Which Jobs Scale and Why?*

Seth Benzell
Chapman University, Stanford University, and Massachusetts Institute of Technology

Erik Brynjolfsson
Stanford University and NBER

Ruyu Chen
Stanford University

September 2024

Abstract

While some jobs, like CEOs', involve tasks that can change the productivity of the entire organization, others, like piece-rate manufacturing workers, do not. In this paper, we measure which jobs scale and why using a sample of large-scale administrative data from the payroll processing company ADP, covering between 14.7 and 16.0 million employees in each year at 168,371 firms across 308,960 establishments from 2017 to 2023. The median worker in a occupation-firm-year is paid 8.6% more at an firm with a payroll twice as large. To evaluate different theories of this premium, we investigate heterogeneity in wage scaling by job characteristics. We find wages scale the most with firm size for workers in managerial, decision making, and abstract tasks, and less for workers in routine and manual tasks. Wage scaling is also stronger for workers who are at high percentiles of wage in their occupation-organization-year and in IT intensive firms. IT intensity at firms is related to larger wage scaling for abstract and top percentile workers, and smaller scaling for routine, manual and low-percentile workers, effects which intensify the direct effect of scaling for these jobs. Taken together, these results are most consistent with 'Span-of-Control'-driven wage scaling. This implies that the disproportionate growth of management-intensive occupations and the rise of large and IT intensive firms will tend to increase aggregate wage polarization.

1 Introduction

Tasks can be categorized as 'scalable', meaning the marginal product of the task increases with the size of the business, or 'non-scalable', meaning it does not. Jobs will differ by the proportion of value they create through scalable and non-scalable tasks. Car windshield installers create value mostly through non-scaling tasks like installing car windshields. Alternatively, CEOs mostly create value by making decisions that impact an entire business. A higher marginal product for setting corporate strategy at a large firm should lead to a higher wage. Indeed, Gabaix and Landier (2008) and Baker and Hall (2004) have analyzed the relationship of CEO compensation to firm size to understand the degree to which CEO talent and effort scale.

But it is not only CEOs that create value through scalable tasks. In this paper we use large-scale administrative data to measure heterogeneity in wage scaling with organization size by job characteristic. The data comes from the payroll processing company ADP. It covers over 15 million employees every year across 168,371 firms at 308,960 establishments from 2017 to 2023.

Controlling for year, industry and occupational fixed effects, a median worker is paid an average of 8.6% more at an firm twice as large. This finding of a large average size-wage premium is not novel.

*We are grateful to our data partner (ADP) for sharing data with us. We gratefully acknowledge feedback and comments from participants at the Stanford Digital Economy Lab Seminar and the Workshop on Information Systems Economics 2024. This research was supported by the Stanford Digital Economy Lab.

In fact, the "large-firm wage premium" was first identified in Moore (1911). However, our unique data set allows us to go further, to test and refine theories of the large-firm wage premium.

We look across jobs to determine which are providing scalable tasks, as measured by the large-firm wage premium. The effect of firm size on wage is larger for workers at a higher wage-percentile in their firm-occupation-year. While median worker in an firm-occupation-year has a 8.6% higher wage in a firm twice as large, the top paid worker in a firm-occupation-year has a 16.9% wage premium.

We find significant variation in wage scaling across occupations. Jobs intensive in leadership and decision making have larger premiums. On the other hand, routine and manual task intensive occupations have a smaller premium. Our main analysis uses wage data at the firm level, and job characteristics at the occupation level from O*NET. Robustness analyses confirms this pattern: (1) in establishments (2) looking at the direct text of individual job responsibilities rather than occupational title and (3) in specifications where individual fixed effects are included.

We also investigate the possibility that the deployment of information and communication technologies (IT) impacts the scalability of tasks. We measure IT use through the firms' historical hiring of IT workers. More IT intensive establishments or firms might be able to better leverage the insights of its best employees, or might pay more to avoid weak links in a complex production function. We find that use of IT has a positive effect on wage scaling for highest-paid workers, but smaller or even negative effects for scaling for workers lower in the wage distribution. Similarly, we find IT increases the scaling of abstract workers' wage with firm size, while decreasing it for manual and physical intensive workers. These results indicate that IT tends to exacerbate already detected patterns in which jobs have wage scaling with firm size.

We interpret our findings through task scaling theories. An important reason some tasks scale is that they involve creating or analyzing information. A good idea can make an entire firm more productive. Information has economies of scale, because ideas only need to be created once to be used by all. We focus on two distinct types of potential informational economies of scale: span-of-control and innovation. Span-of-control refers to orders or instructions, while innovation refers to codified knowledge. These different economies of scale differ both in the type of information involved and how it is conveyed. We also consider the theory that task scaling arises through close complementarities of labor inputs in complex production. We refer to this as the 'O-ring' theory. Overall our results suggest the span-of-control version of information economies of scale are the most important for the large firm wage premium.

Informational economies of scale driven by innovation would entail new trade secrets, branding designs, or production processes. These types of innovations are associated with individuals in creative, science and engineering occupations. The information is typically transmitted in patents, manuals, prototypes, software programs and similar artifacts. Our data shows that jobs in the creative, science, and engineering fields do not see much difference in scaling with firm size compared to average jobs. This finding suggests that most innovation economies of scale cross firm boundaries. It is consistent with small innovative firms (e.g. in pharmaceuticals) paying large wages to innovators with the plan of having any successful ideas being acquired and marketed by a larger firm.

Another informational economy of scale is 'span-of-control'. A worker's span-of-control are the employees and divisions of the firm that a worker can directly command, guide, or otherwise help or hinder in their pursuit of company goals (Drucker, 2007). Executives and managers are core examples of occupations with large span-of-control. But workers in any occupation can potentially exert span-of-control by being exemplars in their tasks, handling exceptional cases, and mentoring their colleagues. The information in span-of-control takes the form of briefings, orders, and instructions. If span-of-control is a cause of the large-firm wage premium, the premium should be larger for occupations calling for leadership or decision making skills, as well as for workers at a higher wage percentile in their occupation-firm group.

One factor uniting each of these mechanisms for task scalability is that each are likely to be strengthened by IT technologies. IT increases both the information available to top workers and their ability to command, monitor, and teach others – allowing for increases in the maximum potential span-of-control. IT is associated with more complex production technologies and therefore firms investing in them might have a more O-Ring style production function, although properly applied IT may help with error detection and modularity, reducing the reliance on weak links. Finally, IT also increases the ability for ideas to be stored and propagated across a firm, potentially increasing the importance of informational economies of scale.

Overall, our evidence is most consistent with span-of-control driving which jobs scale and why, while not ruling out a role for other mechanisms. An implication of our findings is that within-occupation inequality as well as within-firm inequality, in management careers and IT intensive firms, is set to continue increasing. This is because of the long-run secular trend in firm concentration (Autor et al., 2020), which has increased the top 0.1% of US corporations' share of sales from about 40% in 1960 to over 60% today (Kwon et al., 2024). Further, decision making intensive jobs have increased by 2.07% as a share of the workforce over the last decades from 4.84% in 2004 to 6.91% in 2023.[1]. Because the average wage of management occupations is about 2.1 times that of the average job,[2] this within-occupation polarization is likely contributing to the rise of aggregate inequality as well. This is consistent with the findings of Song et al. (2018) which finds an important role for increasing wage variance within mega firms in explaining aggregate inequality growth.

2 Related Literature

Understanding the sources and mechanisms of task scaling is important in light of the rise of 'mega firms'. US corporations have seen strong increases in the rate of profit (Barkai and Benzell, 2024) and concentration (as measured by top percentile asset or sales shares; (Kwon et al., 2024)) since the 1970s. (Autor et al., 2020) connects these two observations, arguing that technological changes have helped the most productive firms expand and grow, simultaneously raising profits (at the expense of labor's share) and concentration without necessarily harming competition. The fact that these megafirms are quantitatively different and more productive than smaller firms suggests that workers in these firms perform different tasks. If workers providing some tasks are paid differently in firms of different sizes, the rise of megafirms will also have implications for the wage distribution.

There has been considerable research tying organization size to worker compensation. The observation of a positive correlation between firm size and wage goes back to at least Moore (1911). An influential modern study on this correlation is Brown and Medoff (1989). Pooling together workers of all types, they find that both firm and establishment size correlate positively with pay, including when both are included in a specification. They also find suggestive evidence that approximately half of the differential is explained by higher skilled and quality laborers at the larger firms. In a Handbook of Labor Economics chapter, Oi and Idson (1999) confirm these findings, finding a role for both unobservable and observable skill in the wage-size premium, and finding the premium is larger in the US than other developed countries.

[1] Authors' calculations from U.S. Bureau of Labor Statistics (2023)

[2] Occupational Employment and Wage Statistics 2023, calculated by the author

In Bloom et al. (2018) and Song et al. (2018), the authors analyse US labor income inequality within and across firms using near universal Social Security Administration data covering 1978-2013. While this data does not distinguish workers by occupation, the authors did make two findings relevant to our study. First, while they confirm the existence of a large firm wage premium, they find it to have shrunk over the period of study. Second, they are able to attribute a large share of the increase in US labor income inequality to an increase in the variance of log wages in mega firms. These mega firms, defined as the roughly 750 firms with more than 10,000 employees, saw within firm inequality rise by more than four times that of other firms. Another recent paper looking at the role of at the role of mega firms on the wages, Bessen et al. (2023) find that local labor markets with a higher share of large-firm job postings have higher wages in managerial, STEM and sales jobs, and reduced startup growth rates.

2.1 Task Scaling Theories of the Large Firm Wage Premium

One reason that larger firms might have higher paid workers is that the performance of workers in large firms is more important, in the sense of having a larger marginal product. We dub these theories of the firm size premium as task scaling theories.

Under this theory, the output of a worker can be decomposed into two parts: the share that is due to their non-scalable inputs and the share that comes from complementing the establishment or firm as a whole. Their wage might be written as:

$$w = f(\psi_0 S, \psi_1) \tag{1}$$

where the wage w is some increasing function of the scale of the firm S. It is also a function of parameters ψ_0 and ψ_1: the first corresponding to the relative importance of scale, and the latter to the relative importance of some 'direct' effect, unmediated by size. At this point we are taking no stance on why wages scale with firm size. If the most talented workers assortatively match to the largest firms, and workers are paid their marginal products, then the large firm wage premium would be the composite of two effects – their skill distribution and the strength of complementarity between skill and firm size.

For a given worker at a business of size S, we can decompose their wage into:

$$w = \gamma w + (1 - \gamma)w \tag{2}$$

with

$$\gamma = \frac{\beta log(S)}{w} \tag{3}$$

and

$$(1 - \gamma) = \frac{D}{w} \tag{4}$$

Where w is the worker's wage, $\gamma \in [0, 1]$ is the share of output coming from scalable tasks, D is the worker's productivity at a business of minimum size (a one-person business), and β is a term that gives how much the worker's wage scales with firm size. Throughout this paper we will be concerned with estimating heterogeneity in β for jobs of different types.

Within an occupation, a worker might have some tasks with a marginal product that scales with firm size, and some that do not. Table 1 gives examples of how γ might differ by task for a selection of occupations.

CEO is an occupation that is focused on scalable tasks. Two papers that investigate scalability in

Table 1: How Different Tasks May Scale with Firm Size

	Marginal Product Scaling by Firm Size	
	$\gamma = 0$	$\gamma = 1$
Elasticity of Marginal Product to Firm Size	Is invariant to firm size	Scales proportionally
Occupations	**Occupational Tasks**	
1. Executives and Managers	1. Redecorate office	1. New corporate strategy
2. Salespersons	2. Checkout counter	2. Closing a large B2B sale
3. Factory Workers	3. Installing a windshield	3. Discovering a dangerously faulty part
4. Doctors	4. Treating a patient	4. Writing a guide for residents

the case of CEOs are Baker and Hall (2004) and Gabaix and Landier (2008). Baker and Hall (2004) notes in a theory model that the elasticity of a CEO's productivity to firm outcomes may differ: for example, the mistake of wasting money on a corporate jet is invariant to firm size, but the return to reorganizing a firm may scale proportionality to firm size. They use this insight to understand the structure of CEO incentives.

This model of scalability also motivates the analysis of Gabaix and Landier (2008), who measure CEO wage scaling with firm size, and, assuming perfect assortative matching, decompose this scaling into a complementarity effect and a talent distribution effect. We replicate their finding of an over 20% increase in CEO compensation at a firm which is twice as large.

2.2 Information Economy of Scale Theories of Task Scaling

One important mechanism by which some tasks may scale is informational economies of scale. A good idea can improve the productivity of additional people in a firm or establishment at no marginal cost. We divide informational economies of scale into two categories: innovation economies of scale and span-of-control.

Innovation we use here to refer to the scalable tasks provided by creatives, scientists, and engineers. When these workers come up with a new branding, or technical innovation, or widget variation, these products may be more valuable at larger firms. On the other hand some innovations can be monetized outside of the firm. For example, a small innovative startup might develop a patent which is then sold to a larger firm to product and market at scale. To the extent this latter effect is important, it would tend to limit the role of innovation economies of scale in the large firm wage premium.

The other informational economy of scale is span-of-control. This is the mechanism that suggests particular attention for CEOs and managers. The span of managerial responsibility, and the closely related concept of span-of-control, is defined as "the number of people whom one superior can assist, teach and help to reach" goals Drucker (2007). This number is not a constant for a certain job-title, but rather varies by the nature of the business, the expertise of the superior, and the technologies employed. At a larger establishment there is a higher upper bound on the number, wage, and importance of the employees and departments one might control. Therefore, span-of-control theory would suggest that managers, and especially top managers, should see a large share of their wage come from scalable tasks. Caliendo et al. (2015) find that workers that are higher in the firm hierarchy earn higher wages and that larger firms have more layers of hierarchy, which is consistent with span-of-control driving wage scaling.

Another reason for task scaling, which is not exactly an informational economy of scale, is close complementarity in production functions. O-ring production functions are those in which the quality of final output is heavily dependent on the quality of the worst input. The name is drawn from the Space Shuttle Challenger disaster, in which the failure of a single part – an O-ring – led to catastrophic mission failure. O-ring production leads to assortative matching of employees to firms and countries Kremer (1993). Oi and Idson (1999) argue that one cause of a wage premium at larger firms may be the need for precise rule following, and "conformance to common work rules which result in paying rents to infra-marginal team members" and note that larger firms are the first to adopt certain precise novel technologies. A notable difference between span-of-control theories and O-ring theories is that O-ring theories would be applicable to all members of an organization, while span-of-control would disproportionately apply to leaders and exemplars.

2.3 IT and the Large Firm Wage Premium

ICTs increase the ease with which information is shared. It therefore may increase the large firm wage premium for tasks that are scaling because of informational economies of scale. Innovations will be more perfectly and costlessly copied, and managers will be able to influence and assist more subordinates with a greater amount of information available.

ICTs may also have heterogenous effects on the large firm wage premium for workers at different skill percentiles. Rosen (1981) proposes the idea that ICTs have led to 'superstar' markets with extreme inequality. The most popular performers have their content digitally reproduced endlessly, leading to high returns, while previously the market may have supported many bards, lecturers, and musicians at an intermediate wage. If ICTs allow a small group of leaders to leverage their expertise more precisely or across a larger group of followers, we would expect ICTs to cause more inequality both within an occupation-firm (i.e. the creation of 'internal superstar markets') and create stronger scalability for leadership intensive occupations across all firms. Relatedly, Piore and Sabel (1984) discuss certain types of innovation as separating production from conception. If ICTs boost the productivity of highest-paid workers' scalable tasks while routinizing and making non-scalable the work of others, it would tend to increase inequality in the largest firms. Indeed we find that ICTs have a positive effect on wage scaling for high wage percentile workers, and a negative effect for low wage percentile workers. This is consistent with mega firms' use of ICTs driving the increase in their contribution to wage inequality noted by Song et al. (2018).

Some research exists on the impact of ICTs on managerial span of control. Pinsonneault and Kraemer (1997) and Pinsonneault and Kraemer (1993) find that IT aids manager productivity, but only reduces the number of middle-managers in firms with already centralized decision-maker authority. The research on skill-biased technical change also suggests these technologies have increased inequality by negatively impacting middle-income occupations, which were disproportionately routine.

Finally, IT may also allow for better matching between top workers and high impact positions. Any of these effects would tend to increase the value of guaranteed high-quality labor inputs for the largest firms.

While previous papers have looked at the role of new technologies in promoting inequality, these have largely focused on the decline of middle-income jobs rather than the rise of top-percentile inequality. Skill-biased technological change is thought to have contributed to the decline of routine, middle income jobs. A representative paper Cortes et al. (2017) found that routine workers saw their wages and employment decrease: from 40.5% of U.S. workers were in routine jobs in 1979, and only 31.2% in 2014. Similarly, Autor et al. (2020) find that globalization and technological changes are leading to increasing inequality in firm sizes, concentrating more production in firms that are highly productive, have high markups, and low labor shares. In that study the authors do not have the data to evaluate the claim wages per worker within these "superstar firms" has increased. However, if the large firm wage premium holds, more concentrated industries will tend to create more top-percentile inequality.

2.4 Alternate Theories of the Large Firm Wage Premium

There are also alternate theories of the large-firm wage premium. One is firm specific human capital. On this theory, Lynch and Black (1998) find workers at larger firms tend to receive more training. A recent cross-country study confirmed this finding overall, but noted that the effect is driven by manufacturing firms, with service firms having a less monotonic relationship between scale and wage (Berlingieri et al. (2018)).

We have theorized that scalability of many tasks may have increased due to diffusion of ICTs. But there are alternate theories of why wage inequality may have increased. For example, changes in norms or laws over the same period may have boosted the wages of those already privileged, or those in growing and successful firms. A representative paper advocating this theory is Piketty et al. (2014). In this paper, the authors develop a model in which executives bargain harder for pre-tax monetary compensation when top marginal tax rates are low. Theoretically and empirically, they find that countries that cut their top tax rates have higher top-percentile pre-tax incomes but not higher economic growth. Under this line of thinking it was the US's decision around the 1980s to deregulate industries and cut top tax rates that led to the rise in top-percentile incomes. Over the same period, and for the same reasons, Stelzner (2017) documents dramatic decreases in strikes, increasingly anti-labor decisions by the National Labor Review Board, and a large rise in permanent replacement labor. Each of these changes might tend to reduce the bargaining power of and profit sharing with lower and middle wage workers.

3 Data

We utilize high-frequency administrative payroll data from ADP, one of the world's leading payroll processing firms, to perform an analysis on occupational wage and firm size. This rich dataset includes information on payroll transactions such as payment dates and amounts for each worker, characteristics of employees and employers, and administrative data on individuals listed on the payroll. The data is provided regardless of whether the employee received payment during the current pay period, which differs across firms.

Employers are reported at the ADP firm level, where a firm of ADP may possess one or multiple "Payroll Accounts", each considered equivalent to a business establishment, that is, a specific company work location. We measure firm size, both in the dimension of payroll and total employment. An establishment is a single physical location where business operations or services are conducted, usually focusing on a specific set of activities or a product line, and is identified by one NAICS code. In contrast, a firm refers to a business entity that operates under a single legal structure and can consist of multiple establishments spread across different locations, each potentially having multiple NAICS codes. We provide empirical evidence at both the establishment level and the firm level, but focus on firms for the main text.

Our study observes payroll transactions from over 25 million workers and their employers at the monthly level between 2017 and 2023. We restrict attention to workers between 18-70, who are paid more than the full time federal minimum wage, at firms with at least 10 employees in the year, which gives us about 15 million employees. After these restrictions, our analysis data set covers between 14.7 and 16.0 million employees depending on the year.

We conduct our main analyses at the firm-occupation-year level. This gives us about 12.5 million firm-occupation-years with at least one worker. We first aggregate the monthly data at the annual level by each individual and their specific employer and occupation. We only include months in which the employee is working at the firm. Subsequently, for each ADP Payroll Account, we calculate the average monthly payroll for each O*NET occupation in a given year. The O*NET occupation is defined using job titles, job descriptions, industry of the client, location, and other information, and this classification is conducted by ADP's own research team.

We also conduct a complementary analysis at the individual level, to allow for individual fixed effects. In this data we have 173,717,169 observations at the individual-firm-occupation-year level, as summarized in Panel D of Table 2.

Our study also uses complementary data from O*NET on occupational definitions and characteristics. One characteristic we use are O*NET work context, which define the types of tasks done by each occupation. Additionally, we categorize occupations by their routine, abstract, and manual task intensity, following (Autor and Handel, 2013). We also employ data from the US Census on commuting zone boundaries and population.

Table 2: Summary Statistics

VARIABLES	(1) # Obs.	(2) Mean	(3) Std. dev.	(4) Min	(5) Median	(5) Max
Panel A: Firm-Occupation-Year Level						
Log of firm employment size	9,606,247	5.30	1.31	2.40	5.19	11.34
Log of firm payroll size	9,606,247	16.41	1.49	9.61	16.35	23.70
Log number of establishments in a firm	9,606,247	0.65	0.89	0.00	0.00	7.26
Log of wages for 1st percentile workers	9,606,247	8.32	0.74	7.06	8.32	18.06
Log of wages for 25th percentile workers	9,606,247	8.47	0.67	7.06	8.44	18.06
Log of wages for median workers	9,606,247	8.61	0.66	7.06	8.57	18.06
Log of wages for 75th percentile workers	9,606,247	8.75	0.69	7.06	8.69	18.95
Log of wages for 90th percentile workers	9,606,247	8.84	0.72	7.06	8.77	18.95
Log of wages for 95th percentile workers	9,606,247	8.88	0.75	7.06	8.80	18.95
Log of wages for 99th percentile workers	9,606,247	8.91	0.78	7.06	8.82	18.95
Log of wages for highest-paid workers	9,606,247	8.92	0.79	7.06	8.83	19.93
Firm employment size	9,606,247	606.95	2,212.37	10.00	178.42	84,142.63
Firm payroll size	9,606,247	48,645,666.81	2.03E+08	14,894.99	12,548,692.00	1.96E+10
Number of establishments in a firm	9,606,247	3.85	14.81	1.00	1.00	1,427.00
Wages for 1st percentile workers	9,606,247	5,866.43	54,233.49	1,160.00	4,105.26	7.01E+07
Wages for 25th percentile workers	9,606,247	6,450.78	54,255.94	1,160.00	4,604.78	7.01E+07
Wages for 50th percentile workers	9,606,247	7,386.50	55,273.51	1,160.00	5,247.35	7.01E+07
Wages for 75th percentile workers	9,606,247	8,982.92	79,520.29	1,160.00	5,968.20	1.71E+08
Wages for 90th percentile workers	9,606,247	10,339.23	95,761.16	1,160.00	6,439.34	1.71E+08
Wages for 95th percentile workers	9,606,247	11,086.46	100,717.60	1,160.00	6,629.20	1.71E+08
Wages for 99th percentile workers	9,606,247	12,225.92	110,001.96	1,160.00	6,795.65	1.71E+08
Wages for highest-paid workers	9,606,247	13,054.88	148,336.51	1,160.00	6,827.01	4.53E+08
BDS number of firms at 2-digit NAICS	9,606,247	513,765.72	197,626.80	5,518.00	606,410.00	732,814.00
BDS number of establishments at 2-digit NAICS	9,606,247	660,421.59	245,889.27	17,995.00	784,643.00	994,644.00
Panel B: O*NET Occupational Level						
Abstract task score (normalized)	898	0.37	0.27	0	0.11	1
Routine task score (normalized)	898	0.43	0.31	0	0.45	1
Manual task score (normalized)	898	0.12	0.12	0	0.71	1
Panel C: ADP Job Level[3]						
Focus on Creating Information	6,494	0.20	0.25	0.00	0.00	0.78
Focus on Managing/Processing Information	6,494	0.62	0.23	0.00	0.78	0.89
Interpersonal Tasks Frequency	6,494	0.66	0.24	0.11	0.67	1.00
Physical Proximity Requirement	6,494	0.57	0.29	0.11	0.67	1.00
Manual Labor Requirement	6,494	0.36	0.34	0.00	0.22	1.00
Remote Work Feasibility	6,494	0.37	0.31	0.00	0.33	1.00
Negative Consequences of Mistakes	6,494	0.70	0.17	0.00	0.67	1.00
Involvement in Innovation	6,494	0.26	0.23	0.00	0.22	0.89
Impact on Firm's Overall Value	6,494	0.46	0.23	0.00	0.44	1.00
Percent of Employees Affected	6,494	0.28	0.23	0.00	0.20	1.00
Number of direct reports	6,494	1.70	3.22	0.00	0.00	35.00
Panel D: Individual-Year Level						
Age	173,717,169	40.09	13.25	18.00	39.00	70.00
Monthly wages	173,717,169	6375.19	91735.11	1160.00	3974.091	4.53E+08
Log of monthly wages	173,717,169	8.37	0.75	7.06	8.29	19.93

4 Theoretical Framework

Inspired by (Baker and Hall, 2004) and (Gabaix and Landier, 2008) we contemplate a worker having tasks with a 'scalable' component as well as a 'direct', additive, effect on firm output that does not depend on firm scale.

A worker's log wage in firm i, of size S_i, is the sum of their marginal product in scalable and non-scalable tasks:

$$log(w_{i,j,o,t}) = log(S_{i,t})C_{i,o,p}T_{i,j,o} + D_{o,i,t} \tag{5}$$

where o indexes the occupation type, and j indexes the individual potential job/worker. $T_{i,j,o}$ is the individual's talent. $C_{i,o,p}$ is how complementary talent is to firm size for this occupation. p indicates what skill percentile the worker is in at their occupation at the firm. So the term $log(S_{i,t})C_{i,o,p}T_{i,j,o}$ captures the contribution of scalable tasks to the employee's wage. $D_{o,t}$ is the *direct* contribution to output of a worker at a given occupation and time to a firm of size 0.[4]

The wage of a worker is increasing in the size of the firm they work at, the degree of complementarity between a workers' talent and the firm size, and the workers' talent.[5]

This model motivates the following reduced form regression:

$$log(w_{i,j,o,t}) = \beta log(S_{i,t}) + \eta X_{o,i,j,t} + \alpha \tag{6}$$

This regression asks how wage in a firm scales with firm size after controlling for a vector of fixed effects $X_{o,i,t}$. Connecting this regression to equation 2, we see that if γ is larger for a certain job (holding firm size fixed) we should estimate a larger β in the above regression.

We are primarily interested in estimates of β and how this varies across occupations and firm types. Therefore we sometimes restrict attention to specific occupation o or firm-occupation wage percentiles p. We also investigate specifications of the form:

$$log(w_{i,j,o,t}) = \beta log(S_{i,t}) + \eta_0 X_{o,i,j,t} + \eta_1 Y_{i,o,t} + \beta_1 Y_{i,o,t} log(S_{i,t}) + \alpha \tag{7}$$

where Y is some job characteristic we think might effect scalability. β_1 tracks how much stronger or weaker scaling is for occupations with the characteristic Y. Variation in β_1 will track variation in C, the degree of complementarity between talent and firm outcomes, under two assumptions:

- There is no variation in the degree of assortative matching and skill distribution across occupations and firm types
- The fixed effects included sufficiently control for regional, occupational, and industry level variation in the non-scaling contribution of an occupation to output. In other words, that $D_{o,i,j,t}$ is not correlated with $T_{o,i,j,t}$ after controlling for fixed effects (hence the j index is dropped).

5 Results

In this section we perform variations of a regression of the log of wage on the log of establishment size and other controls, as well as interactions. These regressions omit firm fixed effects, to avoid overcontrolling, but standard errors are clustered at the establishment level to allow for heteroskedasticity and serial correlation in wages within establishments. We primarily measure establishment size through total payroll, but also consider employment size.

[4] This model directly follows Gabaix and Landier (2008) with two exceptions: the inclusion of a direct, non-scaling, output effect (which is arguably captured in his regressions through the intercept term); we also omit a decreasing/increasing returns to scale term as a power of S which they include in some specifications.

[5] While we do not take a stance on the exact wage setting process, a reasonable assumption is that the wage is increasing in the marginal product of the worker. This arises both in a perfectly competitive model, as well as the bargaining model of Gabaix and Landier (2008)

5.1 Wage Scaling by Occupation, Within-Firm Wage Percentile, and Business Size Definition

In this subsection we provide overall results on how wages scale with market size for different occupations, wage percentiles, and notions of market size.

$$\log(Wage_{ijsmt}) = \alpha + \beta \log(Size_{ismt}) + \delta_t + \mu_s + \pi_o + \varepsilon_{ijsmt} \tag{8}$$

where $\log(Wage_{ijsmt})$ is the average monthly income of the top paid employee in occupation j in a given establishment i from industry s located in market m in year t. $\log(Size_{ismt})$ is the log of firm size measured by total payroll or the log of firm employment size, depending on the specification. δ_t is year fixed effect, and μ_s is industry fixed effect at two-digit NAICS level. π_o is a 6-digit O*NET occupational fixed effect. Standard errors are robust and clustered by firm.

Table 3 report how wages in an occupation-firm-year scale with firm size, for workers at two different wage percentiles and with two measures of firm size. The first two columns report how the highest-paid worker's wage scale with firm payroll and employment size, respectively. The second pair of columns report the same for the median worker. Figure 1 shows these scaling estimates for a wider range of wage percentiles,[6] showing the scaling effect scales monotonically with percentile. Point estimates for the coefficients presented in this table can be found in appendix table A11.

As both the table and figure demonstrate, our first key finding is the strong evidence of a substantial firm wage premium across all measures of firm size.

Our second main result is that the wage premium is estimated to be monotonically larger for workers at higher percentiles in their occupation-firm-year. In other words, occupational wages are more spread out at larger firms, and this effect is larger than would be implied by constant scaling at all percentiles. Notably, measuring firm size by payroll, even 1st percentile workers see their wages scale with firm size.

The next type of heterogeneity we are interested in is how wage scaling varies across different occupations. Figure 2 reports scaling estimates, but restricting the regression to each of thirty-two most common and representative occupations. Estimates of β from these regressions are presented. Wages for median workers in an occupation-firm-year are used.

The result in this figure, which is confirmed in subsequent analysis, is that leadership and supervisory positions scale more and routine and manual occupations scale less. Science and engineering occupations scale an intermediate amount, as well as certain service occupations. All selected occupations see positive median wage scaling with firm payroll, as shown in the bottom part of the figure. The upper part of the figure reports the same results for the highest-paid workers, and a very similar pattern holds, but with larger scaling point estimates for each occupation.

Also notable in this set of figures is that scaling is estimated to be larger, and R-squared is higher, when payroll is used as the measure of firm size, rather than employment. We therefore focus on firm payroll as our main measure of firm size moving forward.

[6] We use the 'approxQuantile' method in PySpark to calculate the wage percentile. This method employs the Greenwald-Khanna algorithm to compute approximate percentiles efficiently. For fewer than 100 workers in a given establishment-occupation-year, the method determines the position of the desired percentile within the sorted wage data. If the position is not an integer, it interpolates between the nearest data points to estimate the percentile value accurately. For example, with 5 software developers in a given establishment and year, the 99th percentile will be identical to the maximum wage.

Table 3: Wage scaling with different measures of firm size

	(1)	(2)	(3)	(4)
	log(wages for highest-paid workers)		log(wages for median workers)	
Log of firm payroll size	0.169***		0.086***	
	(0.001)		(0.001)	
Log of firm employment size		0.132***		0.041***
		(0.001)		(0.001)
Observations	9,606,243	9,606,243	9,606,243	9,606,243
Adj.R^2	0.495	0.448	0.473	0.446
Year FE	Y	Y	Y	Y
2-digit NAICS FE	Y	Y	Y	Y
O*NET occupational FE	Y	Y	Y	Y

Notes: The dependent variable is the logarithm of the monthly total taxable income for the highest-paid workers and median workers within their respective occupations in a firm on an annual basis. The independent variables include the logarithm of the firm total payroll and the logarithm of firm employment size. The regression model includes fixed effects for year, 2-digit NAICS category, and 6-digit O*NET occupation classification. The regression is weighted by the number of firms classified under 2-digit NAICS, using data from the Business Dynamics Statistics (BDS) dataset. Robust standard errors clustered at firm level in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

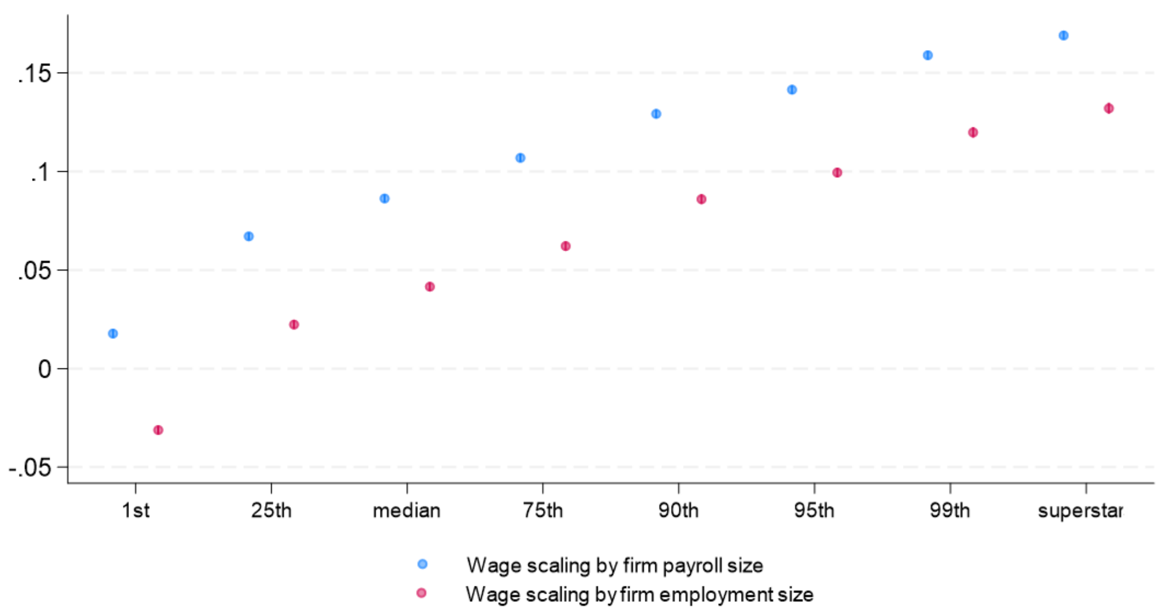

Figure 1: Wages as a function of firm size for different percentile workers

Notes: Estimates of β in regressions following equation 8 where the log of the wage for workers at different firm-occupation wage percentiles is explained by the log of firm size as measured by payroll and employment. 95% confidence intervals displayed. Point estimates underlying this figure can be found in Table A11.

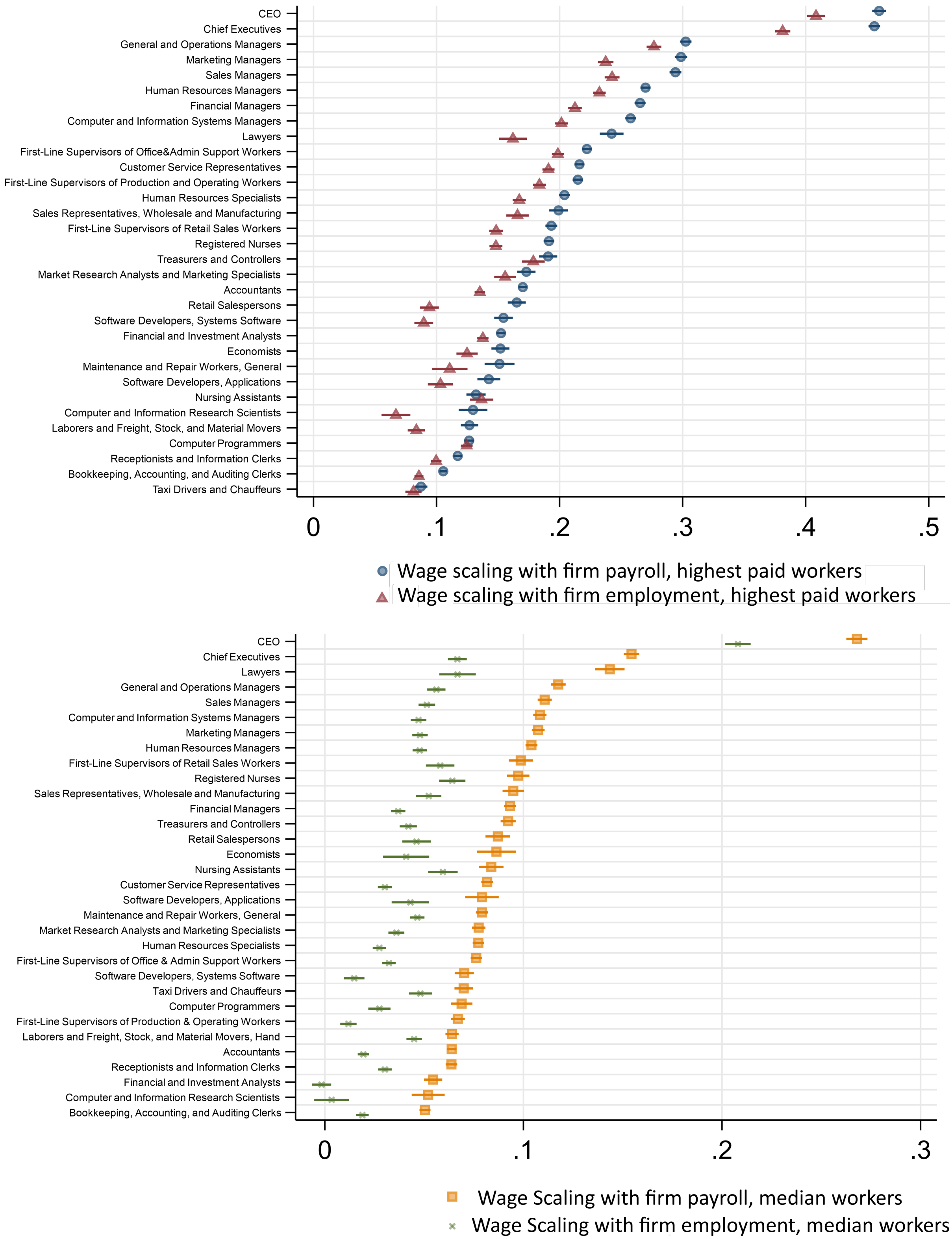

Figure 2: Wage scaling for highest-paid and median workers with two measures of firm size

The figure reports the coefficients of firm size β from equation (8) derived from 32 separate regressions for 32 selected common occupations, with 95% confidence intervals. The dependent variable is the logarithm of the monthly total taxable income for the highest-paid workers in the upper part of the figure and for median workers in the lower part. Independent variables include the logarithm of either firm total payroll or firm employment size, depending on the specification. Fixed effects for year and 2-digit NAICS category are included.

5.2 Wage Scaling and Job Task Characteristics

In this section we investigate how scaling is different for occupations of different types.

Occupation-Firm Level Analysis of Wage Scaling by Task Characteristic

At the occupation-firm-year level, the regressions follow equation 9.

$$\begin{aligned} \log(Wage_{ijsmt}) &= \alpha + \beta \log(Size_{ismt}) + \eta_1 OccCharacteristics^{occ_j} \\ &+ \beta_2 \log(Size_{ismt}) \times OccCharacteristics^{occ_j} + \delta_t + \mu_s + \varepsilon_{ijsmt} \end{aligned} \tag{9}$$

Table 4 reports regressions of the log of the wage on the log of firm size (measured through employment or payroll), occupational task intensity (normalized between 0 and 1), and their interaction. Occupational task intensity indexes are created from O*NET data following Autor and Handel (2013). The direct effect of occupational characteristics on output is absorbed by the occupational fixed effects and are omitted.

We are mainly interested in the term β_2, which reports the effect of an occupational characteristic on scaling. The top half of the table uses the logarithm of the highest-paid worker in an occupation-firm-year as the outcome, while the bottom uses the logarithm of the wages for the median workers as the outcome. As can be seen, scaling is stronger for workers in abstract task-intensive occupations and weaker for occupations that require more manual and routine tasks.

Table 5 performs an analogous analysis, but the interaction term is a binary classification for whether the occupation is 'creative' or 'R&D' related. As can be seen, R&D workers' wages scale only slightly more with firm size than the average job, while creative workers are not significantly different at all. This is consistent with the idea that innovation economies of scale are not trapped within the firm, and that rather small firms can benefit from innovations by licensing the new ideas to other companies.

For a more granular look at what occupational characteristics drive wage scaling, Figure 3 investigates what types of 'work contexts' have the fastest scaling. Work contexts are derived from O*NET, and all occupations are rated from 0 (least intensive) to 1 (most intensive). The figure reports estimates of β_2, the coefficient on the interaction between work context and the log of establishment payroll in explaining the superstar wage. As can be seen, decision making and leadership contexts have the largest, most positive coefficients, while manual, unpleasant, repetitive, and machinery proximate tasks have negative coefficients.

Our raw data, in addition to occupational titles, also includes a short text description of a job. Because there are many more unique job descriptions than titles, using information from this text gives within-occupation occupational variation in occupational tasks. Additional details of data construction are provided in appendix A.2.

Figure 4 reports the results of a regression of the median and highest-paid wage in a firm-occupation-year on firm size, a job characteristic, and their interaction. The coefficient on the interaction is reported. Other controls included are fixed effects for year, 2-digit NAICS category, and 6-digit O*NET occupational code. The results show that the characteristics that see the most wage growth with firm size are "negative consequences of mistakes" and "impact on firm's overall value", while physical proximity requirements and manual work requirements predict negative wage scaling with firm size.

Table 4: Effects of Occupational-level Task Score on Wage Scaling with Firm Size

	(1)	(2)	(3)	(4)
***Panel A:* Log of Wages for highest-paid Workers**				
Log of firm size	0.174***	0.134***	0.220***	0.194***
	(0.001)	(0.001)	(0.001)	(0.001)
Log of firm size × Abstract task score		0.085***		
		(0.002)		
Log of firm size × Routine task score			-0.157***	
			(0.001)	
Log of firm size × Manual task score				-0.287***
				(0.004)
Observations	9,974,733	9,673,093	9,673,093	9,673,093
Adj.R^2	0.492	0.497	0.501	0.497
***Panel B:* Log of Wages for Median Workers**				
Log of firm size	0.086***	0.074***	0.102***	0.090***
	(0.001)	(0.001)	(0.001)	(0.001)
Log of firm size × Abstract task score		0.025***		
		(0.001)		
Log of firm size × Routine task score			-0.054***	
			(0.001)	
Log of firm size × Manual task score				-0.053***
				(0.003)
Observations	9,974,733	9,673,093	9,673,093	9,673,093
Adj.R^2	0.474	0.476	0.477	0.476
Year FE	Y	Y	Y	Y
2-digit NAICS FE	Y	Y	Y	Y
O*NET occupational FE	Y	Y	Y	Y

Notes: The dependent variable is the logarithm of the monthly total taxable income for the highest-paid workers and median workers within their respective occupations in a firm on an annual basis. The independent variables include the log of firm size (measured through total payroll or employment size), occupational task intensity (normalized between 0 and 1), and their interaction. Occupational task intensity indexes are created from O*NET data following Autor and Handel (2013). The regression model includes fixed effects for year, 2-digit NAICS category, and 6-digit O*NET occupation classification. The regression is weighted by the number of firms classified under 2-digit NAICS, using data from the Business Dynamics Statistics (BDS) dataset. Robust standard errors clustered at firm level in parenthesis.*** p<0.01, ** p<0.05, * p<0.1

Table 5: Effects of wage scaling with firm size for R&D and creative workers

Dependent variable is the log of wages for highest-paid workers	(1)	(2)	(3)
Log of firm size	0.137***	0.138***	0.137***
	(0.001)	(0.001)	(0.001)
R&D workers	0.060**		
	(0.030)		
Log of firm size × R&D workers	0.003*		
	(0.002)		
Creative workers		-0.226***	
		(0.034)	
Log of firm size × Creative workers		0.001	
		(0.002)	
Creative workers			-0.151***
			(0.024)
Log of firm size × R&D or creative workers			0.008***
			(0.001)
Observations	9,974,736	9,974,736	9,974,736
Adj.R^2	0.121	0.122	0.120
Year FE	Y	Y	Y
2-digit NAICS FE	Y	Y	Y

Notes: The dependent variable is the logarithm of the monthly total taxable income for the highest-paid workers within their respective occupations in a firm on an annual basis. The independent variables include the log of firm size (measured through total payroll), a dummy variable indicating whether the worker is R&D related or creative worker, and their interaction. The regression model includes fixed effects for year and 2-digit NAICS category, weighted by the number of firms classified under 2-digit NAICS, using data from the Business Dynamics Statistics (BDS) dataset. Robust standard errors clustered at firm level in parenthesis.*** p<0.01, ** p<0.05, * p<0.1

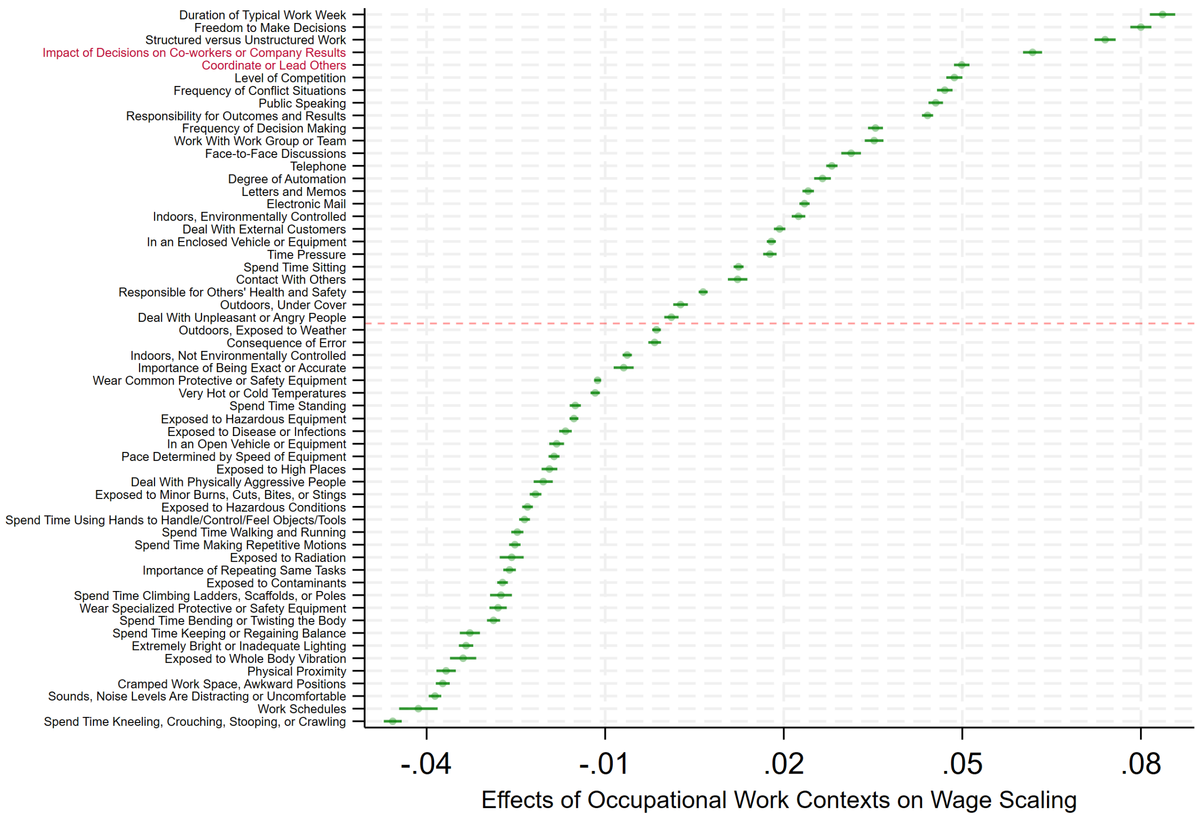

Figure 3: Effect of O*NET work context on wage scaling with firm payroll derived from 57 separate regressions

Notes: The figure reports the coefficients of the interaction term, β_2, from equation 9 derived from 57 separate regressions. The dependent variable is the logarithm of the monthly total taxable income for highest-paid workers and within within their respective occupations in a firm on an annual basis. The independent variables include the logarithm of the firm total payroll, O*NET work context score, and their interaction term. The regression model includes fixed effects for year, 2-digit NAICS category, and 6-digit O*NET occupation classification. The regression is weighted by the number of firms classified under NAICS2, using data from the Business Dynamics Statistics (BDS) dataset.

Figure 4: Wage scaling with firm size and job-level feature score.

Notes: This figure shows the impact of job-level characteristics on wage scaling for the highest-paid worker and median in an occupation-firm-year. Regression models include fixed effects for year, 2-digit NAICS category, and 6-digit O*NET occupational code.

Individual Level Analysis of Wage Scaling by Task Characteristic

While our main results are at the firm-occupation-year level, as robustness we conduct a series of analyses at the individual level. In this data, we observe the individual-level unique identifier, age, and their monthly average wage ranking within their firm-occupation-year.

Table 6 reports the results of a regression of log wage on log firm payroll and the interaction of an occupational task measure with log firm payroll. This analysis is conducted at the individual level, including all workers irrespective of their wage percentiles. Some specifications include firm fixed effects.

In table 6 column (4), the coefficient of 0.023 for the interaction between log of firm size and abstract task score suggests that within the same firm, workers performing higher-level of abstract tasks see their wages increase more with firm size compared to workers in other types of jobs, after controlling for time-invariant firm characteristics. Moreover, this coefficient is both larger and more significant than the corresponding interaction term in column (1), where firm fixed effects are not included. For routine and manual tasks, we observe negative impact on wage scaling with firm size. This finding is consistent with our results for both the highest-paid and median workers, further supporting the conclusion that wage scaling differs significantly by the nature of the tasks performed.

An individual level analysis also allows for individual fixed effects. Table 7 reports the results of a regression of firm size on wages, and an interaction of the wage with occupational task intensities. As can be seen, when individual fixed effects are included, there is still a positive and significant interaction between firm scaling and abstract task intensity, and a negative one for routine task intensity.

Across the individual level results the most surprising one is that there is a negative point estimate of the interaction between firm size and task abstractness in the individual data set, when no firm fixed effects are included. However, it is still clear that the interaction between abstract task intensity and firm size is still significantly greater than the interaction with routine or manual tasks. Relative to these other task orientations, the finding is still that abstractly-intense occupations have the strongest scaling with firm size.

Table 6: Wage scaling with firm size and occupational-level task score, individual level analysis.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable is log of wages for each worker	No Firm FE			with Firm FE		
Log of firm size	0.071***	0.074***	0.070***	0.016***	0.025***	0.021***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.002)	(0.003)
Log firm size×Abstract task score	−0.022			0.023*		
	(0.015)			(0.009)		
Log firm size×Routine task score		−0.051***			−0.017***	
		(0.009)			(0.003)	
Log firm size×Manual task score			−0.076***			0.012
			(0.022)			(0.009)
Num. obs.	169,365,043	169,365,043	169,365,043	169,365,043	169,365,043	169,365,043
Adj. R^2	0.344	0.345	0.344	0.552	0.552	0.552
O*NET FE:	Y	Y	Y	Y	Y	Y
2-digit NAICS FE:	Y	Y	Y	Y	Y	Y
Year FE:	Y	Y	Y	Y	Y	Y
Firm FE:	N	N	N	Y	Y	Y

Notes: The dependent variable is the logarithm of average monthly wages for each worker in a given occupation and firm on the annual basis. We include every employee who have above the federal minimum wage, aged 18-70, in active status in the data. The independent variables include the log of the wage on the log of firm size (measured through employment or payroll), occupational task intensity (normalized between 0 and 1), and their interaction. The regression model includes fixed effects for year, 2-digit NAICS category, and 6-digit O*NET occupation classification. The regression is weighted by the number of firms classified under 2-digit NAICS, using data from the Business Dynamics Statistics (BDS) dataset. Robust standard errors clustered at firm level in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Wage scaling with firm size and occupational-level task score, individual level fixed effects.

	(1)	(2)	(3)
log of wages for each worker	with Individual FE		
Log of firm size	0.028***	0.039***	0.034***
	(0.004)	(0.002)	(0.003)
Log firm size×Abstract task score	0.026**		
	(0.010)		
Log firm size×Routine task score		−0.022***	
		(0.003)	
Log firm size×Manual task score			0.001
			(0.009)
Num. obs.	169,365,043	169,365,043	169,365,043
Adj. R^2	0.630	0.630	0.630
O*NET FE:	Y	Y	Y
2-digit NAICS FE:	Y	Y	Y
Year FE:	Y	Y	Y
Individual FE:	Y	Y	Y

Notes: The dependent variable is the logarithm of average monthly wages for each worker in a given occupation and firm on the annual basis. We include every employee who have above the federal minimum wage, aged 18-70, in active status in the data. The independent variables include the log of the wage on the log of firm size (measured through employment or payroll), occupational task intensity (normalized between 0 and 1), and their interaction. The regression model includes fixed effects for year, 2-digit NAICS category, 6-digit O*NET occupation classification, and individual (defined by hashed SSN and birthyear). The regression is weighted by the number of firms classified under 2-digit NAICS, using data from the Business Dynamics Statistics (BDS) dataset. Robust standard errors clustered at firm level in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

5.3 Wage Scaling and ICTs

In the final analysis section we are interested in how ICTs influence scaling. IT intensity at a firm is measured by the fraction of the payroll paid to IT workers in the previous year in a given firm. Alternatively, we use the payroll paid to IT developers. These regressions follow equation 10.

$$\log(Wage_{ijsmt}) = \alpha + \beta \log(Size_{ismt}) + \eta IT_{i,t-1} + \beta_2 \log(Size_{ismt}) \times IT_{i,t-1} + \delta_t + \gamma_s + \varepsilon_{ijsmt} \quad (10)$$

Table 8 shows how IT intensity at the firm affect wage scaling for workers at different wage percentiles in a given firm-occupation-year. IT intensity significantly increases wage scaling for top earners, particularly for the highest-paid (0.166) and 99th percentile (0.171). This is consistent with the "superstar" markets theory proposed by Rosen (1981), where ICTs enable top performers to capture higher returns by leveraging their expertise more efficiently. In contrast, IT intensity reduces wage scaling for workers at the bottom of the wage distribution, with a negative coefficient for the 1st percentile (-0.035), consistent with the idea of separating conception from production as discussed by Piore and Sabel (1984).

Table 8: Firm-level IT investment and wage scaling by wage percentiles

Dependent variable is log of wages for:	(1) **Highest paid**	(2) 99th	(3) 95th	(4) 90th	(5) 75th	(6) **Median**	(7) 25th	(8) 1st
Log of firm size	0.172***	0.162***	0.144***	0.131***	0.108***	0.086***	0.067***	0.016***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
IT investment (lagged)	-2.172***	-2.232***	-1.630***	-1.137***	-0.425	0.113	0.486*	1.244***
	(0.356)	(0.367)	(0.363)	(0.349)	(0.322)	(0.297)	(0.285)	(0.319)
Log of firm size×IT investment (lagged)	0.166***	0.171***	0.137***	0.108***	0.065***	0.033*	0.010	-0.035*
	(0.022)	(0.023)	(0.023)	(0.022)	(0.020)	(0.018)	(0.018)	(0.020)
Observations	7,986,870	7,986,870	7,986,870	7,986,870	7,986,870	7,986,870	7,986,870	7,986,870
Adj.R^2	0.501	0.499	0.498	0.502	0.511	0.512	0.473	0.435
Year FE	Y	Y	Y	Y	Y	Y	Y	Y
2-Digit NAICS FE	Y	Y	Y	Y	Y	Y	Y	Y
6-Digit O*NET FE	Y	Y	Y	Y	Y	Y	Y	Y

Notes: The dependent variable is the logarithm of the monthly total taxable income for employees at different wage percentiles within their respective occupations and firms on an annual basis. Independent variables include log firm size (annual payroll), the fraction of payroll to IT workers from the previous year, and their interaction. The regression model includes fixed effects for year, 2-digit NAICS category, and 6-digit O*NET occupational code. The regression is weighted by the number of firms in each 2-digit NAICS industry using data from the Business Dynamics Statistics (BDS) dataset. Robust standard errors are clustered at the firm level and are shown in parentheses.

Table 9 displays results from the three-way interaction of firm size, occupational task intensity, and firm IT intensity. As can be seen, the three way interaction with abstract tasks is positive, indicating that IT further increases scaling for abstract workers. However, the three way interaction with manual and routine tasks is negative, showing that IT decreases scaling for these tasks. For all three task measures, the effect of ICTs is to accelerate the effect of size on wage.

Table 9: The three-way interaction of ICT, firm size, and occupational task intensity

Dependent variable is the log of wages for highest-paid workers	(1)	(2)	(3)
Log of firm size	0.141***	0.216***	0.191***
	(0.001)	(0.001)	(0.001)
IT investment (lagged)	-4.620***	-6.484***	-7.197***
	(0.675)	(0.556)	(0.564)
Log of firm size×IT investment (lagged)	0.299***	0.466***	0.490***
	(0.042)	(0.035)	(0.035)
Log of firm size×Abstract task score	0.068***		
	(0.002)		
Abstract task score×IT investment (lagged)	-0.317		
	(0.778)		
Log of firm size×Abstract task score×IT investment (lagged)	**0.108****		
	(0.048)		
Log of firm size×Routine task score		-0.154***	
		(0.002)	
Routine task score×IT investment (lagged)		3.565***	
		(0.621)	
Log of firm size×Routine task score×IT investment (lagged)		**-0.235*****	
		(0.038)	
Log of firm size×Manual task score			-0.278***
			(0.004)
Manual task score×IT investment (lagged)			23.979***
			(2.220)
Log of firm size×Manual task score×IT investment (lagged)			**-1.233*****
			(0.136)
Observations	7,847,602	7,847,602	7,847,602
Adj.R^2	0.502	0.507	0.503
Year FE	Y	Y	Y
2-Digit NAICS FE	Y	Y	Y
6-Digit O*NET FE	Y	Y	Y

Notes: The dependent variable is the logarithm of the monthly total taxable income for highest-paid workers within their respective occupations and firms on an annual basis. Independent variables include log firm size (annual payroll), the fraction of payroll to IT workers from the previous year, task intensity score, and their interactions. The regression model includes fixed effects for year, 2-digit NAICS category, and 6-digit O*NET occupational code. The regression is weighted by the number of firms in each 2-digit NAICS industry using data from the Business Dynamics Statistics (BDS) dataset. Robust standard errors are clustered at the firm level and are shown in parentheses.

6 Discussion

We investigate how wages scale with firm size across different occupations, slices of the wage distribution, and exposure to IT. Our primary method is log-linear regressions of wages on market size. We replicate the findings of Gabaix and Landier (2008) of a large CEO wage scaling with firm size, on a much larger set of firms. We also replicate the common finding of large-firm wage premium across all workers.

Across occupations we find significant heterogeneity in scalability. Leadership and decision making intensive occupations as well as abstract task intensive occupations scale more. Routine and manual intensive occupations scale less. Within occupations it is the top workers in a firm-percentile who scale the most. Across firms we find scaling is strongest in firms and industries with large IT investments, and this disproportionately increasing the scaling of those in top percentiles of their occupation-establishment.

We interpret these findings through the lens of different task scaling theories. Table 10 summarizes the anticipated effects of organization size on firm wage for jobs of different types according to each of these theories. Our findings are consistent with scalability driving the large-firm wage premium, and in particular the mechanism of span-of-control.

Table 10: Comparison of Scaling Theory and Impact on Types of Jobs

Theory of Scaling	**Types of Jobs with Disproportionate Large-Firm Wage Premium**				
	Top Percentile Firm-Occupation Jobs	**Low Percentile Firm-Occupation Jobs**	**Leader and Decision Maker Jobs**	**Creative and R&D Jobs**	**Effect of ICTs**
Span-of-Control	+	-	+	?	+
O-Ring Complementarity	-	+	?	?	?
Innovation Economy of Scale	+	?	?	+	+

Our findings are important for understanding the past and future of inequality, especially within-occupation wage inequality. Over time, as the share of employment in jobs with large spans-of-control and IT exposed firms and careers has grown, the superstar dynamics we observe within occupations may be contributing to peoples' innate sense of growing inequality. As firms become more polarized in size, this too may contribute to felt and objective inequality.

What's next for which jobs scale and why? While AI systems may change the impact of ICTs on scaling, preliminary evidence suggests it will have a similar one, with AI being particularly good at automating and augmenting the worst workers in a firm-occupation, while the quality of this AI being determined by being trained on the best workers in a firm-occupation.

References

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen, "The fall of the labor share and the rise of superstar firms," *The Quarterly Journal of Economics*, 2020, *135* (2), 645–709.

Autor, David H and Michael J Handel, "Putting tasks to the test: Human capital, job tasks, and wages," *Journal of labor Economics*, 2013, *31* (S1), S59–S96.

Baker, George P. and Brian J. Hall, "CEO Incentives and Firm Size," *Journal of Labor Economics*, 2004, *22* (4), 767–798.

Barkai, Simcha and Seth G. Benzell, "70 Years of US Corporate Profits," *Journal of Corporate Finance*, 2024, *87*, 102622. Received 9 July 2023, Revised 18 June 2024, Accepted 19 June 2024, Available online 1 July 2024, Version of Record 8 July 2024.

Berlingieri, Giuseppe, Sara Calligaris, and Chiara Criscuolo, "The productivity-wage premium: Does size still matter in a service economy?," in "AEA Papers and Proceedings," Vol. 108 American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203 2018, pp. 328–333.

Bessen, James E., Felix Poege, and Ronja Röttger, "Competing for Talent: Large Firms and Startup Growth," Technical Report 4673494, Bocconi University Management Research Paper 2023. Available at SSRN: https://ssrn.com/abstract=4673494 or http://dx.doi.org/10.2139/ssrn.4673494.

Bloom, Nicholas, Fatih Guvenen, Benjamin S Smith, Jae Song, and Till von Wachter, "The disappearing large-firm wage premium," *AEA Papers and Proceedings*, 2018, *108*, 317–322.

Brown, Charles and James Medoff, "The employer size-wage effect," *Journal of political Economy*, 1989, *97* (5), 1027–1059.

Caliendo, Lorenzo, Ferdinando Monte, and Esteban Rossi-Hansberg, "The anatomy of French production hierarchies," *Journal of Political Economy*, 2015, *123* (4), 809–852.

Cortes, Guido Matias, Nir Jaimovich, and Henry E Siu, "Disappearing routine jobs: Who, how, and why?," *Journal of Monetary Economics*, 2017, *91*, 69–87.

Drucker, P.F., *The Practice of Management* Classic Drucker collection, Butterworth-Heinemann, 2007.

Gabaix, Xavier and Augustin Landier, "Why has CEO pay increased so much?," *The quarterly journal of economics*, 2008, *123* (1), 49–100.

Kremer, Michael, "The O-ring theory of economic development," *The quarterly journal of economics*, 1993, *108* (3), 551–575.

Kwon, Spencer Y., Yueran Ma, and Kaspar Zimmermann, "100 Years of Rising Corporate Concentration," *American Economic Review*, 2024, *114* (7), 2111–2140.

Lynch, Lisa M and Sandra E Black, "Beyond the incidence of employer-provided training," *ILR Review*, 1998, *52* (1), 64–81.

Moore, Henry Ludwell, *Laws of wages: An essay in statistical economics*, Macmillan, 1911.

Oi, Walter Y. and Todd L. Idson, "Chapter 33 Firm size and wages," in "Handbook of Labor Economics," Vol. 3 of *Handbook of Labor Economics*, Elsevier, 1999, pp. 2165–2214.

Piketty, Thomas, Emmanuel Saez, and Stefanie Stantcheva, "Optimal taxation of top labor incomes: A tale of three elasticities," *American economic journal: economic policy*, 2014, *6* (1), 230–271.

Pinsonneault, Alain and Kenneth L Kraemer, "The impact of information technology on middle managers," *Mis Quarterly*, 1993, pp. 271–292.

— **and** — , “Middle management downsizing: An empirical investigation of the impact of information technology,” *Management science*, 1997, *43* (5), 659–679.

Piore, Michael J. and Charles F. Sabel, *The Second Industrial Divide: Possibilities for Prosperity*, New York, NY: Basic Books, 1984. Recommended Citation: Piore, Michael J. and Sabel, Charles F., ”The Second Industrial Divide: Possibilities for Prosperity” (1984). Faculty Books. 171. https://scholarship.law.columbia.edu/books/171.

Rosen, Sherwin, “The economics of superstars,” *The American economic review*, 1981, *71* (5), 845–858.

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter, “Firming Up Inequality*,” *The Quarterly Journal of Economics*, 10 2018, *134* (1), 1–50.

Stelzner, Mark, “The new American way—how changes in labour law are increasing inequality,” *Industrial relations journal*, 2017, *48* (3), 231–255.

U.S. Bureau of Labor Statistics, May 2023.

A Appendix

A.1 Data Appendix

We selected the top 1% highest-paid employees within a given occupation and year across all firms, totaling 1,740,888 employees. We exclude the firms with below 10 employees and then divided them into 20 bins based on payroll size, The upper figure A5 illustrates the distribution of these top 1% paid employees across the 20 bins, and the bottom figure shows the distribution of the top 5% highest-paid employees across the same 20 bins.

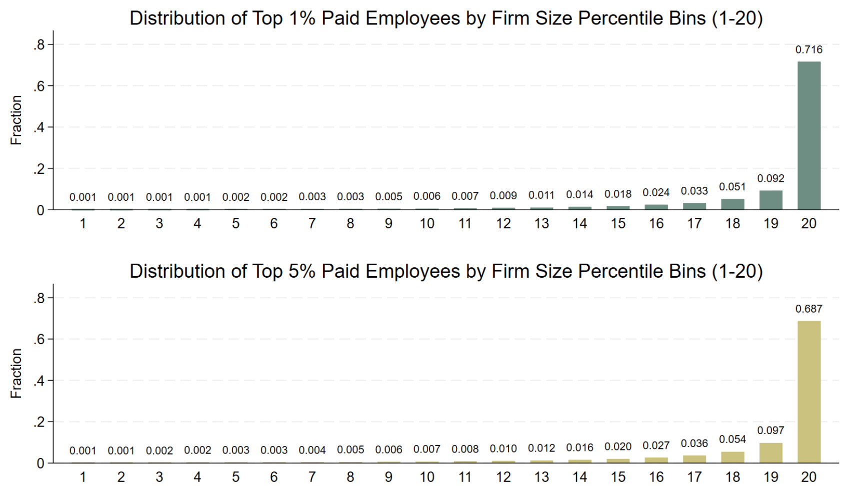

Figure A5: Distribution of Top Paid Employees Across 20 Firm Payroll Size Bins

A.2 Extracting Task Characteristics from Job Descriptions for Figure 4

We calculate the monthly average wage for each individual-firm-occupation within a given year. Our dataset contains over 171 million observations. Each individual is uniquely identified by a combination of hashed Social Security Number (SSN) and birth year. For each individual, we observe demographic information such as age, gender, and job titles. The dataset includes more than 6,494 distinct job titles across over 900 O*NET occupations, with each job accompanied by a detailed description constructed by the data provider's research team.

We utilized the ChatGPT-4 model to estimate the scores of various job features based on these textual descriptions of jobs. For each job description, GPT-4 was prompted to provide responses to a set of 12 questions to assess specific job features. We queried the model three times per job, and about 96% of the responses were identical or very similar. For these, we averaged the scores. Significant discrepancies were reviewed for accuracy. The prompts used for querying GPT-4 are shown below:
Temperature = 0

```
prompt = """
Job Description:
{job_title} {job_description}
Please answer the following questions about the job description provided.
Use only numbers for your answers, and separate each answer with a comma.

1. To what extent is the focus of this job on creating information, ideas, or
   insights that can be widely shared throughout the organization? (Answer 1-10)
2. To what extent is the focus of this job managing, processing, or
   representing information? (Answer 1-10)
3. To what extent does this job require interpersonal tasks, i.e., interacting
```

```
   with people frequently? (Answer 1-10)
4. To what extent does this job require physical proximity to people, equipment,
   or objects in the workplace? (Answer 1-10)
5. To what extent does this job require manual labor or physical tasks? (Answer 1-10)
6. To what extent can this job be done remotely, without being physically present
   in the workplace? (Answer 1-10)
7. To what extent are there negative consequences if someone makes a mistake in
   this job? (Answer 1-10)
8. To what extent does this job involve inventing, developing, or refining new
   scientific, engineering, or artistic ideas? (Answer 1-10)
9. To what extent do the decisions or actions of people in the job affect the
   overall value of the firm (versus only having a local effect)? (Answer 1-10)
10. What percent of the firm's employees are affected by the decisions of someone
    in this job? (Answer 0-100)
11. In a typical firm, how many people report directly to this individual?
    (Answer with a number)
12. Which primary category does this job belong to: Informational (1),
    Interpersonal (2), or Physical (3)?

Please provide your answers in a single line, with numbers only, separated by commas.
"""
```

These estimates show considerable variation within occupations. Figure A6 shows the distribution of scores for the characteristic "negative consequences of mistakes" for 32 selected occupations, as representative.

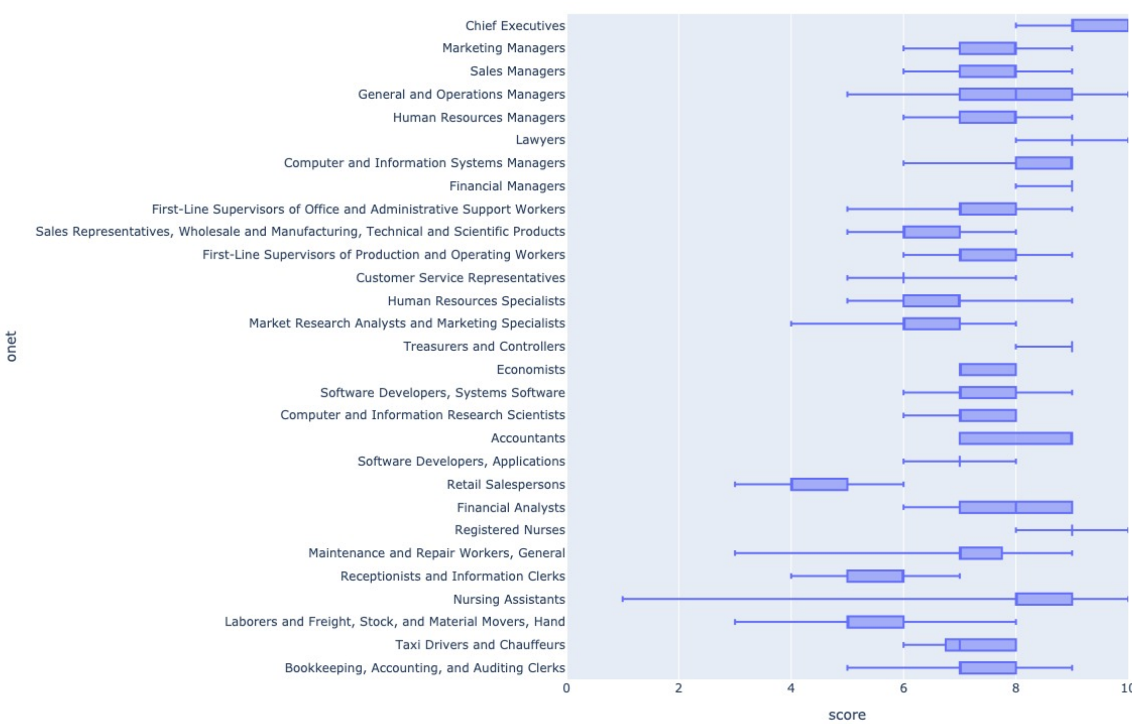

Figure A6: Box and whisker plot showing minimum, 25th percentile, 50th percentile, 75th percentile, and maximum values for the estimated importance of "Negative consequences of mistakes" for 32 selected occupations.

A.3 Additional Tables Relating to Main Text

In Table A11, we present a more granular analysis by using finer grid divisions, allowing for a comprehensive exploration of the relationship between firm size and wage distribution across different percentiles within a specific occupation. Our analysis reveals a notable trend: the scaling effects are more pronounced among higher-paid workers. Specifically, as the size of an establishment grows, the wages of workers in the upper percentiles of the pay scale within an occupation increase at a faster rate compared to those in the lower percentiles.

Dependent variable is the log of wages by percentile	(1) p1	(2) p25	(3) median	(4) p75	(5) p90	(6) p95	(7) p99	(8) highest-paid
Panel A:								
Log of payroll	0.0178***	0.0671***	0.0863***	0.1069***	0.1293***	0.1415***	0.1590***	0.1690***
	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0012)	(0.0011)	(0.0012)
Observations	9,606,727	9,606,727	9,606,727	9,606,727	9,606,727	9,606,727	9,606,727	9,606,727
R-squared	0.440	0.4754	0.5115	0.5083	0.4970	0.4921	0.4927	0.4948
Panel B:								
Log of employment size	-0.0311***	0.0224***	0.0416***	0.0622***	0.0860***	0.0995***	0.1199***	0.1321***
	(0.0011)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0013)	(0.0014)	(0.0014)
Observations	9,606,727	9,606,727	9,606,727	9,606,727	9,606,727	9,606,727	9,606,727	9,606,727
Adj. R^2	0.4412	0.4572	0.4836	0.4733	0.4557	0.4483	0.4467	0.4478
Year FE	Y	Y	Y	Y	Y	Y	Y	
2-digit NAICS FE	Y	Y	Y	Y	Y	Y	Y	
6-digit O*NET FE	Y	Y	Y	Y	Y	Y	Y	Y

Table A11: Table providing estimates underlying figure 5.1

A.4 Additional Results on Wage Scaling by Establishment Size Instead of Firm Size

In this section, we present the results of the three proposed estimation equations using establishment-level data. One additional variable we have is the location of each establishment. Since workers are dispersed at the establishment level than at the firm level, we calculate the most concentrated commuting zone based on each worker's home zipcode and use it as a proxy for market size.

Table A13 provides results of estimates wage scaling for highest-paid workers in each establishment-occupation-year in our data. These are regressions where the log wage of an employee is explained by one or two measures of market size, as well as industry, year, and occupation fixed effects across all occupations. As can be seen there is a robust relationship between both measures of market size and wage, whether or not attention is restricted to a balanced sample of establishments. Column 5, which includes both measures of market size as explanatory variables, shows that wages across all occupations increase 10.8% on average in an establishment twice as large, and increase 6.1% on average in a commuting zone twice as large. These estimates of how wages scale with market size are only slightly attenuated when both are included, suggesting that these are measuring separate dimensions of scaling, rather than both being proxies for the same phenomenon.

Table A14 also reports estimates of wage scaling, but with the dependent variable now being the median paid workers in each establishment-occupation-year. The analysis reveals a similar pattern of scaling with different market sizes, although the magnitudes of the effects are smaller.

In Table A15, we restricted our sample to single establishment firms only, and the results are qualitatively similar for wage scaling among both highest-paid workers and median workers.

Variable	Obs	Mean	Std. dev.	Min	Max
Size of employment	12,543,425	289.970	987.142	10.00	122,527.50
Log of employment size	12,543,425	4.789	1.201	2.398	11.716
Annual payroll of an establishment	12,543,425	2.29E+07	1.07E+08	1.31E+04	1.58E+10
Log of annual payroll of an establishment	12,543,425	15.714	1.490	9.481	23.480
Total population in a commuting zone	11,057,081	4.65E+06	5.15E+06	2.12E+03	1.87E+07
Log of total population in a cz	11,057,081	14.667	1.347	7.658	16.743
Average monthly wages for highest-paid workers	12,543,425	11,330.47	297,388.20	1,160.00	9.60E+08
Average monthly wages for median workers	12,543,425	6,818.71	52,883.28	1,160.00	6.79E+07
Average monthly wages for average workers	12,543,425	7,547.53	63,310.06	1,160.00	8.63E+07
Log of average monthly wages for highest-paid workers	12,543,425	8.831	0.769	7.057	20.683
Log of average monthly wages for median workers	12,543,425	8.522	0.664	7.057	18.033
Log of average monthly wages for average workers	12,543,425	8.596	0.675	7.057	18.273
Fraction of IT workers in an establishment	12,543,425	0.022	0.049	0.000	1.000
Fraction of IT developers in an establishment	12,543,425	0.010	0.035	0.000	1.000
Normalized score for abstract tasks	12,172,655	0.470	0.289	0	1
Normalized score for routine tasks	12,172,655	0.280	0.274	0	1
Normalized score for manual tasks	12,172,655	0.067	0.086	0	1
BDS number of establishments at 2-digit NAICS	11,822,377	506,745.1	284,238.1	17,744	1,005,580
2-digit NAICS code	11,822,377	49.809	14.638	11	81

Table A12: Establishment-Level Summary Statistics

Dependent variable is the log of wages for highest-paid workers	(1)	(2)	(3)	(4)	(5)
Log of total payroll	0.164***			0.159***	
	(0.001)			(0.001)	
Log of employment size		0.109***			0.108***
		(0.001)			(0.001)
Log of commuting zone population			0.065***	0.047***	0.061***
			(0.001)	(0.001)	(0.001)
Observations	11,822,374	11,822,374	10,433,673	10,433,673	10,433,673
R-squared	0.495	0.443	0.430	0.500	0.453
Year FE	Y	Y	Y	Y	Y
NAICS2 FE	Y	Y	Y	Y	Y
Region FE	Y	Y	Y	Y	Y
O*NET FE	Y	Y	Y	Y	Y

Table A13: Wage scaling for highest-paid workers with establishment size and market size

Notes: The dependent variable is the logarithm of the monthly total taxable income for highest-paid workers in a given establishment-occupation-year. The independent variables include the logarithm of the establishment annual payroll), the logarithm of employment size, and the logarithm of the population in the relevant commuting zone. The regression model includes fixed effects for year, 2-digit NAICS category, regional location, and O*NET occupation classification. The regression is weighted by the number of establishments classified under NAICS2, using data from the Business Dynamics Statistics (BDS) dataset. Robust standard errors clustered at establishment level in parenthesis. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Dependent variable is log of wages for median workers	(1)	(2)	(3)	(4)	(5)
Log of total payroll	0.099***		0.095***		
	(0.001)		(0.001)		
Log of employment size		0.040***		0.039***	
		(0.001)		(0.001)	
Log of commuting zone population			0.042***	0.052***	0.053***
			(0.001)	(0.001)	(0.001)
Observations	11,822,374	11,822,374	10,433,673	10,433,673	10,433,673
R-squared	0.506	0.474	0.513	0.484	0.480
Year FE	Y	Y	Y	Y	Y
NAICS2 FE	Y	Y	Y	Y	Y
Region FE	Y	Y	Y	Y	Y
O*NET FE	Y	Y	Y	Y	Y

Table A14: Wage scaling for median workers with establishment size and market size

Note: The dependent variable is the logarithm of the monthly total taxable income for median workers in a given establishment-occupation-year. The independent variables include the logarithm of the establishment annual payroll, the logarithm of employment size, and the logarithm of the population in the relevant commuting zone. The regression model includes fixed effects for year, 2-digit NAICS category, regional location, and O*NET occupation classification. The regression is weighted by the number of establishments classified under NAICS2, using data from the Business Dynamics Statistics (BDS) dataset. Robust standard errors clustered at establishment level in parenthesis. *** p<0.01, ** p<0.05, * p<0.1

	(1)	(2)	(3)	(4)
	log(wages for highest-paid workers)		log(wages for median workers)	
Log of establishment payroll	0.167***		0.140***	
	(0.002)		(0.001)	
Log of employment size		0.076***		0.054***
		(0.002)		(0.002)
Observations	4,822,339	4,822,339	4,822,339	4,822,339
Adj.R^2	0.154	0.100	0.170	0.117
Year FE	Y	Y	Y	Y
2-digit NAICS FE	Y	Y	Y	Y
Region FE	Y	Y	Y	Y
6-digit O*NET FE	Y	Y	Y	Y

Table A15: Wage scaling among standalone establishments

	(1)	(2)	(3)	(4)
	log(wages for highest-paid workers)		log(wages for median workers)	
Log of establishment payroll	0.175***	0.169***	0.109***	0.108***
	(0.001)	(0.001)	(0.001)	(0.001)
AI-intensive industries	0.135***	-0.228***	0.110***	0.042
(NAICS2 = 51, 52, 54)	(0.002)	(0.031)	(0.002)	(0.026)
Log of Est payroll × AI-intensive		0.023***		0.004**
industries		(0.002)		(0.002)
Observations	13,113,492	13,113,492	13,113,492	13,113,492
Adj.R^2	0.479	0.479	0.484	0.484
Year FE	Y	Y	Y	Y
2-digit NAICS FE	Y	Y	Y	Y
Region FE	Y	Y	Y	Y
6-digit O*NET FE	Y	Y	Y	Y

Table A16: Effects of industry-level AI-intensities on wage scaling with establishment size.

Log of wages for highest-paid workers	(1)	(2)	(3)
Log establishment size	0.144***	0.143***	0.144***
	(0.001)	(0.001)	(0.001)
R&D and creative workers	-0.191***		
	(0.023)		
Log establishment size × R&D and creative workers	0.004***		
	(0.001)		
R&D workers		-0.029	
		(0.040)	
Log establishment size × R&D workers		0.006**	
		(0.002)	
Creative workers			-0.152***
			(0.026)
Log establishment size × creative workers			-0.002
			(0.002)
Observations	14,148,300	14,148,300	14,148,300
Adj.R^2	0.166	0.165	0.166
Year FE	Y	Y	Y
2-Digit NAICS FE	Y	Y	Y
2-Digit O*NET FE	Y	Y	Y

Table A17: Effects of superstar wage scaling with establishment size for R&D and creative workers.