a given year. These estimates are based on data from the May panel and the previous five panels. Note that although official estimates are published, they are not a true time series. In year-to-year comparisons of consecutive years, data from approximately 2/3 of units appear in both years. For these units, the wages are updated using the Employment Cost Index, but employment counts are not adjusted.

CMP sample, our goal is to identify the best candidate in the OEWS survey, where the best candidate is defined based on the EIN, NAICS code, geography (state FIPS code), and size (as measured by employment). Loosely speaking, a match occurs if the values of these variables are the same for any two records in the two datasets.

EIN-based matches are in principle exact for single-unit firms. However, even among single-unit EINs, our matches may not be exact for several reasons. First, the two business registers have slightly different criteria for classifying establishments according to single- or multi-unit status.[19] This implies a single-unit CMP establishment may have multiple candidates in the OEWS survey that share the same EIN. Second, the NAICS code may differ between the BLS and Census business registers. This possibility exists because the two agencies use slightly different criteria for classifying establishments into industries. Third, there can be temporal mismatches in the data collected for the establishment because the two surveys may have been conducted at different times and for different reference periods.[20] As described in Section IV.B, the OEWS survey sample scheme is such that three years of OEWS surveys combined produce a representative sample. We match three years of OEWS survey establishments for every one year of CMP establishments, where the years of OEWS survey data are centered on the year of the CMP data. For example, all establishments in the 2014, 2015, and 2016 OEWS surveys would be considered as possible donors for an establishment in the 2015 CMP.

[19] Among other reasons, this discrepancy exists because the timing of single-unit growth into multi-unit, or of multi-unit contraction into single-unit, can be difficult to infer (Chow et al. (2021)).
[20] The reference periods for the ASM and OEWS survey data could differ by up to 18 months.

We require OEWS survey establishments to match on EIN and be of similar size in all steps of our matching procedure. We measure size using employment and calculate the "employment difference" as $|L_{ASM} - L_{OEWS}|/((L_{ASM} + L_{OEWS})/2)$. Our matching procedure is hierarchical in that we prioritize potential donors that match on the most detailed information on industry and geography. We start with the most stringent criteria and then successively relax them.

For establishments with more than 100 employees, the matching criteria are as follows:

(1) EIN, six-digit industry, state, employment difference less than 0.5
(2) EIN, time-consistent six-digit industry, state, employment difference less than 0.5
(3) EIN, six-digit industry, employment difference less than 0.5
(4) EIN, time-consistent six-digit industry, employment difference less than 0.5
(5) EIN, four-digit industry, employment difference less than 0.5

For establishments with fewer than 100 employees, we use an additional employment threshold that depends on the length of time between observations in the OEWS and ASM.[21] To calculate these thresholds, we start with single units whose EINs match exactly. Using these establishments, we calculate the 90th percentile of the absolute difference in employment $|L_{ASM} - L_{OEWS}|$ allowing it to vary by the time between when the establishments are sampled in the OEWS and ASM. The ASM is sampled in March, while the OEWS is sampled in second and fourth quarters. Since we allow for up to two years of timing difference, the length of time between samples can be 90, 270, 450 or 630 days. Using the single-unit establishment matches, we calculate 90th percentile of

[21] We make this modification because small absolute employment changes in small establishments can result in large percentage changes.

the absolute difference in employment for each time length across establishment size categories: 0-4, 5-9, 10-19, 20-49, and 50-99. Then our rules for matching follow the hierarchy as before with slight modification:

(1) EIN, six-digit industry, state, employment difference less than 0.5 OR absolute employment difference <= p90 (time-varying absolute employment difference)
(2) EIN, time-consistent six-digit industry, state, employment difference less than 0.5 OR absolute employment difference <= p90 (time-varying absolute employment difference)
(3) EIN, six-digit industry, employment difference less than 0.5 OR absolute employment difference <= p90 (time-varying absolute employment difference)
(4) EIN, time-consistent six-digit industry, employment difference less than 0.5 OR absolute employment difference <= p90 (time-varying absolute employment difference)
(5) EIN, four-digit industry, employment difference less than 0.5 OR absolute employment difference <= p90 (time-varying absolute employment difference)

In both cases (1-99 and 100+ employees), (1) starts with the original six-digit NAICS codes, whereas (2) is based on our time-consistent codes described in Section IV.A, which are slightly less detailed than the original six-digit codes in some cases due to aggregation where NAICS vintages differ. In (3), we return to our original six-digit codes but relax the geographic requirement, and (4) repeats (3) but instead uses the time-consistent codes. Finally, (5) allows for matches with four-digit industry codes (as well as EIN and size, as in all cases). As mentioned in Section IV.A, we have multiple industry codes that can be used in the matching procedure. Therefore, in each step, we iterate over the three industry codes: starting with the combined NAICS code, then ASM

NAICS code if we found no potential donors with the combined code, then finally LBD NAICS code.

In many cases, there will be multiple potential donors in the OEWS survey that satisfy the same criteria for a match. When this occurs, we break ties by choosing the donor that is closest in size to its CMP establishment. When multiple donors are of the same size, our second tiebreaker is to choose the donor closest in survey year to the CMP establishment.[22] Finally, in cases where both employment and survey year are the same, we randomly choose a donor from among those that meet all the criteria. See Appendix A for an example that illustrates these steps.

The result of the process is that each donor chosen from the OEWS survey is at least from the same EIN, four-digit industry, and size as its CMP recipient. This builds a dataset of CMP observations for which we have information on the occupation distribution from the OEWS survey. We believe our current approach balances match quality with sample size requirements.

D. Final Analysis Sample

The matching procedure detailed above yields a total of approximately 333,000 manufacturing observations between 2001 and 2020, all of which have information about the occupation distribution as well as measures of productivity.

Our final analysis sample incorporates the following modifications. First, we use the vintage-consistent NAICS code from the LBD, instead of the codes used for matching, because we need consistent codes to remove industry and year effects from

[22] Recall that for one year of CMP establishments, we consider potential matches from three years of OEWS survey establishments because of the OEWS survey sampling scheme.

all relevant variables by demeaning (removing industry-year effects).[23] Second, we focus on manufacturing because it is the only sector for which we can construct TFP and the establishment level. Finally, we create inverse propensity score weights, because our linked data are not a representative sample of manufacturing establishments. We estimate a logistic regression to predict the probability of being included in the linked dataset using information on industry, size, and payroll. The inverse of the fitted value from this regression yields the inverse propensity score weight (PW).[24]

Table 1. Descriptive statistics of employment in manufacturing

Summary statistic	OEWS (weighted)	CMP (PW)	LINKED (unweighted)	LINKED (PW)
Mean	32.5	53.1	221.5	61.9
Standard deviation	167.5	204.8	543.1	253.6

Notes: OEWS survey weights account for the probability of selection, the fact that six panels of data are combined to form the full sample, and differences in employment totals between the sample and the QCEW frame. PW refers to inverse propensity weights. CMP refers to the combined ASM, CM, and LBD data. Source: Authors' calculations based on OEWS and CMP.

Table 1 shows descriptive statistics of the establishment size from the OEWS survey, the CMP, and the linked dataset. The differences in the mean and standard deviation of employment help us highlight the potentially different sample characteristics across the three datasets: employment moments are largest in the linked data without using PW, much smaller in the CMP data, and smaller still in the OEWS survey data. With PW, the linked data have patterns broadly similar with the CMP data.

To get a better sense of the distributional differences across the datasets, Table 2 shows the standard deviations of the demeaned variables that we use in our analysis.

---

[23] The vintage-consistent industry codes used are the same industry codes used to create the publicly available DiSP.
[24] Details of the PW construction can be found in Cunningham, et al. (2023).

Employment dispersion after demeaning exhibits the same the patterns as those in Table 1. Task/skill/occupation variation in the linked sample is smaller relative to the OEWS survey sample size without PW but becomes closer with PW. Dispersion in measures of productivity, earnings-per-worker, and capital intensity are similar in the CMP and linked data, especially with propensity score weighting.

Tables 1 and 2 show that the linked dataset retains the basic properties of the OEWS and CMP data, which gives us confidence in the results of the analysis that follows.

Table 2. Standard deviations of key variables

Key variables	OEWS (weighted)	CMP (PW)	LINKED (unweighted)	LINKED (PW)
Employment	164.2	198.6	507.6	239.8
Analytical	0.435		0.3118	0.3320
Interpersonal	0.501		0.3190	0.3866
Physical	0.510		0.3802	0.4331
Routine cognitive	0.501		0.3189	0.3885
Routine manual	0.718		0.4800	0.5486
TSU	0.153		0.1078	0.1147
TSB	0.189		0.1392	0.1486
Log(TFP)		0.4808	0.4952	0.5073
Log(LP)		0.7472	0.7283	0.7461
Log(Earnings-per-Worker)		0.3743	0.4024	0.4793
Log(Capital/Labor)		1.1710	0.9603	1.0963
Production worker share			0.2022	0.2407
STEM share			0.0777	0.0734
Management share			0.0684	0.0885

Notes: OEWS survey weights account for the probability of selection, the fact that six panels of data are combined to form the full sample, and differences in employment totals between the sample and the QCEW frame. PW refers to inverse propensity weights. Industry-year effects are removed. Sample sizes in thousands: 593 (OEWS), 999 (CMP), and 333 (LINKED). CMP refers to the combined ASM, CM, and LBD data. Source: Authors' calculations based on OEWS and CMP.

IV. Relationship between productivity and occupations, tasks, and skills

Through a series of empirical exercises, we highlight the importance of allowing for more complex relationships when trying to understand the relationship between productivity and occupations, tasks, and skills.

A. A Simple Dispersion Exercise

We start by considering a dispersion accounting exercise where we make a simple multiplicative adjustment to the labor input used in the DiSP data (total hours) as shown in equation (7).

$$\log TFP_{et} = logQ_{et} - \alpha_K logK_{et} - \alpha_L \log\left(Z_{et}L_{et}\right) - \alpha_M logM_{et} \qquad (7)$$

where *Q* is real output measured as deflated revenues, *K* is real productive capital stock, *M* is the deflated value of expenditures on intermediate inputs (materials, resales, contract work, electricity, and fuels), *Z* is a normalized version of TSB (TSU) and *L* is total hours. The parameters $\alpha_K, \alpha_L,$ and α_M are factor elasticities measured by the share of expenditures of each input in total cost in each six-digit NAICS industry. For more details on the construction of these variables, see Cunningham et al. (2023).

In constructing *Z*, we normalize TSB (TSU) so it has a mean of one in each industry-year cell. We calculate mean TSB (TSU) by four-digit industry-by-year, then divide each establishment's TSB (TSU) by the industry-year mean value. To adjust total hours, we multiply total hours by this normalized measure of TSB (TSU), yielding a labor input measure that incorporates task-skill intensity in a simple manner.

Table 3. Accounting for dispersion (IQR) in (log) TFP, linked sample

Labor input	PW	AW
Total Hours	0.456	0.496
Total Hours × TSB	0.456	0.494
Total Hours × TSU	0.456	0.494

Notes: PW refers to inverse propensity weights. AW refers to activity weights, where activity weights are employment multiplied by PW weights. Source: Authors' calculations based on OEWS and CMP.

Table 3 shows the average IQRs over our sample period for the IQR of log TFP. In column 1, we report results using just inverse propensity weights (PW) for sample adjustment. In column 2, we report results using activity weights (AW), where AW are PW multiplied by employment. Using this simple approach to convert labor input into efficiency units using the TSB or TSU task/skill measures does not reduce measured dispersion.

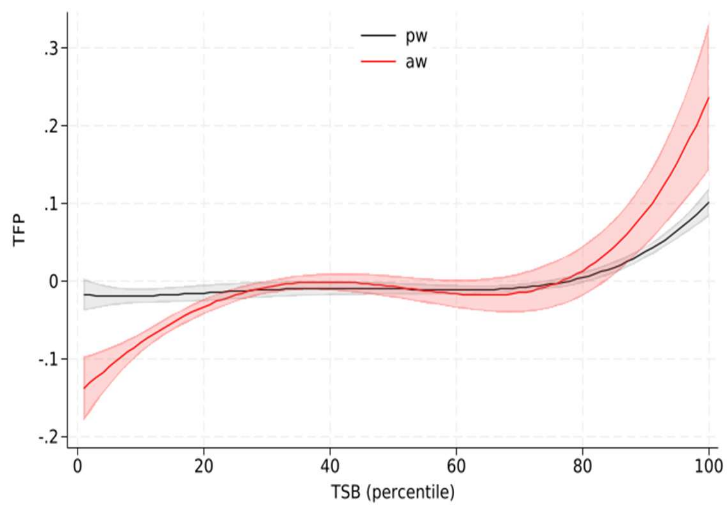
However, this simple exercise imposes the same factor elasticities across all establishments in the same industry. Relatedly and distinctly, this exercise uses a simple one-dimensional adjustment to account for skill and tasks. Establishments that are organized differently likely have different production technologies that are not well captured by this simple Cobb-Douglas specification with only a one-dimensional multiplicative adjustment of the labor input.

This simple exercise does not take into account the rich potential interaction of tasks and other inputs in the type of task content of production approach to productivity as in the seminal work of Acemoglu and Restrepo (2019b). We are not prepared to implement this type of approach in this measurement-oriented paper. Instead, in the sections that follow, we explore the relationship between productivity and the occupational and task mix in a multi-dimensional, nonlinear manner within industries.

B. Importance of Nonlinearities

We now examine the potential for nonlinearities in the relationship between TFP and the measures of skills, tasks and occupations. After sweeping out industry by year effects, we compute percentiles of each task and occupation measure. In turn, we compute the average (log) TFP for each percentile. For disclosure avoidance reasons, we report the results of these exercises fitting a quartic relationship relating average TFP to the percentile ranking of the task/skill/occupation measure. We conduct these exercises separately using PW and AW. Comparing the results using the two weighting methods gives us insight into the differences between large and small establishments.

Figure 1. Relationship between TFP and TSB

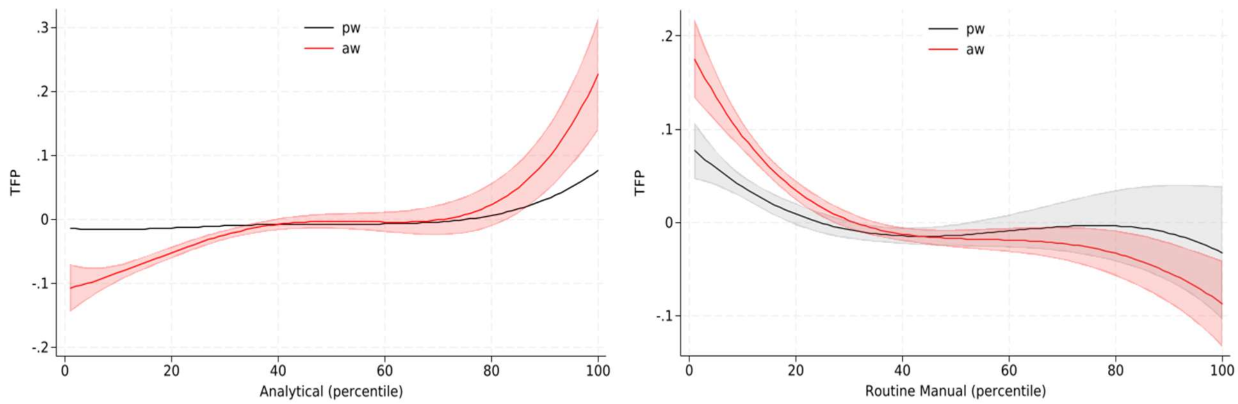


Notes: The label “pw” refers to inverse propensity score weights and label “aw” refers to activity weights.

Figure 1 shows the relationship between TFP and TSB across the TSB percentiles. Notably there is a highly nonlinear relationship, especially when we use AW. In the top quintile, TFP rises rapidly with TSB using either weighting method, but the increase is much faster using AW. In the graph using AW, TFP rises rapidly over the lowest 20 percentiles as well. Together, these imply that much of the nonlinearity is due to large establishments.

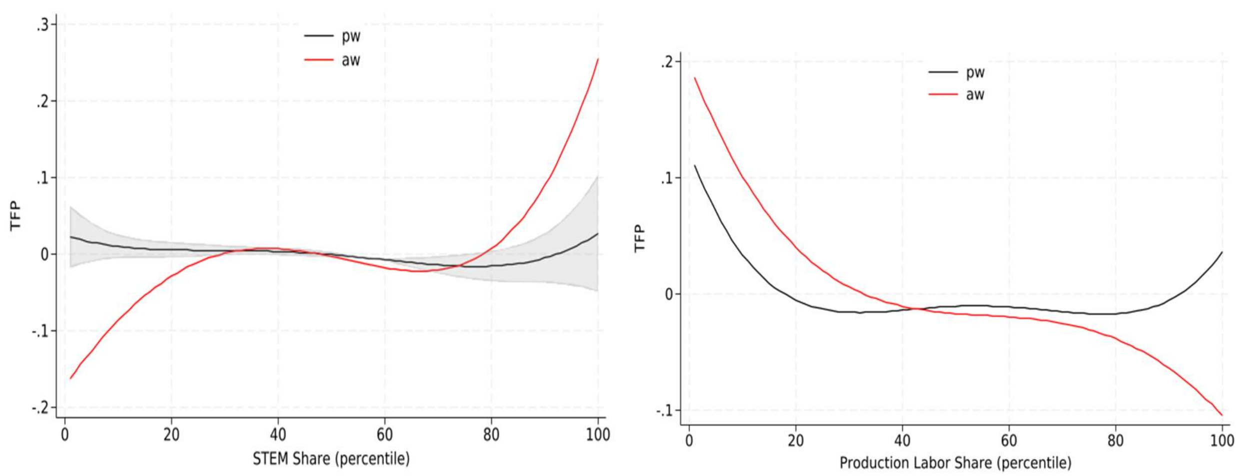
Figure 2 shows analogous relationships between TFP and the O*NET analytical share (left panel) and the routine manual share (right panel). The analytical share results in Figure 2 mimic those for TSB. The patterns for routine manual are the mirror image with TFP declining with routine manual over the bottom quintile and top quintile, but with relatively little relationship in the middle of the distribution. Again, these figures imply it is the large establishments that are driving the nonlinearities in the tails.

Figure 2. Relationship between TFP and ONET analytical and routine manual shares



Notes: The label “pw” refers to inverse propensity score weights and label “aw” refers to activity weights.

Figure 3. Relationship between TFP and STEM and production worker shares



Notes: The label “pw” refers to inverse propensity score weights and label “aw” refers to activity weights. Standard error bands for the aw graph in left panel and for both graphs in the right panel not available for disclosure reasons.

Figure 3 provides additional evidence of significant nonlinear relationships between TFP and the occupational mix within industries. The STEM occupational share is positively related to TFP especially in the tails of the distribution—with evidence for a relationship in the lower tail for larger establishments. The mirror image holds for the production worker share at establishments.

The nonlinear patterns we see in Figures 1–3 suggest complex relationships between variation in measured TFP and the task and occupational mix across establishments. We explore these relationships further below. But we first take a closer look at the relationship between TSB and occupations both between and within industries.

C. A Closer Look at Occupations

From Figures 1–3, it is clear that the relationships between TFP and TSB as well as individual task and occupational shares are nonlinear. Our first step is to look at how production differs by TSB quintile by looking at employment in three broad occupation groups: production workers, STEM workers, and management. We then examine the sources of variation in employment shares by quintile.

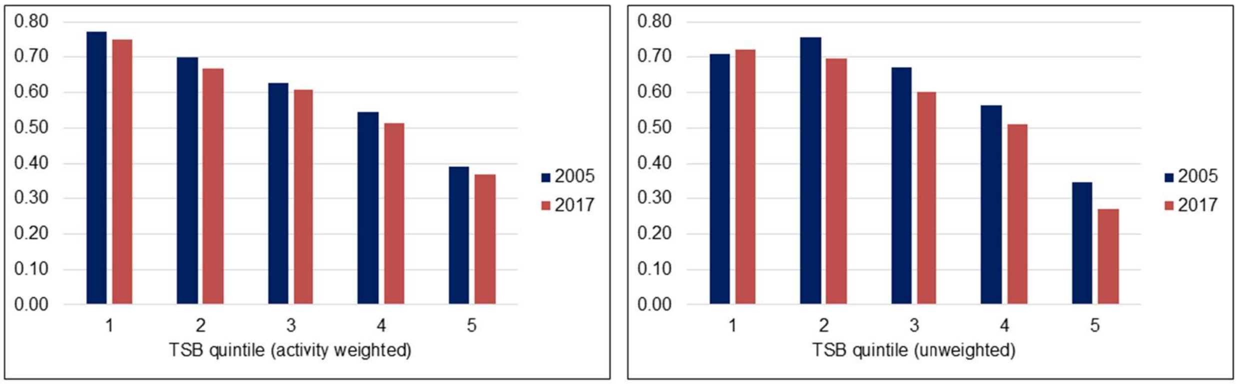
For these exercises, we use the OEWS research dataset used in Blackwood et al. (2025). This dataset includes all establishments in the QCEW. Occupation data from the OEWS are used for OEWS respondents and are imputed for establishments not in the OEWS.[25] The advantage of this approach over using just OEWS data is that each

[25] See Appendix A for a brief description of the dataset construction and Blackwood et al. (2025) and Dey, Piccone and Miller (2019) for detailed descriptions.

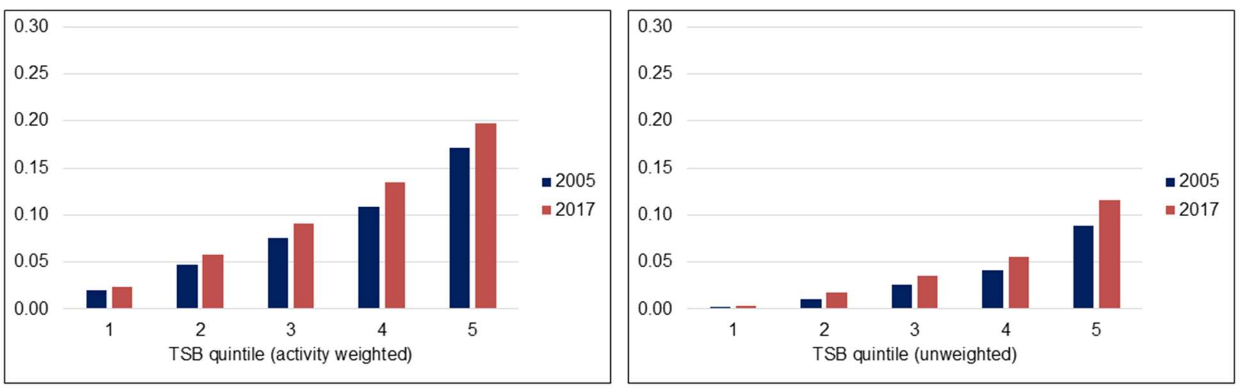
establishment has a weight of one.[26] The imputed values in the research dataset are a combination of OEWS data and data from the QCEW.

Figure 4. Mean occupation share of employment by TSB quintile

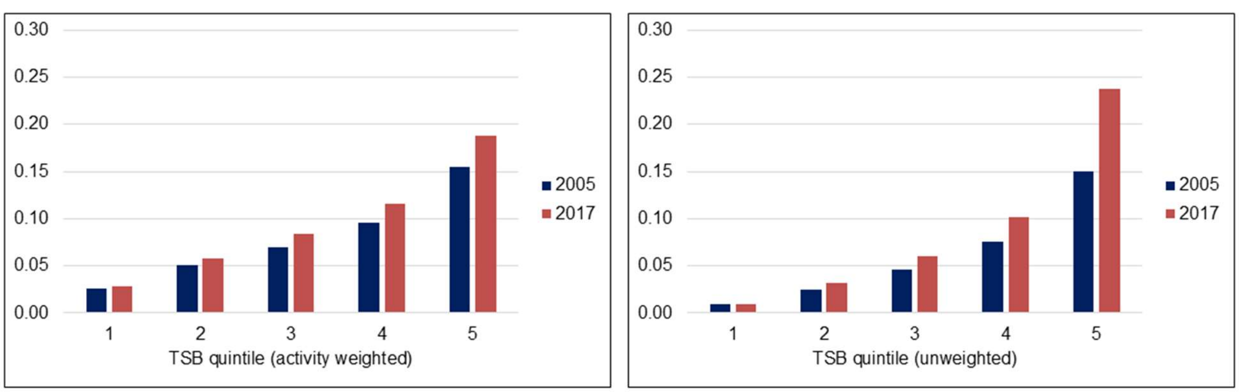
A. Production worker



B. STEM worker



C. Management

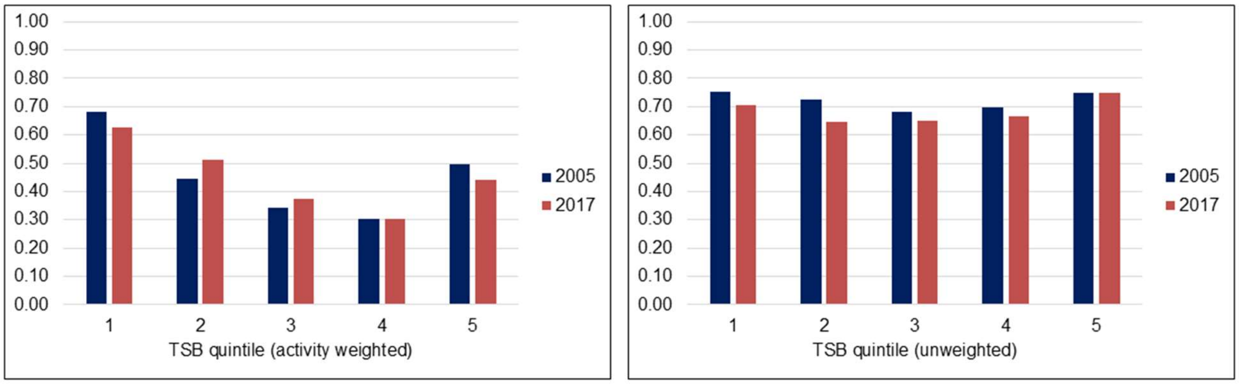


Source: OEWS survey, authors' calculations

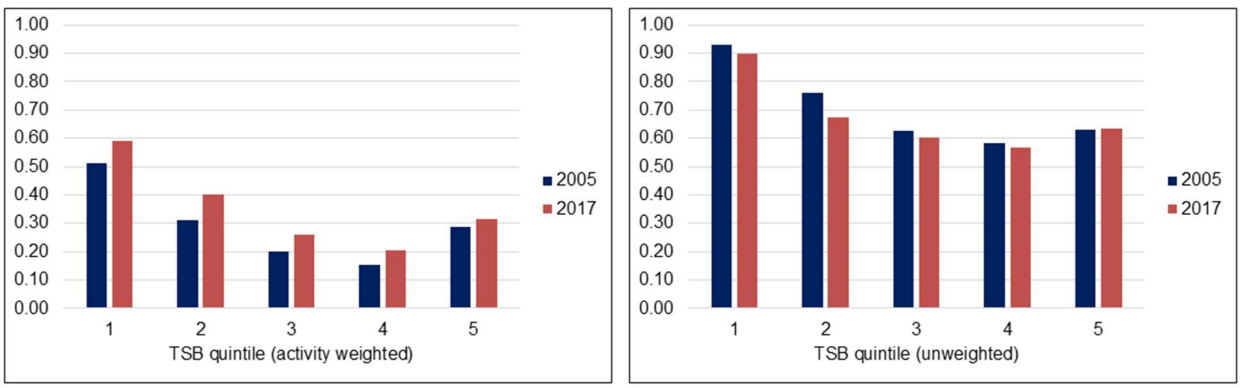
[26] The sample design of the OEWS is geared toward measuring employment, which complicates the construction of establishment weights.

Figure 5. Within-industry variation in occupation share of employment by TSB quintile

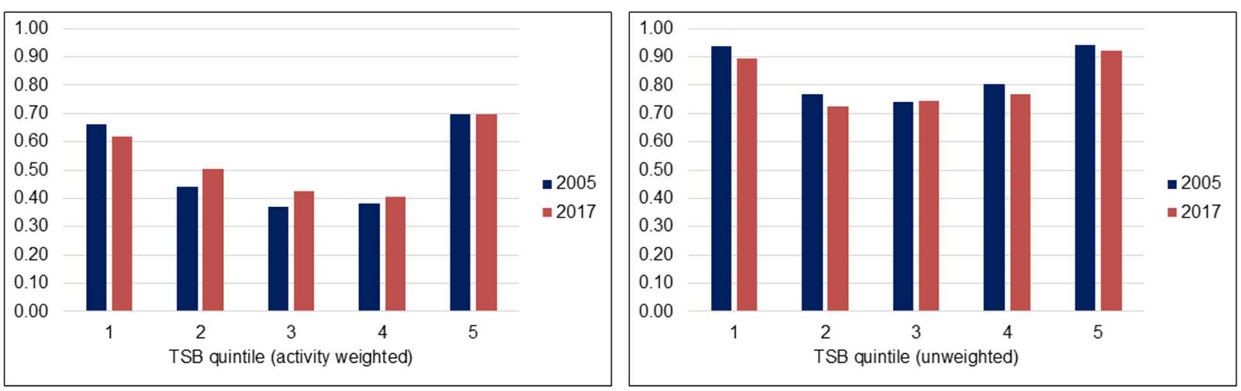
A. Production worker



B. STEM worker



C. Management



Source: OEWS survey, authors' calculations

The six panels in Figure 4 show the mean occupation shares of production workers, STEM workers, and management by quintile on an activity-weighted and unweighted basis. Looking at the activity-weighted graphs on the left, we see that the relationship between the occupational shares and TSB quintiles are as expected, with higher TSB establishments employing relatively fewer (lower-paid) production workers and relatively more (higher-paid) STEM and management employees (note the difference in scales for the STEM and management graphs). Comparing these graphs to the unweighted graphs on the right provides insight into the difference between large and small establishments. The higher employment shares of STEM and management employees in the weighted graphs implies that, within TSB quintiles, larger establishments employ relatively more of these workers.

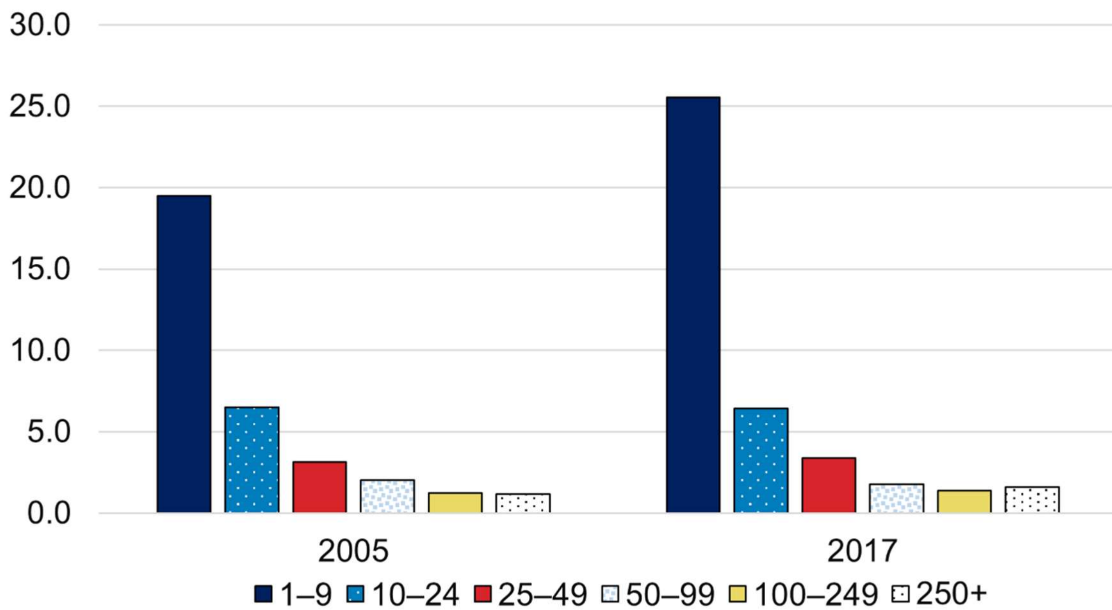
As noted in Blackwood et al. (2025), there are many occupational distributions that are consistent with a given level of TSB. We would expect there to be between-industry variation in employment shares. But before turning to the results that incorporate OEWS data into the dispersion statistics, we take a closer look at occupational differences across establishments using these unlinked OEWS data.

The graphs in Figure 5 show the within-industry share of total variation in employment shares for the three occupation groups.[27] Starting with the activity-weighted graphs, for all three occupation groups, the within-industry share of total variation is greatest in the first quintile—from around 50 percent for STEM occupations to nearly 70 percent for production workers. All three occupation groups exhibit a general U-shape, with troughs around the third and fourth quintiles. The within-industry share of variation in employment shares is lowest for STEM occupations, which should not be too surprising given that there is a lot of industry variation in technology intensity. For all

[27] Within each TSB quintile, we regress the occupation share on industry dummy variables. The figures show values of $(1 - R^2)$ from each regression, which are the within-industry shares of total variation in employment shares.

three occupation groups, the within-industry share of total variation in employment shares is greater in the unweighted graphs, again reflecting the difference between large and small establishments. That the within-industry share is greater in the unweighted graphs suggests that much of the within-industry variation is due to small establishments.

Figure 6. Percent of establishments with zero production workers by employment size class



Notes: Estimates are unweighted. Source: OEWS survey, authors' calculations

Together, these figures point to significant heterogeneity in how establishments organize production—even using these broad occupation groups. As expected, there is significant variation by TSB quintile and across industries, but there is also significant variation in employment shares within industries. Thus, establishments in the same industry do things very differently.

Table 4. Percent of establishments with zero production workers for 13 four-digit industries with the highest percent (SOC definitions)

Industry	25–99 workers (2005)	25–99 workers (2017)	100+ workers (2005)	100+ workers (2017)
Pharmaceuticals	14.1	6.5	1.1	2.2
Machinery	2.2	6.7	4.7	4.1
Computer	17.2	25.3	25.4	22.0
Telecommunications equipment	17.1	17.9	7.1	13.0
AV equipment	7.1	7.4	0.0	7.8
Semiconductor	8.7	4.5	1.3	6.1
Instruments	4.6	9.7	1.7	8.4
Magnetic & optical equipment	28.1	9.5	16.5	38.2
Aerospace	4.6	9.2	3.7	4.3
Dairy	5.3	2.5	0.0	0.0
Tobacco	11.4	12.1	0.0	0.0
Concrete/cement	11.9	13.4	4.5	3.4
Autos/trucks	12.5	0.0	3.9	0.0

Source: OEWS survey, authors' calculations.

Given the apparent differences between large and small establishments with respect to the employment share of production workers, we investigate these shares along a different dimension—the percent of establishments that have zero production workers by size category. Establishments with zero production workers are an interesting extreme version of a manufacturing establishment—consistent potentially with the so-called rise of "factoryless establishments."[28] While further research is required, these are establishments that potentially have automated away their production workers. Moreover, with these fractions varying across establishments within the same industry this is a potentially important indicator of differences in ways of doing business across establishments in the same industry.

[28] Bernard and Fort (2015) examined factoryless goods producers using Wholesale Trade data. They identify three types of production-related activities: pre-production, production, and post-production.

Figure 6 shows the percent of establishments that have zero production workers. Not surprisingly, the largest percent is among smaller establishments. The most likely explanation is that production employees may not be coded as production because they do other activities in these very small establishments. Still, even among larger establishments (25+ employees), there is a non-trivial percent that employs no production workers.[29]

Finally, Table 4 shows the percent of establishments with zero production workers for the 13 industries with the highest percent. To avoid issues with occupation distinctions being less well defined in small establishments, we focus on establishments with 25+ employees. Most of these industries are high-tech, although there are some non-tech industries as well (bottom four rows). Among high-tech industries, some of the highest numbers show up in industries such as Computers, Telecom Equipment, and Magnetic & Optical Equipment. The numbers are quite large, and in some cases changed substantially between 2005 and 2017.

D. A More General Accounting for TFP Dispersion

We now return to examining the relationship between occupations, tasks, skills, and productivity. The analysis above suggests a simple (log) linear relationship between TFP and TSB within industries is inadequate. The objective in this section is to determine how much within-industry (log) TFP variation can be explained by these industry-specific nonlinear relationships. We do this using standard regression analysis focusing on the adjusted R-squared as a metric. This analysis is descriptive and not

[29] We plan to explore this avenue in the next iteration of this paper—for example, recreating Figures 4 and 5 excluding the very smallest establishments.

causal as TFP and the measures of skills/tasks and occupations are all endogenous. Still, this analysis provides insights into how indicators of businesses “doing business differently” (through their skill/task and occupational share differences) accounts for dispersion in measured (log) TFP. We examine this using PW and AW. In all cases, the dependent and explanatory variables have had industry-year effects swept out.

Tables 5A and 5B present results for all industries. We start out by considering each skill, task, and occupation measure separately in a linear fashion. We then allow the coefficient to vary by industry, and then to consider specifications with groups of the skill, task and occupation measures (with and without industry interactions). Finally, we present results from a “full” specification where all measures are included with both linear, quadratic, and cross terms all of which are permitted to vary by industry. The results using PW and AW are in Table 5A and 5B, respectively.

Several interesting patterns emerge. First, without industry interactions, bivariate relationships are statistically significant with expected signs but with virtually no explanatory power. For these specifications, results using AW yield more explanatory power. The bivariate linear results are consistent with our simple decomposition exercise above yielding relatively little explanatory power. Second, when we permit industry-specific relationships even maintaining (log) linear relationships, explanatory power rises notably—especially when we use AW. Third, the “full” specification (last line) yields an adjusted R-squared of 0.16 for the PW case and 0.22 in the AW case. We are also struck by how much the explanatory power increases with the addition of the broad occupational shares (production worker, STEM share, and management share) even after controlling for the TSB, TSU and O*NET measures (the second to last line). Although our final specification is far from parsimonious, the changes in the adjusted R-squared as we add more variables and interactions provide some guidance about what matters in explaining productivity dispersion.

Table 5A. **The relationship between the distribution of productivity (TFP), occupations, tasks, and skills (all industries), propensity weighted**

		Coef.	SE	Adj. R^2	Adj. R^2 (indINT)
Univariate	TSU	0.133	0.022	0.0009	0.0116
	TSB	0.155	0.018	0.0021	0.0110
	Routine man.	-0.037	0.005	0.0016	0.0105
	Routine cog.	-0.037	0.006	0.0008	0.0075
	Physical	-0.046	0.006	0.0015	0.0152
	Interpersonal	0.035	0.007	0.0007	0.0072
	Analytical	0.054	0.008	0.0012	0.0102
	Production share	-0.058	0.011	0.0007	0.0183
	STEM share	0.115	0.040	0.0003	0.0091
	Management share	0.084	0.034	0.0002	0.0058
Multivariate	TSB + TSU + O*NET			0.0027	0.0512
	TSB + TSU + O*NET + occ. shares			0.0032	0.066
	Polynomial(TSB + TSU + O*NET)				0.1109
	Polynomial(TSB + TSU + O*NET + occ. shares)				0.1603

Notes: N=333,000. Dependent variable is log (TFP) demeaned by industry and year. All task/occupation measures demeaned by industry and year. The columns titled "Adj. R^2 (indINT)" refer to regressions in which the explanatory variables are interacted with four-digit industry fixed effects.

Table 5B. **The relationship between the distribution of productivity (TFP), occupations, tasks, and skills (all industries), activity weighted**

		Coef.	SE	Adj. R^2	Adj. R^2 (indINT)
Univariate	TSU	0.575	0.023	0.0155	0.0975
	TSB	0.446	0.019	0.0161	0.0938
	Routine man.	-0.111	0.005	0.0102	0.0593
	Routine cog.	-0.069	0.006	0.0015	0.0118
	Physical	-0.125	0.006	0.0074	0.0663
	Interpersonal	0.089	0.006	0.0025	0.0484
	Analytical	0.197	0.008	0.0153	0.0950
	Production share	-0.306	0.012	0.0124	0.0819
	STEM share	0.782	0.036	0.0215	0.0855
	Management share	0.513	0.034	0.0041	0.0184
Multivariate	TSB + TSU + O*NET			0.0199	0.1391
	TSB + TSU + O*NET + occ. shares			0.0247	0.1590
	Polynomial(TSB + TSU + O*NET)				0.1797
	Polynomial(TSB + TSU + O*NET + occ. shares)				0.2181

Notes: N=333,000. Dependent variable is log (TFP) demeaned by industry and year. All task/occupation measures demeaned by industry and year. The columns titled "Adj. R^2 (indINT)" refer to regressions in which the explanatory variables are interacted with four-digit industry fixed effects.

We investigate these patterns further by re-estimating these regressions separately for establishments in high-tech and low-tech industries (Tables 6A, 6B, 7A and 7B).[30] The skill, task and occupation shares account for considerably more variation in the high-tech industries. Even without nonlinearities or multivariate specifications, individual components (e.g., TSB and STEM share) yield an adjusted R-squared of about 0.20 in the activity-weighted specifications with nonlinearities that the adjusted R-squared rises above 0.10. Still, in the "full" specification, the adjusted R-squared is 0.15 in the

[30] The high-tech group contains the following four-digit NAICS codes: 3241, 3251, 3252, 3254, 3332, 3333, 3336, 3339, 3341, 3342, 3343, 3344, 3345, 3346, 3353, and 3364. All other industries are classified as low-tech.

specification using PW and 0.23 in the specification using AW. with industry interactions. With the “full” specification, the adjusted R-squared for high-tech is 0.35. For the low-tech industries, it is only in the specifications with nonlinearities that the adjusted R-squared rises above 0.10. Still, in the “full” specification, the adjusted R-squared is 0.15 in the specification using PW and 0.23 in the specification using AW.

Table 6A. **The relationship between the distribution of productivity (TFP), occupations, tasks, and skills (high-tech industries), propensity weighted**

		Coef.	SE	Adj. R^2	Adj. R^2 (indINT)
Univariate	TSU	0.330	0.063	0.0049	0.0235
	TSB	0.303	0.043	0.0086	0.0236
	Routine man.	-0.078	0.013	0.006	0.0167
	Routine cog.	-0.087	0.021	0.0033	0.0074
	Physical	-0.061	0.015	0.0026	0.0159
	Interpersonal	0.046	0.022	0.001	0.0074
	Analytical	0.110	0.021	0.0047	0.0202
	Production share	-0.162	0.028	0.0054	0.0273
	STEM share	0.158	0.064	0.0012	0.0216
	Management share	0.038	0.082	0.000	0.0117
Multivariate	TSB + TSU + O*NET			0.0106	0.0596
	TSB + TSU + O*NET+ occ. shares			0.016	0.0864
	Polynomial(TSB + TSU + O*NET)				0.1347
	Polynomial(TSB + TSU + O*NET + occ. shares)				0.2073

Notes: N=49,500. Dependent variable is log (TFP) demeaned by industry and year. All task/occupation measures demeaned by industry and year. The columns titled “Adj. R^2 (indINT)” refer to regressions in which the explanatory variables are interacted with four-digit industry fixed effects. The high-tech group contains the following four-digit NAICS codes: 3241, 3251, 3252, 3254, 3332, 3333, 3336, 3339, 3341, 3342, 3343, 3344, 3345, 3346, 3353, and 3364. All other industries are classified as low-tech.

Table 6B. **The relationship between the distribution of productivity (TFP), occupations, tasks, and skills (high-tech industries), activity weighted**

		Coef.	SE	Adj. R^2	Adj. R^2 (indINT)
Univariate	TSU	1.088	0.076	0.0472	0.2082
	TSB	0.715	0.038	0.0473	0.1995
	Routine man.	-0.203	0.012	0.0276	0.1111
	Routine cog.	-0.216	0.024	0.0066	0.0137
	Physical	-0.229	0.016	0.0220	0.1244
	Interpersonal	0.301	0.022	0.0134	0.1070
	Analytical	0.355	0.019	0.0450	0.2074
	Production share	-0.663	0.032	0.0456	0.1640
	STEM share	0.910	0.049	0.0502	0.1814
	Management share	0.807	0.084	0.0093	0.0343
Multivariate	TSB + TSU + O*NET			0.0570	0.2540
	TSB + TSU + O*NET + occ. shares			0.0683	0.2798
	Polynomial(TSB + TSU + O*NET)				0.3046
	Polynomial(TSB + TSU + O*NET + occ. shares)				0.3531

Notes: N=49,500. Dependent variable is log (TFP) demeaned by industry and year. All task/occupation measures demeaned by industry and year. The columns titled “Adj. R^2 (indINT)” refer to regressions in which the explanatory variables are interacted with four-digit industry fixed effects. The high-tech group contains the following four-digit NAICS codes: 3241, 3251, 3252, 3254, 3332, 3333, 3336, 3339, 3341, 3342, 3343, 3344, 3345, 3346, 3353, and 3364. All other industries are classified as low-tech.

Table 7A. The relationship between the distribution of productivity (TFP), occupations, tasks, and skills (low-tech industries), propensity weighted

		Coef.	SE	Adj. R^2	Adj. R^2 (indINT)
Univariate	TSU	0.080	0.021	0.0003	0.0078
	TSB	0.103	0.019	0.0009	0.0071
	Routine man.	-0.026	0.005	0.0008	0.0085
	Routine cog.	-0.026	0.006	0.0004	0.0075
	Physical	-0.042	0.007	0.0013	0.0149
	Interpersonal	0.032	0.007	0.0006	0.0072
	Analytical	0.038	0.008	0.0006	0.007
	Production share	-0.028	0.011	0.0002	0.0155
	STEM share	0.061	0.042	0	0.0052
	Management share	0.097	0.037	0.0003	0.0039
Multivariate	TSB + TSU + O*NET			0.0016	0.0485
	TSB + TSU + O*NET +occ. shares			0.0019	0.0595
	Polynomial(TSB + TSU + O*NET)				0.1032
	Polynomial(TSB + TSU + O*NET + occ. shares)				0.145

Notes: N=283,500. Dependent variable is log (TFP) demeaned by industry and year. All task/occupation measures demeaned by industry and year. The columns titled "Adj. R^2 (indINT)" refer to regressions in which the explanatory variables are interacted with four-digit industry fixed effects. The high-tech group contains the following four-digit NAICS codes: 3241, 3251, 3252, 3254, 3332, 3333, 3336, 3339, 3341, 3342, 3343, 3344, 3345, 3346, 3353, and 3364. All other industries are classified as low-tech.

Table 7B. **The relationship between the distribution of productivity (TFP), occupations, tasks, and skills (low-tech industries), activity weighted**

		Coef.	SE	Adj. R^2	Adj. R^2 (indINT)
Univariate	TSU	0.233	0.019	0.0013	0.0173
	TSB	0.617	0.016	0.0135	0.039
	Routine man.	0.056	0.005	0.0014	0.0288
	Routine cog.	-0.049	0.006	0.0005	0.0365
	Physical	0.110	0.007	0.0029	0.0315
	Interpersonal	0.121	0.006	0.0031	0.0197
	Analytical	0.101	0.007	0.0021	0.0163
	Production share	-0.072	0.011	0.0004	0.0323
	STEM share	1.220	0.034	0.0132	0.0399
	Management share	0.482	0.033	0.0018	0.0141
Multivariate	TSB + TSU + O*NET			0.0431	0.1345
	TSB + TSU + O*NET + occ. shares			0.0548	0.1599
	Polynomial(TSB + TSU + O*NET)				0.178
	Polynomial(TSB + TSU + O*NET + occ. shares)				0.2283

Notes: N=283,500. Dependent variable is log (TFP) demeaned by industry and year. All task/occupation measures demeaned by industry and year. The columns titled "Adj. R^2 (indINT)" refer to regressions in which the explanatory variables are interacted with four-digit industry fixed effects. The high-tech group contains the following four-digit NAICS codes: 3241, 3251, 3252, 3254, 3332, 3333, 3336, 3339, 3341, 3342, 3343, 3344, 3345, 3346, 3353, and 3364. All other industries are classified as low-tech.

The analysis in this section combined with the prior sections highlights the enormous between-establishment, within-industry variation in measures of skill, task and occupations that are correlated with the between-establishment, within-industry variation in TFP. The fraction of the variance accounted for by skill, task, and occupation variation is higher for larger establishments (based on the activity-weighted results). This may reflect genuine differences in the role of these differences for large establishments, but it may also reflect that large establishments have more defined differences in these dimensions. An important feature of the findings is that the relationship of these skill, task, and occupation measures to TFP varies substantially by industry and is nonlinear. Moreover, these relationships are more pronounced in high-tech industries.

V. Concluding remarks

Measured productivity differences among establishments are ubiquitous. Apart from true differences in efficiency, measured dispersion can be due to unobserved differences in organizational characteristics, input responsiveness, markups, measurement error, and production function specification. In addition, unobserved differences in inputs—for example, capital and labor characteristics and/or composition—are also subsumed in the productivity residual.

In this paper, we examine the relationship between establishment-level TFP within industries and measures of the skill, task, and occupational differences across these same establishments. We construct establishment-level measures of occupation composition, skill, and task intensity using data from the OEWS survey and the O*NET, and then link these to the CMP establishment-level productivity data. We match the OEWS survey data and CMP data using a hierarchical algorithm that prioritizes information on EINs, narrowly-defined industry, and geography—a significant challenge given differences in establishment identifiers across the two data sources.

Our empirical results indicate that there are highly nonlinear, industry-specific relationships between measured TFP and the skill, task, and occupational measures within industries. It is evident that businesses are organized differently via the information on the skill, task, and occupational differences. These organizational differences are related in a complex manner with measured TFP variation.

The standard approach to measure and study within-industry variation in productivity between establishments is to assume that all establishments use the same production technology. The benchmark is often a Cobb-Douglas specification with industry-level factor elasticities. Our results highlight that this approach is inadequate. Understanding measured productivity variation across establishments within industries requires understanding how businesses organize themselves differently in terms of their mix of skills, tasks, and occupations. There are many challenges going forward.

First, it will be important to understand the driving forces underlying these differences. A potentially fruitful next step will be to integrate measures of technology adoption (e.g., automation) into the linked CMP-OEWS data. This integration will allow us to explore how these observable differences are related to the skill/task/occupation differences we identify. This would permit us to begin exploring the hypotheses in the work of Acemoglu and Restrepo (2018a, 2018b, 2019a, 2019b, 2020) that emphasizes the complex relationship between automation, tasks and factors of production such as capital and labor.

Second, there is a large literature studying the creative destruction process using establishment-level differences in measured productivity. Consistent with canonical models of firm dynamics, measured productivity differences across establishments are closely connected to growth and survival dynamics. The latter patterns indicate that measured productivity differences have a systematic relationship with key outcomes. However, the differences in organizational structure in terms of skills, tasks, and occupations raise questions about how to think about these well-documented connections between productivity and reallocation dynamics. Third, and relatedly, our results suggest progress needs to be made to specify parsimonious but much richer production specifications for studying establishment-level variation in measured productivity.

This is an ambitious to-do list and this paper has only taken initial steps. The insights from these first steps suggest that the integrated CMP-OEWS data infrastructure has great potential for making progress on these issues. The research and development activities that led to the integrated CMP-OEWS were possible through the continued collaboration between the two statistical agencies. We hope this paper also serves to highlight the benefits of this collaboration.

References

Acemoglu, Daron and David Autor. 2011. “Skills, Tasks and Technologies: Implications for Employment and Earnings.” In *Handbook of Labor Economics*, vol. 4, Part B, edited by Orley Ashenfelter and David Card, 1043–171. Elsevier.

Acemoglu, Daron, and Pascual Restrepo. 2018a. “Modeling Automation.” *AEA Papers and Proceedings* 108: 48–53.

Acemoglu, Daron, and Pascual Restrepo. 2018b. “The Race Between Machine and Man: Implications of Technology for Growth, Factor Shares, and Employment.” *American Economic Review* 108 (6): 1488–542.

Acemoglu, Daron, and Pascual Restrepo. 2019a. “Artificial Intelligence, Automation and Work.” In *The Economics of Artificial Intelligence: An Agenda*, edited by Ajay Agrawal, Joshua Gans, and Avi Goldfarb, 197–236.

Acemoglu, Daron and Pascual Restrepo. 2019b. “Automation and New Tasks: How Technology Displaces and Reinstates Labor.” *Journal of Economic Perspectives* 33 (2): 3–30.

Acemoglu, Daron, and Pascual Restrepo. 2020. “Robots and Jobs: Evidence from US Labor Markets.” *Journal of Political Economy* 128 (6): 2188–244.

Autor, David H., Frank Levy, and Richard J. Murnane. 2003. “The Skill Content of Recent Technological Change: An Empirical Exploration.” *The Quarterly Journal of Economics* 118 (4): 1279–333.

Bernard, Andrew B. and Teresa C. Fort. 2013. “Factoryless Goods Producing Firm,” American Economic Review: Papers & Proceedings 105(5): 518–523.

Blackwood, G. Jacob, Lucia S. Foster, Cheryl A. Grim, John Haltiwanger, and Zoltan Wolf. 2021. "Macro and Micro Dynamics of Productivity: From Devilish Details to Insights." *American Economic Journal: Macroeconomics* 13 (3): 142–72.

Blackwood, G. Jacob, Cindy Cunningham, Matthew Dey, Lucia S. Foster, Cheryl A. Grim, John Haltiwanger, Rachel Nesbit, Sabrina Wulff Pabilonia, Jay Stewart, Cody Tuttle, and Zoltan Wolf. 2025. “Opening the Black Box: Task and Skill Mix and Productivity Dispersion.” In *Technology, Productivity, and Economic Growth*, edited by Susanto Basu, Lucy Eldridge, John Haltiwanger, and Erich Strassner, NBER.

Chow, Melissa, Teresa C. Fort, Christopher Goetz, Nathan Goldschlag, James Lawrence, Elisabeth Ruth Perlman, Martha Stinson, and T. Kirk White. 2021. “Redesigning the Longitudinal Business Database.” Center for Economic Studies Discussion Paper No. 21-08. https://www.census.gov/library/working-papers/2021/adrm/CES-WP-21-08.html.

Cunningham, Cindy, Lucia Foster, Cheryl Grim, John Haltiwanger, Sabrina Wulff Pabilonia, Jay Stewart, and Zoltan Wolf. 2023. “Dispersion in Dispersion: Measuring Establishment Level Differences in Productivity.” *Review of Income and Wealth* 69 (4): 999–1032. https://doi.org/10.1111/roiw.12616.

Dey, Matthew, David S. Piccone, Jr., and Stephen M. Miller. 2019. "Model-Based Estimates for the Occupational Employment Statistics Program." *Monthly Labor Review* August.

Employment and Training Administration, U.S. Department of Labor. 1991. *The Revised Handbook of Analyzing Jobs*.

Handel, Michael. 2016. "The O*NET Content Model: Strengths and Limitations." *Journal of Labor Market Research* 49:157–76.

Iranzo, Susana, Fabiano Schivardi, and Elisa Tosetti. 2008. "Skill Dispersion and Firm Productivity: An Analysis with Employer-Employee Matched Data." *Journal of Labor Economics* 26 (2): 247–85. https://doi.org/10.1086/587091.

Stoyanov, Andrey and Zubanov, Nick. 2022. "Skill Complementarity in Production Technology: New Empirical Evidence and Implications" *German Economic Review* 23 (2): 233–74. https://doi.org/10.1515/ger-2020-0102

Syverson, Chad. 2004. "Market Structure and Productivity: A Concrete Example", *Journal of Political Economy*, 112(6): 1181-1222

Syverson, Chad. 2011. "What Determines Productivity?" *Journal of Economic Literature* 49 (2): 326–65.

Zoghi, Cindy. 2007. "Measuring Labor Composition: A Comparison of Alternate Methodologies." In *Labor in the New Economy*, edited by Katherine G. Abraham, James R. Spletzer, and Michael Harper, University of Chicago Press.

A. Data Appendix

Details on CM-ASM Industry Codes

As noted in the main text, an establishment's industry code is an integral part of the matching procedure, as detailed below, and it is therefore important to have the best possible chance of matching an establishment's NAICS code to the code that the BLS would assign. Accordingly, we use four different NAICS codes for our CMP dataset: the LBD code, the ASM code (only available for manufacturing observations), the VC code, and a "combined" code created by combining information from the CM, ASM, and LBD. We create our combined code as follows: (1) for establishments that are surveyed by the CM, we use the industry code from the CM year that is closest to the reference year; (2) if no CM code is available, we use the ASM code; and (3) if no ASM code is available, we use the LBD code. We think this combined code most closely aligns with the timing of industry code updates in the OEWS survey.[31] Finally, we make a "time-consistent code" correction to the LBD, ASM, and combined codes. The correction aggregates six-digit codes to five-digit codes in cases where there is consolidation or other changes in the classification between different NAICS vintages. Note that this correction differs from the VC code approach taken by Chow et al. (2021). Our time-consistent codes do not aim to put everything in terms of the 2017 classification vintage, but instead to simply aggregate any codes that disappear or are broken up between vintages so that we can abstract from vintage differences. We describe further below how we use these seven versions (ASM, time-consistent ASM, LBD, time-consistent

[31] The OEWS occasionally updates industry codes based on the information collected from each establishment's answers to the survey. When this occurs, the OEWS industry code will differ from that in the Quarterly Census of Employment and Wages (QCEW), which is the BLS business register.

LBD, combined, time-consistent combined, and VC) of NAICS codes in our matching procedure.

Matching Procedure Example

Consider a CMP establishment within a given EIN and suppose that there are twelve potential donors in the OEWS that have the same EIN. In step 1, we start by using the combined NAICS code. If none of the twelve candidates is in the same six-digit industry code as defined by our combined NAICS measure, we then check for agreement using the six-digit industry ASM NAICS code. And if there is still no match, we use the LBD NAICS code. If there is still no agreement, we move to step 2, allowing for possible mismatches in industry vintage changes by using instead the time-consistent combined, ASM, and LBD NAICS in a similar iterative manner. The algorithm continues through step 5 or until a match is found. If there is no match in step 5, the observation is not used.

In this example, suppose we identify three candidate OEWS donors that match in Step 1. These are very high-quality matches, but we need to narrow them down to one final donor. Among these three candidates, we first look for the one most similar in size to the CMP establishment. Suppose that eliminates one potential donor, but the two other candidates have the same employment. We next compare the years in which those donors were surveyed. If the CMP year were 2013, we would be evaluating potential matches from the 2012, 2013, and 2014 OEWS surveys. Suppose both donors were surveyed in 2013, so that this tiebreaker does not help us narrow down our candidates. The final step would be to randomly choose one of the two remaining candidates to be our preferred donor. By following our step-by-step matching and tiebreaking processes, we have identified one match out of the original twelve candidates.

*The O*NET Data*

The O*NET data are collected from workers in targeted occupations at establishments and contain over 275 variables that describe each occupation. The O*NET database is sponsored by the Employment and Training Administration of the Department of Labor and is collected through the National Center for O*NET Development and the Research Triangle Institute. O*NET first began surveying job holders in 2001. Prior to that, past Dictionary of Occupational Titles data, collected sometimes decades earlier by job analysts visiting workplaces, were recoded into O*NET variables. Because new surveying was rolled in gradually, the first O*NET completely based on surveys was released in 2008. O*NET re-surveys occupations on a rolling basis over a five-year period. The number of respondents per occupation varies, and respondents are randomly selected to answer a subset of the questionnaire. The value of a particular O*NET variable is the average response over the jobholders who answered that question, so within-occupation variation cannot be observed. See Handel (2016) for more about the history of O*NET as well as its strengths and weaknesses.

Acemoglu and Autor (2011) use 16 of these variables corresponding to the five task categorizations described in the text: non-routine cognitive (analytical), non-routine (interpersonal), routine cognitive, routine manual, and non-routine manual physical. Non-routine cognitive (analytical) includes analyzing data/information, thinking creatively, and interpreting information for others. Non-routine cognitive (interpersonal) includes establishing and maintaining personal relationships; guiding, directing, and motivating subordinates; and coaching/developing others. Routine cognitive includes importance of repeating the same tasks, importance of being exact or accurate, and structured vs. unstructured work (reverse). Routine manual includes tasks where the pace of work is determined by speed of equipment, controlling machines and processes, and tasks requiring repetitive motions. Non-routine manual physical includes operating vehicles, mechanized devices, or equipment; tasks where workers use their hands to handle, control, or feel objects, tools, or controls; manual dexterity; and spatial orientation. (See page 1163 of Acemoglu and Autor (2011).) The O*NET-SOC occupational categories are aggregated to SOC categories, and each variable is scaled and then standardized to mean zero and standard deviation one using employment weights from the OEWS survey. The five indexes are created by summing the standardized variables for each task category, which are then once again normalized.

The OEWS Research Dataset

Because the sample design of the OEWS is geared toward measuring employment, creating establishment weights is complicated. As an alternative to reweighting, we use a research dataset that is a modified version of the dataset developed by Dey, Piccone, and Miller (2019). This research dataset supplements the OEWS data by imputing occupation data for the entire Quarterly Census of Employment and Wages (QCEW), which is the BLS business register and is the sample frame for BLS establishment surveys. The main advantage of this approach is that all establishments are represented and have a weight of one.

The imputation process involves two stages, a matching stage where potential donors are identified and a selection stage where the best donor is selected. The process is hierarchical, where the conditions for finding acceptable matches are sequentially relaxed. At the most detailed level of the hierarchy, a donor and frame unit will match on industry (six-digit NAICS), ownership (private or type of government), state, and county and will have similar employment levels. As the process continues through the hierarchy, geography is relaxed first and then ownership. It is not until late in the process, after most of the frame units have already found an acceptable donor, that industry and employment proximity are relaxed. The matching stage often results in multiple potential donors. To preserve variance, the selection of a particular donor from the set of acceptable matches is random. Wages are adjusted to account for differences by MSA and industry. In contrast to the published statistics, the research dataset centers the sample on the reference year instead of using data from the five panels prior to May of the reference year. For example, under this approach, the sample for May 2017 is constructed using data from the following panels: May 2018, November 2017, May 2017, November 2016, May 2016, and November 2015. This results in a nationally representative sample centered on May 2017. To avoid overlap, these “year samples” are constructed at three-year intervals. This effectively assumes the occupational mix within an establishment is fixed over the three-year interval.