Centre for Human Capital and Productivity. CHCP Working Papers

Economics Working Papers Archive

2017

2017-06 Job Tasks, Time Allocation, and Wages

Ralph Stinebrickner

Todd Stinebrickner

Paul Sullivan

Job Tasks, Time Allocation, and Wages

by

Ralph Stinebrickner, Todd Stinebrickner and Paul Sullivan

Western

Centre for Human Capital and Productivity (CHCP)

Working Paper Series

Department of Economics
Social Science Centre
Western University
London, Ontario, N6A 5C2
Canada

Job Tasks, Time Allocation, and Wages[*]

Ralph Stinebrickner[†]
Berea College

Todd Stinebrickner[‡]
University of Western Ontario

Paul Sullivan[§]
American University

July 2017

Abstract

This paper studies wage determination using the first longitudinal dataset containing job-level task information for individual workers. Novel quantitative task measures detail the amount of time spent performing People, Information, and Objects tasks at different skill levels. These measures suggest natural proxies for on-the-job human capital accumulation and provide new insights about wage determination. Current job tasks are quantitatively important, with high skilled tasks being paid double the rate of low skilled tasks. There is no evidence of learning-by-doing for low skilled tasks, but strong evidence for high skilled tasks. Current and past high skilled information tasks are particularly valuable.

1 Introduction

The tasks that a worker performs on-the-job are likely to play a central role in the determination of wages. The notion that worker productivity, and therefore wages, will depend directly on current job tasks is supported empirically by a large literature that explores the existence of wage differentials across employers (Manning, 2003), industries (Gibbons and Katz, 1992), occupations (Heckman and Sedlacek, 1985), and worker skill levels (Acemoglu and Autor, 2011; Autor and Dorn, 2013). In addition, as emphasized by the learning-by-doing model of human capital formation (Becker, 1964), one's productivity today will also depend on past tasks.[1]

Recently, the desire to understand wage determination at a primitive level has motivated research that attempts to characterize the tasks performed on jobs using measures of the type found in the Dictionary of Occupational Titles (DOT) or the newer Occupational Information Network (O*NET).[2] This paper takes advantage of unique data that address fundamental data limitations present in recent research by directly measuring the amount of time that workers spend on specific job tasks over their careers. These data allow us to provide new evidence about the roles that both current and past work activities play in the determination of wages. In particular, we provide empirical evidence on the intuitively appealing hypothesis, raised in a theoretical context by Gibbons and Waldman (2004), that learning how to perform job tasks may be an important component of human capital accumulation.

The most fundamental data limitation present in recent research is that no longitudinal surveys have collected task information directly for individual workers. In the absence of such data, it is necessary to use the DOT or O*NET databases to impute job tasks for workers. Because these data sources provide information only at the occupation level, all workers in the same (typically three digit) occupation must be assigned the same tasks. This type of imputation is not particularly appealing from a conceptual standpoint; Sanders and Taber (2012) note that, from the perspective of the theory of specific human capital, the primary usefulness of categorizing jobs

---

[1]Motivated by this theory, a vast empirical literature examines questions such as whether human capital tends to be general and transferable across all jobs (Mincer, 1974), or specific to occupations or industries. See Neal (1995) and Parent (2000) for studies of industry specific human capital, and Kambourov and Manovskii (2009) and Sullivan (2010b) for evidence of occupation specificity. Gibbons and Waldman (2004) and Lazear (2009) develop theoretical models that use job tasks to understand concepts such as the distinction between general and specific human capital, and the structure of promotions.

[2]See, for example Poletaev and Robinson (2008), Gathmann and Schönberg (2010), Yamaguchi (2012, 2015), Sanders (2017), Lise and Postel-Vinay (2015), Speer (2016), and Roys and Taber (2016).

by occupation is that occupations serve as observable proxies for the true task requirements of jobs. Further, recent evidence suggests that this type of imputation matters in practice; taking advantage of rare cross-sectional data that allow job tasks to be observed at the individual job-level, Robinson (2011) and Autor and Handel (2013) show conclusively that tasks vary significantly within three-digit occupations.[3] As a result, imputing a worker's current tasks based only on her occupation is likely to create substantial amounts of measurement error in job tasks. Given the persistence in one's occupations over time, this measurement error problem may be further exacerbated in a longitudinal setting where it is natural to impute a cumulative measure of tasks for use as a proxy for human capital accumulated over the career. This type of consideration has led to a recognition that individual-level task information is especially crucial when studying issues that require longitudinal data. For example, when describing the (unrealized) possibility of extending the 2006 Skills Survey (UK) to collect longitudinal task information, Green et al. (2006) write: "such a longitudinal study...would enable a huge leap in our understanding of the processes of skills development..."

Other fundamental data limitations arise because the questions used to describe tasks in the DOT or O*NET do not necessarily correspond closely to conceptual or empirical frameworks of interest. One issue is that the questions are typically of a qualitative nature (important, very important, etc.). A second issue is that existing task measures may not provide insight into *why* a task is important because they often do not distinguish between the skill level at which a task is performed and the amount of time that is spent on the task. Of particular salience, these issues imply that it is not possible to quantify the precise object that is arguably of ultimate importance - the relationship between wages and the amount of time that is spent on well-defined job activities. For example, a finding that changing a particular task from "not important" to "important" has a sizable effect on earnings may not be entirely satisfying if one cannot ascertain whether more time was spent performing the task at the same level or whether the same amount of time was

---

[3]Since it has been well established that tasks vary within occupations, our paper is based on the premise that the underlying task measures are the relevant objects of interest, rather than occupations. Robinson (2011) uses the one year (1971) of CPS data in which an analyst assigned DOT tasks to jobs at the individual level to show that within three digit occupations there is a large amount of variation in the tasks that are performed on jobs. Using individual level information collected as part of the Princeton Data Improvement Initiative, Autor and Handel (2013) also find substantial variation in tasks within an occupation and show that this information is an important predictor of earnings even after conditioning on occupation. To the best of our knowledge, The Skills Survey (UK) and new LISA survey (Canada) are the only other two datasets that provide cross-sectional information about job tasks at the level of an individual worker-job.

spent performing the task at a higher level.[4] Throughout the paper we also discuss a variety of other concerns related to measurement, specification, and interpretation that arise because of these issues. As an example of particular relevance in the longitudinal context here, it may be difficult to find an appealing way to aggregate yearly qualitative measures of task importance into a cumulative measure of the task importance over a career.

This paper takes advantage of data from the Berea Panel Study (BPS), which followed two cohorts of students very closely from the time of college entrance past the age of thirty. The BPS task measures, which are described in detail in Section 2, are in the spirit of some of the task information available in the DOT, describing the manner in which tasks relate to People, Information, and Objects.[5] Related to the first fundamental data limitation described above, a crucial feature of the data collection is that the task information comes directly from the worker, providing a rare measure of what the worker is actually doing on-the-job. The fact that this information was collected yearly allows us to construct the first longitudinal dataset containing job-level task information for individual workers, with it being particularly valuable that the dataset documents a period starting at the very beginning of workers' careers. However, the features of our data collection that are perhaps the most innovative were motivated by our desire to also address the other fundamental data limitations described in the previous paragraph. A unique feature of our task measures comes from collecting time allocation information related to the task measures. While the conceptual usefulness of knowing how individuals divide their work time among various job tasks has been recognized (Gathmann and Schönberg, 2010), direct information of this type is not available from existing data sources. The BPS collected data detailing the percentage of work time in a year that is spent in the People, Information, and Objects task categories. Further, within each of these general task categories, respondents report the percentage of time that is spent performing different sub-tasks, which are ordered by skill. These data allow us to construct the six quantitative task variables that we utilize: the proportion

---

[4]In practice, factor analysis is often used to summarize a large number of DOT or O*NET questions in terms of a smaller number of factors (e.g., cognitive or manual). The same issues remain in this case. The weights characterizing the importance of cognitive and manual tasks for a particular occupation are informative about the importance of cognitive and manual tasks in that occupation relative to other occupations, but it is difficult to interpret the information from the factor models in terms of concrete job activities. Further complicating interpretation, it is typically the case that some DOT and O*NET questions used to identify the factor are more related to task level and others are more related to how frequently a task is performed.

[5]We have condensed the number of categories within People, Information, and Objects relative to the DOT, and have also written the categories so that they are ordered in terms of the level of the tasks being performed.

of one's work time that is spent on high skilled People tasks, low skilled People tasks, high skilled Information tasks, low skilled Information tasks, high skilled Objects tasks, and low skilled Objects tasks. For each worker, summing a particular task variable over time, after weighting by hours worked, provides a direct measure of task-specific work experience at each point of the career.

Our results take advantage of up to ten years of labor market data. The descriptive results in Section 3 represent the first time it is has been possible to quantify, using explicit time measures, how workers are spending time on particular jobs. Of particular importance for the interpretation of a variety of results in the paper, we find that, while there is substantial heterogeneity in the tasks that workers perform, almost all workers devote at least some time to each task in each year. That is, even for this sample of college graduates, it is difficult, for example, to entirely avoid performing low skilled People tasks. Taking advantage of the longitudinal nature of our data, we find that the mean amount of time spent on information tasks and the mean amount of time spent on high skilled tasks increase significantly over time.

Section 4 transitions from examining how time is spent on jobs to determining how this time allocation is related to wages. In Section 4.1 we explore OLS relationships between wages, jobs tasks, and task-specific work experience. Given the strong relationships that are found, we introduce, in Section 4.2, a simple conceptual framework to discuss issues of interpretation that arise when attempting to estimate the returns to job tasks. The importance of the new BPS task data is highlighted by enriching the conceptual model of Autor and Handel (2013), who stress the potential importance of taking into account the role of comparative advantage (Heckman and Sedlacek, 1985; Roy, 1951). We begin by discussing why explicitly quantifying the time allocated to tasks at different skill levels is useful when adopting the standard empirical assumption that task prices are constant in the market of interest.

Next, we examine the benefits of our longitudinal data for addressing the selection biases that will be present if workers sort into jobs on the basis of unobserved ability. When workers with higher unobserved ability to perform task k tend to sort into jobs that require more time spent on task k, differences in wages across two workers could arise either because of differences in the *amount* of task k that is performed or because of differences in *ability* to perform the task. If the latter is important, cross-sectional estimators of the return to spending more time on task k would

be biased because the two workers would not obtain the same wages even if they performed the same amount of the task.

The quantitative nature of our novel time allocation measures allows us to difference observations over time to take advantage of within-person variation in job tasks. In the standard case where unobserved ability is one-dimensional, this differencing removes the confounding effect of permanent unobserved ability altogether. However, differencing can also be valuable in the more general case where abilities are multi-dimensional, so that sorting into tasks takes place on the basis of comparative advantage. In this case, even though differencing does not remove the influence of unobserved abilities, we discuss why it is reasonable to view fixed effects as estimating a parameter of substantial interest: the average return, or effect, of having workers spend more time on a particular task. This interpretation is possible because the nature of our task measures allows us to observe that the large majority of individuals spend time on each task (i.e., they spend non-zero time in high and low skilled People, Object, and Information tasks) and are also observed changing tasks over time. Thus, a contribution of our data collection is that it helps us move away from the typical comparative advantage scenario in which most agents are not observed participating in all "sectors."

A comparison of Ordinary Least Squares (OLS) and fixed effects (FE) estimates reveals the patterns one might expect - the estimated role of tasks is smaller when employing the fixed effects estimator which has the goal of addressing potential sorting on the basis of ability. However, despite these differences, the big picture takeaways from the estimation of both OLS and fixed effects, which rely on very different identification strategies, are remarkably similar. Our first main finding is that current job tasks play a large role in determining wages. In particular, there is a large wage premium for performing high skilled work: high skilled tasks are paid roughly double the rate of low skilled tasks. Our second main finding is that there is no evidence of learning-by-doing for low skilled tasks, but there is strong evidence of learning-by-doing for high skilled tasks. Our third main finding is that not all types of high skilled tasks (i.e., People, Information, and Objects) play the same role in the determination of wages. We find that working in high skilled people tasks in the present and/or in the past plays an important role, but that that the influence of working in high skilled information tasks in the present and/or in the past is especially large. We illustrate the benefits of our quantitative measures and put our results into

5

a further "lifecycle" context by examining how hypothetical changes in tasks over the first ten years of a worker's career would affect the worker's log wage in the tenth year. As an example of our findings, shifting 10 percent of work time from low skilled people tasks to high skilled information tasks in each of the first ten years leads to a 22 percent increase in the tenth year wage, with 70 percent of this increase due to the accumulation of task-specific experience during the first nine years and the remainder due to the direct effect of performing different tasks in the tenth year.

From a methodological standpoint, our results indicate that longitudinal data on tasks is crucial for understanding a variety of important issues, including those related to wage determination and wage growth.[6] In terms of direct guidance about survey questions, our time allocation measures allow the type of work required by jobs to be characterized by quantitative task measures that have clear interpretations. Among their many benefits, the measures lead to natural proxies for on-the-job human capital accumulation (task-specific experience), which appear to be an improvement over existing proxies (such as years of work experience). In addition, differentiating tasks by skill level is crucial in order to understand wage differences across workers.

2 Data

This section provides general information about the Berea Panel Study (Section 2.1) and explains how job tasks are measured in this dataset (Section 2.2).

2.1 The Berea Panel Study

Designed and administered by Todd Stinebrickner and Ralph Stinebrickner, the Berea Panel Study (BPS) is a longitudinal survey that was initiated to provide extremely detailed information about the college and early post-college periods. The project involves surveying students who entered Berea College in the fall of 2000 and the fall of 2001 approximately sixty times from the time of college entrance through 2014. In this paper, we focus on understanding the evolution of earnings for graduates, by taking advantage of post-college surveys that were collected annually after students left school. More than ninety percent of all graduates in the two BPS cohort years

---

[6]A number of influential papers in recent years have employed task-based approaches to investigates the causes of changes in the wage structure and the extent of polarization in the U.S. labor market. See, for example, Autor, Levy, and Murnane (2003); Autor, Katz, and Kearney (2009); Autor and Dorn (2013).

Table 1: Descriptive Statistics

Variable	Mean	Std. Dev.
Black	0.15	0.36
Female	0.64	0.48
College GPA	3.17	0.46
Has children	0.29	0.46
Wage	$15.42	9.48
Hours of work	39.58	10.65
Log-wage growth over career	0.49	0.51
Employed	0.87	0.34
Pr(employed │ male)	0.91	0.29
Pr(employed │ female)	0.85	0.36
Pr(employed │ male without child)	0.89	0.31
Pr(employed │ female without child)	0.90	0.30
Number of people	528	
Ave. observations per person	6.24	

Notes: Wages converted to 2005 dollars using the CPI.

completed one or more of these annual surveys, and the response rate on these surveys leveled off at approximately eighty-two percent until 2011 before declining slightly. To avoid the need to impute crucial information, our analysis uses all yearly observations from the time of graduation until an individual first fails to complete a post-college survey. For the 528 individuals in our sample, the maximum number of yearly observations is 10 and the average number of yearly observations is 6.2.

Important for the notion that the basic lessons from our study of one school are pertinent for thinking about what takes place elsewhere, Berea operates under a standard liberal arts curriculum, students at Berea are similar in academic quality to students at schools such as The University of Kentucky (Stinebrickner and Stinebrickner, 2008), and outcomes such as major choice at Berea are similar to those found in the NLSY by Arcidiacono (2004). However, even putting aside the obvious issue of data collection feasibility, there are benefits of studying one school. In particular, the ability to hold school quality constant is beneficial for a variety of reasons, including that it makes academic measures such as college GPA directly comparable across individuals.[7]

This paper is made possible by merging the unique survey data from the early portion of

---

[7]For previous work that has used the BPS to study other issues (see e.g., Stinebrickner and Stinebrickner 2003; 2006; 2008; 2012; 2014).

7

an individual's working life with detailed administrative data. The administrative data provides basic demographic information as well as academic information that is necessary to characterize one's human capital at the time of workforce entrance. Table 1 shows that 64 percent of the sample is female, 15 percent of the sample is black, and the average college grade point average (GPA) is 3.17. Pooling all observations across the full sample period, the average hourly wage, in 2005 dollars, is $15.42, and the average work week is 39.58 hours with 43 percent of jobs reporting exactly a 40 hour week. It is worth noting that in this sample of recent, young college graduates, the labor market experiences of males and females are similar across many dimensions. For example, both men and women have high employment rates (91% men, 85% women), with the small difference between the groups being explained almost entirely by children (childless men and women both have employment rates of approximately 90 percent). Further, the results in Section 4 show that the difference between male and female wages is not statistically significant at conventional levels after controlling for observable variables.

2.2   Measuring Job Tasks in the BPS

The task information associated with the job that a worker holds in a particular year comes from BPS survey Question C, which is shown in Appendix A. A unique component of Question C is the time allocation measures associated with job tasks. Question C4 contains the time allocation measures that indicate the percentage of total work time that is spent on the People, Information, and Objects task categories. Questions C1, C2, and C3, respectively, contain the time allocation measures that indicate the percentages of time that are spent on each specific sub-task within the People, Information, and Objects task categories, respectively. Defining the first two sub-tasks (1 and 2) within each of the People (C1), Information (C2), and Objects (C3) task categories as low skilled and the last two sub-tasks (3 and 4) as high skilled, these questions allow us to compute the percentage of total work time in a year that is spent on each of the three task categories, at each of the two skill levels. That is, they allow us to compute the amount of time devoted to high skilled People tasks, low skilled People tasks, high skilled Information tasks, low skilled Information tasks, high skilled Objects tasks, and low skilled Objects tasks.

In terms of notation, for each task k, $k \in (\text{people}, \text{information}, \text{objects})$, let $\tau^h(k)$ represent the fraction of time on-the-job (in a particular year) that a worker spends performing task k at

a high skill level (h), and let $\tau^l(k)$ represent the fraction of time spent performing task k at a low skill level (l). The vector of all task requirements of a job is denoted by the six element vector $T = \{\tau^s(\text{people}), \tau^s(\text{information}), \tau^s(\text{objects}),\ s = l, h\}$. For notational convenience, we sometimes abbreviate the general task categories as follows: $P = \text{people}$, $I = \text{information}$, $O = \text{objects}$.

Relevant for the question of whether students are able to understand questions such as C1, C2, C3, and C4, these questions are similar in spirit to BPS questions that elicited beliefs (expectations) about grade performance (and other outcomes) by asking respondents to assign "percent chances" to a set of mutually exclusive and collectively exhaustive grade categories. As a result, respondents had received classroom training related to similar types of questions and had answered similar types of questions frequently in the past, with both exit interviews and internal consistency checks confirming a good understanding of these questions. Nonetheless, given the experimental nature of our approach, one of our methodological objectives is to see whether confidence can be gained in these conceptually valuable, but previously untested, survey questions.

3   New Facts About Job Tasks

A descriptive examination of the task data is useful because it represents the first time that it has been possible to quantify, using explicit time measures, how individuals are spending time on their particular jobs.

### 3.1   The Distribution of Job Tasks

This section presents descriptive results obtained by pooling data over the ten year longitudinal sample. We begin by examining the distribution of the amount of time that workers spend performing People, Information, and Objects tasks, for the moment abstracting away from the distinction between high and low skilled tasks. Figure 1 shows the density function of $\tau^l(k) + \tau^h(k)$ for each of the three general task categories k. The unit of measurement for the horizontal axes in these figures is the fraction of time spent performing each task. Figure 1 shows that, on average, workers spend 51 percent of their time on people tasks (i.e., $\text{mean}\left(\tau^l(P) + \tau^h(P)\right) = 0.51$), 34 percent of their time on information tasks, and only 15 percent of their time on objects tasks. The standard deviations of tasks indicate that there is significant variation across jobs

Figure 1: The Distribution of People, Information, and Object Tasks
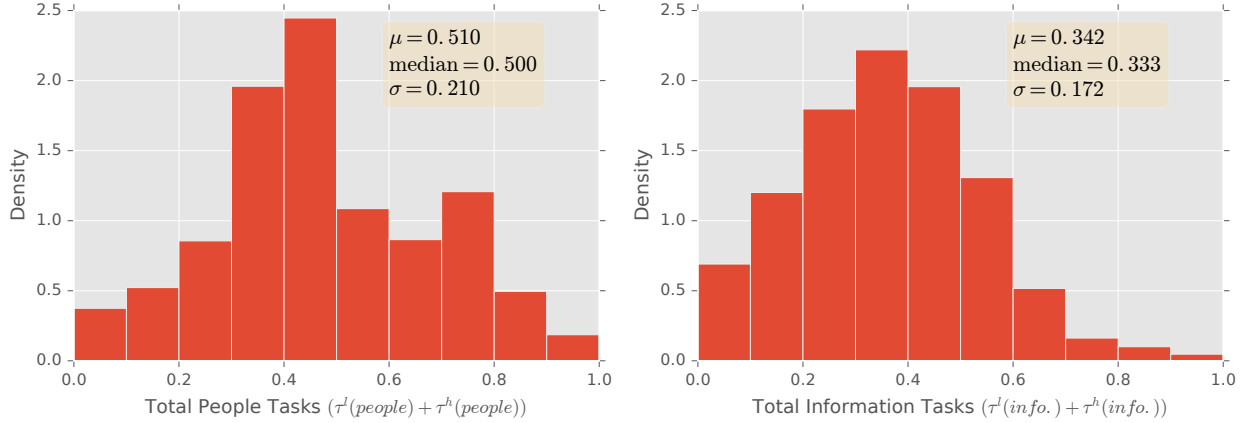
(a) Histogram of People Tasks

(b) Histogram of Information Tasks
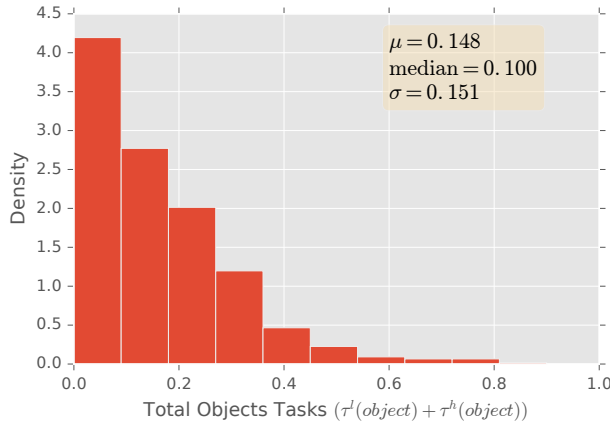
(c) Histogram of Objects Tasks

in the fraction of time spent performing each task, with people tasks varying the most ($\sigma = 0.21$), objects tasks varying the least ($\sigma = 0.15$), and information tasks lying in between the two extremes.

In addition to the overall amount of time spent working on people, information, and objects tasks, the BPS data distinguish between high and low skilled work within each category. Figure 2 shows that within each category, the distributions of high and low skilled tasks are in general quite similar. Averaged across all jobs, an approximately equal amount of time is spent performing each task at the high and low skill levels $\left(\tau^l(k) \simeq \tau^h(k), \ k = P, I, O\right)$. These aggregate statistics alone do not speak to the mix of low and high skilled tasks present at the level of individual jobs. We find that jobs frequently require performing different amounts of low and high skilled People tasks, but that there is comparatively less dispersion in the combinations of low
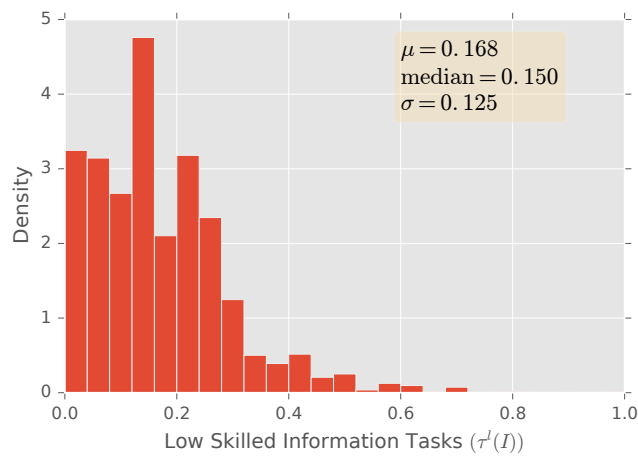
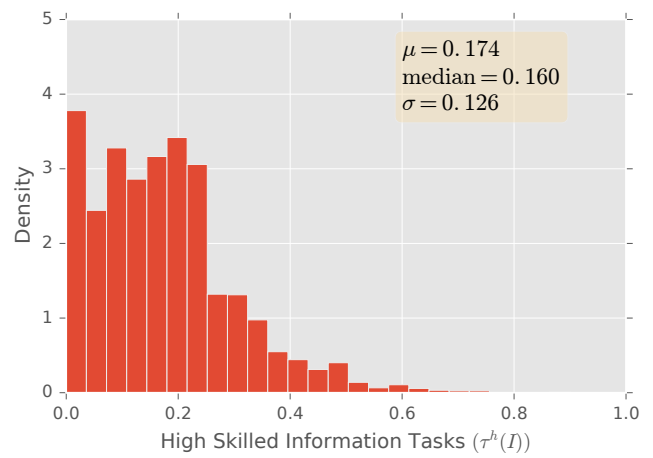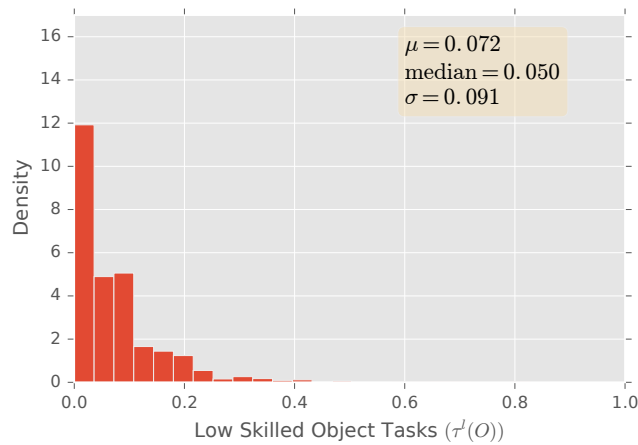Figure 2: The Distribution of High and Low Skilled Tasks
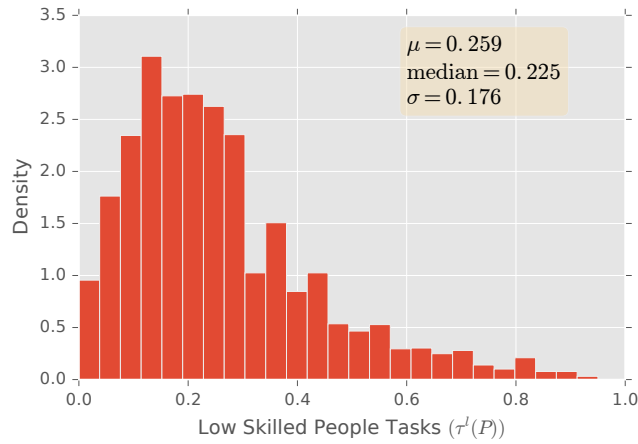
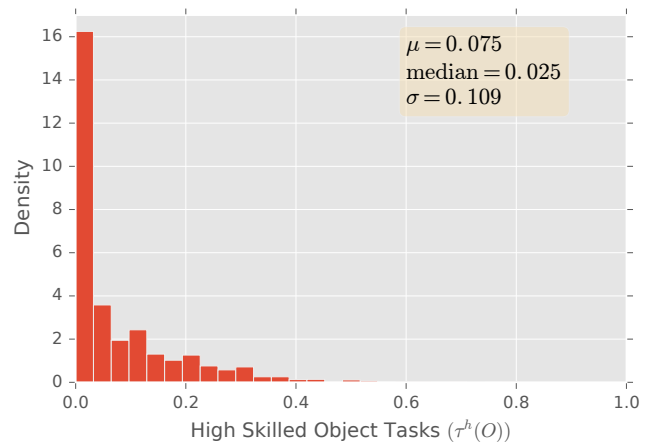(a) Low Skilled People
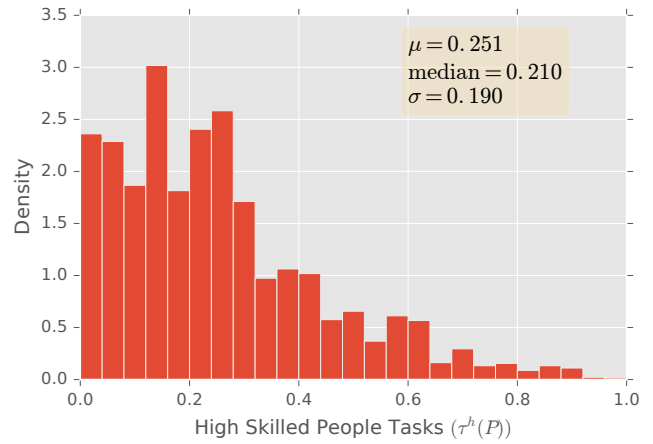
(b) High Skilled People

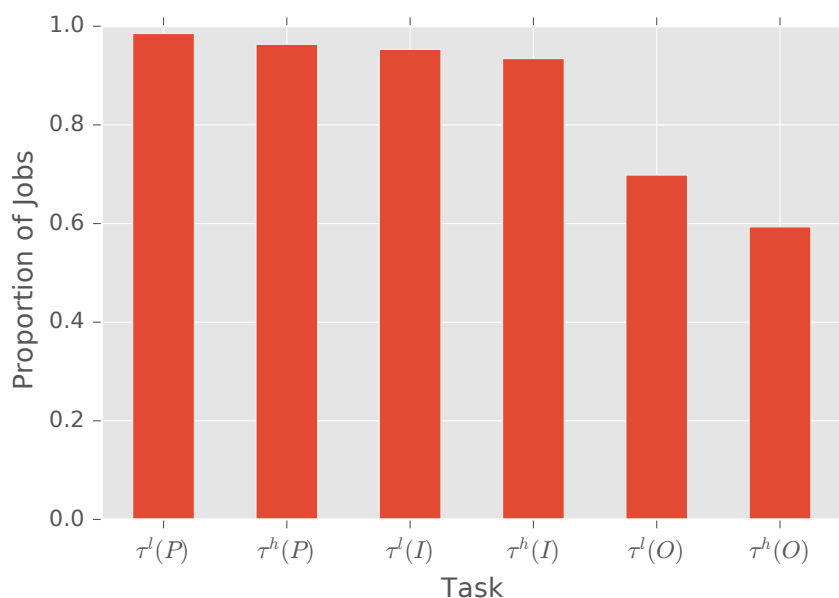(c) Low Skilled Information

(d) High Skilled Information

(e) Low Skilled Objects

(f) High Skilled Objects

Figure 3: Proportion of Jobs Requiring a Particular Task

and high skilled Information tasks required by jobs. For example, 35% of jobs require spending an equal amount of time on high skilled and low skilled Information tasks, but only 21% of jobs require spending an equal amount of time on high skilled and low skilled People tasks.

The statistics presented up to this point reveal substantial variation in the job tasks performed by workers. One interesting feature of the data is the strong tendency for jobs to require performing at least some of each task. For each of the six tasks, Figure 3 shows the fraction of jobs that require performing some of that task (i.e., the fraction of jobs with $\tau^s(k) > 0$). This figure shows that it is very difficult to avoid performing low skilled people tasks, with approximately 99 percent of all jobs requiring at least some low skilled interaction with people. Similarly, the percentages shown in Figure 3 for high skilled people tasks, low skilled information tasks, and high skilled information tasks are each greater than 95 percent. Perhaps not surprisingly, given our sample of college graduates, workers are less likely to spend at least some time performing objects tasks. However, it is interesting to note that even in this highly educated sample, approximately 70 percent of jobs require low skilled objects tasks, and 60 percent require high skilled objects tasks.

Naturally, the percentages in Figure 3 tend to be even higher when we take advantage of the longitudinal aspect of the data to examine the proportion of individuals who ever perform a particular task at some point over their career (i.e., the fraction of individuals who report

$\sum_{t=1}^{T} \tau_{it}^s(k) > 0$, where T is the last year the person is observed). For example, we find that 80 percent of individuals report performing high skilled objects tasks ($\tau^h(O)$) at some point over their career and that 82 percent of individuals report performing low skilled objects tasks ($\tau^l(O)$) at some point over their career. Thus, there is a strong tendency across both jobs and individuals to perform at least some of each task. Section 4.2 discusses how this feature of the data is very useful when estimating the wage returns to different job tasks.

3.2 Task-Specific Work Experience Over the Career

Up to this point, the descriptive analysis of job tasks has primarily pooled all data over the ten year longitudinal sample. This section focuses instead on using the panel aspect of the data to discover new facts about how workers spend their time on the job over their careers. Specifically, we now turn to a description of the distribution of task-specific work experience.

Using the unique features of our panel data, we are able to compute the first measures describing the cumulative amount of time that individuals have spent on different tasks over the career. These proxies for human capital incorporate the task variables, which are measured in fraction of time units, along with data on the hours worked by each worker. Specifically, the cumulative amount of time that individual i at time t has spent performing each of the three tasks (people (P), information (I), objects (O)) at each skill level (high (h) and low (l)) in the past is given by $h_{it}^s(k) = \sum_{j=1}^{t-1} \tau_{ij}^s(k)\omega_{ij}^s(k)$, $s \in (h,l)$, $k \in (P,I,O)$, where $\tau_{ij}^s(k)$ is the fraction of time that individual i spends performing task k at skill level s in time j and $\omega_{ij}^s(k)$ is a weight derived from the hours that person i works in time j. The hours weight is $\omega_{ij}^s(k) = \text{hours}_{ij}/40$, where hours_{ij} represents the hours worked by worker i on her job at time j. The weights are normalized in this manner so that $\omega_{ij}^s(k) = 1$ indicates that a worker works a forty hour week. We make use of the hours data based on the premise that the amount of learning-by-doing depends on the time allocated to each task, rather than simply the percentage of time spent on each task.[8]

Table 2 shows descriptive statistics for task-specific work experience in the sixth survey year. The measures are denominated in full-time years of work units.[9] For example, the first entry of Panel C of Table 2 shows that the average worker accumulates 1.382 years of experience working

---

[8]All of our key empirical results are robust to eliminating the hours weight, which is perhaps not surprising given that 43 percent of all jobs in the data report exactly a 40 hour work week.

[9]Defining a 40 hour work week as full time employment.

Table 2: The Distribution of Task-Specific Work Experience in Year 6

		Statistics				
		Mean	Median	Stand. Dev.	5th perc.	95th perc.
Level of Aggregation	(1)	(2)	(3)	(4)	(5)	(6)
(A) Task Category						
People	$h^l(P) + h^h(P)$	2.738	2.572	1.203	1.119	5.016
Information	$h^l(I) + h^h(I)$	1.835	1.774	0.824	0.600	3.286
Objects	$h^l(O) + h^h(O)$	0.820	0.620	0.735	0.005	2.204
(B) Skill Level						
Low	$h^l(P) + h^l(I) + h^l(O)$	2.654	2.653	1.032	1.042	4.213
High	$h^h(P) + h^h(I) + h^h(O)$	2.739	2.675	1.236	0.896	5.045
(C) Disaggregated Tasks						
Low skilled people	$h^l(P)$	1.382	1.218	0.750	0.442	2.835
High skilled people	$h^h(P)$	1.356	1.102	0.965	0.307	3.363
Low skilled info.	$h^l(I)$	0.897	0.852	0.475	0.209	1.715
High skilled info.	$h^h(I)$	0.938	0.840	0.598	0.161	2.051
Low skilled objects	$h^l(O)$	0.374	0.287	0.359	0.000	0.981
High skilled objects	$h^h(O)$	0.445	0.241	0.522	0.000	1.471

Notes: Entries are statistics for task-specific work experience in the sixth survey year.

with people at a low skill level. The figures are consistent with the findings from Section 3.1. Considering the disaggregated tasks, very few people specialize strongly by primarily performing any one task at any one skill level; column 6 shows that each of the ninety-fifth percentiles are considerably less than 6.0, the maximum amount of experience that a worker could attain by fully specializing in one task over the first six years of the career.[10] However, at the same time, potentially important for explaining issues such as wage dispersion, column 4 shows that the variation across workers in accumulated tasks is quite substantial.

Neither Figure 2 nor Table 2 provides evidence about how the amount of time spent on tasks changes over the first ten years of workers' careers. Taking advantage of the longitudinal nature of our data, Table 3 shows that there exist substantial differences in task usage between the first two survey years (1-2) and the last two survey years (9-10). The Task Category section of the table shows that the mean amount of time spent on information tasks increases significantly over time, primarily because the mean amount of time spent on people tasks decreases significantly over time. The Skill Level section of the table shows that the mean amount of time spent on high skilled tasks increases significantly over time. Together, these findings are of course reflected in

[10]While not shown, the shapes of the histograms for the six disaggregated task-specific work experience measures were found to be very similar to the shapes of the histograms in Figure 2.

Table 3: Changes in Mean Tasks over Time

| Level of Aggregation		Years 1-2	Years 9-10	Difference	t-Statistic
	(1)	(2)	(3)	(4)	(5)
(A) Task Category					
People	$\tau^l(P) + \tau^h(P)$	0.532	0.499	-0.033	-2.595
Information	$\tau^l(I) + \tau^h(I)$	0.307	0.358	0.051	4.920
Objects	$\tau^l(O) + \tau^h(O)$	0.161	0.143	-0.018	-1.899
(B) Skill Level					
Low	$\tau^l(P) + \tau^l(I) + \tau^l(O)$	0.524	0.488	-0.036	-2.754
High	$\tau^h(P) + \tau^h(I) + \tau^h(O)$	0.476	0.512	0.036	2.754
(C) Disaggregated Tasks					
Low skilled people	$\tau^l(P)$	0.296	0.244	-0.052	-4.686
High skilled people	$\tau^h(P)$	0.236	0.256	0.019	1.715
Low skilled info.	$\tau^l(I)$	0.154	0.171	0.016	2.214
High skilled info.	$\tau^h(I)$	0.153	0.187	0.034	4.467
Low skilled objects	$\tau^l(O)$	0.074	0.074	-5.67×10^{-6}	-0.001
High skilled objects	$\tau^h(O)$	0.087	0.069	-0.018	-2.582

Notes: Entries in columns (2) and (3) are mean fractions of time spent on tasks at each level of aggregation. Column (4) shows the difference between columns (2) and (3). The t-statistic for the difference is shown in column (5).

the substantial changes over time shown in the Disaggregated Tasks section of the table. Most notably, the mean amount of time spent on low skilled people tasks decreases by 18% over time, and the mean amount of time spent on high skilled information tasks increases by 22% over time.

4 The Relationship Between Current Tasks, Past Tasks, and Wages

This section transitions from examining how time is spent on jobs to determining how this allocation is related to wages. We begin in Section 4.1 by exploring OLS relationships between wages, job tasks, and task-specific work experience. In Section 4.2 we introduce a conceptual framework, which builds on the recent work of Autor and Handel (2013), to highlight the issues that arise in the interpretation of the types of strong empirical relationships that we find. This analysis points to the value of fixed effects estimators that take advantage of our longitudinal data to difference observations across time. Of particular importance, fixed effects can yield estimates of useful population parameters even in the presence of selection into job tasks based on comparative advantage. Section 4.3 shows the fixed effects results.

Figure 4: Tasks and Human Capital Proxies by Wage Decile

(a) Tasks by Wage Decile   (b) Human Capital by Wage Decile
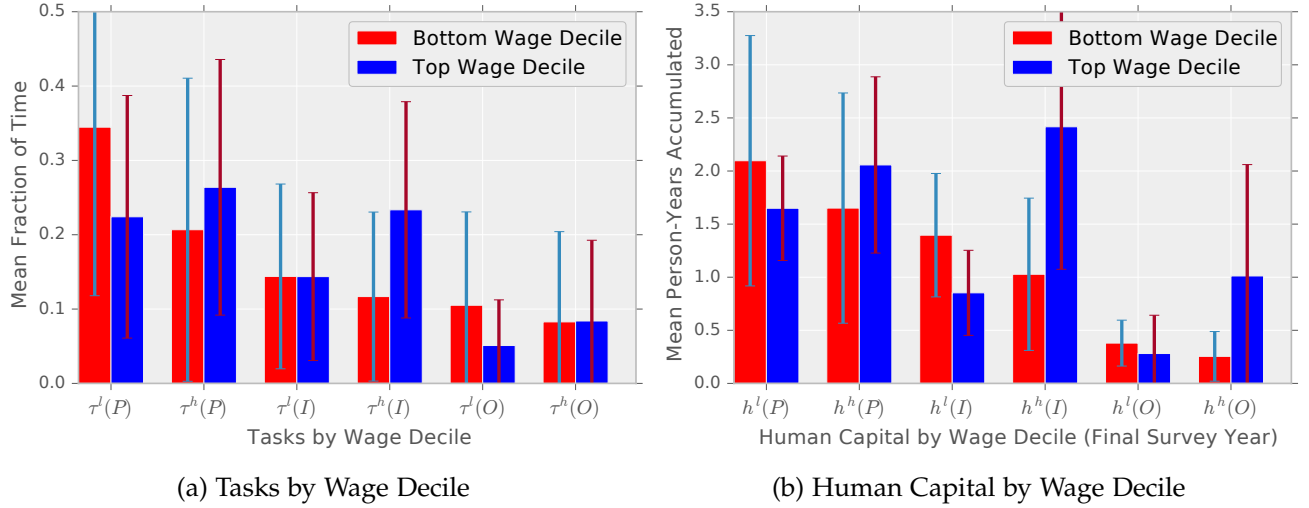
4.1 Tasks and Task-Specific Work Experience Across the Wage Distribution

The graphs in Figure 4 illustrate bivariate relationships between wages and job tasks (Figure 4a), and wages and task-specific work experience (Figure 4b). Pooling across all workers and all years, Figure 4a depicts the mean fraction of time spent on tasks for the bottom and top wage deciles. Perhaps most notable, when compared to jobs in the top decile of the wage distribution, jobs in the bottom decile require approximately 50 percent more low skilled people tasks, twice as much low skilled objects tasks, and half as much high skilled information tasks. Figure 4b shows how task-specific work experience in the final survey year varies across the wage distribution. Most notable here, the gaps in low skilled task usage that were seen between the bottom and top wage deciles in Figure 4a are not as large (people) or are not present at all (objects). On the other hand, the large gap that was observed for high skilled task usage (information) in Figure 4a is also observed in Figure 4b and a new gap has emerged for high skilled objects tasks. Thus, our descriptive results in Figure 4 suggest that both current and past tasks may play an important role in determining wages, with high skill past tasks potentially being more important than low skill past tasks.

OLS provides a way to delve deeper into the relationships in the data. The large majority of the human capital literature has focused on human capital produced during the formal schooling period. In recognition of a central assumption in this literature - that observable academic measures are reasonably viewed as direct measures of human capital at the time of entrance

to the workforce - column 1 of Table 4 shows a regression of log-wages on college GPA and demographic characteristics. The coefficient on GPA is quantitatively large and statistically significant (t-stat of 3.76). Nonetheless, the fact that the R^2 in column 1 is only 0.014 reinforces the message in much prior research that understanding how human capital accumulates in the post-schooling period is of great importance for understanding the evolution of earnings (Ben-Porath, 1967; Mincer, 1974; Katz and Murphy, 1992; Heckman, Lochner, and Taber, 1998).

Unfortunately, in the post-schooling period, it is typically not possible to observe either inputs into the production of human capital or any direct measures of human capital (Bowlus and Robinson, 2012). In practice, empirical researchers attempt to proxy for a worker's human capital using observable characteristics that are related to wages, with experience measures (overall experience or experience in different occupations) playing a prominent role given their time-varying nature. Column 2 of Table 4 shows that an overall experience measure is significant when added to the wage regression, and the new R^2 is 0.059.

The six measures of task-specific work experience, $h_{it}^s(k), s \in (h, l), k \in (P, I, O)$, described in Section 3.2 allow a more complete and in-depth way to characterize the activities that a person has performed during her career. Column 6 of Table 4 adds these six task measures to the wage regression along with the five measures that characterize the tasks a person is currently performing on the job.[11] The task parameters in the wage equation are measured relative to the omitted task category of low skilled objects tasks, $\tau^l(O)$.[12] Confirming that the new variables constructed from our time allocation measures contain substantial content, the R^2 of the regression that includes the new measures is 0.112 (column 6).

An appealing and unique aspect of both the $\tau^s(k)$ measures and the $h^s(k)$ measures is that they have have very clear interpretations. For example, $\tau^l(P) = 0.20$ means that a person spends 20 percent of her time in year t performing people tasks at a low skill level. Similarly, $h^h(I) = 1.5$ means that a person has spent a total of 1.5 years in the past performing high skilled information tasks.

The estimates related to current tasks shown in column 6 of Table 4 make considerable intuitive sense, with two main takeaways. First, performing objects tasks has a particularly large

---

[11]Quadratic terms for the task-specific experience variables are not included because for these relatively young workers, we did not find evidence of curvature in the experience-wage profile.

[12]There are only five task measures included in the regression because by definition the six "fraction of time" task measures must sum to one for each job.

Table 4: Regression of Log-Wage on Tasks and Human Capital

	OLS						FE
Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(A) Tasks							
$\tau^l(people)$			0.293		0.331	0.350	0.079
			(0.186)		(0.161)	(0.162)	(0.168)
$\tau^h(people)$			0.676		0.664	0.657	0.497
			(0.164)		(0.154)	(0.153)	(0.179)
$\tau^l(information)$			0.515		0.522	0.524	0.413
			(0.189)		(0.173)	(0.172)	(0.159)
$\tau^h(information)$			1.463		1.118	1.103	0.693
			(0.185)		(0.150)	(0.149)	(0.147)
$\tau^h(objects)$			0.530		0.441	0.427	0.167
			(0.213)		(0.183)	(0.183)	(0.202)
(B) Human Capital							
$h^l(people)$				-0.018	0.015	0.019	-0.026
				(0.027)	(0.025)	(0.025)	(0.030)
$h^h(people)$				0.028	0.008	0.008	0.040
				(0.017)	(0.018)	(0.017)	(0.021)
$h^l(information)$				-0.029	-0.021	-0.016	0.052
				(0.044)	(0.041)	(0.041)	(0.043)
$h^h(information)$				0.283	0.199	0.189	0.134
				(0.038)	(0.036)	(0.035)	(0.047)
$h^l(objects)$				-0.114	-0.030	-0.028	0.136
				(0.073)	(0.069)	(0.070)	(0.087)
$h^h(objects)$				0.066	0.084	0.077	0.052
				(0.052)	(0.054)	(0.054)	(0.071)
(C) Other							
experience		0.054					
		(0.005)					
female	-0.089	-0.089				-0.048	
	(0.039)	(0.039)				(0.035)	
black	0.028	0.038				0.008	
	(0.045)	(0.045)				(0.044)	
college GPA	0.158	0.158				0.100	
	(0.043)	(0.042)				(0.040)	
constant	2.139	1.961	1.959	2.442	1.883	1.596	2.079
	(0.133)	(0.132)	(0.152)	(0.021)	(0.134)	(0.186)	(0.134)
R^2	0.014	0.059	0.063	0.080	0.107	0.112	0.090
Individuals	528						
Total observations	3,295						

Notes: Columns (1)-(2) estimated by OLS. Column (7) shows fixed effects (within) estimates.

The task coefficients (panel A) are measured relative to the omitted task $\tau^l(objects)$.

Experience is measured in years.

Standard errors are clustered by individual.

negative association with wages, especially if these tasks are performed at a low skill level. Second, performing tasks at high skill levels rather than low skill levels is associated with large wage increases. To see the quantitative importance of these facts, consider a person who is spending at least half of her current hours working in low skilled objects tasks (i.e, $\tau^l(O) > 0.50$). With respect to the wage penalty for low skilled objects tasks, holding constant the person's past work, moving half of her current work hours from low skilled objects tasks to low skilled people tasks is associated with an increase in earnings of 17.5 percent ($\frac{0.350}{2} \times 100\%$). With respect to the large payoff for high skilled work, moving half of her current work hours from low skilled objects tasks to high skilled people tasks (instead of low skilled people tasks) leads to an almost doubling of this wage increase to 32.8 percent ($\frac{0.657}{2} \times 100\%$). Similarly, moving half of her current work hours from low skilled objects tasks to low skilled information tasks is associated with an increase in earnings of 26 percent ($\frac{0.524}{2} \times 100\%$), and moving half of her current work hours from low skilled objects tasks to high skilled information tasks (instead of low skilled information tasks) leads to a more than doubling of this wage increase to 55 percent ($\frac{1.103}{2} \times 100\%$). As some final evidence of the wage premium for high skilled tasks, the predicted increase in wages from moving half of her current work from low skilled objects tasks to high skilled objects tasks is 21 percent ($\frac{0.427}{2} \times 100\%$).

Turning to the six proxies for task-specific human capital, $h^s(k), s \in (h, l), k \in (P, I, O)$, that describe a person's past activities, perhaps the most striking result in column 6 of Table 4 is the role of accumulated time spent performing high skilled information tasks. The estimated coefficient on $h^h(I)$ of 0.189 suggests that working in high skilled tasks may be very important for creating one's human capital; a worker with one extra person-year of experience performing high skilled information tasks earns 18.9 percent more than a worker with no such work experience, and the associated t-statistic is 5.4. The estimated coefficient on $h^h(O)$ of 0.077 is also quantitatively large, although neither this estimated effect nor the estimated effect of $h^h(P)$ is significant at conventional levels (t-statistics of 1.42 and 0.47, respectively). In contrast, there is no evidence that the amount of time spent performing tasks at low skill levels in the past has any relationship with current wages. The point estimates associated with the three low skilled tasks are economically small ($h^l(P) = 0.019$, $h^l(I) = -0.016$, and $h^l_t(O) = -0.028$) and the associated t-statistics are each below 1.0.

It is worth noting that the OLS estimates are based on a fairly parsimonious specification, as are the fixed effects estimates discussed in Section 4.3. We experimented with richer alternative specifications that allowed for interactions between current tasks ($\tau's$) and task-specific experience ($h's$) variables. To give one example, it is natural to wonder whether or not the return to the cumulative amount of time spent performing a particular task in the past depends on the current task requirements of a job (Gathmann and Schönberg, 2010). Unfortunately, given our relatively small sample size and the large number of potential interactions that could exist between the eleven current task and task specific experience variables shown in Table 2, we did not find these specifications to be particularly informative. Thus, roughly speaking, our results indicate the overall effects of accumulating experience in particular tasks in the past, but do not provide information about whether heterogeneity in these effects might arise from differences in current tasks.

4.2 Conceptual Framework: Wages and Tasks

Given the strong relationships found in Section 4.1, it is worthwhile to use a simple conceptual framework to discuss issues of interpretation that arise when attempting to estimate the returns to job tasks. The value added of the BPS tasks measures is highlighted by enriching the conceptual model of Autor and Handel (2013), who stress the potential importance of taking into account the role of comparative advantage, as in the extensive literature that considers estimation and identification of the Roy (1951) model.[13] In particular, we adapt the Autor and Handel (2013) framework to incorporate the direct measures of time spent on tasks that are a novel feature of the BPS data.

Comparative advantage arises when workers have different efficiencies (i.e., abilities) at performing each of the K job tasks present in the labor market. To help fix ideas, note that our empirical application has K = 6 tasks corresponding to high and and low skilled People, Information, and Objects tasks.[14] The efficiency endowment of worker i is $\Theta = \{\theta_{i1}, ..., \theta_{iK}\}$, where, as in Autor and Handel (2013), θ_{ik} is the amount of task k that person i can complete in a given time interval (e.g., per hour).[15] Each job j requires a worker to devote a specified amount of time

---

[13]See, for example, (Heckman and Sedlacek, 1985; Heckman and Honoré, 1990; Heckman and Sedlacek, 1990; Gould, 2002; Heckman and Taber, 2010)

[14]To simplify notation, in this section we do not explicitly distinguish between high and low skilled tasks.

[15]We initially abstract away from worker learning about ability to perform different job tasks, but return to this

τ_{jkt} to each of the $k = 1, ..., K$ job tasks in time period t. Thus, the total units of job task k that are performed by person i in time t, which we often refer to as the total number of efficiency units of task k, is given by $\tau_{jkt}\theta_{ik}$. Letting p_k represent the productivity of each efficiency unit of job task k, letting ε_{ijt} represent an idiosyncratic shock to output, and allowing each worker to possess an endowment of general ability μ_i reflecting differences in productivity that are not directly tied to job tasks, the output of worker i in job j in time period t is

$$y_{ijt} = exp\left(\mu_i + \sum_{k=1}^{K} p_k[\tau_{jkt}\theta_{ik}] + \varepsilon_{ijt}\right). \tag{1}$$

Assuming workers are paid their marginal product, the log wage is

$$\log(w_{ijt}) = \mu_i + \sum_{k=1}^{K} p_k[\tau_{jkt}\theta_{ik}] + \varepsilon_{ijt}. \tag{2}$$

The majority of the discussion in this section assumes, in effect, that the task efficiencies (θ_{ik}) are: 1) constant across time and 2) known by both the worker and by employers. Regarding the first assumption, Appendix B expands our conceptual framework to explore how time-varying task efficiencies could arise because of learning-by-doing. Regarding the second assumption, one potential concern is that individuals and employers could learn about comparative advantage (θ_{ik}) over time as workers perform different tasks.[16] Like much previous literature, we do not dwell on this issue. However, it is important to acknowledge that the possibility of this type of learning suggests a need to be cautious when interpreting our estimates, and in the remainder of this section we provide further discussion about this issue where relevant.

A standard objective of empirical work is to estimate the market price of each task in the labor market, p_k, $k = 1, ..., K$. This requires an assumption that p_k is constant across jobs. While Autor and Handel (2013) describe the reasons that the assumption of constant prices is not generically appealing, two unique features of our data are crucial for increasing its plausibility as a rough

---

important issue at the end of the section. Sanders (2017) uses the O*NET data to estimate a structural model of occupational choice in which workers learn about their skills at performing job tasks.

[16]Gibbons et al. (2005) provide a detailed analysis of an empirical model of learning and comparative advantage within the context of sectoral choice. It is also worth noting that estimating an empirical model of learning is quite difficult in the present context, where there are six separate task efficiencies that individuals could potentially learn about over time. Specifically, one efficiency parameter for each of the three task categories (People, Information, and Objects) at both skill levels (high and low).Sanders (2017) estimates a structural model that includes both task-specific human capital and learning about ability to perform tasks. Guvenen et al. (2017) investigate the extent of skill mismatch using a framework that incorporates occupation-level job task data and learning about skills.

approximation in our environment. First, our collection of time allocation measures allows τ_{jkt} to appear explicitly in the efficiency unit measure, $\tau_{jkt}\theta_{ik}$. To see why this is important for the constant price assumption, note that the situation without explicit time measures is consistent with the standard data sources analyzed in the growing empirical literature on tasks. When time allocation data is not available, the unobserved time spent on a task, τ_{jkt}, is subsumed into the "price" of θ_{ik}. Specifically, as in Autor and Handel (2013), whose model takes into account the realities of traditional data, wages depend on $\sum_{k=1}^{K} p_k^* \times \theta_{ik}$, where $p_k^* = p_k\tau_{jkt}$. This implies that in the absence of measures of time spent on tasks, the "price" (p_k^*) must by definition vary across jobs to the extent that different jobs involve spending varying amounts of time on each task.

The second feature of the BPS data that increases the plausibility of the constant price assumption is that the data clearly identifies the skill level at which a task is performed. Thus, a particular price in Equation 2 corresponds to a particular task being performed at a specific skill level. This is helpful because, for example, two identical workers spending exactly the same amount of time on people tasks would likely be paid differently if one was primarily a supervisor (high skill) while the other was primarily being supervised (low skill).

A large body of research on the determination of wages has viewed ability as one-dimensional (i.e., $\theta_{ik} = 1 \, \forall \, (i, k)$). The primary identification concern expressed in this literature is that explanatory variables may be correlated with general ability (μ_i). A concern about bias arising through this avenue is relevant for our regressor of interest τ_{jkt}; a variety of unobserved employer and employee considerations might lead high ability workers to end up in different types of jobs than lower ability workers. In this setting, a natural approach is to take advantage of a fixed effects estimator. However, data limitations have limited the usefulness of this approach in the task context. Given that tasks are not measured at the level of individual jobs in existing longitudinal data, researchers typically assign all workers in a particular occupation the same tasks from the DOT or O*NET. However, in this scenario, changes in tasks will be ruled out as an explanation for any of the substantial changes in earnings that are known to exist for a person within an occupation. Further, existing task data is frequently characterized in a qualitative manner, and it is not obvious how to interpret "differenced" ordinal measures. Thus, our collection of the first longitudinal, individual-level task information, when combined with our new time

allocation measures, provides a unique opportunity to implement a fixed effects estimator in this context.

The potential benefits of differencing are also readily apparent when we return to a scenario where θ_{ik} varies across i and k, so that task efficiencies are multi-dimensional and heterogeneous across individuals. A primary identification concern in this comparative advantage scenario is that individuals with a higher efficiency at task k (i.e., higher values of θ_{ik}) will tend to select into jobs in which they spend more time performing task k (i.e., jobs with higher values of τ_{jkt}). This implies that workers who spend different amount of times on task k would not likely receive the same wages even if they spent the same amount of time on task k. As a result, given that θ_{ik} is unobserved, the OLS estimator for the effect of τ_{jkt} on wages, which exploits this cross-sectional variation, will suffer from an omitted variables bias.

The previous discussion suggests the potential benefits of utilizing the within-person variation that the fixed effects estimator is based on. Consider a version of our model in which there are two time periods and a single task ($K = 1$). The difference in wages between two time periods t and t' for a particular person is given by $\log(w_{ijt}) - \log(w_{ijt'}) = p_k \theta_{ik}(\tau_{jkt} - \tau_{jkt'}) + (\varepsilon_{ijt} - \varepsilon_{ijt'})$. Under the assumption of exogenous mobility presented by Abowd et al. (1999), $E(\varepsilon_{ijt} - \varepsilon_{ijt'} | \tau_{jkt} - \tau_{jkt'}) = 0$, and the average value of $(\log(w_{ijt}) - \log(w_{ijt'}))/(\tau_{jkt} - \tau_{jkt'})$ across all workers is equal to $p_k E(\theta_{ik})$. Note that, unlike what is seen with fixed effects in the one-dimensional ability scenario, fixed effects in the comparative advantage scenario does not remove the influence of θ_{ik} altogether. Nonetheless, the parameter $p_k E(\theta_{ik})$ is of substantial interest, representing the average return, or effect, of having a person spend more time on task k.

It is worth noting that this interpretation would become more complicated if we moved away from the assumption that the task efficiencies θ_{ik} are known by both the worker and by employers in all periods, because in this case an individual's wage would depend on what the market believes about θ_{ik} at a particular point in time. Further, in this type of learning environment, changes in beliefs about θ_{ik} could potentially create a spurious correlation between the within-person changes in wages and the within-person changes in tasks that are utilized by the fixed effects estimator. For example, this would be the case if, when the market learns that a worker's ability to perform a particular task is higher than previously believed, the worker chooses to

perform more of this task and employers pay the worker more for all units of this task that she performs.

As in Abowd et al. (1999), the exogenous mobility assumption does not imply that individuals move between jobs in an unconditionally random manner. Rather, it is an assumption that an individual decides whether to accept a new job offer by considering the tasks she would perform (τ_{ikt}, $k = 1, ..., K$), how good she is at performing these tasks (θ_{ik}, $k = 1, ..., K$), and the prices (p_k, $k = 1, ..., K$). This assumption is consistent with much recent research (see, e.g., Yamaguchi (2012) and Autor and Handel (2013)), which implicitly assumes that selection due to comparative advantage into job tasks is of primary importance.[17],[18] Nonetheless, as always, the interpretation that follows from the exogenous mobility assumption should be viewed with an appropriate degree of caution given that this assumption likely will not be literally true.

The interpretation of the fixed effects estimator as the average return to spending more time on a task depends critically on whether agents are observed performing each task during their careers. This message is front and center in the literature interested in identifying and estimating models of comparative advantage. In the typical context where comparative advantage is relevant, workers will frequently not be observed in all sectors. For example, in an occupational choice context, workers are usually observed working in only white collar occupations or only blue collar occupations (but not both). Roughly speaking, in this situation, a fixed effects estimator will capture the average return for only the subset of individuals who participate in both sectors. This is problematic because the average return for this subset of workers may be very different from the average return for the overall population. Thus, in a typical comparative advantage scenario, researchers are forced to specify and estimate the process by which selection takes place.

In contrast, the view of our fixed effects estimator as an informative average return is supported by our descriptive results in Section 3, which indicate that the large majority of workers in our data are observed spending time on all tasks (and have changes over time in these tasks). We note that our ability to uncover the degree to which workers participate in different "sectors"

---

[17]Exceptions include Sullivan (2010a), and Taber and Vejlin (2013). Bonhomme et al. (2016) describe the central role of the exogenous mobility assumption in recent research and develop an empirical framework based on matched worker-firm data that allows for endogenous mobility.

[18]Under the exogenous mobility assumption, even if tasks tend to be systematically related to unobserved wage determinants of jobs ($\varepsilon's$) in the population, our estimates would have a useful interpretation; they would capture the average total wage effect of moving to a job with different tasks rather than capturing a pure task effect.

comes from the nature of our unique time allocation measures. As such, a contribution of our project comes from the role that our data collection efforts play in allowing a useful interpretation of an estimator that does not require parameterizing and directly estimating the selection process.[19]

4.3 Fixed Effects Wage Equation Estimates

In this section we do not dwell on whether our fixed effects results should be interpreted under the one-dimensional ability scenario or the comparative advantage scenario that were discussed in Section 4.2. Instead we simply keep in mind that, subject to the caveats/assumptions stated in Section 4.2, in either of these scenarios a large estimated effect of tasks suggests that what a person does on her job plays an important role in wage determination.

Column 7 of Table 4 shows person-level fixed effects (within) estimates of the log wage equation. Beginning with the current task information, a 37 percent decrease in the coefficient on $\tau^h(I)$ is observed as we move from OLS (Column 6) to fixed effects (Column 7). This decrease is consistent with the hypothesis that the OLS estimator is upward biased due to a positive correlation between unobserved worker ability and the tendency to perform high skilled information tasks. Nonetheless, the estimated wage effect of working in a job with high skilled information tasks remains quantitatively large and statistically significant. The estimated coefficient on $\tau^h(I)$ of 0.693 is the largest of the five task coefficients and the associated t-statistic is greater than 2.5. Similarly, the fixed effects estimated coefficients on $\tau^h(P)$ and $\tau^l(I)$ are 24 percent and 21 percent smaller than the OLS estimates, but still indicate economically large and statistically significant wage premia for performing high skilled people tasks and low skilled information tasks relative to low skilled objects tasks (point estimates of 0.497 for $\tau^h(P)$ and 0.413 for $\tau^l(I)$, and p-values of 0.006 and 0.010). Perhaps the biggest differences between the fixed effects and OLS results are seen in the remaining two variables, with both the coefficient on $\tau^l(P)$ and the coefficient on $\tau^h(O)$ falling by over 50 percent. Nonetheless, the general takeaway from the "current tasks" estimates in column 7 of Table 4 is similar to that found using OLS, despite the fact that the fixed effect estimates are based solely on within-person variation in tasks and wages.

---

[19]Of course, directly modeling the difficult selection problem has the benefit of potentially allowing one to examine a variety of population parameters other than the parameter discussed here (Heckman et al., 1998; Keane and Wolpin, 1997).

The spirit of our results on the importance of task-specific work experience is also remarkably robust to controlling for fixed effects. Similar to what was seen with $\tau^h(I)$, a decrease of 30 percent in the estimated coefficient on $h^h(I)$ is observed as we move from OLS to fixed effects. Nonetheless, the estimated effect of accumulating experience in high skilled information tasks remains quantitatively and statistically important, with one person-year of $h^h(I)$ being estimated to increase wages by approximately 13 percent and the associated t-statistic having a value of 2.85. The fixed effects estimates also reveal that accumulating time in high skilled people tasks increases future wages, with $h^h(P)$ having an estimated coefficient of 0.040 and a p-value of 0.059. Also similar in spirit to the OLS results, none of the estimated low skill task-specific experience coefficients are statistically different from zero at conventional levels, although the point estimate for $h^l(O)$ is quite large but imprecisely estimated. Thus, the fixed effects estimates support the intuitively appealing notion that there is considerable scope for workers to become more productive through learning-by-doing as they perform high skilled tasks, with accumulated experience in high skilled information tasks being particularly important. However, there is little evidence that performing low skilled tasks today is a path to higher future wages.

4.4 Decomposing Wage Changes: Current Tasks vs. Task-Specific Experience

A number of influential papers have demonstrated the value of the task approach to labor markets by using factor analysis to merge occupation-level task data with individual-level occupational choice and wage data (Poletaev and Robinson, 2008; Yamaguchi, 2012). In this approach, the data describing occupations in a source such as the DOT is distilled into a small number of factors (such as cognitive and manual). Higher factor scores are typically interpreted as more "complex" tasks. However, within this framework it is inherently difficult to identify concrete changes in job activities that affect wages.[20] In addition, there is no natural way to directly translate the factor scores into a cumulative measure of tasks performed over the career. In contrast, the novel structure of the BPS data allows us to quantify the wage effects of changing the amount of time spent performing well-defined tasks over the career. Further, we decompose the wage effects into two components: 1) human capital accumulation (past tasks), and 2) the direct (current-period) effect of tasks on wages.

---

[20]Of course, applying factor analysis in this manner is an excellent way to distill the very high dimensional DOT data into information that has been shown to be quite useful in empirical work.

Table 5: Decompositions of Wage Changes from Shifting Time Spent on Tasks

Permanent Shifts from 25th to 75th Percentile of High Skilled Task

Row	(1) Task Shift	(2) Inter-Quartile Range	Components of Year Ten Log-Wage Change		
			(3) Current Tasks	(4) Human Capital	(5) Total
1	$\tau^l(P) \to \tau^h(P)$	0.239	0.100	0.158	0.258
2	$\tau^l(I) \to \tau^h(I)$	0.160	0.045	0.130	0.175
3	$\tau^l(O) \to \tau^h(O)$	0.100	0.017	-0.084	-0.067

Notes: Each row shows the wage change resulting from the shift in tasks listed in column 1. Column 2 lists the amount of time that is shifted between tasks. The hypothetical change in tasks occurs in each year, over a ten year period. Column 5 shows the total log-wage change in year 10: $w(t = 10|task\ shift) - w(t = 1)$. Columns 3 and 4 decompose the total change into the Current Task ($\tau's$) and Human Capital ($h's$) components. All log-wage changes are based on the fixed effects estimates shown in column 7 of Table 2.

The results in column 7 of Table 2 can be broadly summarized by three main findings: 1) current job tasks play a large role in determining wages, with high skilled tasks being paid roughly double the rate of low skilled tasks, 2) there is no evidence of learning-by-doing for low skilled tasks, but there is strong evidence of learning-by-doing for high skilled tasks, and 3) not all types of high skilled tasks play the same role in the determination of wages, with current and past high skilled people tasks being important, but current and past high skilled information tasks playing an especially prominent role. Table 5 puts these results into further "lifecycle" context by examining how hypothetical changes in tasks over the first ten years of a worker's career would affect the worker's log wage in the tenth year. When describing the table, it is useful to first define new notation for the estimated returns to both current tasks and task specific experience. In particular, let $\beta_\tau^s(k)$, $s \in (h, l)$, $k \in (P, I, O)$ represent the fixed effects estimate of the returns to time spent on tasks (τ).[21] Analogously, let $\gamma_h^s(k)$, $s \in (h, l)$, $k \in (P, I, O)$ represent the fixed effects estimate of the returns to task-specific experience (h).

The main findings 1) and 2) above, which highlight the importance of past and present high skilled tasks, are apparent in the three rows of Table 5. The first row shows the change in the log wage in the tenth year for a worker who, in each of the first ten years, experiences a hypothetical move from the 25th to the 75th percentile of the distribution of high skilled people tasks.[22] Row

---

[21]Recall that the coefficient on low skilled objects tasks is normalized to zero.

[22]Expressing the change in tasks in terms of percentiles is useful because column 2 of Table 5 reveals that there are fairly large differences in the inter-quartile ranges across the high-skilled people, information, and object task categories. In other words, what one might consider a large change in job tasks varies considerably across the task categories.

1, column 2 shows that the movement from the 25th to the 75th percentile involves increasing $\tau^h(P)$ by 0.239 units in each of the ten years, and we assume that this time is shifted away from low skilled people tasks (thus, the notation $\tau^l(P) \to \tau^h(P)$ used in row 1, column 1 of the table). This re-allocation of time from low skilled to high skilled people tasks impacts wages through two distinct channels shown in columns 3 and 4 of the table. On the one hand, current tasks affect wages directly through the task component of the wage equation, $\sum \beta^s_\tau(k)\tau^s(k)$ (column 3). On the other hand, changing current-period tasks alters a worker's accumulated time spent on tasks, which affects wages through the human capital portion of the wage equation, $\sum \gamma^s_h(k)h^s(k)$ (column 4). Column 3 of Table 5 shows that the "Current Tasks" component of the log wage change is equal to $0.239 \times (\beta^h_\tau(P) - \beta^l_\tau(P)) = 0.100$, where 0.239 is the per-year increase in $\tau^h(P)$ and $(\beta^h_\tau(P) - \beta^l_\tau(P))$ is the estimated premium for performing high skilled people tasks over low skilled people tasks. Column 4 shows the "Human Capital" component of the log wage change stemming from the returns to task-specific work experience. This component of the change in log wages is equal to $0.239 \times (\gamma^h_h(P) - \gamma^l_h(P)) \times 10 = 0.158$, where 0.239 is the per-year change in tasks, $(\gamma^h_h(P) - \gamma^l_h(P))$ is the premium for accumulated experience performing high skilled people tasks, and 10 is the number of years over which tasks change. Finally, column 5 of Table 5 shows that a permanent move from the 25th to the 75th percentile of $\tau^h(P)$ results in a total increase in wages of 25.8%.

The second row of Table 5 quantifies the effect of shifting time from low to high skilled information tasks $(\tau^l(I) \to \tau^h(I))$. Moving from the 25th to the 75th percentile of high skilled information tasks requires increasing $\tau^h(I)$ by 0.160. This change is 25 percent smaller than the change in time required to move between the same percentiles of the distribution of high skilled people tasks (row 1). The current tasks component of the wage increase is 0.045 log-points (column 3), and the task-specific experience component of the wage change is 0.130 (column 4). The total wage change, shown in column 5, is equal to 0.175, which again suggests that performing high skilled tasks is a key determinant of earning high wages. The third row of Table 5 decomposes the wage effects of shifting time from low to high skilled objects tasks $(\tau^l(O) \to \tau^h(O))$. The total log wage change resulting from this shift is -0.067, although it is worth noting that the estimated object task and human capital parameters in column 7 of Table 4 are by and large imprecisely estimated.[23]

---

[23]Despite our third main finding - that current and past high skilled information tasks play an especially prominent

5 Conclusion

This paper presents new empirical evidence on the relationship between job tasks and wages by taking advantage of novel data that allow a natural way to characterize current and past job activities at a primitive level. The data from the Berea Panel Study represent the first longitudinal dataset containing job-level task information for individual workers. Further, novel quantitative task measures detail the amount of time spent performing People, Information, and Objects tasks at different skill levels. We find that current job tasks are quantitatively important, with high skilled tasks being paid double the rate of low skilled tasks. There is no evidence of learning-by-doing for low skilled tasks, but strong evidence for high skilled tasks. Current and past high skilled information tasks are particularly valuable. Our new methods for the collection of task information are of importance given that a variety of recent data initiatives have been motivated, in part, by a desire to collect better measures of what workers are doing on their jobs, .

---

role in the determination of wages - the total effect in row 2 (information) is smaller than the total effect in row 1 (people) because low skilled information tasks pay substantially more than low skilled people tasks. In results not shown, a log wage increase of 0.113 arises in the tenth year when, in each of the first ten years, a worker experiences a hypothetical increase in $\tau^h(I)$ of 0.10 and an accompanying decrease in $\tau^h(P)$ of 0.10 ($\tau^h(P) \to \tau^h(I)$). Similarly, we find that a shift of 10 percent of work time from low skilled people tasks to high skilled information tasks ($\tau^l(P) \to \tau^h(I)$) leads to a total log-wage increase of 0.221.

Appendix A: Job Task Survey Questions

Question C: How does your JOB1 require you to relate to PEOPLE and INFORMATION?

- Question C1: Below are 4 ways that you may interact with PEOPLE on a job.

  1. Following instructions from others such as supervisors or directly serving the needs of customers or animals.
  2. Persuading others about a company product/service or point of view (e.g. sales) or entertaining others.
  3. Supervising others or instructing/teaching others.
  4. Exchanging ideas/information/opinions or negotiating with others to make decisions or formulate policies.

  - Think about the TOTAL time that you spend **interacting with PEOPLE** as part of your JOB1. Indicate what percentage of the total time is spent interacting in each of the four ways. **Note:** Each percentage should be between 0 (the item plays no role) and 100 (all interactions are from the one item) **and the four items should sum to 100.**

- Question C2: Below are 4 ways that you may interact with INFORMATION on a job.

  1. Entering data; typing documents written by others; posting information etc.
  2. Gathering or classifying information/data and performing simple calculations using data.
  3. Analyzing data/information
  4. Using data analysis done by yourself/others to develop knowledge/solutions and make important decisions.

  - Think about the TOTAL time that you spend **interacting with INFORMATION** as part of your JOB1. Indicate what percentage of the total time is spent interacting in each of the four ways. **Note:** Each percentage should be between 0 (the item plays no role) and 100 (all interactions are from the one item) **and the four items should sum to 100.**

- Question C3: Below are 4 ways that you may interact with OBJECTS on a job.

  1. Working with or moving objects or operating a machine in a way that requires only a small amount of judgment.
  2. Working with or moving objects or operating a machine in a way that requires a moderate amount of judgment.
  3. Working with or moving objects or operating a machine in a way that requires a large amount of judgment.
  4. Working with or moving objects in a way that judgment is extremely important; or having full responsibility for planning or setting up machines or processes.

  - Think about the TOTAL time that you spend **interacting with OBJECTS** as part of your JOB1. Indicate what percentage of the total time is spent interacting in each of the four ways. **Note:** Each percentage should be between 0 (the item plays no role) and 100 (all interactions are from the one item) **and the four items should sum to 100.**

- Question C4: Now think about your TOTAL job responsibilities on your JOB1. Indicate the percentage of your responsibilities that involve interacting with PEOPLE, INFORMATION, and OBJECTS, respectively. Each percentage should be between 0 and 100 and the three percentages should sum to one.

Appendix B: Extended Conceptual Framework

This appendix expands the conceptual framework presented in Section 4.2 to consider how past tasks, in addition to current tasks, might affect wages. Taking advantage of the explicit time allocation measures in our data, we consider a framework where past tasks are summarized by accumulated task-specific work experience. Specifically, the output of worker i in job j in time period t is

$$y_{ijt} \;=\; exp\left(\mu_i + \sum_{k=1}^{K} p_k[\tau_{jkt}f_k(\theta_{ik}, h_{ikt}, \Gamma)] + \sum_{k=1}^{K} \beta_k h_{ikt} + \varepsilon_{ijt}\right). \tag{3}$$

This equation expands on the framework presented in Section 4.2 in two ways. First, and most straightforwardly, it allows task-specific work experience to directly affect wages through the term $\sum_{k=1}^{K} \beta_k h_{ikt}$. Second, it allows the worker's efficiency endowment at performing task k, $f_k(\theta_{ik}, h_{ikt}, \Gamma)$, to depend on initial ability, θ_{ik}, accumulated task-specific work experience, h_{ikt}, and a vector of parameters, Γ.

In the absence of data limitations, in principal it would be possible to estimate a flexible functional form for the relationship between accumulated task-specific experience and task efficiency $(f_k(\cdot))$. However, given the high dimensional nature of our task measures, and the limited sample size, it seems reasonable to consider a linear relationship, $f(\theta_{ik}, h_{ikt}, \Gamma) = \theta_{ikt} + \gamma_k h_{ikt}$, where the parameter γ_k reveals how accumulated experience at task k affects the worker's efficiency at performing task k. Under this specification equation 3 becomes

$$\begin{aligned} y_{ijt} &= exp\left(\mu_i + \sum_{k=1}^{K} p_k[\tau_{jkt}f_k(\theta_{ik}, h_{ikt}, \Gamma)] + \sum_{k=1}^{K} \beta_k h_{ikt} + \varepsilon_{ijt}\right) \\ &= exp\left(\mu_i + \sum_{k=1}^{K} p_k[\tau_{jkt}(\theta_{ikt} + \gamma_k h_{ikt})] + \sum_{k=1}^{K} \beta_k h_{ikt} + \varepsilon_{ijt}\right) \\ &= exp\left(\mu_i + \sum_{k=1}^{K} p_k\tau_{jkt}\theta_{ikt} + \sum_{k=1}^{K} p_k\gamma_k(\tau_{jkt}h_{ikt}) + \sum_{k=1}^{K} \beta_k h_{ikt} + \varepsilon_{ijt}\right), \end{aligned} \tag{4}$$

which yields a wage equation in which log-wages depend on current job tasks $(\tau_{jkt})_{k=1}^{K}$, accu-

mulated task-specific experience $(h_{ikt})_{k=1}^K$, and their interactions $(\tau_{jkt}h_{ikt})_{k=1}^K$,

$$\log(w_{ijt}) = \mu_i + \sum_{k=1}^K p_k \theta_{ikt}(\tau_{jkt}) + \sum_{k=1}^K \gamma_k p_k(\tau_{jkt}h_{ikt}) + \sum_{k=1}^K \beta_k h_{ikt} + \varepsilon_{ijt}.$$

As discussed in Section 4.1, we estimated several versions of this model, and did not find specifications allowing for interactions between current tasks and task specific experience to be particularly informative. Perhaps this is not surprising, given the high dimensional nature of the task measures. In addition, the specification of our wage equations is constrained by the relatively small sample size of 528 individuals, with an average 6.24 observations per individual. However, if future data collection efforts obtain similar task measures on a larger scale, it would certainly be interesting to estimate more general relationships between current tasks, past tasks, and wages.

References

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999). High wage workers and high wage firms. *Econometrica 67*(2), 251–333.

Acemoglu, D. and D. Autor (2011). Skills, tasks and technologies: Implications for employment and earnings. *Handbook of labor economics 4*, 1043–1171.

Arcidiacono, P. (2004). Ability sorting and the returns to college major. *Journal of Econometrics*, 1–45.

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

Autor, D. H. and M. J. Handel (2013). Putting Tasks to the Test: Human Capital, Job Tasks, and Wages. *Journal of Labor Economics 31*(2), S59–S95.

Autor, D. H., L. F. Katz, and M. S. Kearney (2009). The Polarization of the U.S. Labor Market. *The American Economic Review 96*(2), 189–194.

Autor, D. H., F. Levy, and R. J. Murnane (2003). The Skill Content of Recent Technological Change: An Empirical Exploration. *Quarterly Journal of Economics 118*(4), 1279–1333.

Becker, G. S. (1964). Human capital theory. *Columbia, New York*.

Ben-Porath, Y. (1967). The Production of Human Capital and the Life Cycle of Earnings. *The Journal of Political Economy 75*(4), 352 – 365.

Bonhomme, S., T. Lamadon, and E. Manresa (2016). A distributional framework for matched employer employee data. *Working Paper*.

Bowlus, A. J. and C. Robinson (2012). Human capital prices, productivity, and growth. *The American Economic Review*, 3483–3515.

Gathmann, C. and U. Schönberg (2010). How general is human capital? a task-based approach. *Journal of Labor Economics 28*(1), 1–49.

Gibbons, R. and L. Katz (1992). Does unmeasured ability explain inter-industry wage differentials? *The Review of Economic Studies 59*(3), 515–535.

Gibbons, R., L. Katz, T. Lemieux, and D. Parent (2005). Comparative Advantage, Learning, and Sectoral Wage Determination. *Journal of Labor Economics 23*, 681–724.

Gibbons, R. and M. Waldman (2004). Task-Specific Human Capital. *American Economic Review: Papers and Proceedings 94*(2).

Gould, E. (2002). Rising Wage Inequality, Comparative Advantage, and the Growing Importance of General Skills in the United States. *Journal of Labor Economics 20*(1), 105–147.

Green, F., D. Gallie, A. Felstead, and Y. Zhou (2006). Guide to the 2006 skills survey. https://www.kent.ac.uk/economics/documents/gfg/GUIDE%20TO%20THE%202006%20SKILLS%20SURVEY_revised.pdf.

Guvenen, F., B. Kuruscu, S. Tanaka, and D. Wiczer (2017). Multidimensional skill mismatch. *Working Paper*.

Heckman, J., H. Ichimura, J. Smith, and P. Todd (1998). Characterizing selection bias using experimental data. *Econometrica 66*(5), 1017–1098.

Heckman, J. J. and B. E. Honoré (1990). The Empirical Content of the Roy Model. *Econometrica 58*(5), 1121–1149.

Heckman, J. J., L. Lochner, and C. Taber (1998). Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents. *Review of economic dynamics 1*(1), 1–58.

Heckman, J. J. and G. Sedlacek (1985). Heterogeneity, aggregation, and market wage functions: an empirical model of self-selection in the labor market. *The journal of political economy*, 1077–1125.

Heckman, J. J. and G. L. Sedlacek (1990). Self-Selection and the Distribution of Hourly Wages. *Journal of Labor Economics 8*(1), S329–S363.

Heckman, J. J. and C. Taber (2010). Roy model. *Microeconometrics, New Palgrave Economics Collection*.

Kambourov, G. and I. Manovskii (2009). Occupational specificity of human capital*. *International Economic Review 50*(1), 63–115.

Katz, L. F. and K. M. Murphy (1992). Changes in Relative Wages, 1963-1987: Supply and Demand Factors. *Source: The Quarterly Journal of Economics 107*(1), 35–78.

Keane, M. and K. Wolpin (1997). The career decisions of young men. *Journal of political Economy 105*(3), 473–522.

Lazear, E. (2009). Firm-Specific Human Capital: A Skill Weights Approach. *Journal of Political Economy 117*(5), 914–940.

Lise, J. and F. Postel-Vinay (2015). Multidimensional Skills, Sorting, and Human Capital Accumulation. *Working Paper* (2002), 1–40.

Manning, A. (2003). *Monopsony in motion: Imperfect competition in labor markets*. Princeton University Press.

Mincer, J. (1974). *Schooling, Experience, and Earnings*. NBER.

Neal, D. (1995). Industry-specific human capital: Evidence from displaced workers. *Journal of labor Economics*, 653–677.

Parent, D. (2000). Industry-specific capital and the wage profile: Evidence from the national longitudinal survey of youth and the panel study of income dynamics. *Journal of Labor Economics 18*(2), 306–323.

Poletaev, M. and C. Robinson (2008). Human capital specificity: evidence from the dictionary of occupational titles and displaced worker surveys, 1984–2000. *Journal of Labor Economics 26*(3), 387–420.

Robinson, C. (2011). Occupational mobility, occupation distance and specific human capital. Technical report, CIBC Working Paper.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford economic papers 3*(2), 135–146.

Roys, N. and C. Taber (2016). Skills prices, occupations and changes in the wage structure.

Sanders, C. (2017). Skill Accumulation, Skill Uncertainty, and Occupational Choice. *Working Paper*, 1–44.

Sanders, C. and C. Taber (2012). Life-cycle wage growth and heterogeneous human capital. *Annual Review of Economics 4*(1), 399–425.

Speer, J. D. (2016). Pre-market skills, occupational choice, and career progression. *Journal of Human Resources*.

Stinebrickner, R. and T. Stinebrickner (2008). The effect of credit constraints on the college drop-out decision: A direct approach using a new panel study. *American Economic Review 98*(5), 2163–84.

Stinebrickner, R. and T. Stinebrickner (2014). A major in science? initial beliefs and final outcomes for college major and dropout. *The Review of Economic Studies*, rdt025.

Stinebrickner, R. and T. R. Stinebrickner (2003). Working during school and academic performance. *Journal of Labor Economics 21*(2), 473–491.

Stinebrickner, R. and T. R. Stinebrickner (2006). What can be learned about peer effects using college roommates? evidence from new survey data and students from disadvantaged backgrounds. *Journal of public Economics 90*(8), 1435–1454.

Stinebrickner, T. and R. Stinebrickner (2012). Learning about Academic Ability and the College Dropout Decision. *Journal of Labor Economics 30*, 707–748.

Sullivan, P. (2010a). A Dynamic Analysis of Educational Attainment, Occupational Choices, and Job Search. *International Economic Review 51*(1), 289–317.

Sullivan, P. (2010b). Empirical Evidence on Occupation and Industry Specific Human Capital. *Labour economics 17*(3), 567–580.

Taber, C. and R. Vejlin (2013). Estimation of a Roy / Search / Compensating Differential Model of the Labor Market.

Yamaguchi, S. (2012). Tasks and heterogeneous human capital. *Journal of Labor Economics*, 0–55.

Yamaguchi, S. (2015). Changes in returns to task-specific skills and gender wage gap. *Available at SSRN 2035833*.