Enhancing Team Productivity through Shorter Working Hours: Evidence from the Great Recession

Ruo Shangguan*
Waseda University

Jed DeVaro†
California State University, East Bay

Hideo Owan‡ §
Waseda University

December 31, 2020

Abstract

When output demand drops during recessions, employers decrease labor inputs by cutting workers and/or hours. Cutting hours that (pre-recession) were excessive might increase labor productivity, given an inverted-U-shaped hours-productivity profile; less worker exhaustion implies higher effort-per-hour. Teams exhibit a concurring effect, where labor reallocation causes hours to be concentrated among top performers after total hours decrease. The adjustment process is examined using single-firm Japanese data on construction design projects. A theoretical model is proposed to analyze within-team labor allocation. Its parameters are calibrated with data to quantify its predictions. Regression results derived from the actual data match the model's predictions.

Keywords: team production, labor productivity, working hours, allocation of labor, labor changes during recessions, Great Recession, global financial crisis

JEL classification: J23, J24, M50, M54

---

*Institute for Empirical Research in Organizational Economics. Email: roy1312@gmail.com
†Departments of Economics and Management. Email: jed.devaro@csueastbay.edu
‡School of Political Science and Economics. Email: owan@waseda.jp
§Helpful comments from John Pencavel, Kensuke Teshima, Sachiko Kuroda, Yuta Toyama and seminar participants at the 2020 Organizational Economics Workshop (Sydney, Australia), Waseda University and Hitotsubashi University are gratefully acknowledged.

1 Introduction

There's an old saying, "Never let a good crisis go to waste." One interpretation is that a deep recession offers employers the opportunity to reinvent themselves, possibly for the better, while at the same time the big changes in employer and worker behavior that occur during a recession offer researchers the opportunity to gain deeper insights into the operation of the workplace. A universal feature of recessions is that firms must decrease labor inputs in response to decreased demand for products and services, but the approach for cutting back on labor varies with the institutional environment. In a labor market with relatively light institutional constraints, like the U.S., much of the adjustment occurs in the form of layoffs and firings. In labor markets that are more institutionally constrained and in which layoffs and firings are rarer, like Japan, much of the adjustment occurs in terms of hours and the termination of non-standard contract workers.[1]

Those two approaches for reducing labor inputs can have different implications for labor productivity. In the U.S. context, where the threat of getting fired during a recession is very real, Lazear *et al.* (2016) analyzed individual worker productivity data from a single firm during 2006 to 2010 and found evidence that workers' efforts increased during the 2008-2009 financial crisis. The underlying theoretical argument for that result has an efficiency wage flavor, *à la* Shapiro and Stiglitz (1984), namely that workers' worsening outside options and fear of losing their jobs in the wake of widespread layoffs and firings spur them to invest greater effort. That mechanism is likely to be less potent in settings like Japan, where firings are relatively rare. In those settings, even though a recession worsens workers' outside options, this would not affect their effort investments much, given that their jobs are secure.

How might productivity respond, then, to a significant cutback in hours? A potential answer is offered in Pencavel (2015), using data on women working in manufacturing plants to produce artillery shells for the British military during the First World War. That study provided evidence of an inverted-*U*-shaped hours-productivity relationship in which a worker's productivity falls when working hours become very long.[2] An inter-

[1]Japan's Labor Contract Law prohibits "abuse of the right to dismiss", which basically means that the firm cannot lay off its workers unless it has made reasonable efforts to avoid that action. An implication is that the firm generally cannot lay off its workers when many workers are working overtime, because in that case the court would order the firm to reduce working hours to standard working hours before reducing its employment. The following newsletter contains a concise explanation of how "abuse of the right to dismiss" is defined.

https://www.jurists.co.jp/sites/default/files/newsletter_pdf/newsletter_1701_labor_employment_law.pdf

The preceding argument, however, applies only to regular workers. Terminating contracts with workers hired under fixed-term contracts is not considered an "abuse of the right to dismiss".

[2]See DeVaro (2020) for a theoretical model that yields this inverted *U* as a prediction and for empirical

pretation is that there is substitution between working hours and effort per hour; when hours become very long effort per hour falls, perhaps due to workers' fatigue and exhaustion.[3] This implies that productivity could increase during a recession, as long work schedules are reduced.[4]

Thus, Lazear *et al.* (2016) and Pencavel (2015) offer different explanatory mechanisms for why productivity (i.e., effort per hour) may increase during a recession. In Lazear *et al.* (2016), the mechanism is increased effort per hour even when hours remain fixed, as workers hustle to avoid getting canned. In Pencavel (2015), the mechanism is increased effort per hour in the wake of a reduction in working hours, because workers have more energy and are less exhausted due to their shorter work schedules.[5] A priori, of the two narratives, the latter would appear to be more relevant to Japan, both because hours are long in Japanese firms and because the risk of getting fired is low.

The present study's goal is to provide a deeper understanding of the preceding narratives in the context of team production rather than individual production, via empirical analysis of a unique data set of construction projects in a Japanese architectural and engineering consultancy firm during the years 2004 to 2016, a time span that includes the global financial crisis from December 2007 to June 2009. Specifically, the hours-productivity profile is analyzed for construction project design teams. The focus on teams is novel. It is also important given that team production is a common aspect of many production settings. Team production introduces two additional elements into an analysis of the hours-productivity relationship. The first is complementarity among the outputs of individual team members; interdependence among team members' outputs is, in fact, one of the main reasons why employers organize production in teams. In the presence

evidence of this pattern in recent worker-employer matched panel data from the U.K. that span a wide variety of occupations and industries.

[3]Pencavel (2016a) used the same data as in Pencavel (2015) to provide further evidence consistent with exhaustion. Specifically, workers who worked long hours in a given week were found to have lower productivity in the subsequent week. More generally, there is a large literature on the negative health effects of long working hours, including injuries and accidents, cardiovascular disease, reduced cognitive function, and diminished well-being of the household that may induce stress as worklife balance deteriorates. In one recent example, Sato *et al.* (2020) provide evidence that long working hours, as well as night and weekend work, are associated with a deterioration in mental health. See chapter 6 of Pencavel (2018) for a survey of the literature. All of the preceding consequences of long working hours can plausibly contribute to diminished productivity of the afflicted workers.

[4]If working hours are excessively long before recession (as in many Japanese firms, including the one investigated in the present study), why do Japanese firms/workers choose such inefficient locations on the hours-productivity profile, far to the right of the inflection point? One well-accepted view among economists is that, due to high adjustment costs of labor (i.e., firms need to hoard labor during recessions), firms underemploy workers during normal times and require them to work long hours to meet demand (Kuroda and Yamamoto 2013).

[5]Pencavel (2015) was not focused on changes induced by recessions, per se. But the argument presented here is a natural implication of his finding of an inverted-U-shaped hours-productivity profile.

of complementarities, an increase in an overworked employee's individual productivity due to reduced working hours can be amplified via a ripple effect that improves the productivity of other team members. The second element concerns how employers allocate labor hours within teams of workers with heterogeneous abilities. The marginal productivity of an additional hour that is assigned to a team of a given size depends on which team member is assigned that hour. As assigned hours increase to meet demand for the firm's output, the time constraints of the team's highest-ability workers begin to bind, which requires the employer to assign further hours to team members of lesser ability.

Sections 3 through 5 present the analysis, which begins by documenting a productivity increase coinciding with and following the crisis and by showing that the downward adjustment in the labor input during the crisis occurred more for working hours than for employment. The average team-level productivity is found to increase by 7.6% after the crisis. Evidence on the within-team allocation of working hours reveals a considerable concentration of working hours. That is, a small number of workers contribute the bulk of the team's working hours. That observation motivates a new theoretical framework for understanding within-team labor allocation. The model's workers, who differ in their abilities and time endowments, are assigned working hours based on their absolute advantages in production, and they are allocated to tasks based on their capacities. The most able workers are assigned hours first. When product demand overwhelms those workers' capacities, additional hours are assigned to less able workers, which decreases average team productivity.

The model's parameters are calibrated with data to quantify the magnitudes associated with the model's predictions. The calibrated model generates an average team-level productivity increase of 7.3% after the crisis, which is statistically indistinguishable from that found in the real data. Moreover, average simulated worker-level productivity increases by 3.1%, which only explains 42% of the team-level productivity increase. Therefore, the results suggest that labor reallocation and worker complementarity play important roles in explaining team-level productivity changes. Additionally, the calibrated model successfully generates several patterns that are quantitatively similar to those found in the data, including: (1) the labor share concentrates more heavily with the team's top 2 workers, and team size decreases after the crisis; (2) the productivity improvement is greater for larger teams and less productive teams; (3) larger teams exhibit lower average productivity because weaker workers join teams when more hours are needed than the stars can handle.

In further empirical evidence concerning overtime, increases in within-team overtime shares are found to be associated with decreases in team productivity. The magnitude of those decreases tends to be highest (lowest) for the team members who contribute the most (least) hours to the project. A negative impact of overtime is also found in a two-stage least squares (2SLS) regression using industry demand as the instrument for the team-level overtime share. Finally, more overtime hours are found to be positively associated with more frequent design defects.

The setting offers an ideal laboratory for investigating the productivity effects of long working hours. One reason is that Japan is famous for long hours.[6] Moreover, layoffs and firings are less common in Japan than in the U.S., so reductions in labor inputs must befall other margins like overtime hours, which is exactly what happened during the crisis. From the standpoint of a single firm, the crisis is an exogenous event that provides the variation necessary to identify the effect of interest. Focusing on a particular firm and industry also holds constant the considerable firm and industry heterogeneity that would otherwise complicate the interpretation of results in a broader sample. Given its strong sensitivity to the business cycle, construction is an ideal industry for studying firm responses to business cycle shocks.

This study contributes by providing evidence outside of the oft-studied (in the teams literature) manufacturing sector, especially in the knowledge-intensive professional jobs where teamwork is becoming the norm--design, R&D, consulting, accounting, auditing, academic research, etc. Within-team heterogeneity in hours arises in such settings because team members can work in different places, at different times, and for different durations.[7] Economists have been unable to study productivity in such occupations given their idiosyncratic outputs. The present setting and data facilitate productivity measurement and analysis because the production process is standardized enough so that the total labor required to complete each job is predictable, and the value of the output is fixed on each project before teamwork commences. Consequently, productivity depends only on total inputs.

The focus on within-team heterogeneity in working hours is new. The underlying theoretical framework is based on the idea that heterogeneity in hours is a consequence of heterogeneity in team members' individual abilities, where the employer assigns the most able workers to tasks first, followed by the less able workers. Although team com-

[6] As noted in Yamamoto (2016), "The length of work hours in Japan stands out among industrialized nations. According to the International Labour Organization (ILO) statistics and other sources, the percentage of workers working long hours (defined as at least 49 hours per week) in recent years is about 10%-16% in North America and Europe, but 22% in Japan."

[7] In contrast, manufacturing jobs often require workers to be physically and temporally proximate. On an assembly line, for example, complementarities are achieved only when the team members are physically present at the same time, so within-team heterogeneity in working hours (regardless of heterogeneity in abilities) is limited or nonexistent.

position is endogenous in the present framework, the available talent pool of candidates is randomly drawn, which implies significant variation in the distribution of available skill levels. This team formation procedure in the model creates a negative correlation between heterogeneity and productivity because less productive teams tend to add more workers from the lower part of skill distribution. It is shown that the real data and those simulated from the calibrated model exhibit a similar pattern–team heterogeneity in skills and team size are negatively correlated with productivity.

There is a related literature on team diversity and productivity. Although many dimensions of heterogeneity have been explored, the present study investigates the productivity implications of worker heterogeneity in ability and, consequently, in assigned working hours. There are theoretical rationales for both positive and negative team-level productivity effects.[8] Hamilton *et al.* (2012) discuss gains from task coordination and peer learning. Productivity improves when workers' skill levels and task difficulties are optimally matched or when more experienced workers share their knowledge with less experienced ones. When team members are peers who compete with each other for advancement within the organization, additional implications are derived. Classic tournament theory (Lazear and Rosen 1981) predicts that heterogeneity in ability depresses incentives, which would hurt team performance.[9] In contrast, the market-based tournament model of Gürtler and Gürtler (2015) shows that the opposite prediction can arise.[10] Empirical evidence favors a positive effect of heterogeneity in ability on team performance.[11]

2 Data and Measures

The data come from a large Japanese architectural and engineering consultancy firm and include personnel records (from 2011 to 2016) and project management data (from fiscal

[8]There are also theoretical rationales for both positive and negative team-level productivity effects on dimensions of heterogeneity other than individual ability (e.g., various demographic characteristics). The positive view is that diversity broadens the set of perspectives and approaches that team members bring to the table, which fosters creativity, scope for complementarities, and ultimately high group performance. The negative view, which is supported by the preponderance of the evidence (Mannix and Neale 2005), is that diversity induces communication challenges and social divisions that hurt group performance. See Lazear (1999) for discussions in line with the positive view and Lang (1986), Kandel and Lazear (1992) for discussions on the negative view.

[9]The reason is that the high-ability workers do not need to exert much effort because they are likely to win regardless, and the low-ability workers do not exert much effort because their chances of winning are low regardless.

[10]The idea is that winning a promotion against a competitive pool characterized by a wide range of talent causes competing employers in the labor market to update their beliefs about the winner's ability to a greater extent than if the worker had prevailed over a level playing field. Workers anticipate large prizes from promotion due to this larger updating, which creates a strong incentive to exert effort to try to win the prize. See also Deutscher *et al.* (2020).

[11]See Hamilton *et al.* (2003, 2012),Franck and Nüesch (2010), Parrotta *et al.* (2014), and Garnero *et al.* (2014).

years 2004 to 2016). The empirical analysis is also informed by in-person interviews that the authors conducted with seven of the firm's managers and by other less formal communication with the firm's human resource managers.[12] The personnel records cover all employees, including dispatched or contract workers who may be included in the project management data, and include education, salary, and hierarchical ranks that are classified into three levels (manager, senior architect, and junior architect). Projects consist of multiple phases, called jobs, and the job is the unit of observation.[13] Contracts are negotiated separately for each job in a project, with contract terms set before the job begins. Throughout the analysis, workers, jobs, and time periods are indexed by i, j, and t. To differentiate within multi-year time periods, years are indexed by y, and months are indexed by m.

The sequence of jobs in a particular construction project might be as follows: initial planning, schematic design, structural design, detailed design, technical design and engineering, and supervision of the construction process. Each job is performed in a team of about 14 workers and is assigned to a chief manager who is the top person fully responsible for the job and who bears a penalty in the event of quality problems. Chief managers may also join teams in the capacity of expert consultants, to provide technical guidance. The chief manager's responsibilities include identifying one or more team leaders, usually senior architects, to lead daily operations. Teams also have junior architects who execute tasks (e.g., drawing pictures after the details of the design are confirmed).

The data include two kinds of jobs. External jobs are profit-center jobs that generate revenue. Internal jobs are cost-center jobs that mainly entail administrative responsibilities. Given its use of a productivity measure based on revenue, Rev_j, the present study focuses on external jobs, though it should be noted that internal responsibilities may also contribute to overtime hours. Revenue, costs (both labor and non-labor), and other characteristics are observed for each job. Finer components of nonlabor costs are also observed, including material/traveling costs and three types of outsourcing costs. The project management data also include information on the client's identity and industry, type and size of the building being designed, location where the work is conducted, phase of work, contractor selection method, etc.

Information about the entire industry comes from the annual survey conducted by the Ministry of Land, Infrastructure, Transport and Tourism (MLIT).[14] For each client industry and each type of construction, the survey reports the annual amount of total

[12]The seven managers were selected on the basis of the manager effects estimated in Shangguan and Owan 2019, which also contains further details about the data.

[13]Usage of the word "job" here differs from that in either the personnel economics literature or the forthcoming theoretical model. In the context of the data, a job is a phase of a longer-term project.

[14]http://www.mlit.go.jp/statistics/details/kkoji_list.html

orders received by the 50 largest construction companies. In the present analysis, the survey data are aggregated to the industry level and then connected to the data described above using the industry information recorded for each job.

2.1 Hours and overtime

Interviews with the firm's managers revealed that the managers allocate tasks across workers and set workers' hours.[15] They also conduct regular internal meetings to communicate about the status of each worker on each project, so as to make better subsequent labor allocation decisions. The data on working hours are available at the worker-job-month level. They are from the project management data and reported by the workers for internal accounting purposes.[16]

Let h_{ijm} denote the total working hours of worker i on job j in month m, and let h_{im} denote the total working hours of worker i across all jobs in month m. Worker i's overtime hours in month m are defined as $OT_{im} = \max(0, h_{im} - h^*_m)$, where h^*_m is 8 times the number of business days in month m. The exact day on which the labor input transaction occurred is unobserved. Therefore, a worker's overtime hours cannot be allocated to each job exactly, and an assignment rule is needed. To calculate overtime for each job, each worker's monthly overtime is allocated to each job by her total working hours on each job (i.e., the weight for allocating total overtime is $\frac{h_{ijm}}{\sum_j h_{ijm}}$, so that $OT_{ijm} = \frac{h_{ijm}}{\sum_j h_{ijm}} OT_{im}$). Finally, the overtime share for each job-year is defined as follows, where the summation is over all workers i and months m for each year y: $\frac{OT_{jy}}{h_{jy}} = \frac{\sum_{i,m} OT_{ijm}}{\sum_{i,m} h_{ijm}}$. Total hours on job j, and total overtime hours on job j, are defined by $h_j = \sum_{i,m} h_{ijm}$ and $OT_j = \sum_{i,m} OT_{ijm}$.

[15] The fact that hours are clearly employer assigned in this setting eliminates an identification problem that plagues the literature on working hours, i.e., do observed hours reflect workers' preferences or employers' preferences? As discussed in Pencavel (2016b) that identification question received attention in the 1960s and 1970s (e.g., Feldstein 1968, Rosen 1969, Abbott and Ashenfelter 1976) but was then largely forgotten for more than four decades as the empirical literature became dominated by labor supply models that implicitly resolved the preceding question in favor of workers' preferences. The present study's operating assumption of employer-determined hours is particularly appropriate given the focus on hours variation over the business cycle. As Pencavel (2016b, p. 18) notes, "A role for employers' preferences in the determination of hours of market work that appears to be widely acknowledged concerns business cycle movements in hours. A well-established pattern is that hours are pro-cyclical and, moreover, that movements in hours precede turning points in business activity (production, sales, and new orders). A cut in the working hours of employees tends to be among employers' first reactions to the over-accumulation of inventories and to the weakening of new orders."

[16] One might wonder whether workers have an incentive to re-allocate their hours to another job. If a job incurs too many costs, workers might re-allocate their hours to another high-margin job, so as to please their bosses. The management, however, strongly discourages such behavior. The interviewed managers revealed that large-scale manipulation was infeasible, but workers might shift hours from a low-revenue job to a high-revenue job for the same client.

2.2 Productivity

Job-level revenue is the main output measure. Since detailed labor inputs are also observed, a natural choice for the dependent variable that measures team productivity is the ratio between revenue and total working hours. However, typically some of the work is outsourced to third parties, whose labor inputs are not observed in the data. To render the input and output measures compatible, revenue is adjusted using each job's outsource costs. Let O_j be the outsource ratio for each job, which is calculated as the ratio between outsource costs, $OutsourceCost_j$, and job-level total costs, $Cost_j$. "Adjusted revenue" is then defined as $Adj\_Rev_j = Rev_j\left(1 - O_j\right) = Rev_j - \frac{Rev_j}{Cost_j} OutsourceCost_j$. In this definition, $OutsourceCost_j$ is adjusted for the markup before subtracting it from revenue. The job-level productivity measure that serves as the main dependent variable is $\frac{Adj\_Rev_j}{h_j}$, or the ratio of adjusted revenue to total working hours.[17]

Revenue is determined before the job starts and mainly depends on job characteristics such as job content, client industry, building type, floor size, etc. The relationship between revenue and the preceding categorical varaibles is relatively stable across years, as will be shown subsequently. Thus, any variation in the productivity measure conditional on the job characteristics comes from the denominator, h_j. Therefore, it is natural to treat h_j as a proxy for the (negative) productivity measure. Given that $\frac{Adj\_Rev_j}{h_j}$ is left skewed and that it depends on the size of the job (in terms of revenue), lnh_j is used as the dependent variable, and $lnAdj\_Rev_j$ is included as a control in the forthcoming regression analysis. In such regressions, the coefficients of the independent variables can be interpreted as (negative) productivity effects.

2.3 Other variables

$Area_j$ denotes the floor area (measured in square meters) of the building being designed in job j. $TeamSize_j$ denotes the number of workers engaged on job j. $Defect_j$, which is only available for jobs starting in 2011 or later, is a dummy equaling 1 if a defect was detected after delivery for job j and its chief manager was penalized for the error. $JobContent_j$ denotes a categorical variable (with 22 categories) that controls for the type of service

[17] Another way to adjust for outsourcing is to define $Adj\_Rev_j = Rev_j - O_j$,but in that case productivity goes to infinity as the outsource ratio goes to 100%. The firm avoids outsourcing too many tasks because of its hidden costs, such as difficulty in coordination, loss of control over time management, loss of learning opportunities, etc. The definition of adjusted revenue accounts for such hidden costs.

in each job j.[18] Ind_j denotes a vector of 39 dummies indicating the client's industry.[19] $IndOrder_{jy}$ indicates the industry-level total orders for the client's industry defined for job j in year y, obtained from the MLIT survey. Let $AfterCrisis_j$ denote a dummy equaling 1 if the starting year for job j occurred in the immediate post-crisis period (i.e., years 2010, 2011, or 2012).

2.4 Sample selection and summary statistics

Only jobs with at least one million Japanese yen are included. Since that revenue threshold is rather low across jobs, this restriction essentially excludes failed jobs that do not generate any revenue. Jobs with $Area_j = 0$ are excluded, which tend to be consulting jobs that differ in nature from design jobs. Jobs are required to be completed so that the labor input records are complete. Jobs with all costs outsourced are dropped, since for those there is essentially no valid labor input data. Although the data cover the period from 2004 to 2016, only jobs that started from 2004 to 2013 are included, to avoid right censoring. A typical job lasts longer than one year, and a big one could last three years. Similarly, jobs that started before 2004 are excluded and are expected to be longer jobs. The need to drop observations beyond 2013 is clear from Figure 1, which plots the average duration of jobs by their starting years. A sharp drop is observed in 2014 as a consequence of right censoring.

Table 1 reports summary statistics for all variables used in the analysis.

3 Productivity in Adjustments of Labor Inputs During the Crisis

This section documents the key empirical patterns that motivate the subsequent analysis. Section 3.1 documents productivity changes during and surrounding the crisis. Section 3.2 provides evidence that the downward adjustment in the labor input that occurred during the crisis targeted working hours rather than employees.

[18]The top 10 categories of $JobContent_j$ cover 96.1% of the number of jobs and 98.3% of revenue in the sample. Ordered in terms of revenue, they are: Construction documentation (32.5%), Design/Construction supervision (27.0%), Construction supervision (16.6%), Design development (12.9%), Other (4.1%), Schematic design (2.1%), Other planning (1.0%), Planning & development management (0.9%), Construction supervision consulting (0.7%), Design/Construction supervision consulting (0.5%).

[19]The top 10 client industries cover 65.4% of the number of jobs and 70.1% of the revenue in the sample. They are: Real-estate (21.0%), Education (10.1%), Financial/insurance (9.4%), Transportation (6.2%), Other public interest organizations (5.2%), Municipal government (4.0%), Electronics (3.9%), Others (3.5%), Medical related organizations (3.4%), and Service industry (3.3%).

3.1 Productivity changes during the crisis

Figure 2 plots the average adjusted revenue per hour – the main productivity measure that serves as the dependent variable – for jobs that were started in the given year and completed by the end of the sample period. The trend decreases until 2008 and then increases until 2012. The pattern observed in Figure 2 is presumably generated by changes in demand. Figure 3 shows the total orders received annually, in trillions of Japanese yen, by 50 large construction companies. It reveals that total demand decreased starting in 2008 and recovered to its pre-crisis level after 2013. The observed time lag is understandable given that planning and design precedes construction.

Let $t-1$ denote the pre-crisis period (2005, 2006, and 2007), and let t denote the post-crisis period (2010, 2011, and 2012). Column 1 of Table 2 reports the difference of the weighted average productivity across jobs, $\Delta A_t \equiv \ln\left(\frac{A_t}{A_{t-1}}\right)$, where $A_t = \sum_j \frac{h_j}{\sum_j h_j}\frac{Adj\_Rev_j}{h_j}$. The estimate is 5.1%. That productivity change, like those plotted in Figure 2, is unconditional. As explained in Section 2.2, the productivity change is estimated conditional on the job characteristics defined in Section 2.3, with lnh_j as the dependent variable and $\ln Adj\_Rev_j$, as a control variable:

$$\ln h_j = \beta_0 + \beta_1 1\left\{AfterCrisis_j\right\} + \beta_2 \ln Adj\_Rev_j + \sum_{k=3}^{41} \beta_k Ind_{jk} + \sum_{k=42}^{63} \beta_k JobContent_{jk} + \varepsilon_j. \tag{1}$$

Note that $-\beta$ is interpreted as the productivity change, $\Delta\hat{A}_t$, resulting from the crisis. The estimate of that productivity change is 7.6%, as reported in column 2 of Table 2.[20]

A potential concern is that revenue is not a perfect measure of output, as it also incorporates price changes that may obscure productivity changes.[21] The markup, or spread between the selling price and the production cost, likely fell in response to the drop in demand that coincided with the crisis. That fall may at least partly explain why revenue per hour decreased in 2008, when the shock of the crisis had the largest impact. Following the same logic, the increase in revenue per hour that followed the crisis may at least partly reflect a recovering markup instead of an improving production technology.

To assess the role of the markup in the business cycle, Figure 4 plots average adjusted revenue per job for jobs that were started in the year indicated on the horizontal axis and

[20] If the pre-crisis and post-crisis definitions are both shortened by a year (i.e., 2006-2007 and 2010-2011), ΔA_t = 0.064, and $\Delta\hat{A}_t$ = 0.089 with standard error 0.027. If they are both lengthened by a year (i.e., 2004-2007 and 2010-2013), ΔA_t = 0.036, and $\Delta\hat{A}_t$ = 0.034 with standard error 0.020. The tradeoff is that smaller bandwidths reduce the sample size, whereas longer ones increase the risk that events unrelated to the crisis may cloud the picture.

[21] The problem is well known and widespread in productivity analysis, as discussed in Syverson (2011).

completed by the end of the sample. If the change in the markup is the major force that drives productivity, then the average adjusted revenue per job should decrease during the crisis and increase after the crisis. Figure 4 reveals the opposite pattern, suggesting that the productivity increase is not driven by a changing markup.[22]

3.2 Adjustments of labor inputs (hours versus workers) during the crisis

As explained in footnote 1, Japan's Labor Contract Law prohibits "abuse of the right to dismiss". An example of such an abuse would be laying off workers while the firm's other workers have substantial overtime. Thus, it is reasonable to expect the design company to respond to the crisis by slashing working hours before reducing employment.[23] Figure 5a plots the average monthly overtime hours, excluding zero overtime observations.[24] Figure 5b plots the share of workers, within each month, who do not have positive overtime. Figure 5c plots the share of overtime out of total working hours. The decreasing trends in Figure 5a and Figure 5c and the increasing trend in Figure 5b are consistent with a downward adjustment in working hours during the crisis. An additional striking feature of Figure 5a is that hours did not return to their pre-crisis levels after the economy rebounded. In fact, a further drop occurred in 2016, coinciding with an action by the government to restrict overtime hours.[25]

Next, a measure of the job-level overtime share is constructed using a two-step approach. First, the following regression is estimated, in which the dependent variable is the overtime share (as defined in section 2.1) for each job year, y:

$$\frac{OT_{jy}}{h_{jy}} = \beta_0 + \sum_{k=1}^{3} \gamma_k IndOrder^k_{jy-1} + \sum_{k=1}^{3} \alpha_k IndOrder^k_{jy} + \boldsymbol{X}'_j \boldsymbol{\beta} + u_{jy}, \tag{2}$$

where industry total orders (by 50 large construction companies) are included as indepen-

[22]A regression analysis further controls for $\ln Area_j$, and dummies for industry and job content show that there is no evidence of significant changes of adjusted revenue per job during the sample period. Given that the regression controls for detailed job characteristics, the result further confirms that a changing markup is unlikely to drive the productivity change.

[23]The firm would also likely terminate temporary workers and reduce outsourcing.

[24]The months that have significantly lower overtime mostly correspond to August, December, and January, when there are longer public holidays in Japan.

[25]Specifically, Prime Minister Shinzo Abe launched the Council for the Realization of Work Style Reform. At its first meeting in September 2016, Abe ordered the Council to focus on nine areas of "work style reform" in the immediate future. The third of these was, "improvement of long work hours such as by considering a regulatory limit on overtime work." For further details, see https://japan.kantei.go.jp/97_abe/actions/201609/27article2.html

dent variables in years y and $y-1$. Quadratics and cubics of those independent variables are also included because the increase in the amount of the order is likely to take more than one year to finish. The vector of job characteristics, $\boldsymbol{X}_j$, includes $Area_j$ and dummies for industry and job content. Second, for each job j, the job-level overtime share is predicted from industry demand fluctuations, as a weighted average across years, using annual total working hours as weights:

$$OTShare\_Ind_j = \sum_y \left(\frac{h_{jy}}{h_j}\right)\left(\frac{\tilde{OT}_{jy}}{h_{jy}}\right), \tag{3}$$

where $\frac{\tilde{OT}_{jy}}{h_{jy}}$ is predicted from Equation (2), and the summation is over the active years (i.e., those in which some workers are spending positive hours) for each job. Figure 6 plots $OTShare\_Ind_j$, averaged across jobs, for each starting year. The clear discontinuity in 2009 is consistent with the evidence in Figure 5 that working hours dropped during the crisis.

The preceding evidence reveals that the crisis caused a downward adjustment in hours, but what about in the number of workers? Workers' fear of job loss in the U.S. is a key ingredient in Lazear *et al.* (2016), which predicts higher effort levels from the workers whose jobs are spared and who want to reduce their probability of being fired into a weak labor market. As noted in footnote 1, however, the institutional landscape in Japan makes firing workers more difficult. The change in employment during the crisis is plotted in Figure 7, which shows the number of workers (i.e., the count of worker identification numbers in the labor input data) participating in "external" (i.e., revenue generating) jobs. Despite the overall increasing trend, a decrease is observed starting in 2009 and extending to 2013. Comparing Figure 7 to Figure 5a reveals that the hours adjustment, which started as early as 2007-2008, preceded the employment adjustment, which did not begin until late 2009.

To further explore the employment reduction, the change in the number of workers was decomposed into the numbers of hirings and separations, by calculating the number of workers participating in "external" jobs during the current year but not during the past year, and the number of workers present in the past year but not in the current year. Figure 8 plots the rates of hiring and separation, where both numbers are normalized by the number of total active workers (i.e., those who spend some hours on external jobs) in the past year. The plot shows that the first reversal of the sign of the slope in Figure 7 is largely driven by a lower hiring rate, instead of a higher separation rate. The decomposition sheds further light on the timing of the employment adjustment. Even though, as mentioned, the hours adjustment preceded the employment adjustment, the hiring rate dropped precipitously starting in 2008 and continued dropping for the next two years.[26]

The personnel data are available starting from 2011, so the workers who leave the company can be identified given that those workers appear in the personnel data and labor input data of "external" jobs in year $y-1$ but not in the personnel data of year y. Table 3 reveals the important point that most of the separations are not those of regular workers. The annual separation rate is only about 1 to 2 percent among the regular workers who participate in external jobs.[27] Overall, the data show that the risk of being fired is low for regular workers and that it did not substantially change during the crisis.[28]

In summary, the evidence from this firm suggests that, as anticipated in Japan, the downward adjustments to the labor input that are needed in the wake of a drop in product demand are made mostly on hours rather than on employment. To the extent that reductions in force happen, they target temporary or contract workers whose attachment to the firm is weak. Such "marginal workers", who are the first to be let go in downturns and the first to be hired during recoveries, are likely also the marginal participants in joining project design teams when team size is increased. Their lack of firm-specific and team-specific human capital makes them less productive than their "regular" peers, consistent with the model of Section 4.1.

3.3 Labor Allocation Patterns Within Teams

Table 4 illustrates the within-team allocation of working hours. The rows are listed in descending order by the team members' hours contributions, with the highest-ranked worker (i.e., the one who contributes the most hours) listed first.[29] The table reveals a

[26]The drop in the separation rate that occurred in 2008 and 2009 may capture a drop in voluntary separations, reflecting workers' reluctance to jump ship to enter a weak job market, as argued in Lazear *et al.* (2016).

[27]There is a difference in the definition of separation in Figure 8 and Table 3. The separation rate reported on the vertical axis of Figure 8 is the share of workers leaving the production of the external jobs, while the separation rate reported in Table 3 is the share of workers leaving the company. Comparing the product of the separation rate reported in Figure 8 and the share of regular workers (first column of Table 3) and the separation rate reported in the second columns of Table 3 shows that the separation rate reported in Figure 8 slightly overstates the rate of leaving the company, because some regular workers shift from external to internal work instead of leaving the company.

[28]In a similar analysis that includes workers who only participate in "internal" jobs (i.e., cost center, administrative jobs), it is found that the drop in the hiring rate is even larger, and the share of regular workers among those who leave is substantially lower.

[29]For example, consider teams of size 4, as indicated in the fourth column. The table shows that 57% of a 4-person team's hours are contributed by the top-ranking worker, whereas 25% are contributed by the second-ranked worker. The third and fourth-ranked workers contribute only 12% and 5%, respectively, of total working hours.

striking concentration of within-team labor allocation. In a team of 5 people, the top worker contributes more hours than the other 4 team members combined. Although the top worker's contribution share of the team's total hours naturally decreases with team size, it remains substantial even in teams as large as 19.

The average (across workers for each month) number of jobs on which a worker spends a positive amount of time is plotted in Figure 9, along with the annual average across all months in each year.[30] A clear decrease in the average number of jobs is observed around 2007 to 2009, and a rebound to its 2006 level does not occur even when demand recovers following the crisis. This suggests the importance of labor assignment: as the number of jobs assigned declines, the team's top performers can cover more of the work, which enhances team productivity.

To summarize, within-team working hours are heavily concentrated, with a small number of team members contributing the bulk of the hours. The average number of jobs on which each worker participates fell with the crisis and did not recover afterwards.

For the purpose of investigating whether changes in labor allocation play a significant role in explaining the increase in team productivity after the crisis, the following regression is estimated:

$$Outcome_j = \beta_0 + \beta_1 1\left\{AfterCrisis_j\right\} + \beta_2 \ln Adj\_Rev_j + \sum_{k=3}^{41} \beta_k Ind_{jk} + \sum_{k=42}^{63} \beta_k JobContent_{jk} + \varepsilon_j, \tag{4}$$

where $Outcome_j$ is measured in the following 5 alternative ways: (1) $\ln TeamSize_j$, (2) l^c_{2j}, the cumulative labor shares of the two top-ranked workers, (3) $\ln h^{(1)}_j$, the logarithm of hours of the top-ranked worker, (4) $Std\left(h_{ij}\right)$, the standard deviation of working hours per day across team members, and (5) $\max_i h_{ij} - \min_i h_{ij}$, the range of working hours per day within the team.

The estimation results are reported in Table 5. The first two rows show that team size decreases, and the labor share becomes more concentrated after crisis. These patterns are consistent with labor reallocation underlying the increase in team productivity. The working hours of the rank-1 worker decrease, as shown in the third row, providing evidence of an increase in efficiency at the worker level.[31] The last two rows reveal that both the standard deviation and the range of within-team hours decrease after the crisis,

[30]The definition of number of jobs includes administrative jobs that do not generate revenue, and those jobs where a worker serves as chief manager but spends zero hours. Those jobs can be identified given that the project management data contain the chief manager's identity. For the jobs in which a worker has some responsibility but does not serve as chief manager, no information is observed if the worker spends zero hours. Therefore, these jobs are excluded in the definition of number of jobs.

[31]An unreported regression also reveals decreases in working hours for workers ranked 2 through 5.

providing further evidence of changes in labor allocation.

4 A Theoretical Model of Labor Assignment Within Teams

Motivated by the empirical evidence from Section 3.3, an organizing framework is now presented, the main purpose of which is to deepen understanding of the within-team assignments of working hours across a team's workers and tasks.

4.1 Model

Consider a single firm (also called the employer) that operates in a production setting consisting of a set of jobs, with j indexing jobs. Each job requires completion of S tasks, indexed by s. The set of tasks, denoted by Ω^S,is the same across jobs. For simplicity, it is assumed that each task can be assigned to at most one worker. The production process is divided into two stages. In stage 1, the output level, Y_j, is determined, and given that value, N_j^{Ini} workers are assigned to job j, where[32]

$$N^{Ini}\left(Y_j\right) = \lfloor \alpha_1 Y_j^{\alpha_0} \rfloor, \tag{5}$$

where $\lfloor x \rfloor$ denotes the largest integer smaller than or equal to x. The parameters α_0 and α_1 are both strictly positive. While α_0 determines the relative number of workers assigned to big and small jobs, α_1 determines the average number of workers assigned to each job. Each worker i takes a draw (ϕ_i, H_i) from the joint distribution of these two variables, where ϕ_i and H_i denote worker i's productivity and time endowment, respectively. The employer observes both parameters.

In stage 2, the employer allocates labor within each job by deciding how many worker hours to assign to each task in that job. The employer chooses that labor allocation by minimizing costs.[33] Let h_{ijs} denote the hours that worker i is assigned on task s of job j. Each unit of the product $\phi_i h_{ijs}$ is referred to as an "effective labor hour", and w denotes the employer's cost per effective labor hour. Each effective labor hour represents the task-specific contribution to job j's total output, Y_j. That contribution is denoted q_{js}, where $q_{js} = \phi_i h_{ijs}$, recalling that an i subscript is omitted on q_{js} as a consequence of the assumption that each task is assigned to at most one worker. These task-specific contributions are aggregated over all S tasks in job j to produce total output for job j, as expressed

[32]The superscript Ini denotes "initial".

[33]Equivalently, the employer can maximize production given the budget.

in the following production function: $Y_j = \prod_{s \in \Omega^S} \left(q_{js}\right)^{\gamma_{js}}$. The positive parameter γ_{js} can be interpreted as the weight of task s on job j. To simplify the analysis, it is assumed that tasks are symmetric,[34] i.e., $\gamma_{js} = \frac{1}{\eta S}$, where $\eta > 0$ is the parameter that controls the returns to scale in the production function. When $\eta < 1$, returns to scale are increasing: increasing the quantity of all tasks by the same proportion increases aggregate output by a larger proportion.

The employer's problem in stage 2, which is to minimize the weighted average of effective labor costs and labor hours, given the output requirement for each job and the task production function, is as follows:

$$\min_{h_{ijs}} (1-\epsilon)\left(\textstyle\sum_s w\phi_i h_{ijs}\right) + \epsilon\left(\textstyle\sum_s h_{ijs}\right) \quad , \tag{6}$$

$$s.t. \quad Y_j = \prod_{s \in \Omega^S} \left(q_{js}\right)^{\gamma_{js}}, \tag{7}$$

$$q_{js} = \phi_i h_{ijs}. \tag{8}$$

Even though w is constant, wages generally differ across workers given that workers are heterogeneous in productivity, ϕ_i. The presence of labor hours in the objective function reflects the employer's preference to deliver the output faster, given total labor costs.

To minimize total costs, it is optimal to exhaust the most productive worker's time first. If that worker were endowed with sufficient time, it would be optimal to have that one worker complete all S tasks.[35] In general, that worker's time constraint binds before all S tasks can be covered. At that point, it is optimal to assign the second most productive worker until that person's time endowment is exhausted, and so on, until all S tasks on job j are covered by a worker. Given the assignment of workers to tasks, the problem's solution yields the following expression for the optimal labor input:

$$h_{ijs} = \frac{Y_j^{\eta}}{\left[\prod_{s \in \Omega_j^S} \left(\frac{\phi_{i(s)}}{(1-\epsilon) w\phi_{i(s)} + \epsilon}\right)^{\frac{1}{S}}\right]\left[(1-\epsilon) w\phi_i + \epsilon\right]}. \tag{9}$$

To simplify the analysis, $\epsilon = 0$ is assumed, implying that the employer cares mostly about

[34]This assumption and the one that each task is assigned to at most one worker are relatively innocuous, because each job can be divided into many small tasks. The assignment algorithm easily generalizes to asymmetric tasks by sorting the tasks in descending order and considering the biggest task first.

[35]The implicit assumption in that case is that the $N_j^{Ini} - 1$ remaining employees are assigned to work on profitable activities other than job j.

labor costs. This yields a simplified expression for h_{ijs}:

$$h_{ijs} = \frac{Y_j^{\eta}}{\phi_i}. \tag{10}$$

Denote the number of tasks assigned to worker i as M_{ij}. Given that tasks are symmetric, total working hours at worker-job level, h_{ij}, is the product between $M_{ij}$and h_{ijs}: $h_{ij} = M_{ij}h_{ijs}$. For productive workers that exhaust their time endowment,

$$M_{ij} = \lfloor \frac{H_i}{h_{ijs}} \rfloor = \lfloor \frac{\phi_i H_i}{Y_j^{\eta}} \rfloor. \tag{11}$$

The model predicts that job-level productivity, A_j, is a ratio between the size effect, determined by the force of returns to scale, and the weighted harmonic mean across worker-level productivities:

$$A_j \equiv \frac{Y_j}{\sum_{s \in \Omega^s} h_{ijs}} = \frac{Y_j^{1-\eta}}{\sum_i \frac{M_{ij}}{\phi_i}}. \tag{12}$$

It is easily verified that this expression implies a complementarity between workers' productivities:

$$\frac{\partial^2 \ln A_j}{\partial \phi_{i_2} \partial \phi_{i_1}} = \frac{M_{i_1 j} M_{i_2 j}}{\phi_{i_1}^2 \phi_{i_2}^2} \left(\sum_i \frac{M_{ij}}{\phi_i} \right)^{-2} > 0, \tag{13}$$

for any worker i_1 and i_2. It is evident from Equation (12) that conditioning on M_{ij}, with a uniform increase in worker productivity, the job-level productivity increases by the same proportion. However, because the labor assignment responds to an increase in ϕ_i by reallocating tasks from less to more productive workers, the increase in A_j exceeds the proportional increase in worker-level productivity. Moreover, the effect of labor allocation is increasing in the amount of tasks shifted, as well as the productivity difference between the workers.

The labor share of worker i in job j satisfies

$$l_{ij} = \frac{h_{ij}}{h_j} = \frac{M_{ij}/\phi_i}{\sum_{i'} \left(M_{i'j}/\phi_{i'} \right)}, \tag{14}$$

where h_j is the total hours of job j . Assuming that worker i exhausts her time endowment with a sufficiently high ϕ_i (i.e., h_{ij}=H_i), when ϕ_i rises, M_{ij} also increases in the same proportion from Equation (11), which implies that $M_{i'j}$ for a marginal worker i' decreases. As a result, l_{ij} will increase and the job-level productivity A_j will improve. This creates a positive correlation between l_{ij} and A_j. On the other hand, with an increase in H_i, although similarly both l_{ij} and A_j will increase, h_{ij} will also increase. This shows that the effects of increasing ϕ_i and H_i are different.

The model simulation aims to replicate this assignment procedure that determines team composition. In the algorithm, not all workers are assigned a positive number of tasks, which reflects the reality that not all candidates who are considered for inclusion on a team will join it. If $\rho_{\phi H} < 0$, then more productive workers tend to have lower time endowments. In this case, the correlation parameter $\rho_{\phi H}$ parsimoniously captures the trade-off between increasing the productivity of the current job and the opportunity cost of not assigning a productive worker to another job.[36]

The model's following two predictions are tested in the subsequent empirical work:

1. Conditional on the output requirement, Y_j, the larger the team size, the lower the productivity. Intuitively, a large team size indicates the introduction of less productive workers.

2. The effect of a uniform increase in worker-level productivity has a larger impact in larger teams. This is because larger teams imply greater variation in worker ability, which tends to amplify the labor reallocation effect.

4.2 Calibration of the model

The model suggests that within-team labor reallocation potentially plays an important role in driving the productivity changes that accompany recessions. A calibration exercise is helpful for quantifying this effect. Simulations are conducted for a pre-crisis sample of jobs starting from 2005 to 2007 and, separately, for a post-crisis sample of jobs starting from 2010 to 2012.

4.2.1 Targeted moments

The following figures describe the key moments that are used to calibrate the model. The left panel of Figure 10 plots the empirical distribution of $\ln \frac{Adj\_Rev_j}{h_j}$, which is the empirical measure of job-level productivity. The standard deviation of $\ln \frac{Adj\_Rev_j}{h_j}$ is 0.76, which exhibits substantial productivity variation: jobs that are 1 standard deviation higher than the mean have 114% higher job-level productivity. The right panel of Figure 10 plots the

[36] This feature allows the model to partially offset the unrealistic feature that team members are randomly assigned to each job.

cumulative distribution function of $\ln \frac{Adj\_Rev_j}{h_j}$. Plotting the cumulative distribution function of a normal random variable with the same mean and standard deviation reveals that the empirical distribution of $\frac{Adj\_Rev_j}{h_j}$ is well described by a log-normal distribution. Figure 11 shows the empirical distribution of $\ln l_j^{(1)}$ in a similar format. Except for the spike at 1, the empirical distribution of $\ln l_j^{(1)}$ fits well with a normal distribution. Finally, Figure 12 plots $\ln Adj\_Rev_j$ on the horizontal axis and $\ln h_j$ on the vertical axis. The points closely align on the 45-degree line, providing indirect support for the decision to adjust the revenue measure by the outsource ratio. The fitted curve, after controlling for industry and job contents, has a slope of 0.996, which is not statistically different from 1. As shown later, this slope provides an important benchmark that regularizes the value of η. Intuitively, if $\eta = 1$, because jobs with larger Y_j require larger teams, thereby resulting in a lower productivity, a slope smaller than 1 is expected. The unit slope in the data, therefore, suggests that $\eta < 1$, and the magnitude of increasing returns to scale exactly cancels the negative effect of a larger team.

4.2.2 Calibration procedure

Worker i's productivity parameter, ϕ_i, and time endowment parameter, H_i, are assumed to be jointly log-normally distributed, i.e.,

$$\ln \phi_i \sim \mathcal{N}\left(\mu_\phi, \sigma_\phi^2\right),$$

$$\ln H_i \sim \mathcal{N}\left(\mu_H, \sigma_H^2\right),$$

where $\rho_{\phi H}$ denotes the correlation between the two variables. The following parameters must be calibrated:

$$\left(\mu_\phi, \sigma_\phi^2, \mu_H, \sigma_H^2, \rho_{\phi H}, \alpha_0, \alpha_1, \eta\right).$$

A sample of simulated jobs is constructed to calibrate the parameter values and assess their implications.[37]

The parameters $\left(\mu_H, \sigma_H^2, \alpha_0, \alpha_1\right)$ can be calibrated separately using data. First, α_0 is

[37] The simulation assumes $S = 100$. Since each task is assigned to at most one worker, the value of S limits the maximum team size. This choice ensures that the maximum team size is consistent with the observed maximum, which is 98. Other than that, the choice of S does not generate real changes in the model. To see that, if S is increased by a factor of 2, and workers' productivities are increased by the same factor, then everything remains the same except that the number of tasks assigned to each worker is doubled.

pinned down by estimating the following regression using the empirical data:

$$\ln TeamSize_j = \beta_0 + \beta_1 \ln Adj\_Rev_j + \sum_{k=2}^{40} \beta_k Ind_{jk} + \sum_{k=41}^{62} \beta_k JobContent_{jk} + \varepsilon_j. \tag{15}$$

Given that $\ln Adj\_Rev_j$ measures output, β_1 measures the elasticity of team size with respect to output size. Estimating this regression using both the pre-and-post-crisis samples yields an estimate of β_1 of 0.399, with standard error of 0.005. Thus, the simulation is based on $\alpha_0 = 0.4$. The value of α_1 is chosen such that the job with highest output size in the sample is given 100 draws. This is motivated by the fact that in the data, the maximum team size is 98. Given that the resource constraints should be near binding for large jobs, the maximum team size is taken as a proxy of the potential candidates being considered for large jobs. As no significant difference in β_1 is found by estimating Equation (15) separately for samples before and after the crisis, the values of (α_0, α_1) are held constant before and after the crisis.

The parameter values of (μ_H, σ_H^2) are pinned down using the empirical average and standard deviation of $h_j^{(1)}$, separately for the pre-crisis and the post-crisis samples. Because the rank-1 workers tend to exhaust their time endowments, the observed working hours are closer to the total time endowment H_i. Ideally, (μ_H, σ_H^2) could be estimated by matching the empirical moments related to working hours. But due to the discrete nature of tasks, a slight change in time endowments does not result in any change in labor assignment, which increases the difficulty of identifying (μ_H, σ_H^2). It is shown later that, in practice, the calibrated parameter values explain the observed working hours rather well.

The time interval is standardized to one day to make the hours comparable across jobs. As a result, in the simulation working hours and revenue for each job are both divided by job duration (measured in days).[38] Let Y_j denote revenue per day of job j. Consistent with the model, the labor assignment of each simulated job j takes Y_j as given. Outputs of simulated jobs are drawn from the empirical distribution of Y_j using the following steps. Firstly, 500 equally-distant points are drawn from the interval $[0.1, 0.99]$. Then for each grid point, the corresponding quantile is drawn from the empirical distribution of Y_j.[39]

[38] The job-level productivity measure is invariant to this standardization because both its numerator and its denominator are divided by the same number. Given the standardization, 4 jobs that do not have a well-defined time length variable are dropped. Because the model is static, the duration of each job is not determined. In the simulation, impose that each job finishes in one day.

[39] Each sample uses only the jobs for which Y_j lies between 1st and 99th percentiles To avoid the noise in the extreme observations. Compared to taking random draws from empirical distribution of Y_j, the uniform approximation of the empirical distribution provides a better coverage and reduces variance in the simulation. See the discussion in Chapter 9 of Train (2009).

To reduce the simulation error, 20 draws are taken for each value of Y_j[40]. Time constraints create the possibility that assigned workers cannot complete the job, and the simulation drops failed jobs from the sample. As a result, the simulated distribution of Y_j can differ from the empirical distribution of Y_j. The empirical distribution of Y_j is included in the target moments to minimize this effect.

Given Y_j, as in the first stage of the model, N_j^{Ini} draws are taken according to equation (5) from the joint distribution of $(\ln \phi_i, \ln H_i)$, under the calibrated values of $\left(\mu_H, \sigma_H^2, \alpha_0, \alpha_1\right)$, and any values of other parameters.[41] Then workers are assigned to tasks as described in the previous subsection. Working hours, job-level productivity, and labor shares are calculated according to equations (10), (12), and (14). Because each draw of Y_j corresponds to an observation of simulated job, a sample of simulated jobs is obtained after applying the above of procedure for all draws of Y_j. Values for $\left(\mu_\phi, \sigma_\phi^2, \rho_{\phi H}\right)$ are chosen by matching the moments calculated using the simulated sample and the corresponding empirical moments. The targeted moments include:

- The distribution of job-level productivity. Since the model predicts that, conditional on Y_j, job-level productivity is determined by the average of worker-level productivity, these moments helps to identify μ_ϕ and σ_ϕ^2.

- The distribution of output size, Y_j. As shown in equations (10) and (12), both the time required to complete one task and the job-level productivity is increasing in Y_j. Therefore, the distribution of Y_j is also informative about the underlying worker-level productivity. Given that Y_j is taken from the empirical distribution, matching the simulated and the empirical distribution of Y_j essentially punishes the optimization algorithm from failing the jobs.

- The distributions of cumulative labor shares up to the worker who ranks 5 in terms of working hours within the team.[42] Labor shares are informative about the correlation parameter $\rho_{\phi H}$. To see this, observe that given the other parameters, if $\rho_{\phi H} > 0$,

[40] To save computation time, only 1 repetition is used at first. This is increased to 20 repetitions after the error stabilizes to a small value.

[41] Drawing from the joint distribution of $\ln \phi_i$ and $\ln H_i$ involves first taking draws a_i and b_i from the standard normal distribution and then using the following matrix multiplication:

$$\begin{pmatrix} \ln \phi_i - \mu_\phi \\ \ln H_i - \mu_H \end{pmatrix} = \begin{pmatrix} \sigma_\phi \sqrt{1 - \rho_{\phi H}^2} & \sigma_\phi \rho_{\phi H} \\ 0 & \sigma_H \end{pmatrix} \begin{pmatrix} a_i \\ b_i \end{pmatrix}.$$

It is easy to verify that the resulting random variables have the desired joint distribution. The draws of a_i and b_i are fixed throughout the simulation process.

[42] The top 5 workers cover most of the labor inputs. In the sample, the top five workers, on average, account for about 88.3% of labor inputs.

more productive workers tend to have greater time endowments. Therefore, labor will be more concentrated compare to the case when $\rho_{\phi H} < 0$. The labor share is targeted instead of team size because a complete profile of the labor share is sufficient to calculate team size, and it contains more information. For example, the labor shares in a two-person team could be split in an infinite number of ways.

Letting Q denote the number of percentiles, the optimization problem to solve is

$$\min_{\mu_\phi,\sigma_\phi,\rho_{\phi H}} \sqrt{\frac{1}{3Q}\sum_q \left[\left(\frac{A_q-\hat{A}_q}{\hat{A}_q}\right)^2+\left(\frac{Y_q-\hat{Y}_q}{\hat{Y}_q}\right)^2+\frac{1}{5}\sum_{k=1}^{5}\left(\frac{l^c_{kq}-\hat{l}^c_{kq}}{\hat{l}^c_{kq}}\right)^2\right]}$$

where A_q is the qth percentile of simulated job-level productivity, $\hat{A}_q$ is the corresponding percentile in the data, Y_q and $\hat{Y}_q$ denote the qth percentiles of the simulated and empirical distributions of output, and l^c_{kq} is the percentile of the rank-k cumulative labor share, i.e., $l^c_{kq} = \sum_{i=1}^{k} l_{iq}$, where l_{iq} is the qth percentile of the rank-i labor share. Of the Q percentiles, those that are matched are the 5th, 6th, 7th, ..., and 95th.[43]

The above procedures are repeated for both the pre-crisis and the post-crisis samples, to calibrate values of $\left(\mu_\phi, \sigma^2_\phi, \mu_H, \sigma^2_H, \rho_{\phi H}, \alpha_0, \alpha_1\right)$ given any value of η. Finally, using the combined simulated data before and after the crisis with the same value of η, $\ln h_j$ is regressed on $\ln Y_j$ to obtain the simulated slope $\frac{\partial \ln h_j}{\partial \ln Y_j}$. η is chosen such that the resulting slope is 1, which is the value suggested in Figure 12. Given that $A_j = \frac{Y_j}{h_j}$, Equation (12) implies that the slope $\frac{\partial \ln h_j}{\partial \ln Y_j}$ is an increasing function of η.[44] Moreover, when $\eta = 1$, $\frac{\partial \ln h_j}{\partial \ln Y_j} > 1$ because the average worker productivity is decreasing in Y_j. A search can therefore be conducted for values of η below 1. The search is conducted over the grids with $\eta = 1,\ 0.99,\ 0.98, ...$, choosing the value that yields a slope of $\frac{\partial \ln h_j}{\partial \ln Y_j}$ that is closest to 1.

4.2.3 Calibration results

Figure 13 illustrates the productivity increase arising from the crisis, by plotting the implied density function of $\ln \phi_i$ before and after the crisis. The increase in worker-level

[43] Given that tasks are discrete in the model, a small change of H_i will not change the minimand. Thus, the objective function has derivatives of zero with respect to μ_H or σ_H, and gradient-based methods are unsuitable for optimization. The Basin-hopping algorithm (Wales and Doye (1997)) is applied to avoid having the optimization routine trapped at a local minimum.

[44] This is true as long as increases in η do not lead to a stronger sensitivity of average worker productivity with respect to Y_j. In the simulation, the slope $\frac{\partial \ln h_j}{\partial \ln Y_j}$ is indeed found to be increasing in η.

productivity happens at the lower end of the distribution, whereas the change is smaller for high-productivity workers. Table 6 shows the values of the calibrated parameters, which reveal that the crisis induces an increase in μ_ϕ and a decrease in σ_ϕ. The calibrated $\rho_{\phi H}$ is negative, consistent with the intuition that more productive workers tend to be time constrained.[45] The calibrated η is 0.84. Given that η is chosen to exactly balance the effect of decreasing average worker productivity, the elasticity of average worker productivity with respect to Y_j is -0.16. The final column shows the minimized value of the objective function, which indicates that, on average, the simulated moments deviate from their empirical counterparts by 5.8% for the pre-crisis sample and by 5.0% for the post-crisis sample.

Figure 14 shows the simulated and empirical cumulative distribution function of each targeted distribution. In all figures, the solid lines plot the cumulative distribution functions of the simulated data, and the dashed lines show the cumulative distribution functions of the empirical data. The first two figures in the first row show that the simulated data fit the distribution of job-level productivity and output size well. The fit of the distributions of labor shares are shown starting from the third figure of the first row to the third figure in the second row. The fits are good except for spikes at zero, which represent the deviation at the smallest quantiles that are not targeted. Despite the fact that the distributions of working hours are not explicitly targeted, the figures starting from the fourth figure in the second row until the last figure in the third row show that the model explains the distributions of working hours reasonably well.

Using the calibrated parameters, the environment before and after crisis can be simulated. By calculating $\ln\left(\frac{\frac{1}{N_t}\sum_{i\in\Omega_t^I}\phi_i}{\frac{1}{N_{t-1}}\sum_{i\in\Omega_{t-1}^I}\phi_i}\right)$, or the log difference of the average ϕ_i, across workers that are assigned at least one task, the average worker-level productivity is estimated to increase by 3.1%. The following regression is estimated to examine the change of variables at the job level:

$$Outcome_j = \beta_0 + \beta_1 1\left\{AfterCrisis_j\right\} + \beta_2 \ln Y_j + u_j, \tag{16}$$

where $Outcome_j$ is measured using the change in productivity, $\ln h_j$, and also in the following 5 alternative ways: (1) $\ln TeamSize_j$, (2) l_{2j}^c, the cumulative labor shares of the two top-ranked workers, (3) $\ln h_j^{(1)}$, the logarithm of hours of the top-ranked worker, (4) $Std\left(h_{ij}\right)$, the standard deviation of working hours per day across team members, and (5)

[45] This can also be understood by observing that from the standpoint of a social planner who is optimally allocating labor, the implicit price of productive workers will be higher because of their higher marginal output.

$\max_i h_{ij} - \min_i h_{ij}$, the range of working hours per day within the team. These 5 measures, which are the same as those used in Table 5, are for the purpose of investigating whether labor reallocation plays a role.

Table 7 reports estimates of changes of the job-level variables after the crisis. To facilitate comparisons, results from the previous regressions using the real data are also reported. The first row shows that job-level productivity increases by 7.6%, compared to 7.3% in the simulation. Even though the estimate is very precise using the simulated data, the difference between the two numbers is not statistically significant. Importantly, because the linkage between output size and job-level productivity is governed by the labor assignment process, the similarity between the job-level productivity increase in the real data and the simulated data provides support for the model.

When comparing the job-level and worker-level productivity increases, the model illustrates that the average worker-level productivity only explains about 42% of the average job-level productivity increase. Therefore, labor reallocation and the complementarity between workers significantly amplifies the worker-level productivity increase. The effect of labor reallocation is also verified in the third and fourth rows of Table 7, where it is shown that $\ln TeamSize_j$ decreases after crisis. These results are consistent with the firm relying on smaller teams of better workers in the wake of a crisis-induced reduction in demand. The fifth and sixth rows of Table 7 show that the labor share is more concentrated both in the real and the simulated data, further supporting the presence of labor reallocation. The calibrated model also successfully generates the decrease in total working hours of the rank-1 worker, $h_j^{(1)}$, as shown in the seventh and eighth rows, and the decrease in the within-team dispersion of hours (as measured by the standard deviation and the range), and shown in rows 9 to 12. Overall, the calibrated model successfully reproduces the qualitative and quantitative patterns in the data, thereby providing strong support for the importance of within-team labor reallocation.

Two tests of the theoretical model are shown next. The first, which explores whether the effect of the financial crisis is higher for larger teams, is achieved via the following regression:

$$\ln h_j = \beta_0 + \beta_1 1\{AfterCrisis\} \times \ln TeamSize_j \qquad (17)$$
$$+\beta_2 1\{AfterCrisis\} + \beta_3 \ln TeamSize_j + \beta_4 \ln Y_j + u_j.$$

If the hypothesis is correct, the coefficient of the interaction term between $AfterCrisis$ and $\ln TeamSize_j$ is expected to be negative. As shown in Table 8, support was found for this prediction in both the real and simulated data.

As shown in Figure 13, the main productivity increase happens for less productive workers. If the labor assignment mechanism is relevant in the data, a bigger productivity improvement should be expected for less efficient teams. The second test examines this hypothesis by estimating the quantile regressions using the specification

$$\ln h_j = \beta_0 + \beta_1 1\left\{AfterCrisis\right\} + \beta_2 \ln Y_j + u_j \tag{18}$$

for different quantiles to test whether the effect of the crisis is larger for higher quantiles.

Table 9 reports the results of quantile regressions, using both real and simulated data. In both cases, there is a clear sorting pattern that the effect of the crisis is higher for larger quantiles, which is consistent with the hypothesis.

Finally, the following regression is estimated to test the correlation between the labor share, team size, and productivity:

$$\ln h_j = \beta_0 + \beta_1 l_j^{(1)} + \beta_2 TeamSize_j + \beta_3 \ln Y_j + u_j, \tag{19}$$

where $TeamSize_j$ is the count of workers who participate in job j, and $l_j^{(1)}$ is the share of hours contributed by the worker ranking 1 in job j. As discussed in the theoretical section, conditional on team size, labor shares and job-level productivity should be positively correlated if the variation is driven mainly by the worker productivity. Although worker ability is not observed in the data, it is verified in the simulation that the rank of labor share is a valid proxy for worker ability. The rank correlation between l_{ij} and ϕ_i is 0.3 in the simulation. On the other hand, conditional on output, larger team size is an indicator of lower efficiency due to the introduction of less productive workers. Table 10 displays the results. The negative sign of team size is consistent with prediction 1 of the model, which states that larger team sizes are associated with lower productivity. The negative sign of $l_j^{(1)}$ is consistent with the fact that, conditional on team size, productivity is higher when within-team working hours are more concentrated.

5 Empirical Analysis of Overtime

The increase in worker-level productivity that followed the crisis is potentially explained by decreases in overtime. Section 5.1 explores heterogeneity in the productivity effect of overtime within teams, and section 5.2 presents evidence concerning product quality, as measured by the defect rate.

5.1 Heterogeneous productivity effects of overtime within teams

In a team setting, the productivity effect of an additional working hour hinges on which team member is assigned that hour. Heterogeneous productivity effects are explored empirically in three steps. First, overtime hours are computed for every worker. Second, for each job j on which at least 5 workers are engaged, the "top 5" workers are ranked in terms of their total job-level working hours. That is, "worker 1" has the highest working hours on job j and is said to have the highest rank, "worker 2" has the second highest hours on job j and the second highest rank, and so on, up through the fifth-ranked worker. Third, a regression is estimated with productivity as the dependent variable and the 5 overtime shares as independent variables, along with controls.

To start, let $OT_j^{(r)}$ and $h_j^{(r)}$ denote the amount of overtime and working hours for job j's worker ranking r in terms of total working hours. Table 11 displays summary statistics for the overtime share for the entire team, and for each of the "top 5" workers.[46] The team-level total share of overtime appears in column 1, and the overtime shares for the five highest ranking workers, i.e., $\frac{OT_j^{(r)}}{h_j^{(r)}}$ appear in columns 2 through 6. Column 1 reveals that overtime hours account for 17% of the total hours worked on job j. As revealed by columns 2 through 6, among the five highest-ranked team members on job j, the share of overtime is similar to the team average but strictly decreasing in rank. The overtime hours of the highest-ranked worker on job j account for 18.8% of the total hours he spends on job j, whereas the overtime hours of the fifth-ranked worker account for only 15.7% of his total hours on job j.

The aforementioned regression is specified as follows:

$$\ln \frac{Adj\_Rev_j}{h_j} = \beta_0 + \sum_{r=1}^{5} \gamma_1^{(r)} \frac{OT_j^{(r)}}{h_j^{(r)}} + \sum_{r=1}^{5} \gamma_2^{(r)} l_j^{(r)} + \boldsymbol{X}_j' \boldsymbol{\beta} + \varepsilon_j, \tag{20}$$

where $l_j^{(r)}$ is the labor share of the rank r worker, $\boldsymbol{X}_j$ includes $\ln Area_j$ and dummies for industry and job content.[47] Labor shares are included to control for the effect of labor allocation. Column 1 of Table 12 displays ordinary least squares (OLS) estimation results of Equation (20), which reveal that increasing the share of each worker's hours that come from overtime work is harmful for revenue per hour. Interestingly, the magnitude of the

[46] Given that step 2 requires ranking workers (by their hours inputs) up through the fifth worker, the sample only includes jobs that have at least five workers.

[47] Using the overtime share alleviates the concern that overtime hours may be allocated to each job in a biased manner. To see this, for worker i in job j in month m, under the allocation rule of overtime, $OT_{ijm} = \frac{h_{ijm}}{\sum_j h_{ijm}} OT_{im}$. Thus, $\frac{OT_{ijm}}{h_{ijm}} = \frac{OT_{im}}{\sum_j h_{ijm}}$, which equals the monthly overtime share.

negative productivity effect is larger for the higher-ranking workers who are working more overtime. Because those workers tend to have a higher share of overtime, the results are consistent with an inverted-U-shaped hours-productivity profile, which predicts that the marginal damaging effect of overtime on productivity is increasing in the level of overtime. Moreover, the labor share coefficients are all positive, which is consistent with the model's prediction that a higher concentration of labor is associated with higher average worker productivity.

Since each worker's hours are determined by the employer, a concern is that overtime hours are assigned to those jobs where emergencies happen, thus creating a correlation between overtime and job-level productivity. If the correlation between overtime and productivity is driven only by emergencies, such as last-minute design changes or errors, then it is likely that the overtime shares of lower-ranked workers who are more (or equally) likely to be affected by such urgent needs are more (or equally) strongly associated with low team productivity. This is the opposite of the sorting pattern observed in the data. To provide further evidence of a causal impact of overtime, the following 2SLS regression is estimated:

$$\ln \frac{Adj\_Rev_j}{h_j} = \beta_0 + \beta_1 \frac{OT_j}{h_j} + \boldsymbol{X}_j' \boldsymbol{\beta} + \varepsilon_j, \tag{21}$$

where $\frac{OT_j}{h_j}$ is the job-level overtime share, and $\boldsymbol{X}_j$ includes $\ln Area_j$ and dummies for industry and job content. In the regression's first stage, the job-level overtime share $\frac{OT_j}{h_j}$ is instrumented by the average overtime share predicted from industrial demand, namely the average of the variable $\frac{\tilde{OT}_{jy}}{h_{jy}}$ used in Equation (3). Although using industry demand as the instrument raises the concern that the exclusion restriction might be violated because of its likely correlation with the markup, unreported regression results (available upon request) confirm that job-level revenue is not correlated with industry demand in the current or preceding year.

Column 2 of Table 12 displays 2SLS regression results for Equation (21). The overtime predicted from industrial demand has a significant negative association with job-level productivity, and the magnitude is economically significant: a 1 percentage point increase in the overtime share decreases team productivity by about 1.4. Column 3 of Table 12 reports the results from the same equation using OLS. The IV estimate is slightly smaller than the OLS estimate, but the difference is small, mitigating concerns about endogeneity.

5.2 Overtime and product quality

Contracts are negotiated separately for each job in a project, so contract terms (and, therefore, total revenue) are determined before the job begins. With a job's total revenue predetermined, the only way in which workers on a job can enhance productivity is to complete their work more efficiently, i.e., in fewer hours. Quality problems – which may be revealed to the client ex post – are not reflected in the predetermined revenue or productivity measures. Product quality is a potentially important consideration, however, in this setting in which there may be repeated interactions between clients and the firm across projects and in which unhappy clients could damage the firm's reputation through bad word-of-mouth.

To address the issue of product quality, penalty records are exploited. These are available at the job-level starting from 2011. Records indicate when a defect in the final product was detected after delivery and what penalty was imposed on the chief manager.[48] Defects are quite rare, with only 85 occurring out of 10,764 valid observations. Define a binary variable, $Defect_j$, equaling 1 whenever a defect occurs and 0 otherwise. The following probit model describes the relationship between overtime and the defect rate:

$$Defect_j^* = \beta_0 + \sum_{r=1}^{5} \beta_1^{(r)} \ln \left(OT_j^{(r)} + 1 \right) + \beta_2 \ln Area_j + \beta_3 \ln Rev_j + \varepsilon_j, \tag{22}$$

where $Defect_j^*$ is a latent variable such that $Prob\left(Defect_j = 1\right) = Prob\left(Defect_j^* \geq 0\right)$.[49]

Table 13 displays the estimation results for the preceding probit. Conditional on size (as measured by square footage) and revenue, higher amounts of overtime for the top-ranking worker (and only for that worker) are associated with a higher defect rate at conventional levels of statistical significance. This result is consistent with the monitoring role of the top-ranked worker who contributes the most hours: when long hours cause that worker to become too tired to detect the mistakes of the other team members, flaws and the resulting penalties are increasingly likely.

[48] There are three penalty levels: (1) *kunkoku,* the lightest one, is a warning that would not affect the prospects of promotion or pay raise; (2) *kaikoku* is a more serious admonition that implies damage to the person's promotion prospects; and (3) *genkyu* is the heaviest penalty that results in a disciplinary pay reduction. The following analysis treats all three levels equally.

[49] Dummies for industry and job content are omitted as controls, given that the estimation fails to converge in their presence. This happens because the dependent variable equals 1 in a relatively small number of cases (85).

6 Discussion

As in any study that is limited to a single firm, it is appropriate to comment on the extent to which the analysis and results might generalize. Potential threats to external validity arise for several reasons; this firm might not be representative of architectural and engineering consultancy firms (even within Japan), the industry itself may be idiosyncratic even if this firm is representative of the industry, the institutional environment is specific to Japan, the global financial crisis might be an idiosyncratic example of a major recession, etc. Such considerations should be borne in mind when interpreting results. This pattern, however, is not specific to the industry or to Japan, as the same pattern is found in Lazear *et al.* (2016), using data from a single U.S. firm offering technology-based services. Nothing stands out as being particularly unusual about the firm, the industry, or the services provided, so while future research in other production settings is clearly desirable, we anticipate that such inquiry should be corroborative.

The external validity issues that seem particularly salient to us pertain to labor market institutions. Specifically, downward adjustments to a firm's labor input are more likely to occur in hours than in employment, given the regulatory environment in Japan that protects workers from being fired. This stands in contrast to the situation in U.S. firms, including the one studied in Lazear *et al.* (2016). While the response of the firm under investigation is, therefore, likely unrepresentative of employer responses in labor markets that are less heavily regulated, this also has an upside. That is, the institutions that limit employment reductions also induce substantial variation in working hours that facilitates identification of the hours-productivity profile in teams. Moreover, even in lightly regulated labor markets in which it is easier than in Japan to fire workers, downward adjustments in hours occur and are typically among the first employer responses in a recession.[50]

A further consideration related to external validity concerns the extent to which working hours are assigned by the employer or chosen by workers. The presumption in this study (as made explicit in the model) is that team members' working hours are assigned by the employer, specifically by the chief manager, rather than chosen by the worker.[51] In alternative production settings, the reverse might be true. Ambiguity concerning which assumption is correct in general was highlighted in Pencavel (2016b). This problem of

[50]See the quote from Pencavel (2016b) in the first footnote of section 2.1.

[51]The exception is the chief manager. That person is supervised by the executive committee, but the committee only influences the decision of assigning jobs to chief managers. Chief managers are often assigned to several jobs that they manage simultaneously, and they can decide how much of their attention to devote to each.

ambiguity is avoided in the present context with single-firm personnel data, in which our interviews with the firm's manager's make clear that hours are assigned to workers by the chief manager. The extent to which the analysis generalizes to alternative production settings in which workers exercise greater autonomy over choosing their hours is unclear.

Optimal within-team labor allocation when team members have heterogeneous abilities requires that individual abilities be at least partially observed by the entity who assigns the hours. The model, for example, assumes that the employer observes ϕ_i, which is worker i's individual productivity. This assumption is not always reasonable in a team setting, and in fact that is a reason why group-based (as opposed to individual-based) incentive contracts are often used in teams.[52] In the present context, it is reasonable to assume that the chief manager possesses information about workers' abilities that is harnessed when assigning hours to workers. This is especially so given that turnover rates at the firm are low, for the institutional reasons previously described. Information about workers' abilities is revealed to the employer from the long job tenures and repeated observations of individual workers on a variety of projects. The data also include subjective ratings of individual workers' overall performance, which can be interpreted as providing the chief manager with at least partial information about ϕ_i. However, it is also likely that a chief manager will not know every worker in the firm well, particularly workers from other departments and with whom the chief manager has never worked. Therefore, there might exist some information friction that prevents the optimal labor allocation in reality.

Finally, note that one implication from the previous findings in Sections 4 and 5 is that, when encountering a heavier workload than expected, a chief manager faces a tradeoff between requesting one more worker or requiring existing members to work overtime. An additional worker will be employed only when the marginal effect of additional overtime hours is sufficienctly damaging. In fact, unreported regression analysis reveals that increasing the overtime hours of the highest-ranking worker tends to be followed by an expansion in team size.

[52] It is not always the case, however, that individual output is hard to measure in team settings. Moreover, group piece rates are sometimes used in teams even when individual output is easily measured. For example, Koret, the garment manufacturing plant analyzed in Hamilton et al. (2003), switched its seamstresses from individual piece-rate pay to a group piece-rate scheme in which they were allowed to self select into teams. At Koret, individual output was easily measured and compensated via individual piece rates prior to the change in the compensation system, which was made for reasons unrelated to the observability of output.

7 Conclusion

This study is the first to analyze the productivity of working hours in teams. Within-team allocation of hours is found to be far from uniform. Most of a team's hours are concentrated among a small number of workers, with the top-ranked worker having the most hours by a considerable margin. The marginal effect on productivity of an increase in the share of overtime hours is also found to vary considerably across team members based on their hours worked. The drop in productivity associated with a greater overtime share is larger for the team's top-ranked worker who works the most hours and smallest for the lower-ranked workers who work fewer hours.

These results are consistent with a model of within-team labor allocation. The importance of the top-ranked team member in terms of that worker's concentration of hours also extends to product quality. Assigning more overtime to the top-ranked worker is associated with an increased incidence of penalties for flaws detected by clients after delivery, whereas such an effect is absent for the team's lower-ranked workers.

This study makes a novel contribution to the literature by analyzing team formation and working hours as jointly determined endogenous choices of the employer, given a workforce of employees who are heterogeneous in their abilities. To make headway on the problem, a simple Cobb-Douglas technology is assumed to describe the team's output. This means that only complementarity among team members' outputs is assumed, and workers' skills only increase their effective labor hours. This simple structure does a remarkably successful job of explaining cyclical movements in team productivity, given that simulations from the calibrated model closely replicate the patterns observed in the real data. Nonetheless, incorporating a more general specification of technology would allow additional functions of teams to be fruitfully addressed that we abstract from in this analysis. Enriching our analytical framework to incorporate such features as problem solving, coordination, and peer learning would lend further credence to the policy and managerial implications of our study.

References

ABBOTT, M. and ASHENFELTER, O. (1976). Labour supply, commodity demand and the allocation of time. *The Review of Economic Studies*, **43** (3), 389–411.

DEUTSCHER, C., GÜRTLER, M., GÜRTLER, O. and DEVARO, J. (2020). Firm choice and career success - theory and evidence. *European Economic Review*, **127**, 103470.

DEVARO, J. (2020). Work schedules.

FELDSTEIN, M. S. (1968). Estimating the supply curve of working hours. *Oxford Economic Papers*, **20** (1), 74–80.

FRANCK, E. and NÜESCH, S. (2010). The effect of talent disparity on team productivity in soccer. *Journal of Economic Psychology*, **31** (2), 218 – 229, the Economics and Psychology of Football.

GARNERO, A., KAMPELMANN, S. and RYCX, F. C. (2014). The heterogeneous effects of workforce diversity on productivity, wages, and profits. *Industrial Relations: A Journal of Economy and Society*, **53** (3), 430–477.

GÜRTLER, M. and GÜRTLER, O. (2015). The optimality of heterogeneous tournaments. *Journal of Labor Economics*, **33** (4), 1007–1042.

HAMILTON, B. H., NICKERSON, J. A. and OWAN, H. (2003). Team incentives and worker heterogeneity: An empirical analysis of the impact of teams on productivity and participation. *Journal of Political Economy*, **111** (3), 465–497.

—, — and — (2012). Diversity and productivity in production teams. In *Advances in the Economic Analysis of Participatory and Labor-Managed Firms*, vol. 13, pp. 99–138.

KANDEL, E. and LAZEAR, E. P. (1992). Peer pressure and partnerships. *Journal of Political Economy*, **100** (4), 801–817.

KURODA, S. and YAMAMOTO, I. (2013). Firms' demand for work hours: Evidence from matched firm-worker data in japan. *Journal of the Japanese and International Economies*, **29**, 57 – 73.

LANG, K. (1986). A language theory of discrimination. *The Quarterly Journal of Economics*, **101** (2), 363–382.

LAZEAR, E. P. (1999). Globalisation and the market for team-mates. *The Economic Journal*, **109** (454), C15–C40.

— and ROSEN, S. (1981). Rank-order tournaments as optimum labor contracts. *Journal of Political Economy*, **89** (5), 841–864.

—, SHAW, K. L. and STANTON, C. (2016). Making do with less: Working harder during recessions. *Journal of Labor Economics*, **34** (S1), S333–S360.

MANNIX, E. and NEALE, M. A. (2005). What differences make a difference?: The promise and reality of diverse teams in organizations. *Psychological Science in the Public Interest*, **6** (2), 31–55, pMID: 26158478.

PARROTTA, P., POZZOLI, D. and PYTLIKOVÁ, M. (2014). Labor diversity and firm productivity. *European Economic Review*, **66**, 144 – 179.

PENCAVEL, J. H. (2015). The Productivity of Working Hours. *Economic Journal*, **125** (589), 2052–2076.

— (2016a). Recovery from work and the productivity of working hours. *Economica*, **83** (332), 545–563.

— (2016b). Whose preferences are revealed in hours of work? *Economic Inquiry*, **54** (1), 9–24.

— (2018). *Diminishing Returns at Work*, vol. 1. Oxford University Press.

ROSEN, S. (1969). On the interindustry wage and hours structure. *Journal of Political Economy*, **77** (2), 249–273.

SATO, K., KURODA, S. and OWAN, H. (2020). Mental health effects of long work hours, night and weekend work, and short rest periods. *Social Science & Medicine*, **246**, 112774.

SHANGGUAN, R. and OWAN, H. (2019). How good managers steer their projects: Using value-added measures of manager quality.

SHAPIRO, C. and STIGLITZ, J. E. (1984). Equilibrium unemployment as a worker discipline device. *The American Economic Review*, **74** (3), 433–444.

SYVERSON, C. (2011). What Determines Productivity? *Journal of Economic Literature*, **49** (2), 326–365.

TRAIN, K. E. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press, 2nd edn.

WALES, D. J. and DOYE, J. P. (1997). Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *Journal of Physical Chemistry A*, **101** (28), 5111–5116.

YAMAMOTO, I. (2016). Can japan's long work hours be corrected? adopting regulations to cap total work hours. *Nihon Keizai Shimbun*.

	count	mean	std	min	25%	50%	75%	max
$\frac{Rev_j}{h_j}$	7167	72,782.690	840,858.096	499.645	9,923.434	15,602.002	29,181.134	38,200,000.000
$\frac{Adj\_Rev_j}{h_j}$	7167	20,316.516	141,253.941	134.187	7,130.503	9,915.976	15,312.884	6,450,000.000
Rev_j	7167	38,493,293.515	84,525,150.665	1,000,000.000	3,100,000.000	9,431,000.000	34,899,800.000	1,148,320,000.000
Adj_Rev_j	7167	26,008,804.329	56,333,084.093	2,841.756	1,890,000.000	5,912,052.898	23,558,266.672	778,876,558.156
h_j	7167	2932.628	6549.442	0.500	143.500	575.000	2684.250	106,801.500
OT_j	7167	543.758	1318.959	0.000	18.027	85.445	437.668	20,077.053
$\frac{OT_j}{h_j}$	7167	0.163	0.077	0.000	0.111	0.158	0.211	0.493
$Area_j$	7167	52,132.564	97,849.847	0.010	4337.405	15,000.000	53,331.619	1,000,000.000
$TeamSize_j$	7167	13.788	13.637	1.000	4.000	9.000	19.000	98.000
O_j	7167	0.298	0.285	0.000	0.050	0.222	0.463	1.000
$Defect_j$	2047	0.013	0.114	0.000	0.000	0.000	0.000	1.000

Note: Summary statistics for all variables in the analysis, as defined in Section 2.

Table 1: Summary statistics

		(1)	(2)
Pre-crisis definition	Post-crisis definition	ΔA_t	$\Delta \hat{A}_t$
$2005 \leq StartYear \leq 2007$	$2010 \leq StartYear \leq 2012$	0.051	0.076^{***}
			(0.022)

Note: ΔA_t is the productivity change before and after the crisis, $\Delta \hat{A}_t$ is the productivity change after controlling for industry and job content fixed effects. Standard errors are reported in parentheses. Statistical significance at the 1% level based on a two-tailed test is indicated by ***.

Table 2: Change of productivity

Year	Share_regular	Separation_rate_regular
2012	0.304	0.019
2013	0.397	0.018
2014	0.296	0.013
2015	0.378	0.011
2016	0.389	0.016

Note: Share of regular workers among those who leave the firm, and the separation rate for regular workers participating in "external" jobs.

Table 3: Composition of workers leaving the firm

Team Size	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Contribution Rank																			
1	1.00	0.76	0.64	0.57	0.52	0.48	0.46	0.43	0.41	0.40	0.37	0.36	0.36	0.34	0.34	0.32	0.31	0.30	0.30
2		0.23	0.25	0.25	0.25	0.24	0.23	0.22	0.22	0.22	0.21	0.20	0.19	0.19	0.20	0.19	0.18	0.19	0.18
3			0.10	0.12	0.13	0.14	0.13	0.14	0.13	0.13	0.14	0.13	0.13	0.13	0.13	0.13	0.12	0.13	0.12
4				0.05	0.07	0.07	0.08	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
5					0.03	0.04	0.05	0.05	0.06	0.06	0.06	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
6						0.02	0.03	0.03	0.04	0.04	0.04	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
7							0.01	0.02	0.02	0.03	0.03	0.03	0.04	0.04	0.04	0.04	0.04	0.04	0.04
8								0.01	0.01	0.02	0.02	0.02	0.03	0.03	0.03	0.03	0.03	0.03	0.03
9									0.01	0.01	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03
10										0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02	0.02
11											0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02
12												0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01
13													0.00	0.00	0.01	0.01	0.01	0.01	0.01
14														0.00	0.00	0.00	0.01	0.01	0.01
15															0.00	0.00	0.00	0.00	0.01
16																0.00	0.00	0.00	0.00
17																	0.00	0.00	0.00
18																		0.00	0.00
19																			0.00

Note: Average share of hours contributed by each worker, conditional on the rank of total working hours and the team size. The apparent zeros are the result of rounding positive numbers.

Table 4: Allocation of working hours across team members

Outcome	Change after crisis	Standard error	No. Obs	Adj. R^2
$\ln TeamSize_j$	-0.040^{***}	(0.016)	4431	0.734
l^c_{2j}	0.016^{***}	(0.005)	3991	0.577
$\ln h^{(1)}_j$	-0.072^{***}	(0.021)	4431	0.818
$Std\left(h_{ij}\right)$	-0.029^{**}	(0.012)	4431	0.315
$\max_i h_{ij} - \min_i h_{ij}$	-0.153^{***}	(0.044)	4431	0.397

Note: Estimation results of equation (4). Sample of data includes jobs with $2005 \leq StartYear \leq 2007$ or $2010 \leq StartYear \leq 2012$. Regressions control for industry and job content fixed effects. Statistical significance at 10%, 5%, 1% level, based on two-tailed tests, is indicated by *, **,***.

Table 5: Change of job-level variables after crisis

Sample	μ_ϕ	σ_ϕ	μ_H	σ_H	$\rho_{\phi H}$	η	error
$2005 \leq StartYear \leq 2007$	10.600	1.256	-0.189	1.247	-0.633	0.84	0.058
$2010 \leq StartYear \leq 2012$	10.851	1.143	-0.193	1.236	-0.632	0.84	0.050

Note: Calibrated parameter values. The error is the minimized value of the objective function.

Table 6: Calibrated parameters

Outcome	Real or simulated	Change after crisis	Standard error	No. Obs	Adj. R^2
$\ln h_j$	Real	-0.076^{***}	(0.022)	4431	0.866
$\ln h_j$	Simulated	-0.073^{***}	(0.009)	19552	0.813
$\ln TeamSize_j$	Real	-0.040^{***}	(0.016)	4431	0.734
$\ln TeamSize_j$	Simulated	-0.044^{***}	(0.008)	19552	0.702
l_{2j}^c	Real	0.016^{***}	(0.005)	3991	0.577
l_{2j}^c	Simulated	0.011^{***}	(0.002)	16909	0.556
$\ln h_j^{(1)}$	Real	-0.072^{***}	(0.021)	4431	0.818
$\ln h_j^{(1)}$	Simulated	-0.058^{***}	(0.008)	19552	0.708
$Std\left(h_{ij}\right)$	Real	-0.029^{**}	(0.012)	4431	0.315
$Std\left(h_{ij}\right)$	Simulated	-0.013^{***}	(0.003)	19552	0.416
$\max_i h_{ij} - \min_i h_{ij}$	Real	-0.153^{***}	(0.044)	4431	0.397
$\max_i h_{ij} - \min_i h_{ij}$	Simulated	-0.108^{**}	(0.022)	19552	0.344

Note: Estimation results of equation (16). Sample of real data includes jobs with $2005 \leq StartYear \leq 2007$ or $2010 \leq StartYear \leq 2012$. Regressions using real data control for industry and job content fixed effects. Statistical significance at 10%, 5%, 1% level, based on two-tailed tests, is indicated by *, **,***.

Table 7: Change of job-level variables after crisis

	$\ln h_j$	Standard error	$\ln h_j$	Standard error
	Real data		Simulated data	
$1\{AfterCrisis\}$ $\times \ln \bar{TeamSize_j}$	-0.042^{**}	(0.018)	-0.040^{***}	(0.007)
$1\{AfterCrisis\}$	-0.061^{***}	(0.020)	-0.057^{**}	(0.007)
$\ln \bar{TeamSize_j}$	0.875^{***}	(0.019)	0.811^{***}	(0.007)
$\ln Y_j$	0.656^{***}	(0.009)	0.495^{***}	(0.005)
No. Obs.	4431		19552	
Adj. R^2	0.913		0.901	

Note: Regression results of equation (17). $\ln \bar{TeamSize_j}$ is re-centered such that it is equal to zero at average team size. Sample of real data includes jobs with $2005 \leq StartYear \leq 2007$ or $2010 \leq StartYear \leq 2012$. Regressions using real data control for industry and job content fixed effects. Statistical significance at the 10%, 5%, and 1% levels based on two-tailed tests is indicated by *, **,***.

Table 8: Correlation between productivity, labor share and team size

Quantile	$\ln h_j$	Standard error	$\ln h_j$	Standard error
	Real data		Simulated data	
0.100	0.003	(0.032)	0.023	(0.019)
0.250	-0.058^{**}	(0.025)	-0.032^{***}	(0.012)
0.500	-0.045^{**}	(0.019)	-0.085^{***}	(0.009)
0.750	-0.103^{***}	(0.018)	-0.130^{***}	(0.010)
0.900	-0.160^{***}	(0.022)	-0.160^{***}	(0.013)
No. Obs	4431		19552	

Note: Quantile regressions (18) using real and simulated data. Sample of real data includes jobs with $2005 \leq StartYear \leq 2007$ or $2010 \leq StartYear \leq 2012$. Regressions using real data control for industry and job content fixed effects. Standard errors are heteroskedasticity robust. Statistical significance at the 5%, and 1% levels based on two-tailed tests is indicated by **,***.

Table 9: Differential impact on productivity

	$\ln h_j$	Standard error	$\ln h_j$	Standard error
	Real data		Simulated data	
$l_j^{(1)}$	-1.214^{***}	(0.055)	-0.843^{***}	(0.019)
$TeamSize_j$	0.028^{***}	(0.001)	0.058^{***}	(0.001)
$\ln Y_j$	0.729^{***}	(0.010)	0.601^{***}	(0.004)
No. Obs.	4431		19552	
Adj. R^2	0.898		0.887	

Note: Estimation results of regression equation (19) using real and simulated data. Sample of real data includes jobs with 2005 $\leq StartYear \leq$ 2007 or 2010 $\leq StartYear \leq$ 2012. Regressions using real data control for industry and job content fixed effects. Statistical significance at the 1% levels based on two-tailed tests is indicated by ***.

Table 10: Testing the model's predictions

	(1)	(2)	(3)	(4)	(5)	(6)
	$\frac{OT_j}{h_j}$	$\frac{OT_j^{(1)}}{h_j^{(1)}}$	$\frac{OT_j^{(2)}}{h_j^{(2)}}$	$\frac{OT_j^{(3)}}{h_j^{(3)}}$	$\frac{OT_j^{(4)}}{h_j^{(4)}}$	$\frac{OT_j^{(5)}}{h_j^{(5)}}$
mean	0.170	0.188	0.172	0.167	0.161	0.157
std	0.068	0.113	0.108	0.108	0.106	0.108
min	0.003	0.000	0.000	0.000	0.000	0.000
25%	0.123	0.104	0.090	0.086	0.079	0.071
50%	0.162	0.181	0.163	0.157	0.154	0.149
75%	0.212	0.263	0.242	0.241	0.233	0.232
max	0.493	0.562	0.587	0.611	0.635	0.525
count	4658	4658	4658	4658	4658	4658

Note: Summary statistics for the overtime shares that appear in regression equation (20).

Table 11: Sample description: overtime share conditional on contribution rank

	(1)		(2)		(3)	
	$\ln \frac{Adj\_Rev_j}{h_j}$	Standard error	$\ln \frac{Adj\_Rev_j}{h_j}$	Standard error	$\ln \frac{Adj\_Rev_j}{h_j}$	Standard error
$\frac{OT_j}{h_j}$			-1.359^{***}	(0.148)	-1.525^{***}	(0.134)
$\frac{OT_j^{(1)}}{h_j^{(1)}}$	-0.431^{***}	(0.086)				
$\frac{OT_j^{(2)}}{h_j^{(2)}}$	-0.326^{***}	(0.081)				
$\frac{OT_j^{(3)}}{h_j^{(3)}}$	-0.206^{***}	(0.079)				
$\frac{OT_j^{(4)}}{h_j^{(4)}}$	-0.195^{**}	(0.083)				
$\frac{OT_j^{(5)}}{h_j^{(5)}}$	-0.072	(0.080)				
$l_j^{(1)}$	0.823^{***}	(0.078)				
$l_j^{(2)}$	0.741^{***}	(0.140)				
$l_j^{(3)}$	1.007^{***}	(0.230)				
$l_j^{(4)}$	1.045^{***}	(0.362)				
$l_j^{(5)}$	0.137	(0.469)				
$\ln Area_j$	0.063^{***}	(0.005)	0.047^{***}	(0.005)	0.048^{***}	(0.005)
Model	OLS		IV		OLS	
No. Obs.	4658		4645		4645	
Adj. R^2	0.121		0.130		0.131	

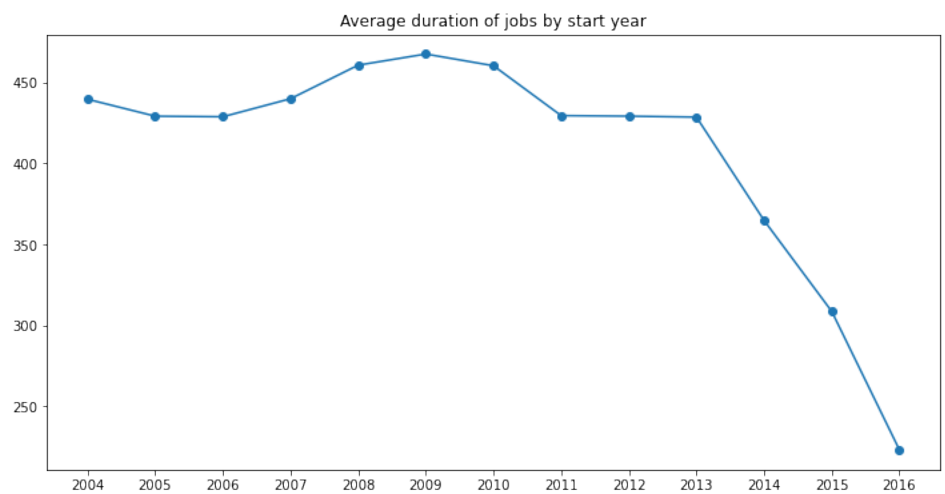
Note: Estimation results for regression equation (20). Standard errors are clustered at the project level. Statistical significance at the 5%, and 1% levels based on two-tailed tests is indicated by **,***.

Table 12: Overtime productivity effects for team members with heterogeneous hours

	$\partial y/\partial x$	Standard error
$\ln\left(OT_j^{(1)}+1\right)$	0.533^{**}	(0.260)
$\ln\left(OT_j^{(2)}+1\right)$	-0.208	(0.223)
$\ln\left(OT_j^{(3)}+1\right)$	0.263	(0.229)
$\ln\left(OT_j^{(4)}+1\right)$	0.153	(0.224)
$\ln\left(OT_j^{(5)}+1\right)$	-0.221	(0.191)
$\ln Area_j$	-0.277^{**}	(0.133)
$\ln Rev_j$	0.346	(0.361)
Pseudo R^2	0.104	
No. Obs.	1454	

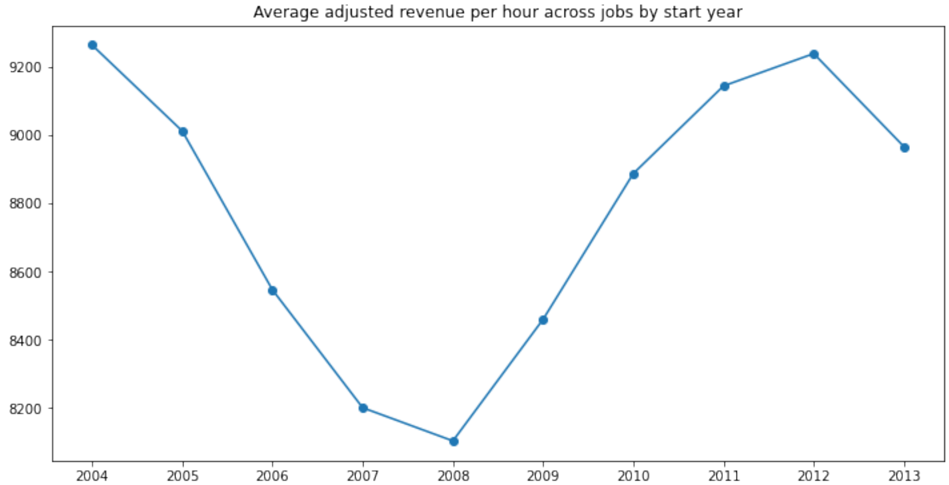
Note: Cell entries in the first column are marginal effects (multiplied by 100) from the probit model (22), computed at the mean of each regressor. Cell entries in the second column are the associated standard errors. Statistical significance at the 5% levels based on two-tailed tests is indicated by **.

Table 13: Penalty probability as a function of workers' overtime



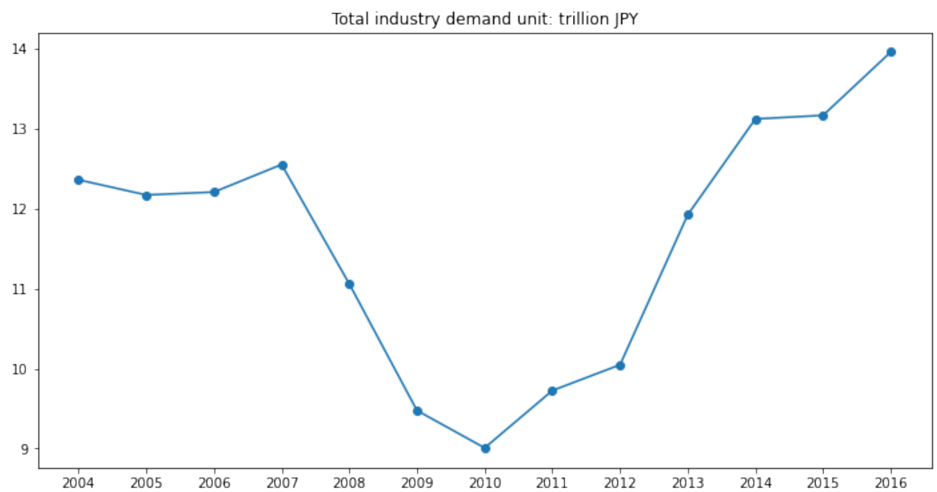
Note: Average duration (unit: days) of jobs that start in each year and are completed before the end of the sample period.

Figure 1: Average duration of jobs, 2004 to 2016



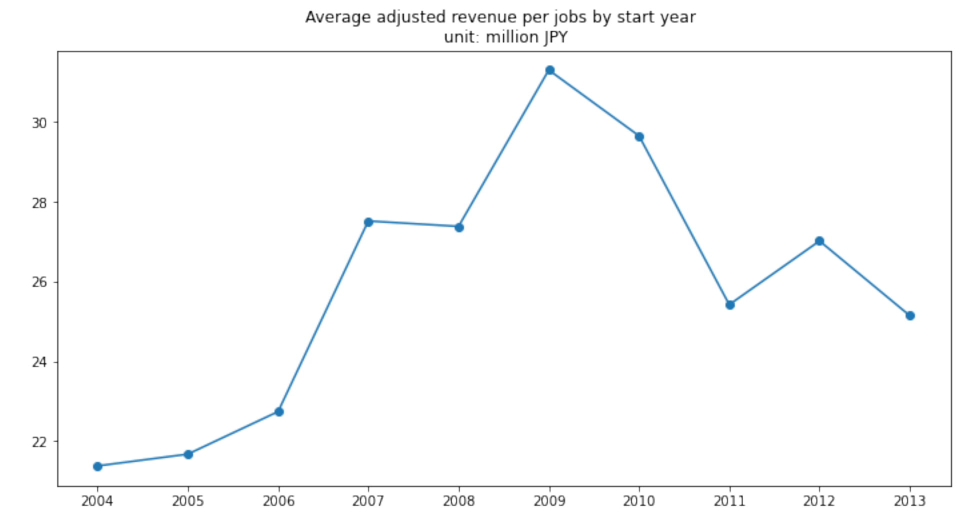
Note: For each year, the figure plots average adjusted revenue per hour, $\frac{Adj\_Rev_j}{h_j}$ (weighted by h_j), for jobs that started in that year.

Figure 2: Revenue per hour, 2004 to 2013



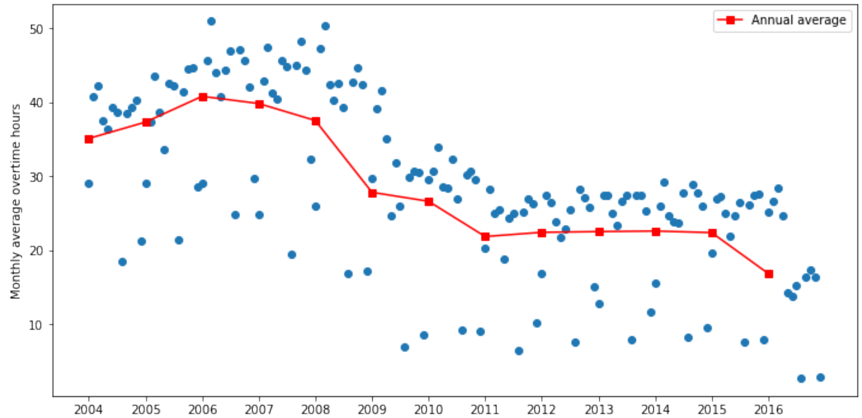
Note: Total amount of orders in the industry survey data each year.

Figure 3: Industry demand, 2004 to 2016

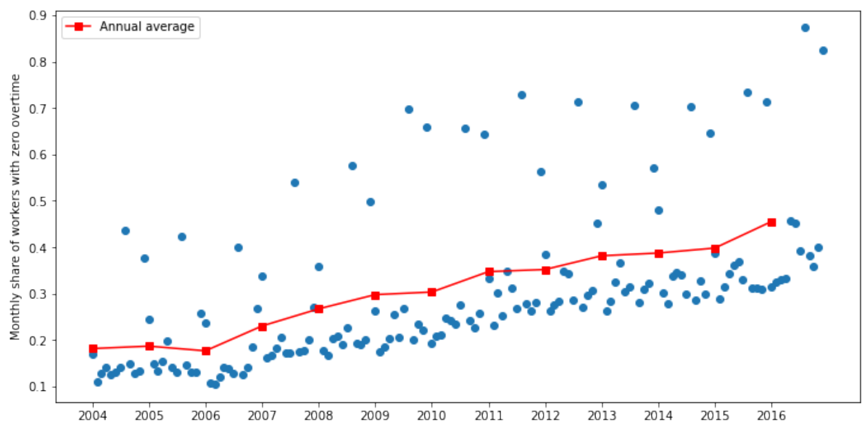


Note: For each year, the figure shows the average adjusted revenue per job.

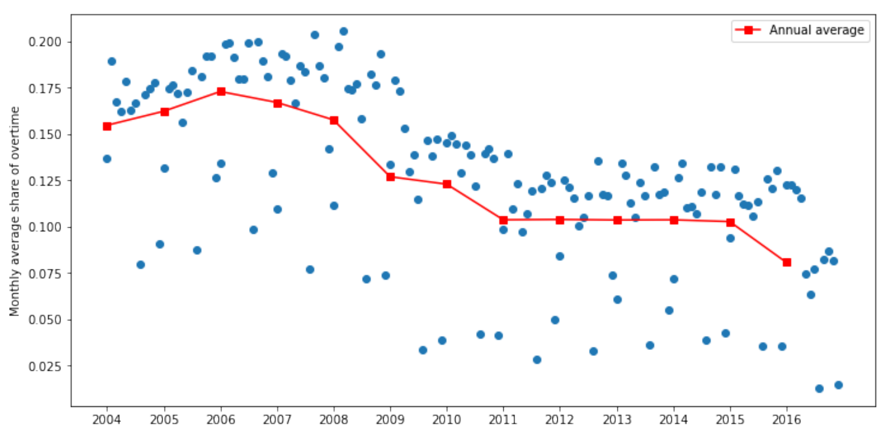
Figure 4: Adjusted revenue per job and number of jobs start in each year, 2004 to 2013



(a) Average overtime



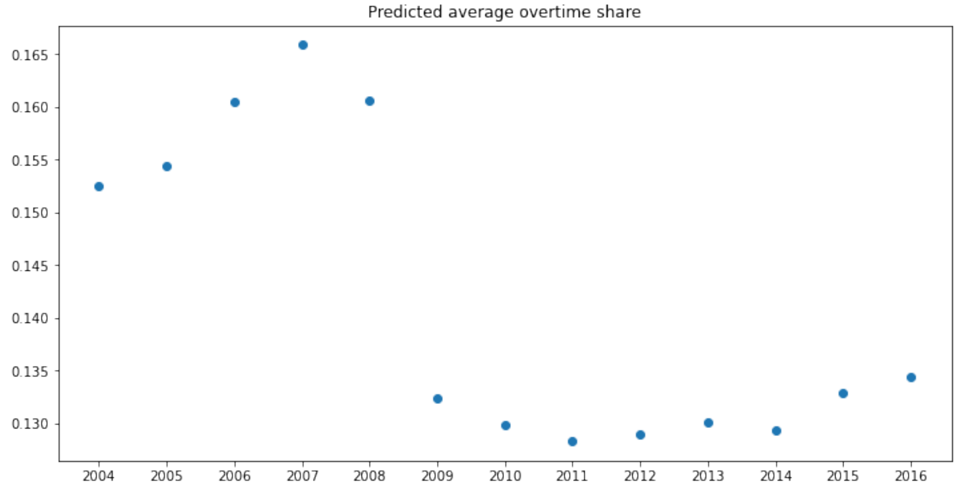
(b) Share of worker with zero overtime



(c) Average share of overtime

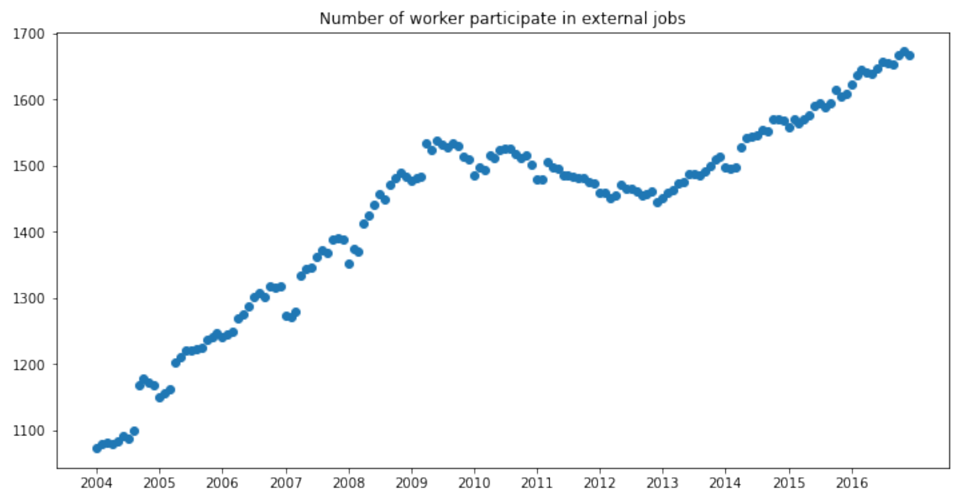
Note: The figure shows how average overtime across workers changes from 2004 to 2016. Panel (a) plots the average overtime hours, panel (b) plots the share of workers with zero overtime, and panel (c) plots the average share of overtime out of total monthly working hours.

Figure 5: Average monthly overtime across time



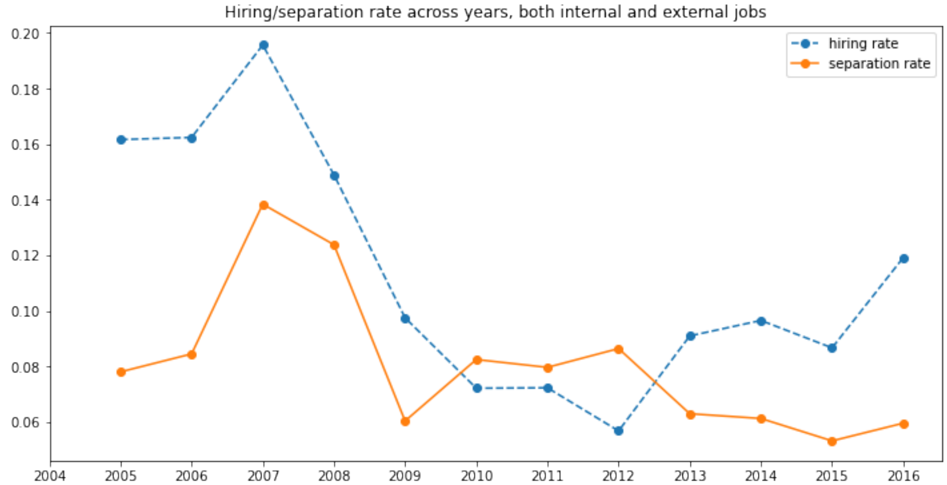
Note: The figure plots $OTShare\_Ind_j$, defined in Equation (3), averaged across jobs, for each starting year.

Figure 6: Average predicted overtime share, 2004 to 2016



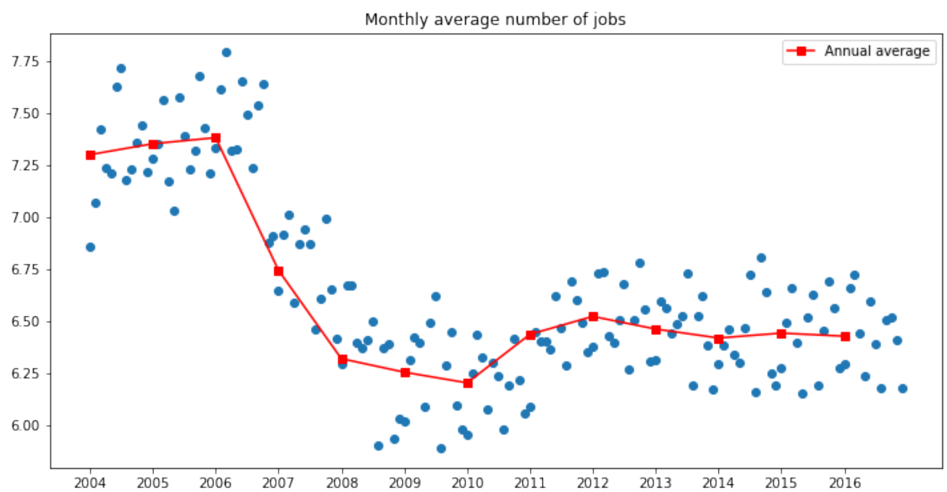
Note: The figure shows the number of workers appearing in the labor input data in each month.

Figure 7: Number of workers, 2004 to 2016



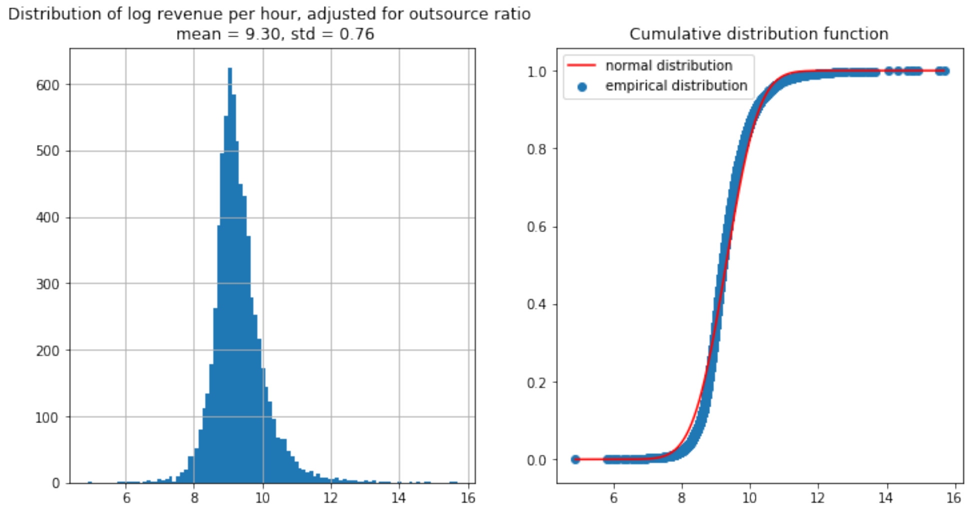
Note: The figure shows the hiring rate and separation rate of workers who participate in external jobs during the period. The hiring rate is the share of workers present in the current year but not in the previous year. The separation rate is the share of workers present in the previous year but not in the current year.

Figure 8: Hiring/separation rate over time



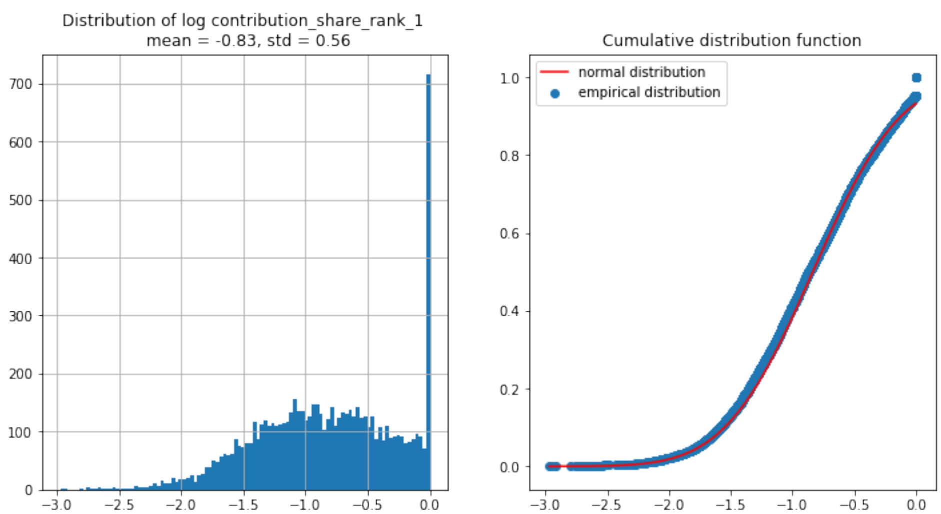
Note: The figure plots the average number of jobs in which each worker participates.

Figure 9: Average number of jobs assigned to each worker



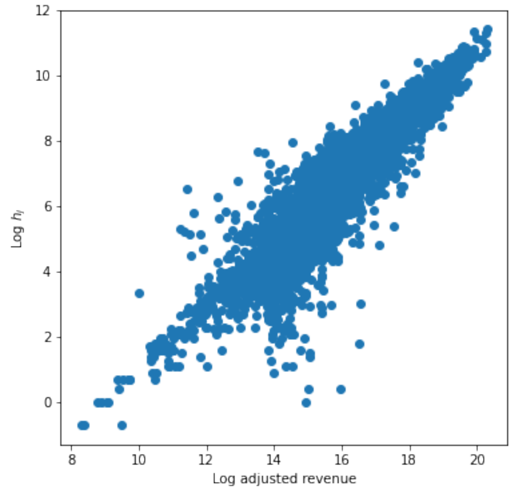
Note: The left panel plots the histogram of $\ln \frac{Adj\_Rev_j}{h_j}$. The right panel plots the corresponding cumulative distribution function. The red line is the theoretical cumulative distribution function of a normal random variable with the same mean and standard deviation.

Figure 10: Distribution of log productivity in data



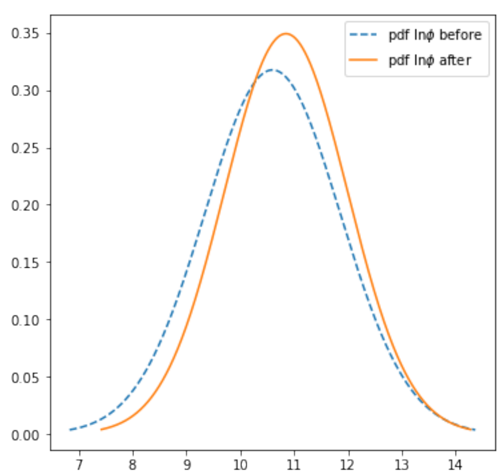
Note: The left panel plots the histogram of $\ln l_j^{(1)}$, or the log of the rank-1 labor share. The right panel plots the corresponding cumulative distribution function. The red line is the theoretical cumulative distribution function of a normal random variable with the same mean and standard deviation.

Figure 11: Distribution of log labor share of rank 1 worker in data



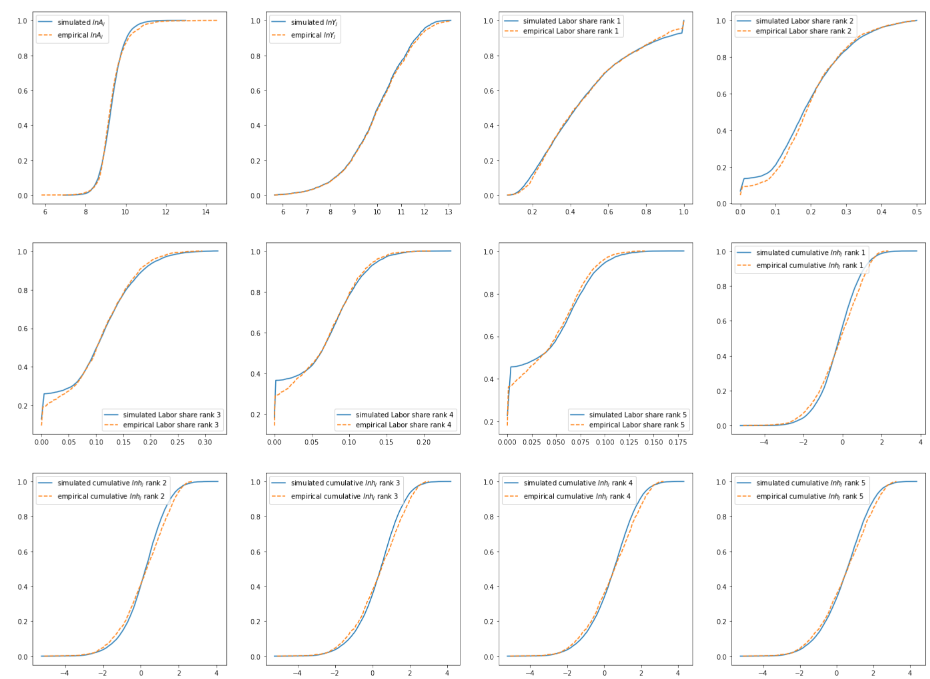
Note: The figure plots $\ln Adj\_Rev_j$ on the horizontal axis and $\ln h_j$ on the vertical axis. The regression line, after controlling for client industry and job content dummies, has a slope of 0.996.

Figure 12: Return to scale



Note: The figure plots the implied distribution of $\ln \phi$, before and after the crisis.

Figure 13: Calibrated PDF of $\ln \phi_i$



Note: The figures plot the empirical distributions and the simulated distributions of key variables. The first two figures in the first row plot the cumulative distribution function of A_j, and Y_j. Starting from the third figure in the first row to the third figure in the second row, the cumulative distribution function of labor shares up to the rank-5 worker are plotted. Starting from the fourth figure in the second row to the last figure in the third row, the cumulative distribution function of labor hours up to the rank-5 worker are plotted.

Figure 14: Model fit, after crisis