IARIW 2024

IARIW 2024

Thursday, August 22 – Friday, August 30

Retail Productivity Dispersion: 1987-2017

Dominic A Smith
(U.S. Bureau of Labor Statistics)
G. Jacob Blackwood
(Amherst College)
Michael Giandrea
(U.S. Bureau of Labor Statistics)
Cheryl Grim
(U.S. Census Burea)
Jay Stewart
(U.S. Bureau of Labor Statistics)
Zoltan Wolf
(New Light Technologies)

Paper prepared for the 38th IARIW General Conference
August 26-30, 2024

Session 2A-1, Output and Productivity Measures in Dynamic Industries I

Time: Friday, August 30, 2024 [14:00- 15:30 GMT]

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 30, 2024

Abstract

The theoretical literature on differences in production technologies across businesses increasingly emphasizes the task content of production. At the aggregate level, accounting for changes in labor composition reduces measured total factor productivity growth and illustrates the contribution of changes in workers' skills to output growth. At the establishment level, we expect that accounting for worker skills and tasks will better reflect differences in how inputs are used in production. This paper advances our understanding of the relationships between job tasks, workers' skills, and productivity by matching occupation data from the Bureau of Labor Statistics Occupational Employment and Wage Statistics survey to productivity data from Census Bureau manufacturing surveys to examine the impact of incorporating task/skill intensity measures into measures of productivity. Our findings indicate that standard productivity indicators are correlated with task/skill measures. However, these correlations depend on the task/skill measure, vary across industries, and depend nonlinearly on task/skill levels. This is intuitive because different types of establishments may have different production technologies that comprise of different tasks, which in turn require differently skilled labor.

JEL codes: D24, J24

Keywords: productivity dispersion, tasks, skills

*Blackwood: Amherst College; Cunningham, Dey, Pabilonia, and Stewart: U.S. Bureau of Labor Statistics; Foster and Grim: Center for Economic Studies, U.S. Census Bureau; Haltiwanger: University of Maryland; Nesbit: University of Maryland and Center for Economic Studies, U.S. Census Bureau; Tuttle: University of Texas at Austin; Wolf: New Light Technologies. Blackwood, Haltiwanger, and Tuttle were also Schedule A 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). Corresponding author: Stewart.Jay@bls.gov

I. Introduction

It is well known that productivity varies across establishments, even within detailed industries. For example, using publicly available productivity dispersion statistics from the Dispersion Statistics on Productivity (DiSP), Cunningham et al. (2023) find that on average an establishment at the 90th percentile of the total factor productivity (TFP) distribution is about 2.9 times as productive as an establishment at the 10th percentile within four-digit NAICS manufacturing industries.[1] Other researchers have reported similar results.[2]

Syverson (2011) reviews possible sources of productivity dispersion, including difficult-to-measure factors such as differences in managerial talent and differences in the quality of labor and other inputs. Cunningham et al. (2023) find that establishment-level characteristics from the firm dynamics literature (i.e., state, age class, and size class) have limited explanatory power for productivity dispersion, which suggests a need to look beyond such standard establishment-level characteristics as sources of productivity dispersion. In this paper, we take that step by focusing on one potential source of heterogeneity unobserved in typical micro datasets: establishment-level differences in the characteristics of workers and tasks.

While it is standard to measure labor input using total hours worked by all workers as in the DiSP data, allowing for skill variation could be important for

---

[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).

productivity measurement because differences in measured productivity could reflect differences in the types of workers the establishment employs and the tasks that they perform.[3] With appropriate measures for skill, we can control for these differences in skills and tasks across establishments and highlight their contribution to measured productivity variation. Our goal is to address this measurement question by integrating establishment-level data from the BLS Occupational Employment and Wage Statistics (OEWS) survey and Census Bureau manufacturing survey data.

In an earlier paper (Blackwood et al., 2023), we explored the conceptual, measurement, and specification issues to be addressed for this integration to be successful. We examined the relationship between within-industry dispersion of productivity measures and within-industry dispersion of task/skill measures for four-digit NAICS manufacturing industries over the 2000–2017 period using seven task/skill indexes constructed using data from the OEWS survey and the Occupational Information Network (O*NET).[4]

Blackwood et al. (2023) developed two establishment-level composite measures that summarize information about the distributions of occupations and tasks/skills. One reflects establishment-level differences in the occupation distribution and is labeled a “bundled” task/skill intensity index (TSB), because the pricing of tasks is bundled through occupations. It is related to, but distinct from, the skill-adjusted labor input measure BLS publishes as part of its official TFP measures.[5] The second composite

---

[3] A few empirical studies allow workers’ skill levels to vary. See Iranzo, Schivardi, and Tosetti (2008) and Stoyanov and Zuanov (2022).
[4] Due to restrictions on data sharing at that time, we could not link the two dataset and instead conducted a parallel analysis.
[5] 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).

measure reflects cross-establishment variation in the task/skill content of occupations, based on five aggregate tasks constructed from the work activities and work-context-importance scales in the O*NET (as described in Acemoglu and Autor (2011)). We refer to this index as an “unbundled” task/skill index (TSU) because it prices the tasks directly regardless of which occupations perform these tasks. TSB and TSU both reflect task/skill differences across establishments as well as 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 that TSB 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 TSU prices the tasks individually and ignores any complementarities between tasks within occupations.

Blackwood et al. (2023) compared the within-industry labor productivity (LP) and TFP dispersion measures from DiSP to the within-industry dispersion in these task/skill measures. They find that TSB and TSU are positively correlated across establishments within industries, with the correlation being higher in high-tech manufacturing industries. TSB is also positively correlated with indexes of analytical task content, interpersonal task content, and the percent of employees in STEM occupations (%STEM), but negatively correlated with the non-routine manual physical, routine manual, and routine cognitive task content. These establishment-level correlations confirm our intuition about the skills required to perform these composite tasks. They also find that higher within-industry productivity dispersion is associated with higher within-industry dispersion of TSB, TSU, the analytical task index, and %STEM. The patterns in this parallel analysis strongly suggest that more productive establishments employ highly skilled workers that perform highly-valued tasks.

In this paper, we expand on the analysis of Blackwood et al. (2023) by matching OEWS survey occupation data to ASM productivity data. The resulting establishment-level linked dataset permits us to interact total hours worked with a multiplier that converts labor hours into efficiency units based on our task/skill intensity measures. This approach yields adjusted versions of our productivity measures that can then be compared with standard productivity measures. We find that adjusting for skill intensity in this manner results in only a small reduction in productivity dispersion across establishments in the same industry. However, we find that we can account for a sizeable fraction (5–10 percent) of the productivity dispersion if we make an analogous adjustment using earnings-per-worker differences across establishments.

We further extend the analysis in Blackwood et al. (2023) by estimating establishment-level correlations between standard productivity measures and our measures of occupations, tasks, and skills. Our findings indicate that these relationships vary across industries and are nonlinear in the sense that the impact of occupations/tasks/skills on productivity is larger in high-tech industries and for the most productive establishments. In contrast, such variation appears to be less important in low-tech industries and establishments that are closer to the average of the productivity distribution. This finding provides context for the result on the skill intensity adjustment to productivity. The contribution of specific skills and tasks varies across establishments in the same industry. Taking such between establishment variation within industries into account is apparently needed in assessing the role of skills and tasks in productivity variation.

The paper proceeds as follows. Section II presents a conceptual framework largely through a review of the literature relating productivity to the skills of workers and the tasks they perform. Section III describes how we measure occupations, tasks, and skills. Section IV discusses our data sources and the matching procedure. Section V presents empirical results, and Section VI concludes and provides an overview of next steps.

II. Background and conceptual framework

Our starting point is the general production function specification:

$$Q_{et} = A_{et} \cdot F(L_{et}, K_{et}, M_{et}) \quad (1)$$

where Q_{et} is output, L_{et} is labor input, K_{et} is capital input, M_{et} is intermediate input, A_{et} is a Hicks-neutral productivity term, and *e* and *t* index establishments and time.

An establishment can have higher productivity than its competitors if it uses inputs more efficiently, or if its production process consists of more advanced tasks (generally) accompanied by more skilled labor. A simple way to model task/skill differences across establishments is to introduce a multiplier, Z_{et}, that converts labor hours into efficiency units based on skills and tasks.[6] The first argument of *F*(.) then becomes $Z_{et}L_{et}$. Thus, A_{et} increases the productivity of all factors of production, while

[6] Gollop et al. (1987) first demonstrate the potential importance of using efficiency units of labor. The methods developed from this early study have been widely adopted by statistical agencies around the world (see Schreyer 2001). BLS uses a related approach to measure total factor productivity (see https://www.bls.gov/opub/hom/msp/home.htm).

Z_{et} affects only the productivity of labor. Section III.D describes how we incorporate this Z_{et} term in practice using our matched data.

Other possible approaches explicitly model the different types of labor and/or different tasks and skills employed by firms. We provide a brief summary of these models in Blackwood et al. (2023). In future work, we will draw on that discussion to help map out a potential link between differences in worker types and productivity.

III. Measuring occupations, tasks and skills

Before discussing how we measure tasks and skills in our data, we summarize our interpretation of the basic concepts, relying on the nomenclature from the Revised Handbook of Analyzing Jobs (Employment and Training Administration (1991)) and Acemoglu and Autor (2011). *Tasks* are activities that when combined with capital and intermediate goods create a good or service and are the true factors of production that we would like to measure. However, because we do not observe time spent in different tasks, we use occupations as proxies. An *occupation* is a job in which “a common set of tasks are performed or are related in terms of similar objectives, methodologies, materials, products, worker actions, or workers characteristics” (Employment and Training Administration, 1991, p. 9). Thus, an occupation can be thought of as a bundle of tasks.

In contrast, *skill* refers to “a worker’s endowment of capabilities for performing various tasks” (Acemoglu and Autor (2011), p. 1045). Skill is commonly thought of as a function of education and experience. Operationally, it is often proxied by some measure of wages projected on observable indicators such as education and experience or, alternatively, wages are projected on occupations as in Acemoglu and

Autor (2011, Figure 10). 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.

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

Our first index 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 for each year in the sample. 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 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.

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

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

where E_n is the set of all establishments, and L_{ej} is the number of employees in occupation *j* at establishment *e*. The counterfactual mean log wage for establishment *e*, $\widetilde{w}_e$, can then be written as:

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

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 provides an index of the 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]

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

Our second task/skill intensity index 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 developed a two-dimensional categorization of tasks based on whether they are (1) routine or non-routine and (2) cognitive or manual. Routine tasks are those that can be described using a set of rules or specifications; non-routine tasks are those that cannot be described in this manner. They further break down non-routine cognitive tasks into analytic and interpersonal tasks. This yields five

[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 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.

categories of tasks: non-routine cognitive (analytical), non-routine (interpersonal), routine cognitive, routine manual, and non-routine manual physical.[8]

The O*NET data are collected from workers in targeted occupations at establishments and contain over 275 variables that describe each occupation.[9] Acemoglu and Autor (2011) use 16 of these variables corresponding to the five task categorizations described above.[10] 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.

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,

---

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

[9] 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 DOT 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.

[10] 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).)

2014, 2017).[11] We merge these five task indexes to OEWS survey wage data by occupation and estimate the following regression of the national occupational mean log wage for each year on these five task indexes:

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

where τ_{jk} is the O*NET measure of task *k* for occupation *j*, and $\bar{w}_{nj}$ is defined as in equation (2).[12] 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:

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

where the summation in square brackets is the total amount of task *k* employed by the establishment and $\hat{\beta}_k$ is the "price" of task *k* estimated from the regression in 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 the index.

---

[11] 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.

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

C. Individual Average Task Indexes

In addition to the two task/skill intensity measures based on counterfactual wages, $\widetilde{w}_e$ and $\widehat{w}_e$, we also develop a set of five task measures based on the average value of the individual O*NET task indexes. Recall, we are using five categories of O*NET tasks: non-routine cognitive (analytical), non-routine (interpersonal), routine cognitive, routine manual, and non-routine manual physical. For each of the five task indexes, we measure an employment-weighted establishment-level average for task index *k* as follows:

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

where $k = 1, \ldots, 5$. Thus, $\bar{\tau}_{ek}$ is the average task *k* content of all jobs in establishment *e*. Again, time subscripts are suppressed for expositional convenience. These measures are constructed for each establishment for each year in our sample.

D. Using Task/Skill Intensity to Adjust Productivity

We consider two productivity measures: LP and TFP. Cunningham et al. (2023) detail our approach to creating these measures using combined ASM, CM, and LBD data. In this paper, we explore the impact of allowing for more flexibility in the labor input (L_{et}). We create adjusted versions of our productivity measures by interacting L_{et} (defined as total hours) with Z_{et} (defined as either TSB or TSU) in order to convert labor hours into efficiency units based on skills and tasks.

We first apply a normalization to TSB and TSU so that they have a mean of one in a given industry-year. 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 that incorporates task-skill intensity. We then use adjusted total hours in standard productivity formulas to calculate log LP and log TFP. For example, establishment-level TFP in logs is measured as follows:

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

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* and *L* are as defined above, and α_K, α_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).

IV. 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 create 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), while longitudinal links are established using information from the Longitudinal Business Database (LBD), which is based on the Census Bureau's Business Register (see Chow

[13] Each year, the CMP team releases a new version of this database; the version used for this paper is Version 7, which covers years 1972–2020.

et al. (2021)). The ASM collects data annually and is a five-year panel of manufacturing establishments updated by births in each year.[14] The CM collects data from all manufacturing establishments, except those that are very small, every five years.[15] The LBD provides high-quality longitudinal links and information on the universe of active non-farm private sector employer establishments. 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 and LBD. Because production functions are calculated industry-by-industry, the most relevant are disagreements in industry codes, which can arise for several reasons. For example, because the ASM and CM industry codes are based on the actual survey responses, whereas the LBD codes are updated with a lag, an establishment that changes its industry in a given year might show up in different industries in the LBD, the ASM, and the CM. As a result, there are three potential "raw" industry codes given by this CMP data: the LBD industry code, the CM industry code, and the ASM industry code. A separate but related issue is a consequence of the changes in the NAICS industry classification system over time, and the differential timing of the implementation of those changes within 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, referred to as vintage consistent (VC) industry code hereafter as in Chow et al. (2021).

---

[14] ASM panels start in years ending in "4" and "9."
[15] The CM is collected in years ending in "2" and "7."

This code aims to pick one vintage of NAICS (in this case, the 2017 vintage) and extend that vintage backward so that all establishments during the sample period have an industry code 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.

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 assigning the same NAICS code to an establishment 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.[16] 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

[16] 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.

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

While this paper focuses on manufacturing establishments, for which TFP can be calculated, 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 mailed to approximately 200,000 establishments in May and November of each year.[17] 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 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.[18] The OEWS survey uses the Office of Management and Budget's

[17] 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.
[18] 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

occupational classification system, the SOC, to categorize workers into over 800 detailed occupations. The SOC system provides much more occupational detail 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. This sample design implies that six consecutive panels can be used to create a representative sample that corresponds to any three-year period. 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.[19] We use this property 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, detailed above. 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 firm (EIN) and the industry (the NAICS code) attached to

---

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.

[19] 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.

each establishment. For each establishment in our augmented CMP data, 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.

Because the EIN is a firm-specific identifier, EIN-based matches are 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.[20] 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.[21] As described in Section IV.B, the OEWS survey sample scheme is such that three years of OEWS surveys combined produce a representative sample. Therefore, our approach to matching is to use 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.

[20] 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; this difficulty is discussed in Chow et al. (2021).
[21] 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. The matching criteria are as follows:

(1) EIN, 6-digit industry, state, employment difference less than 0.5
(2) EIN, time-consistent 6-digit industry, state, employment difference less than 0.5
(3) EIN, 6-digit industry, employment difference less than 0.5
(4) EIN, time-consistent 6-digit industry, employment difference less than 0.5
(5) EIN, 4-digit industry, employment difference less than 0.5

As the hierarchy above shows, Step 1 starts with the original six-digit NAICS codes, whereas Step 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 Step 3, we return to our original six-digit codes but relax the geographic requirement, and Step 4 repeats Step 3 but instead uses the time-consistent codes. Finally, Step 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. Appendix B describes an example to illustrate the steps of the procedure.

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. In future work, we will explore ways to increase the number of matched establishments and will examine the robustness of our results with respect to different specifications of the matching algorithm.

D. *Final Analysis Sample*

The matching procedure detailed above yields a total of approximately 328,000 manufacturing observations between 2001 and 2020, all of which have information about the occupation distribution as well as measures of productivity.

We make several additional modifications to form our final analysis sample. First, our analysis going forward is based on the vintage-consistent NAICS code from the LBD.[23] We mainly use this code to remove industry and year effects from all relevant

[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.
[23] The same industry code used to create the publicly available DiSP.

variables by demeaning (removing industry-year effects). Second, we remove observations whose TFP is lower than the 1st or higher than the 99th percentile in an industry-year cell. Third, our current analysis focuses on the manufacturing sector because we can construct total factor productivity for these observations only. Finally, we create inverse propensity score weights to address concerns that our linked data might not form a representative sample of manufacturing establishments. We do so by estimating 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 (IPW).

Table 1. Descriptive statistics of employment in the three datasets

	OEWS	CMP	LINKED	
	(weighted)	(IPW)	(unweighted)	(IPW)
Mean	32.5	53.1	224.5	142
Standard Deviation	167.5	204.8	540.8	389.1

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. IPW refers to inverse propensity weights. CMP refers to the combined ASM, CM, and LBD data.

Table 1 shows descriptive statistics of the establishment size from the OEWS survey, the CMP, and the linked dataset, respectively. 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, much smaller in the CMP data, and slightly smaller still in the OEWS survey data. These patterns can be explained by the different sampling schemes and weights applied.

To get a better sense of the distributional differences across datasets, Table 2 shows the standard deviations of demeaned variables that will be used in the analysis. Employment dispersion after demeaning follows the patterns of those in Table 1. Task/skill variation in the linked sample is smaller relative to the OEWS survey sample size. This could be because we were not able to match all OEWS observations, and the matched observations tend to be larger establishments. For the same reason, variation in CMP measures of productivity, earnings-per-worker, and capital intensity is smaller after matching, although the reduction is more modest than that for the OEWS variables.

Table 2. Standard deviations of key variables in the three datasets

	OEWS	CMP	LINKED	
	(weighted)	(IPW)	(unweighted)	(IPW)
Employment	164.2	198.6	489	355.9
Analytical	0.435		0.2964	0.3052
Interpersonal	0.501		0.3009	0.3207
Physical	0.510		0.3499	0.3653
Routine cognitive	0.501		0.297	0.3136
Routine manual	0.718		0.4495	0.4704
TSU	0.153		0.1025	0.1055
TSB	0.189		0.131	0.1352
Log(TFP)		0.4808	0.474	0.4754
Log(LP)		0.7472	0.6829	0.6893
Log(Earnings-per-Worker)		0.3743	0.3021	0.3173
Log(Capital/Labor)		1.171	0.8939	0.9628

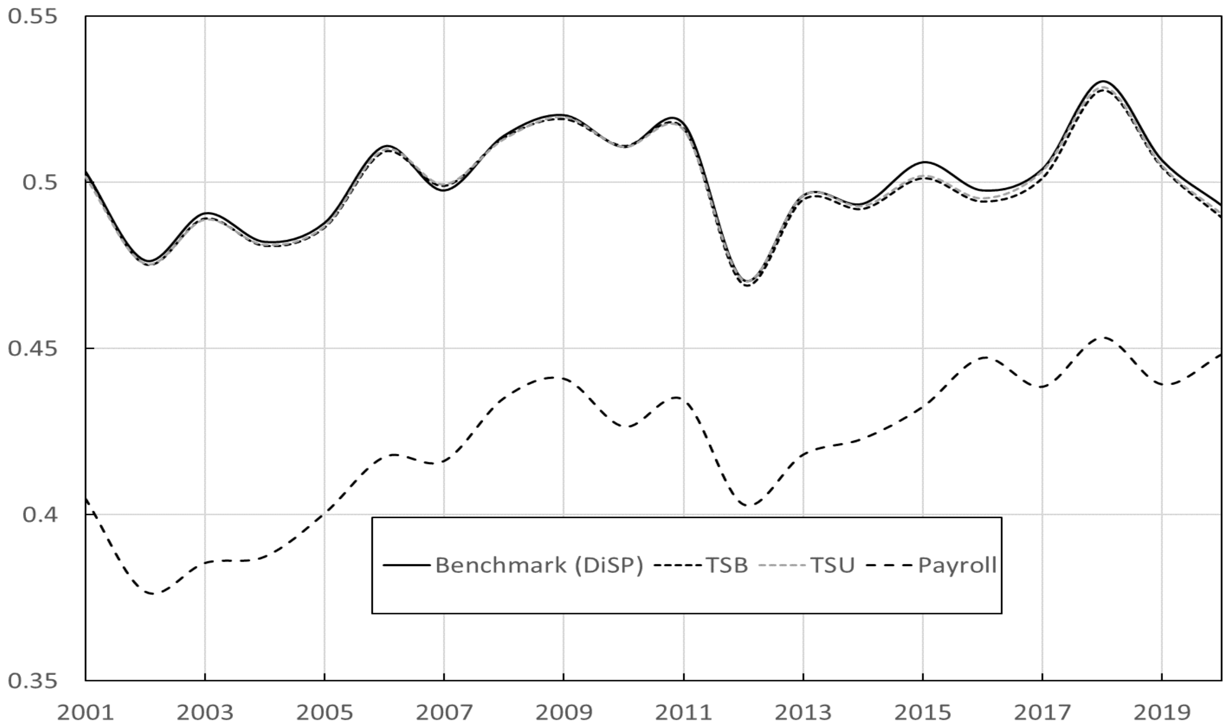
Notes: OEWS survey weights account for the probability of selection, the fact that six panels of data are combined to form the full sample, and differences in employment totals between the sample and the QCEW frame. IPW refers to inverse propensity weights. Industry-year effects are removed. Sample sizes in thousands: 593 (OEWS), 999 (CMP), and 328 (LINKED). CMP refers to the combined ASM, CM, and LBD data.

V. The relationship between productivity, occupations, tasks, and skills

As noted above, linking establishments in the OEWS survey to the ASM allows us to adjust labor input for quality. This translates into modifying equation (1) as discussed in Section II, where TSB or TSU is used to adjust labor input (the Z variable from Section II). We start by examining the impact of this adjustment on the dispersion of the Cobb-Douglas residual in equation (1), measured as the average interquartile range (IQR). The IQR is calculated for four alternative adjustments for tasks and skills: (1) no adjustment (labeled as “Benchmark” and calculated using DiSP methodology); (2) adjusting labor using TSB; (3) adjusting labor using TSU; and (4) using total wages and salaries as the labor measure. The last measure is useful because it controls for between-establishment variation in wages in addition to variation in hours. This adjustment may capture many factors but arguably those factors include differences in worker quality.

Figure 1 shows the time series of the four IQRs. The first implication of this exercise is that converting labor input into efficiency units using the TSB or TSU task/skill measures does not reduce measured dispersion. However, dispersion is 5–10 log points lower when we adjust using payroll (which is essentially adjusting for a measure of establishment-level wages). It is striking that the TSB and TSU measures do not provide much explanatory power whereas adjusting for wages does.

Figure 1. Average interquartile range of log TFP implied by alternative labor measures



Notes: Benchmark (DISP) refers to unadjusted productivity; TSB, TSU, and Payroll refer to productivity calculated using total hours adjusted by the respective measure.

To shed more light on these issues, we now explore the relationship between occupations, tasks/skills, and productivity. To extract basic correlation patterns, we first regress productivity on each of TSB, TSU, and the five O*NET indexes. The partial correlations from these regressions are shown in column (1) of the TFP and LP panels of Table 3. Looking at the left panel, the first entry indicates a positive and statistically significant relationship between TSU and log TFP. The patterns are similar when TSB is used as an explanatory variable. The correlations are negative for routine cognitive, routine manual, and physical skills, while they are positive but closer to zero for analytical and interpersonal skills. All are statistically significant, but none of these skill indexes explain more than 1% of log TFP's variance on their own (column 3 in the left panel of Table 3). The results for LP in the right panel of Table 3 are qualitatively the same, although the magnitudes of the coefficients are larger for most task/skill measures. These findings are broadly consistent with the limited contribution of these measures to productivity dispersion shown in Figure 1.

The explanatory power of these task/skill measures increases by an order of magnitude if the relationships are permitted to vary by industry, i.e., when the explanatory variable is interacted with four-digit industry fixed effects, shown in column 4 of the left and right panels of Table 3. Although these univariate regressions still explain less than three percent of the variation in TFP and less than five percent of the variation in LP (right-hand-side panel in Table 3), the increase in explanatory power suggests that there is meaningful cross-industry heterogeneity in these relationships.

We investigate this further by estimating more flexible specifications, which are shown in the last two rows of the table. Specifically, the row labeled "All" reports the explanatory power of multivariate regressions where all O*NET variables are included (column 3) plus industry interactions (column 4). The row labeled "All with int" reports the explanatory power of regressions where the set of explanatory variables contains all O*NET variables and all of their two-way interactions. That is, the last specification regresses productivity on a second-order polynomial in these variables (column 3) plus industry interactions (column 4). The conclusion from these multivariate models is that the explanatory power increases substantially in these flexible specifications. The explanatory power is qualitatively similar when the dependent variable is LP, shown in the right-hand-side panel of Table 3. However, the explanatory power of the LP regressions is greater, which we discuss below.

Table 3. The relationship between the distribution of productivity, occupations, tasks, and skills

		Dependent variable: Log TFP				Dependent variable: log LP			
		All industries				All industries			
		(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
		Coef.	SE	R^2	Adj. R^2 (indFE)	Coef.	E	R^2	Adj. R^2 (indFE)
Univariate	TSU	0.2303	0.0121	0.0026	0.0272	0.4319	0.0179	0.0044	0.0333
	TSB	0.1892	0.0098	0.0029	0.0266	0.7132	0.0151	0.0196	0.0489
	Routine manual	-0.0576	0.0027	0.0033	0.0175	-0.0259	0.0043	0.0003	0.0303
	Routine cognitive	-0.0559	0.0041	0.0014	0.0124	-0.0493	0.0067	0.0005	0.0266
	Physical	-0.0419	0.0026	0.0017	0.0076	-0.0002	0.0040	0.0000	0.0124
	Interpersonal	0.0382	0.0041	0.0007	0.0134	0.1769	0.0060	0.0068	0.0244
	Analytical	0.0865	0.0042	0.0031	0.0245	0.1622	0.0062	0.0052	0.0310
Multi-variate	All			0.0050	0.0676			0.0329	0.1199
	All with int			0.0126	0.1100			0.0390	0.1699
N=328,000									
		High-tech industries				High-tech industries			
		Coef.	E	R^2	Adj. R^2 (indFE)	Coef.	E	R^2	Adj. R^2 (indFE)
Univariate	TSU	0.5406	0.0303	0.0113	0.0633	1.1590	0.0408	0.0289	0.0892
	TSB	0.3644	0.0206	0.0113	0.0630	1.0340	0.0286	0.0506	0.0985
	Routine manual	-0.0957	0.0068	0.0068	0.0260	-0.1876	0.0095	0.0145	0.0480
	Routine cognitive	-0.1679	0.0125	0.0074	0.0124	-0.0243	0.0170	0.0001	0.0127
	Physical	-0.0317	0.0066	0.0007	0.0196	-0.1923	0.0088	0.0153	0.0243
	Interpersonal	0.0787	0.0115	0.0017	0.0303	0.2805	0.0152	0.0122	0.0398
	Analytical	0.1764	0.0100	0.0106	0.0612	0.4076	0.0137	0.0314	0.0875
Multi-variate	All			0.0193	0.1096			0.0554	0.1283
	All with int			0.0498	0.1663			0.0685	0.1798
N=48,500									
		Low-tech industries				Low-tech industries			
		Coef.	E	R^2	Adj. R^2 (indFE)	Coef.	SE.	R^2	Adj. R^2 (indFE)
Univariate	TSU	0.1472	0.0130	0.0012	0.0139	0.2371	0.0199	0.0013	0.0165
	TSB	0.1200	0.0108	0.0011	0.0132	0.5866	0.0177	0.0123	0.0341
	Routine manual	-0.0479	0.0029	0.0024	0.0143	0.0155	0.0049	0.0001	0.0249
	Routine cognitive	-0.0342	0.0042	0.0006	0.0124	-0.0541	0.0072	0.0007	0.0307
	Physical	-0.0480	0.0030	0.0023	0.0032	0.0533	0.0046	0.0013	0.0089
	Interpersonal	0.0302	0.0043	0.0005	0.0072	0.1564	0.0065	0.0057	0.0198
	Analytical	0.0608	0.0045	0.0016	0.0110	0.0923	0.0070	0.0017	0.0141
Multi-variate	All			0.0043	0.0520			0.0319	0.1174
	All with int			0.0088	0.0889			0.0398	0.1667
N=279,000									

Notes: In each panel, label "All" denotes regressions in which all O*NET variables are included. Label "All with int" denote specifications where the dependent variable is regressed on a second-order polynomial in all O*NET variables. The columns titled "Adj. R^2 (indFE)" refer to regressions in which the explanatory variable is interacted with four-digit industry fixed effects.

To shed more light on the nature of industry heterogeneity, we re-estimate these regressions separately for establishments in high-tech and low-tech industries.[24] Results in the high-tech (low-tech) group are shown in the middle (lower) section of Table 3. Comparing the two sets of results from the TFP regressions, it is clear that tasks and skills matter more for high-tech industries compared with low-tech industries. The magnitudes of almost all correlations for the high-tech industries are at least twice those estimated using the entire sample, while the correlations for the low-tech industries are significantly closer to zero. Comparing the results for LP, we see that the results for high-tech and low-tech industries are more similar to each other, especially in the last two rows.

In our baseline results, we find that skill and task indexes account for a larger fraction of the variation in LP than in TFP. In contrast, for high-tech industries, they account for a larger fraction overall but a similar fraction of variation in LP and TFP. We think these findings deserve further investigation. Differences in results for LP and TFP are potentially driven by the role of capital intensity and the relationship between skill and task measures and capital intensity. That is, capital intensity accounts for variation in LP across establishments and, as discussed below, variation in capital intensity is positively associated with variation in the skill and task measures.

From these regressions, it is clear that: (1) there is meaningful industry-specific heterogeneity in the correlation between TFP and various task/skill indexes, and (2) these relationships may be nonlinear. More generally, simple models provide a better fit

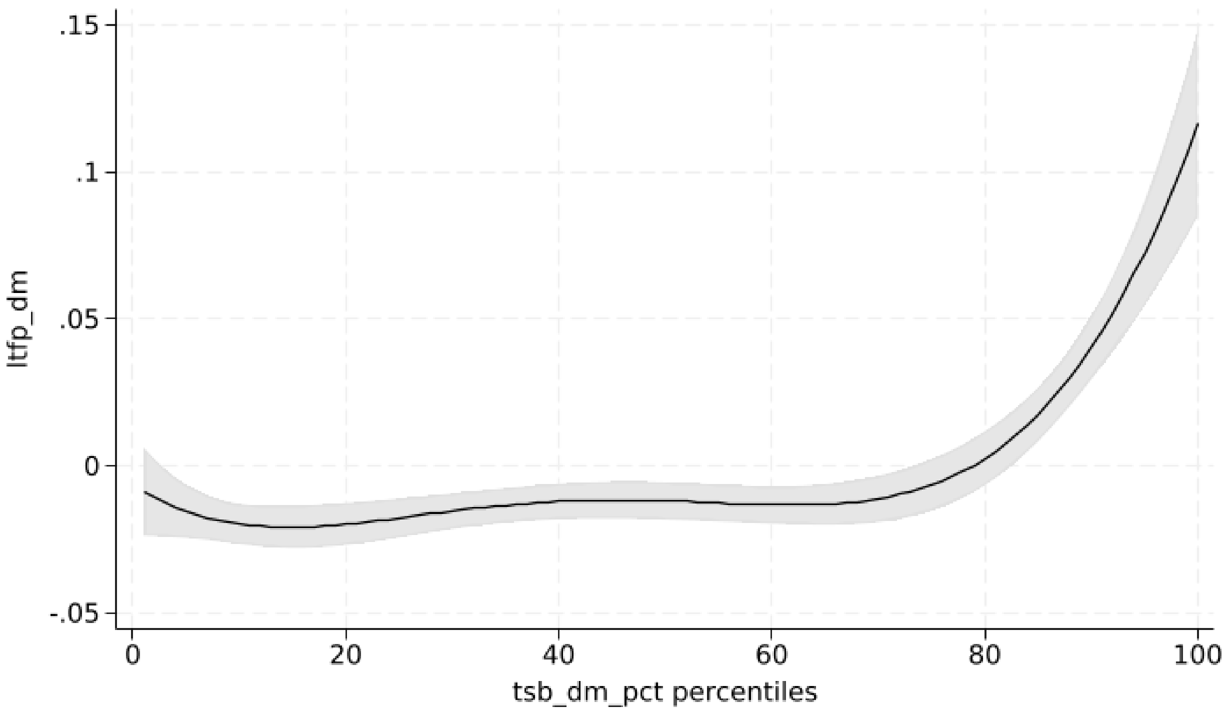
[24] The high-tech group contains the following four-digit NAICS codes: 3241, 3251, 3252, 3254, 3332, 3333, 3336, 3339, 3341, 3342, 3343, 3344, 3345, 3346, 3353, 3364. All other industries are classified as low-tech.

for high-tech industries than for low-tech industries, more complex models fit the data better, and the best fit is attained when including nonlinear terms. These points illustrate cross-industry heterogeneity in the correlations: different industries have different production technologies that comprise different tasks, which in turn require differently skilled labor. In addition, nonlinearities are more likely to be present in certain industries than others.

We repeat the previous analysis using earnings-per-worker (average wage) and the capital-labor ratio as dependent variables. The left-hand-side panels of Table A1 (earnings per worker) in the Appendix show correlations and R^2 values that are qualitatively similar to those for LP, which is not surprising given that in efficient markets, workers are paid their marginal product. The relationship between the capital-labor ratio and occupations/tasks/skills is relevant in the present context because the variation in capital-labor ratio can be interpreted as a simple indicator of technological differences. The results for the capital-labor ratio are interesting because the task/skill measures (fully-interacted) account for about the same fraction of its variation in high-tech as in low-tech industries. This finding is related to our findings on high- and low-tech industries above. While the relationship between capital intensity and skill/task measures is similar in high- and low-tech industries, it may be that the relationship between capital intensity and productivity differs between high- and low-tech industries. This discussion is speculative but highlights an area for future research.

We can gain more insight into the nature of the nonlinearities by estimating local correlations between TFP and some of the occupation/task/skill measures (TSB, Analytical, Routine manual). Specifically, in each percentile of each measure, we calculate the propensity-score-weighted mean TFP and then regress this mean on a quartic function of the percentiles of the index.

Figure 2. Fitted values from regressing log TFP on a quartic function of TSB percentiles



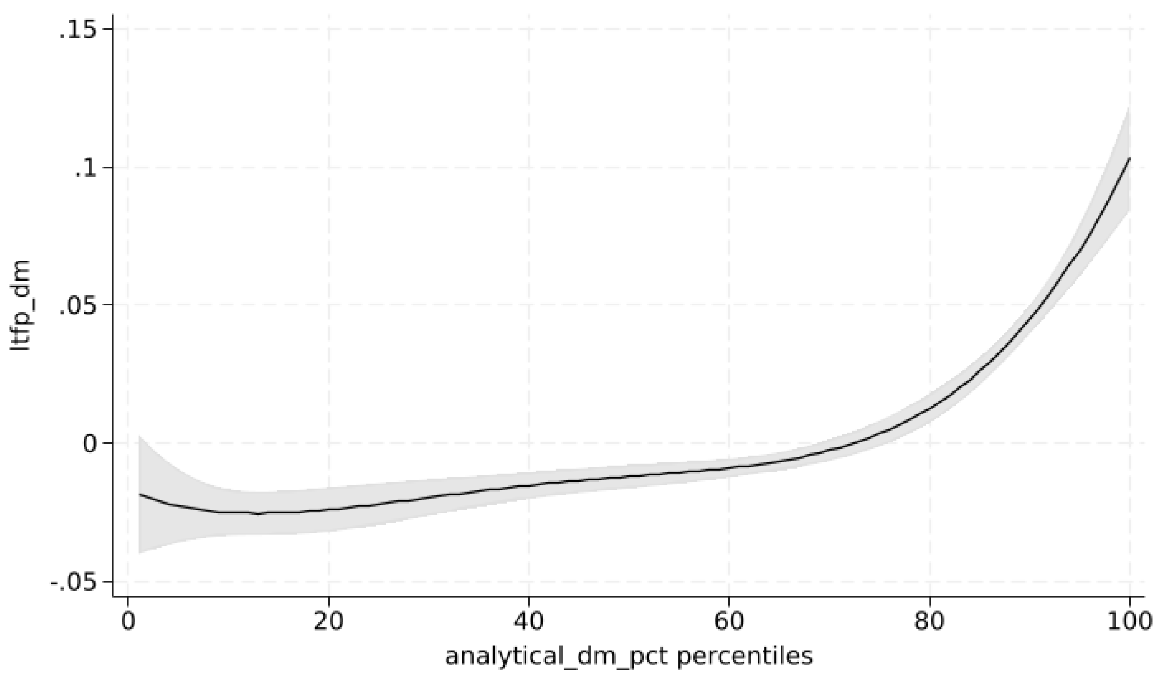
Notes: The shaded area shows the 95% confidence region.

Figure 2 shows log TFP values as predicted by TSB percentiles, where the slope of the prediction captures local estimates of the correlation. The most important implication of this exercise is that the relationship is nonlinear, which means the slope depends on where in the TSB distribution it is measured. To be specific, the slope hovers around zero up to about the 80th percentile. But it is increasingly positive in the top TSB quintile, which means productivity increases disproportionately as TSB increases among establishments with the highest TSB values.

To the extent that TSB reflects variation in occupation-specific national wages (see Section III.A), these patterns confirm that more productive establishments generally employ high-wage occupations, but they also indicate that establishments that employ the highest paying occupations are disproportionally more productive. In other words, controlling for the heterogeneity in TSB matters for the most productive establishments. This implication puts the findings of Table 3 into perspective: the statistically significant positive average correlation between log-TFP and TSB is coupled with a low R^2 *because* that specification uses a single coefficient to describe the relationship and ignores nonlinearity and heterogeneity. This implies that for the bottom 80% of establishments, the correlation is overestimated, while for the top quintile, it is underestimated, which is why the explanatory power of the regression in Table 3 remains under 1%. In addition, Figure 2 along with Table 3 suggest that establishments in the top quintile are *not* randomly distributed across industries, because TSB matters more in high-tech industries than low-tech ones, which is why the adjusted R^2 of models with industry interactions in Table 3 are an order of magnitude higher.

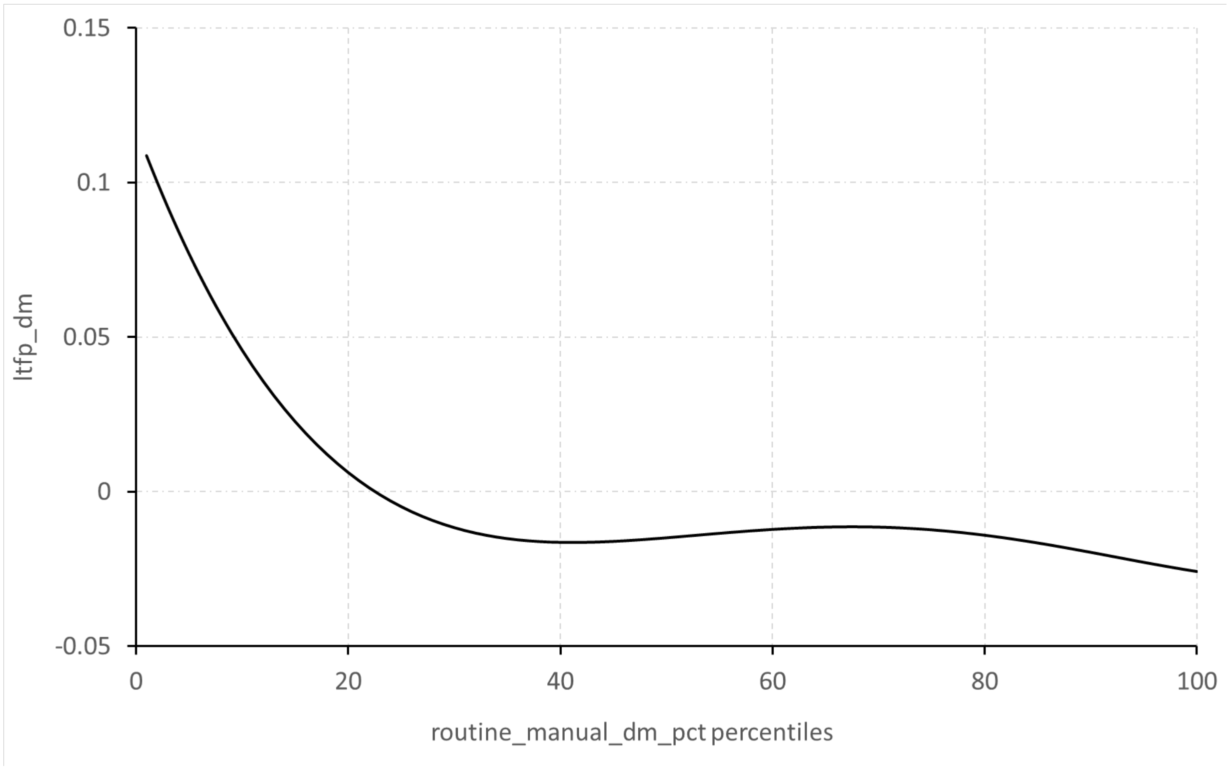
We can gain some additional insight by looking at a couple of our aggregate tasks. Figure 3 and Figure 4 show the analogous relationships between log TFP and analytical and routine manual tasks. In Figure 3, the quartic function indicates a slightly positive and unchanging slope over the first four quintiles, and an increasingly positive slope in the top quintile. Together with the results of Table 3, this means analytical skills are most strongly associated with the most productive establishments in high-tech industries, which are also disproportionately more productive. Figure 4 shows the locally predicted correlation between log TFP and the index of routine manual skills. The graph is consistent with the patterns in Figure 2 in the sense that productivity is generally negatively correlated with this type of skill, but the negative average slope in Table 3 is the result of a steeper negative slope among the most productive establishments and a close-to-zero slope in the majority of the skill distribution. Together, Figures 2–4 imply that controlling for the variation in occupations, tasks and skills likely matters, especially for the most productive establishments.

Figure 3. Fitted values from regressing log TFP on a quartic function of analytical percentiles



Notes: The shaded area shows the 95% confidence region.

Figure 4. Fitted values from regressing log TFP on a quartic function of routine manual percentiles



Notes: The confidence region is not shown because the disclosure of standard errors is delayed for technical reasons.

VI. 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 take a first step toward better understanding the role of labor heterogeneity in this context. Specifically, we look at how the Cobb-Douglas productivity residual is affected if we control for differences in the distributions of occupations, tasks and skills. 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 data. We match the OEWS survey data and CMP data using a hierarchical algorithm that prioritizes information on EINs, narrowly-defined industry, and geography. The uncertainty in our procedure stems from differences in the business registers, industry codes, reference periods, and also the fact that the two data sources have different establishment identifiers. A match is defined as a pair of establishments for which the values of the following variables are the same: EIN, narrowly-defined industry, FIPS state identifiers, and a restriction on the difference in establishment size.

Our empirical results indicate that correcting total hours worked for differences in occupations/tasks/skills—analogous to converting labor into efficiency units—has a small effect on the measured dispersion of Cobb-Douglas residuals. On the other hand, when we study the relationship between these variables directly, we find meaningful and interpretable correlation patterns. The main conclusion is that these correlations vary across industries and that the relationship between TFP and occupations/task/skills is nonlinear, and that the role of occupations/tasks/skills is most obvious for the most productive establishments and in high-tech industries. In contrast, such variation appears to be less important in low-tech industries and establishments that are closer to the average of the productivity distribution. This finding puts the result on the multiplicative correction into perspective: adjusting the labor input of every establishment is not necessarily justified; such an adjustment seems more appropriate where these variables matter.

More work is needed to quantify the extent to which measured productivity is affected if the variation in occupations/tasks/skills are accounted for during production function estimation. This question is relevant because the exact specification of the production function affects the estimated productivity residual. One interesting avenue for future research in this context is to explore the properties of more general specifications. For example, Dinlersoz and Wolf (2023) find that in the presence of advanced technologies that generate complementarities between capital and labor, the Cobb-Douglas residual systematically underestimates productivity if the true production technology is CES. This approach requires identifying industries or groups of establishments within an industry for which it is reasonable to assume that they are different enough from the rest of the establishments because such complementarities exist. Such cluster analysis is beyond the scope of the current paper but may be useful in future research. Alternatively, one could explore more flexible specifications like translog, which is followed by, for example De Loecker et al. (2020) and Foster et al. (2022). This specification approximates the underlying production function using a second-order polynomial in inputs. The translog specification is relevant because it does not require clustering prior to estimation, it is linear in elasticities, and its flexibility may ultimately result in a better fit.

References

Acemoglu, Daron. 1998. “Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality.” *The Quarterly Journal of Economics* 113 (4): 1055–89.

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.

Baily, Martin Neil, Charles Hulten, and David Campbell**.** 1992. “Productivity Dynamics in Manufacturing Plants.” *Brookings Papers on Economic Activity: Microeconomics*: 187–249.

Berman, Eli, John Bound, and Zvi Griliches. 1994. “Changes in the Demand for Skilled Labor within U. S. Manufacturing: Evidence from the Annual Survey of Manufactures.” *Quarterly Journal of Economics* 109 (2): 367–97.

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, 2023. “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.

Card, David and John E. DiNardo. 2002. “Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles.” *Journal of Labor Economics* 20 (4): 733–83.

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. 2021. “Chaos Before Order: Productivity Patterns in U.S. Manufacturing.” *International Productivity Monitor* 41: 138–52.

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.

Decker, Ryan A., John Haltiwanger, Ron S. Jarmin, and Javier Miranda. 2020. "Changing Business Dynamism and Productivity: Shocks versus Responsiveness." *American Economic Review* 110 (12): 3952–90.

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

De Loecker, Jan, Pinelopi K. Goldberg, Amit K. Khandelwal and Nina Pavcnik. 2016. "Prices, Markups, and Trade Reform." *Econometrica* 84 (2): 445–510.

De Loecker, Jan, Jan Eeckhout, and Gabriel Unger. 2020. “The Rise of Market Power and the Macroeconomic Implications.” *The Quarterly Journal of Economics* 135 (2): 561–644.

Dinlersoz, Emin, and Zoltan Wolf. (2023). “Automation, labor share, and productivity: plant-level evidence from U.S. Manufacturing.” *Economics of Innovation and New Technology* 33 (4): 604–626.

Dunne, Timothy, John Haltiwanger, and Kenneth R. Troske. 1997. "Technology and Jobs: Secular Changes and Cyclical Dynamics." Carnegie-Rochester Conference Series on Public Policy, 46 (1): 107–78.

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

Eslava, Marcela, John Haltiwanger, and Nicolas Urdaneta. 2023. "The Size and Life-Cycle Growth of Plants: The Role of Productivity, Demand, and Wedges." Forthcoming *The Review of Economic Studies* https://doi.org/10.1093/restud/rdad056

Fairman, Kristin, Lucia Foster, C.J. Krizan, and Ian Rucker. 2008. “An Analysis of Key Differences in Micro Data: Results from the Business List Comparison Project.” Center for Economic Studies Discussion Paper No. 08-28.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2008. “Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?” *American Economic Review* 98(1): 394–425.

Foster, Lucia, John Haltiwanger, and Cody Tuttle. 2022. “Rising Markups or Changing Technology?” NBER Working Paper No. 20491.

Gollop, Frank M., Barbara M. Fraumeni, and Dale W. Jorgenson. 1987. *Productivity and U.S. Economic Growth*. Harvard University Press.

Haltiwanger, John C., Henry R. Hyatt, Erika McEntarfer, Liliana Sousa, and Stephen Tibbets. 2014. “Firm Age and Size in the Longitudinal Employer-Household Dynamics Data.” Center for Economic Studies Discussion Paper No. 14-16.

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.

Oberfield, Ezra and Devesh Raval. 2021. “Micro Data and Macro Technology.” *Econometrica* 89 (2): 703–32.

Raval, Devesh. 2019. “The Micro Elasticity of Substitution and Non-Neutral Technology.” *RAND Journal of Economics* 50(1): 147–67.

Schreyer, Paul. 2001. “OECD Productivity Manual: A Guide to the Measurement of Industry-Level and Aggregate Productivity Growth.” Paris, Organisation for Economic Co-Operation and Development.

Standard Occupational Classification Policy Committee. 2010. “Attachment C: Detailed 2010 SOC Occupations Included in STEM.” https://www.bls.gov/soc/2010/home.htm.

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.

Wolf, Michael and Dalton Terrell. 2016. “The High-Tech Industry, What Is It and Why It Matters to Our Economic Future.” *Beyond the Numbers: Employment and Unemployment* 5(8), (U.S. Bureau of Labor Statistics, May 2016).

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.

Zolas, Nikolas, Zachary Kroff, Erik Brynjolfsson, Kristina McElheran, David N. Beede, Cathy Buffington, Nathan Goldschlag, Lucia Foster, and Emin Dinlersoz. 2020.

“Advanced Technologies Adoption and Use by U.S. Firms: Evidence from the Annual Business Survey.” NBER Working Paper No. 28290.

A. Appendix Table

Table A 1. The relationship between labor productivity, occupations and tasks/skills

		Dependent variable: Log-earnings-per-worker				Dependent variable: log(capital/labor)			
		All industries				All industries			
		Coef.	SE	R^2	Adj. R^2 (indFE)	Coef.	E	R^2	Adj. R^2 (indFE)
Univariate	TSU	0.6399	0.0089	0.0453	0.0600	0.5788	0.0266	0.0040	0.0169
	TSB	0.7179	0.0077	0.0936	0.1066	0.9920	0.0216	0.0194	0.0361
	Routine manual	-0.1128	0.0020	0.0280	0.0531	-0.0048	0.0064	0.0000	0.0156
	Routine cognitive	-0.0760	0.0030	0.0056	0.0181	0.0369	0.0098	0.0001	0.0113
	Physical	-0.0705	0.0021	0.0109	0.0208	0.0081	0.0059	0.0000	0.0054
	Interpersonal	0.1319	0.0030	0.0178	0.0338	0.1842	0.0093	0.0038	0.0123
	Analytical	0.2264	0.0031	0.0474	0.0627	0.2084	0.0093	0.0044	0.0159
Multi-variate	All			0.0967	0.1328			0.0357	0.0828
	All with int			0.1079	0.1698			0.0402	0.1260
N=328,000									
		High-tech industries				High-tech industries			
		Coef.	SE	R^2	Adj. R^2 (indFE)	Coef.	E	R^2	Adj. R^2 (indFE)
Univariate	TSU	0.9917	0.0191	0.1272	0.1456	1.0900	0.0512	0.0176	0.0453
	TSB	0.7797	0.0137	0.1730	0.1833	1.0789	0.0356	0.0380	0.0637
	Routine manual	-0.1823	0.0045	0.0824	0.1127	-0.1107	0.0125	0.0035	0.0174
	Routine cognitive	-0.1133	0.0071	0.0113	0.0404	0.0063	0.0214	0.0000	0.0053
	Physical	-0.1510	0.0039	0.0568	0.0709	-0.1232	0.0117	0.0043	0.0101
	Interpersonal	0.2036	0.0068	0.0387	0.0690	0.2167	0.0197	0.0050	0.0230
	Analytical	0.3454	0.0063	0.1356	0.1540	0.3860	0.0173	0.0194	0.0459
Multi-variate	All			0.1798	0.2045			0.0528	0.0925
	All with int			0.1986	0.2522			0.0608	0.1368
N=48,500									
		Low-tech industries				Low-tech industries			
		Coef.	E	R^2	Adj. R^2 (indFE)	Coef.	E	R^2	Adj. R^2 (indFE)
Univariate	TSU	0.5457	0.0100	0.0317	0.0412	0.4420	0.0308	0.0022	0.0110
	TSB	0.6935	0.0093	0.0764	0.0897	0.9577	0.0266	0.0157	0.0304
	Routine manual	-0.0951	0.0023	0.0193	0.0400	0.0222	0.0073	0.0001	0.0152
	Routine cognitive	-0.0687	0.0033	0.0047	0.0131	0.0429	0.0110	0.0002	0.0126
	Physical	-0.0535	0.0026	0.0056	0.0097	0.0454	0.0073	0.0004	0.0044
	Interpersonal	0.1178	0.0034	0.0144	0.0260	0.1778	0.0104	0.0035	0.0101
	Analytical	0.1925	0.0035	0.0326	0.0426	0.1577	0.0109	0.0023	0.0098
Multi-variate	All			0.0797	0.1170			0.0328	0.0808
	All with int			0.0904	0.1516			0.0374	0.1238
N=279,000									

Notes: In each panel, label "All'' denotes regressions where all O*NET variables are included. Label "All with int'' denote specifications where the dependent variable is regressed on a second-order polynomial in all O*NET variables. The columns titled "Adj. R^2 (indFE)" refer to regressions in which the explanatory variable is interacted with four-digit industry fixed effects.

B. 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.