The Inner Beauty of Firms

Jacob Kohlhepp*

[Click Here for Latest Version]

August 21, 2024

Abstract

This paper uses millions of task assignments across hundreds of hair salons from management software to study the connection between task assignment within the establishment and productivity. There is significant dispersion in productivity and internal task specialization, and a strong association between the two: The top 25 percent of salon-quarters in terms of specialization are on average 68 percent more productive than the bottom 25 percent. I construct a model where organizationally unique firms choose how to assign tasks to workers with multidimensional skills in product and labor market equilibrium. Heterogeneity in task specialization is micro-founded by costly communication within the firm. I prove the model, including worker skills and firm-specific organization costs, is constructively identified from task assignments, prices and market shares. I consider four counterfactual policies: a diffusion of management practices, a sales tax increase, an increase in market concentration, and immigration. Allowing salons to internally reorganize alters the industry-wide productivity effects of economic shocks, in some cases reversing the direction.

JEL Codes: J2, L2

*Department of Economics, University of North Caroline, Chapel Hill. Contact: kohlhepp@unc.edu
I am grateful for helpful suggestions from Tara Sinclair, Maurizio Mazzocco, Simon Board, Daniel Haanwinckel, Moritz Meyer-ter-Vehn, Alex Bloedel, Robert McDonough, Ichiro Obara, Jay Lu, Felipe Goncalves, Till von Wachter, Coen Teulings, Tyler Ransom, Rodrigo Pinto, Jesse Shapiro, John Kennan, Stepan Aleksenko, Marco Duarte, Andrii Babii, Jonathan Kowarski, and Victoria Barone. I thank Malayvardhan Prajapati and Will Firmin for excellent research assistance. This work received support from the Institute for Humane Studies under grant number IHS017120.

"Of all the things I've done, the most vital is coordinating those who work with me and aiming their efforts at a certain goal." — **Walt Disney**

1 Introduction

There are large differences in productivity across similar firms. Additionally, there are persistent differences in management practices that impact the ability of firms to assign the right task to the right worker. Two natural questions emerge from these observations. First, to what extent is internal task specialization related to individual firm productivity? Second, how is aggregate productivity determined when organizationally unique firms choose how to assign tasks?

These questions are critical for understanding the productivity implications of many economic shocks, including the diffusion of management practices, taxation, immigration and increased market concentration. However, answering these questions poses a dual challenge. Empirically, it is necessary to look inside the black box of the firm and observe how workers are assigned to tasks. Further, this must be done not just for one firm but for many comparable firms. Theoretically, it is necessary to develop a model where organizationally unique firms choose their task assignments in product and labor market equilibrium.

In this paper, I overcome the empirical challenge using novel data from a management software company to show a robust association between task specialization within the firm and productivity. I overcome the theoretical challenge by proposing a model where task specialization within the firm is heterogeneous and endogenous. Tractability, identification and estimation are achieved by modeling organizational differences via firm-specific mutual information-based specialization costs. I use the model to study the industry-wide labor productivity effects of several counterfactual economic shocks. I show that allowing firms to internally reorganize in response to a shock qualitatively changes the productivity implications, sometimes reversing the sign of the change in aggregate labor productivity.

The data document how millions of tasks are assigned to individual workers across hundreds of hair salons. To compare internal task specialization across firms, I develop the task specialization index (which I call the s-index), which indicates how far the observed task assignment is from a generalist benchmark where tasks are assigned randomly to workers.

As documented in other contexts, revenue productivity dispersion across similar establishments is large: The salon-quarter at the 75th percentile is twice as productive as the salon-quarter at the 25th percentile. However, dispersion in the s-index is even larger: The salon-quarter at the 75th percentile is 13 times as specialized as the salon-quarter at the 25th percentile. There is also a positive correlation between the s-index and productivity, even among salons with the same number of employees. The top 25 percent of salon-quarters in terms of specialization are on average 68 percent more productive than the bottom 25 percent.

Dispersion in task specialization and the relationship between task specialization and productivity exist even among firms with similar numbers of employees. By decomposing the link between productivity and the s-index, I find that rather than servicing a greater number of customers, task-specialized salons generate more revenue per customer and have higher customer return rates. This finding indicates that productivity gains manifest via quality differences. Finally, the s-index is positively correlated with other potentially productive management practices including teamwork (assigning multiple workers to the same customer on a single date) and early adoption of software features.

In light of these facts, I propose a model where task specialization within the firm is heterogeneous and endogenous. Workers have multidimensional skills, and wages are worker-specific. Strategic firms engage in Bertrand-style price competition while also choosing the composition of their workforce and the task assignment of each hired worker. Better assignment of workers to tasks increases product quality but requires the firm to spend more time communicating task assignments to workers. Importantly, there are firm-specific specialization costs, because firms differ in their ability to communicate task assignments.

I show that costly communication within the firm provides a microfoundation for the s-index and its positive correlation with productivity. The model delivers firm- and worker-specific task assignments that take a logit-like form. For any fixed wages, firm strategies exist and are essentially unique. Tractability is maintained even in equilibrium because the pricing and task assignment decisions can be separated, and organization costs are modeled using the mutual information, which makes the task assignment decision equivalent to a well-studied problem in behavioral economics (i.e., rational inattention) and computer science (i.e., rate-distortion).

I prove that firm organization costs, worker skills and worker wages are all identified from data which contain task assignments, firm prices and firm market shares. The identification proof is constructive and motivates an estimation procedure which does not require solving the model. In the first step of the procedure, workers are classified into types based on their task assignments, and firm organization costs are obtained by comparing how firms utilize pairs of workers with the same skills. In the second step, wages, skills and other parameters are obtained by solving a linear system of moment conditions. I implement a version of this procedure in order to estimate the model for salons in New York County, NY, Cook County, IL, and Los Angeles County, CA, for 12 quarters between 2018 and 2021.

The estimated model reveals wide heterogeneity across firms in terms of organization costs and across workers in terms of skill sets. Organization costs typically account for between 6% and 16% of the observed price. In Los Angeles and Cook Counties, the highest-wage workers are color specialists, while in New York County, the highest-wage workers are "superstar" generalists, or workers with high skill in several tasks.

I study labor-labor substitution patterns by increasing the wages of different worker types and studying relative labor demand responses at different firms. This analysis reveals that two workers are frequently substitutes at one firm and complements at another. I also study how the labor productivity of a worker changes in response to own and coworker wage increases. Because salons need to reassign the tasks of workers that are laid-off, wage increases of coworkers typically reduce own productivity, a phenomena I term picking up the slack. In contrast, own wage increases encourage salons to more carefully assign workers to tasks, increasing own productivity.

In the model the standard economic intuition that larger, more productive firms should hire workforce with a broader range of skills is not always true: When firm productivity manifests as the ability to assign workers to tasks, an improvement in productivity can either decrease or increase the number of employed worker skill sets.

Finally, I study the productivity implications of four counterfactual economic shocks: a diffusion of management practices, a sales tax increase, an increase in market concentration, and immigration. For each shock, I solve for the new product and labor market equilibrium under two regimes. First, I fix task assignments at their initial configuration, and allow firms to change only the total amount of labor they hire. I can thus quantify the reallocation effects of each shock. Second, I allow firms to fully reorganize. I show that neglecting reorganization within the firm changes the magnitude and in some cases the sign of the change in aggregate labor productivity. This change occurs because labor productivity is determined by task specialization within the firm, and economic shocks change firms' incentives to engage in specialization.

Specifically, the diffusion of management practices lowers the cost of specialization for each firm, increasing specialization and productivity. The immigration of low-wage workers reduces the wage of workers that share their skill set, effectively making a specialized work force cheaper, increasing productivity. Sales tax increases reduce specialization and productivity, because they make it more difficult for firms to pass on the cost of specialization to consumers. Increased market concentration has different effects in different markets, decreasing productivity in Los Angeles County but slightly increasing productivity in Cook County and New York County. Even in cases where aggregate productivity impacts are large, the impacts for individual workers often vary in sign.

These patterns point to a broader message: When internal organization is endogenous and heterogeneous, it is difficult to make broad statements about the distributional and sometimes even the aggregate productivity implications of economic shocks. The reason is that firms differ in how much they overlap with other firms in the labor market in the initial equilibrium, and they differ in their ability to adjust their internal organization to accommodate external shocks. Although firms with lower organization costs will always find it easier to adjust, the way in which all competitors adjust may still disadvantage productive firms, or, worse yet, discourage productive firms from engaging in costly specialization.

This paper contributes to four strands of the literature: the determinants of firm productivity, the task-based perspective of labor markets, talent allocation within the firm, and organizational economics.

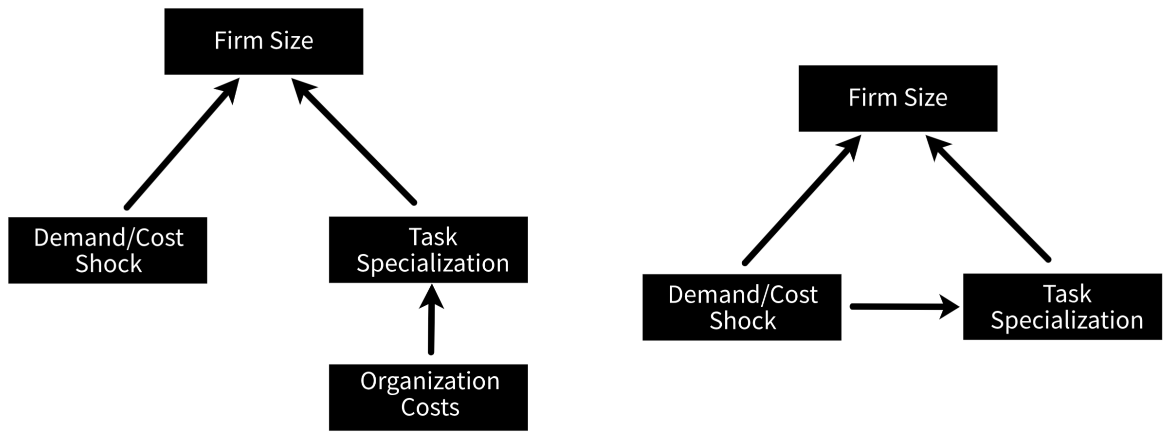
Determinants of Firm Productivity. Past studies have shown that productivity differences across firms are large (Syverson 2004) and can be linked to management practices (Bloom and Van Reenen 2007). Much of existing work focuses on manufacturing firms. This paper finds similar revenue productivity dispersion in service firms, and shows that task specialization accounts for some of this dispersion. Further, firms that engage in task specialization also engage in more sophisticated ways with management software. The differences in specialization, and, in particular, the large number of generalized salons in the U.S. are similar to the patterns Bassi et al. (2023) observe in Ugandan manufacturing firms. The link between productivity and task specialization is consistent with the link between productivity and coordination established by Kuhn et al. (2023).[1]

Task-Based Perspective of Labor markets. I model labor as being divisible into tasks which can be assigned to workers with different skills, a tradition that dates back to at least Sattinger (1975) but has seen growing use since Autor, Levy, and Murnane (2003). I incorporate features present in different parts of the literature, including multidimensional worker types (Lindenlaub 2017, Ocampo 2022), firms with multiple worker types (Freund 2022, Haanwinckel 2023), organization costs (Garicano 2000,Adenbaum 2022, Bassi et al. 2023), and firm-specific task demands (Lazear 2009).

The main innovation relative to prior work is that I allow for firm-specific specialization costs. These costs are micro-founded by costly communication, as in Garicano (2000), and are tied to a primitive that is unrelated to classic demand or cost shocks. These are not fixed costs per occupation, as in Adenbaum (2022). As a result, specialization is not directly caused by demand or cost shocks. Rather, specialization is determined separately and interacts with demand and cost shocks to determine firm size, forming the causal relationships depicted in Panel (a) of Figure 1. In most other models, firms specialize because they expect to be large, leading to the causal relationships depicted in Panel (b) of Figure 1. My model can accommodate firms that are large due to specialization and firms that are large due to other reasons. It can thus rationalize the differences in task specialization I observe among firms of the same size.

1. This paper is also relevant to the literature on the role of firms in wage inequality (Card, Heining, and Kline 2013, Alvarez et al. 2018, Song et al. 2019). My paper positions task assignment as a mechanism through which this effect could operate.

Figure 1: Causal Relationships



(a) This Paper

(b) Others

A secondary methodological contribution is that my model can be identified and estimated without classic linked employer-employee data. In much of the literature, the distribution of worker skills is identified using detailed information on workers, typically, wages, occupations and education. In this paper, worker skills are recovered from task assignment patterns. Such data are becoming increasingly available across a wide range of industries, including pharmaceutical research and restaurants.[2] These are complementary approaches: Wage information allows the researcher to recover absolute advantage and unidimensional forms of comparative advantage across workers, while task assignment information allows the researcher to recover horizontal differences across multidimensional worker skill portfolios.

Talent Allocation within the Firm. A recent body of work has shown that managers play an important role in determining worker productivity within the firm (Haegele 2022, Coraggio et al. 2023, Minni 2023). Consistent with my findings, these effects operate through the assignment of workers to jobs within the firm. While my paper studies the beauty industry, Minni (2023) studies a consumer goods multinational, Haegele (2022) studies a large European manufacturer, and Coraggio et al. (2023) studies Swedish registry data. The diversity of these contexts strengthens the external validity of the each paper's individual findings. This paper also complements this work by providing a way

2. Some examples include TruLab (trulab.com), which is used by vaccine and biotech companies to manage laboratories, and 7shifts (www.7shifts.com), which is used by restaurants.

to integrate management heterogeneity in market equilibrium. I show that doing so qualitatively changes the impact of economic shocks.

Organizational Economics. Finally, the literature in organizational economics provides many reasons why firms may differ in their ability to assign workers to tasks. These include monitoring (Alchian and Demsetz 1972, Baker and Hubbard 2003), relational contracts (Baker, Gibbons, and Murphy 2002), knowledge (Garicano and Wu 2012), coordination (Dessein and Santos 2006), trust (Meier, Stephenson, and Perkowski 2019) and culture (Martinez et al. 2015). This paper contributes to this literature by studying how such organizational differences change industry-wide outcomes. This is done by representing organizational heterogeneity as firm-specific specialization costs in a model where firms strategically interact.

2 Data

This section describes the salon management software data I use in this paper.

2.1 Context and Institutional Details

The data set was obtained from a data sharing agreement I negotiated with a salon management software company. The software facilitates running a beauty business, including scheduling, pricing, payments, inventory, staffing, business reporting, client profiling and marketing. As of July 19, 2022, a monthly subscription has a base price of $175. Although the company also markets its software to spas, tanning salons and massage parlors, hair salons and barbers make up the majority of its clients. For this reason, I analyze only hair salons and barbershops.

The software is sold to beauty businesses throughout the United States, but the data indicate uptake is largest in Los Angeles (where the company was founded) and New York City. An important aspect of the data set is that it allows me to observe the internal organization of salons that are geographically close and therefore likely to be direct competitors in labor and product markets. For example, I observe 10 salons in the Lower Manhattan zip code 10013, which is a 0.55 square mile area.

The data document which stylist is assigned to each task and client, and record the duration of the appointment, the price paid, and a custom text description of each task. If more than one employee is assigned to a single client, this is recorded as multiple entries describing what each employee contributed. Although the data are de-identified, IDs that are unique within a firm allow me to track employees and clients across time within a firm but not across firms.

A sample from the data is provided in Table 1, with IDs replaced with pseudonyms. This sample shows the different ways two salons coordinate employees to meet customer demand. Blake requested a cut, highlights and a treatment at salon 1A. The salon had a single employee, Rosy, perform all three services. Grace requested a cut and a single process (color) at salon 2A. Unlike salon 1A, salon 2A chose to assign each of these tasks to two employees, Tyler and Ben. Both of these salons are in the same zip code. Throughout, I measure labor in units of time using two variables form the data.

Table 1: Salon Activity Data Sample

Firm	Salon	App.	Cust.	Service	Staff	Time Stamp	Price	Duration (minutes)
1	1A	123	Blake	Advanced Cut	Rosy	3/26/2021 16:15	100	72
1	1A	123	Blake	Full Head - Highlights	Rosy	3/26/2021 16:15	243	127
1	1A	123	Blake	Treatment Add On (Olaplex)	Rosy	3/26/2021 16:15	39	72
2	2A	9982	Grace	Women's Cut	Tyler	3/17/2021 11:00	225	43
2	2A	9982	Grace	Single Process	Ben	3/17/2021 11:00	200	77

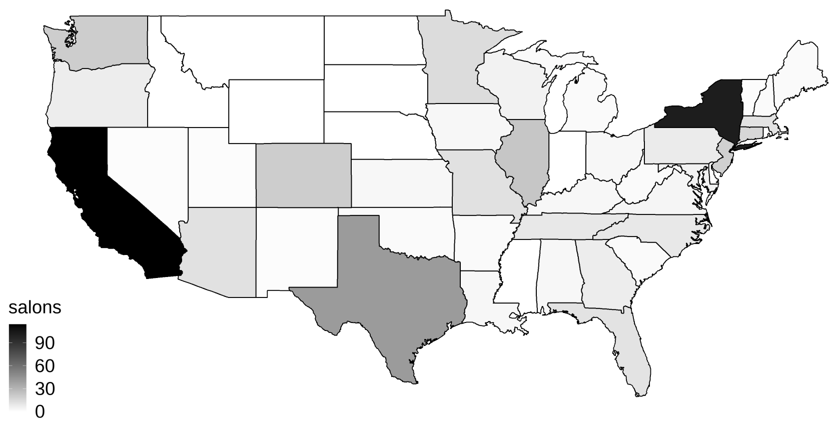
Note: This table is a snapshot displaying two actual appointments at salons in the same zip code from the data used for the estimation. Customer, firm, salon and client IDs are replaced by pseudonyms.

While the data are rich in terms of task content and worker assignments, information about employee compensation is sparse. The software can track some compensation information (tips, commissions and employment relationship, etc.), but these additional functions are not used consistently by client salons, as my discussions with the company and analysis of internal data revealed. As the data set contains 20,560 unique text descriptions of services, I hired a licensed cosmetologist to group the tasks into five mutually exclusive task categories. Appendix Section A.16 details the process. I label each task category based on the main type of task it represents, with the understanding that there is heterogeneity within a category.

2.2 Descriptive Statistics

The data used in this section and in the Stylized Facts section include all observed salon-quarters where revenue per customer is positive. I exclude 2021 Q3, because I observe only part of the quarter. I also exclude an establishment in Kentucky with revenue that is implausibly high. The data contain information on 445 hair salon establishments, which represent 316 unique firms, 9,179 hair stylists, 1,654,233 customers and 10.8 million services performed. Firms first appear in the data when they adopt the management software. Figure 2 illustrates the distribution of salon establishments by state. Although the software is used by salons across the country, users are concentrated in the states of New York and California.

Figure 2: Hair Salons by State



Note: The distribution of unique salon establishments by state. Includes all hair salons observed prior to quarter 3 of 2021.

I aggregate the data to the salon-quarter level for analysis. Descriptive statistics at this level are provided in Table 2.

Table 2: Salon-Quarter Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Revenue	4,599	212,419.70	247,576.60	5.00	2,559,703.00
Employees	4,599	13.42	10.76	1	92
Customers	4,599	1,155.68	1,094.34	1	16,768
Share Haircut/Shave	4,599	0.40	0.23	0.00	1.00
Share Color/Highlight/Wash	4,599	0.38	0.20	0.00	1.00
Share Blowdry/Style/Treatment/Extensions	4,599	0.10	0.12	0.00	1.00
Share Admininstrative	4,599	0.05	0.11	0.00	1.00
Share Nail/Spa/Eye/Misc.	4,599	0.06	0.16	0.00	1.00

Note: Summary statistics for all salon-quarters used for the stylized facts. Excludes 2021Q3 (a partial quarter) and a single outlier salon where revenue appears incorrectly denominated.

The salons in the sample have an average quarterly revenue of $213,201 and an average of 13 employees. Johnson and Lipsitz (2022) studies a sample of salon owners and reports an average annual (not quarterly) revenue of $233,000 and an average of seven stylists. Given this, the sample in this paper should be viewed as a positively selected sample of salons, and the heterogeneity in productivity and specialization observed in this sample is likely underestimating the heterogeneity that would be observed in the universe of U.S. hair salons. Such positive selection is reasonable given that salons must pay a subscription fee to access the software.

In terms of task composition, salons typically spend most of their time on the Haircut/Shave and Color/Highlight/Wash tasks, but there are large differences across salon-quarters. Even though there is variation in the relative intensity of tasks at different salons, most salons offer at least four of the five task categories in a given quarter. Throughout the paper, I refer to the task mix of a salon as the fraction of total time spent on each of the five tasks. I define a salon's price to be its total quarterly revenue divided by its total number of customers. The average price across all saon-quarters in the sample is $200.[3]

3 Stylized Facts

This section presents four stylized facts about the relationship between task assignment within the firm and productivity. These facts require the definition of two concepts which will be used throughout the paper. To begin, denote workers by the index m, firms by the index j, and tasks by the index k.

Definition 1 *A firm's observed organization, denoted by B_j, is a matrix where element $B_j(m,k)$ is the fraction of labor assigned to worker m and task k.*

Given firm j's observed organization, I define a firm's *generalist benchmark* as $G(B_j)(m,k) = \left(\sum_{m'} B_j(m',k)\right)\left(\sum_{k'} B_j(m,k')\right)$. This is the counterfactual task assignment that would be observed if the firm randomly assigned workers to tasks, holding fixed the firm-level marginal distribution of time across tasks and workers. In Figure 3 I provide an example of an observed organization and the corresponding generalist benchmark.

3. I analyze firms as offering a representative differentiated product (service) rather than multiple products (services).

Figure 3: An Organization and the Corresponding Generalist Benchmark

Task Assignment (B_j)

Employee	Tasks: 1	2	3	
A	1/2	0	0	1/2
B	0	1/4	0	1/4
C	0	0	1/4	1/4
Tot.	1/2	1/4	1/4	1

Generalist Benchmark ($G(B_j)$)

Employee	Tasks: 1	2	3	
A	1/4	1/8	1/8	1/2
B	1/8	1/16	1/16	1/4
C	1/8	1/16	1/16	1/4
Tot.	1/2	1/4	1/4	1

Note: The left panel depicts a task-specialized salon, while the right panel depicts the corresponding generalist task assignment. Column sums represent the task mix, and row sums represent the fraction of work performed by each employee.

To study task specialization, I need to first define a notion of within-firm specialization. I measure how far the observed task assignment is from the corresponding generalist task assignment, using the Kullback–Leibler divergence (denoted D_{KL}) as the notion of distance. In Section 4.3, I show that this measure of task specialization can be microfounded as the minimum amount of communication that must occur within the firm in order for it to implement a given task assignment (Shannon 1948).

Definition 2 *The task specialization index* ***(s-index)*** *of a firm j with organization B_j is*

$$I(B_j) := D_{KL}(B_j||G(B_j)) = \sum_{m,k} B_j(m,k) log\left(\frac{B_j(m,k)}{G(B_j)(m,k)}\right).$$

With this measure, I present four stylized facts about productivity and internal organization. Throughout the rest of the paper, task assignments and in particular the s-index are assumed to be measured without error.[4]

Fact 1 *There is large dispersion in labor productivity and internal task specialization.*

I measure labor productivity as total revenue divided by the total duration of all services. Total duration serves as a proxy for utilized labor. Table 3 documents the result. In line with past work in the literature, I find large differences in productivity across otherwise similar salons. Remarkably, the ratio of productivity between the 75th and 25th percentiles is 2-to-1, the same as that documented by Syverson (2004) among manufacturing firms.

4. Appendix Section A.17 provides evidence that measurement error is small.

Table 3: Dispersion of Labor Productivity and Task Specialization

Statistic	N	Mean	Min	Pctl(25)	Median	Pctl(75)	Max
Labor Productivity	4,599	1.81	0.003	1.03	1.38	2.05	42.80
S-index	4,599	0.22	0.00	0.03	0.11	0.41	1.02

Note: The table describes the distribution of labor productivity and task specialization across salon-quarters. Labor productivity is measured as total revenue divided by total utilized labor, expressed in units of dollars per minute.

There are also large differences in internal task specialization, measured by the s-index, among the salons in the sample. The ratio of the s-index between the 75th and 25th percentiles is 13-to-1. The distribution of the s-index, depicted in Figure 4 Panel A, roughly follows a power law, with a large number of generalized salon-quarters and a long right tail of specialized salon-quarters. The majority of this variation comes across salons, rather than within salon over time.

The majority of the variation in both productivity and the s-index is not accounted for by location, time or task composition. The standard deviation in the productivity measure residualized for the task mix, zip code fixed effects and quarter fixed effects is 70% of the raw standard deviation. The standard deviation of the s-index residualized for the task mix, zip code fixed effects and quarter fixed effects is 51% of the raw standard deviation.

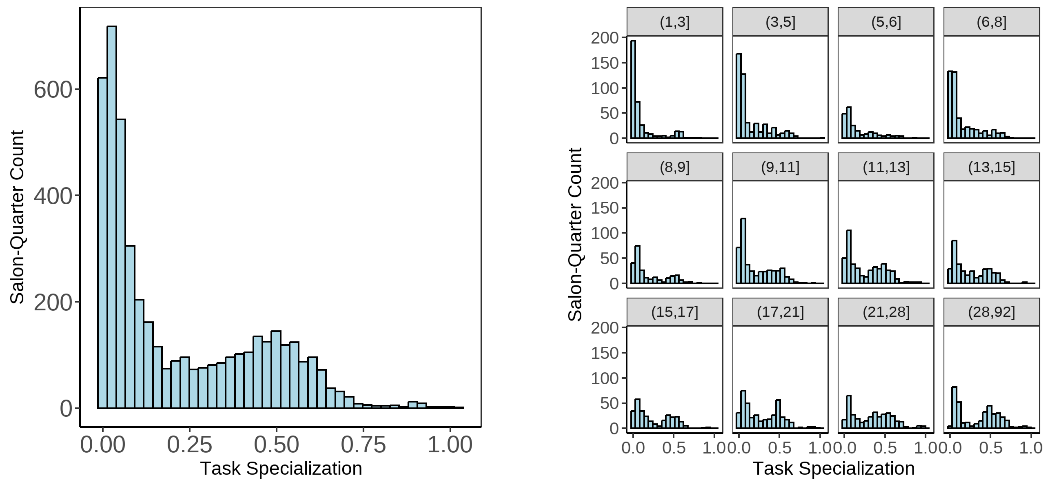
Notably, while the s-index is positively correlated with the number of employees and other measures of firm size, it varies even among firms of the same size. I demonstrate this variation in Figure 4 Panel B, which displays histograms of the s-index among salon-quarters with the same number of employees. There is significant dispersion, and the power law shape of the unconditional histogram exists among many of the firm size bins.

Such heterogeneity among similarly sized firms is difficult to reconcile with traditional models. This is because specialization differences typically arise in these models because firms anticipate being able to spread the fixed costs of specialization across a larger amount of output.

Fact 2 *Task-specialized salons are more productive than generalized salons.*

Table 4 documents this result via regressions of productivity on the s-index, while Figure 5 shows a binned scatterplot. There is a robust positive correlation between labor

Figure 4: Histogram of the Task Specialization Index (S-index)



(a) All Salon-Quarters

(b) By Number of Employees

Note: The figure displays the distribution of the s-index unconditionally and among salon-quarters with similar numbers of employees. The distribution roughly follows a power law, with a large number of generalized salons and a long right tail of specialized salons. The upper bound of the s-index depends on the specific generalist benchmark, but for most salons it is around 1.

productivity and the s-index. A one standard deviation increase in the s-index is on average associated with a 0.11 standard deviation increase in labor productivity. To relate this finding back to the productivity dispersion documented earlier, the top 25 percent most specialized salon-quarters on average generate \$1.08 or 68% more revenue per minute than the bottom 25 percent.

This correlation persists both in magnitude and statistical significance even after controlling for the task mix, zip code fixed effects, time fixed effects and firm size fixed effects (indicator variables for each number of employees). The fact that the coefficient remains similar after zip code fixed effects are included provides reassurance that differences in the customer base are not creating a spurious association between specialization and the revenue-based productivity measure.

The coefficient falls to 0.066 after controlling for interacted firm size and zip code fixed effects (2,290 indicators, see column 5 of Table 4) but remains statistically significant. This finding suggests that while demand-side differences and economies of scale may play a role, they are not the the main drivers of the productivity-specialization association. In Panel B of Figure 5, I also show that there is a positive correlation between task specialization and the productivity measure among salon-quarters with similar numbers of

Table 4: Regressions of Productivity on Task Specialization

Dependent Variable:	Revenue per Minute (standardized)					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
S-Index	0.1099*	0.1091*	0.1019*	0.0999*	0.1059*	0.0663*
	(0.0555)	(0.0549)	(0.0510)	(0.0499)	(0.0508)	(0.0332)
Color Task Mix		0.0233	-0.0506	-0.0235	-0.0063	0.0280
		(0.0619)	(0.1237)	(0.1216)	(0.1144)	(0.1196)
Blowdry Task Mix		0.1511	0.1478	0.2157	0.1933	0.2075
		(0.0884)	(0.1054)	(0.1122)	(0.1128)	(0.2333)
Admin. Task Mix		-0.1813	-0.1473	-0.1267	-0.1181	-0.0402
		(0.1247)	(0.0834)	(0.0895)	(0.0799)	(0.1049)
Nail Task Mix		0.1007	0.1729	0.1274	0.1245	0.0304
		(0.1459)	(0.2574)	(0.2394)	(0.2270)	(0.1390)
Fixed-effects						
Zip			Yes	Yes	Yes	
Quarter-Year				Yes	Yes	Yes
Firm Size					Yes	
Zip-Firm Size						Yes
Fit statistics						
R^2	0.05847	0.06368	0.51221	0.52402	0.53741	0.89597

Clustered (Establishment) standard errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Note: Each column represents a regression of the labor productivity measure on the s-index. From right to left additional controls are included. The preferred specification is column 4; however, columns 5 and 6 illustrate that the association persists even with firm size fixed effects.

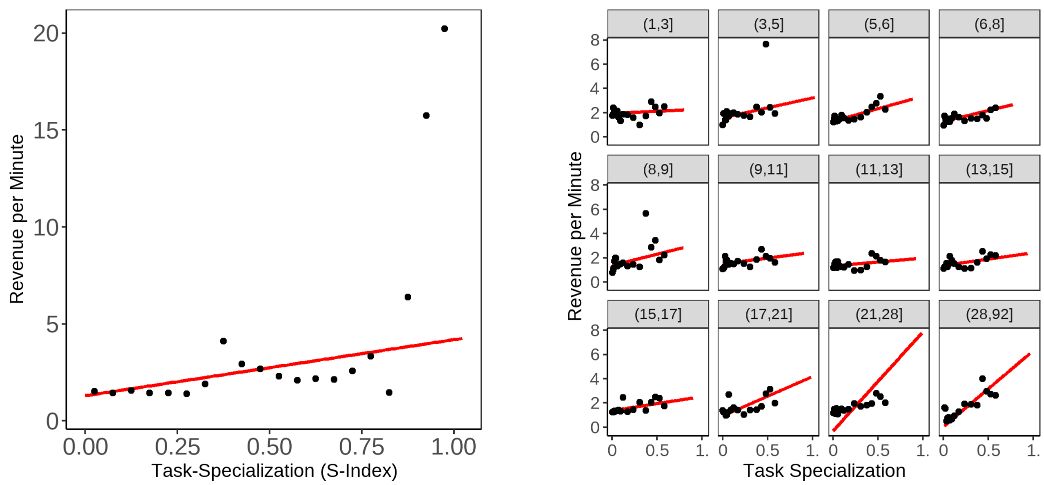
employees.

Fact 3 *Task-specialized salons earn more revenue per customer and have a higher customer return rate compared with generalized salons.*

In Table 5, I study the components of the productivity-specialization association by regressing the number of customers, the revenue per customer and the future return rate of current customers on the s-index. I interpret revenue per customer as a proxy for the price of the average bundle of services purchased at the salon. I interpret the future return rate as reflecting perceived service quality.

After accounting for the number of employees, I find no statistically significant association between the number of customers and task specialization. This finding suggests that while specialized salons may be able to serve more customers because they have

Figure 5: Binned Scatterplot of Productivity and Task Specialization



(a) All Salon-Quarters **(b)** By Number of Employees

Note: Salon-quarters are placed into 5 percentile bins based on s-index, and the average of the revenu-based labor productivity variable is displayed on the y-axis.

more employees, they are not serving more customers, holding fixed the number of employees. However, there is a positive and statistically significant association between revenue per customer and the s-index, as well as the future return rate. This finding is inconsistent with task specialization reducing marginal costs but is consistent with specialization improving service quality.[5]

I now show evidence that the s-index is related to the management practices of a salon.

Fact 4 *Task-specialized salons engage in more teamwork and are earlier adopters of software features compared with generalist salons.*

The s-index captures differences in how particular workers are assigned to tasks in a quarter. These differences include instances of teamwork, where multiple workers combine their skills to serve a single client on a single day. But it also includes cases where two workers perform different types of tasks on different days or with different customers. Although both are examples of specialization, teamwork involves more coordination and interaction, and therefore more deliberate management of workers.

To measure teamwork, I divide the number of customer visits where more than one worker is assigned by all customer visits where more than one service is performed. There

5. I show in Appendix A.15 that under multinomial logit demand and marginal cost reductions, prices and specialization are negatively related.

Table 5: Decomposing the Productivity-Specialization Association

Dependent Variables:	Customer Count	Rev. per Customer	Customer Return Rate
Model:	(1)	(2)	(3)
Variables			
S-Index	-0.0297	0.0320*	0.0864***
	(0.0347)	(0.0124)	(0.0245)
Color Task Mix	-1.280***	0.3936***	0.2372
	(0.2395)	(0.0588)	(0.1505)
Blowdry Task Mix	-1.157***	0.3274*	-0.5652*
	(0.2097)	(0.1408)	(0.2278)
Admin. Task Mix	-0.2951	0.1451*	-0.2547
	(0.4493)	(0.0600)	(0.1826)
Nail Task Mix	0.2979	0.0192	-0.2242
	(0.3236)	(0.0479)	(0.1537)
Fixed-effects			
Quarter-Year	Yes	Yes	Yes
Zip	Yes	Yes	Yes
Firm Size	Yes	Yes	Yes
Fit statistics			
R^2	0.77213	0.65443	0.73598

Clustered (Establishment) standard errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Note: This table examines the relationship between task specialization and the productivity measure (revenue per minute). The covariate of interest and the dependent variable are standardized by their respective standard deviations. The return rate of customers is measured as the fraction of customers that are observed returning in a future quarter.

is a strong positive correlation of 0.71 between teamwork and the s-index, as Appendix Figure A5 shows.

Another way to understand whether the s-index is capturing underlying differences in management is to see whether it is correlated with how firms engage with the software. I compute the time in days it takes a salon to begin using a feature relative to the first salon observed using the feature. The four features I consider are general adoption of the software, tipping, prebooking and requesting staff. I then regress these measures of time to adoption on the s-index. As shown in columns 3–6 in Table 6, on average, task-specialized salons adopted all features earlier than generalized salons, suggesting task-specialized salons are more sophisticated users of the software.

Table 6: Regressions of Management Software Engagement on the S-Index

Dependent Variables:	Teamwork	Service Descriptions	Product Discounts	Software Adopted	Tip Feature	Prebook Feature	Request Feature
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variables							
S-Index	0.6551***	0.1167*	0.1107*	-0.2100***	-0.3066***	-0.2790***	-0.0802*
	(0.0492)	(0.0509)	(0.0461)	(0.0476)	(0.0551)	(0.0482)	(0.0397)
Color Task Mix	-1.199***	1.195***	0.2914				
	(0.2166)	(0.1949)	(0.2164)				
Blowdry Task Mix	-0.1380	0.0532	-0.6734				
	(0.4647)	(0.2388)	(0.3712)				
Admin. Task Mix	-0.6919	0.4318	-0.1898				
	(0.4525)	(0.3541)	(0.2548)				
Nail Task Mix	-0.8357**	0.5059*	0.3665*				
	(0.2791)	(0.2421)	(0.1638)				
Fixed-effects							
Zip	Yes	Yes	Yes				
Quarter-Year	Yes	Yes	Yes				
Fit statistics							
R^2	0.78858	0.74935	0.78898	0.04410	0.08819	0.07965	0.00654

Clustered (Establishment) standard errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Note: Each column represents a regression of a variable related to management practices on the s-index. The variables related to adoption are conducted at the establishment level (not salon-quarter) and therefore do not include fixed effects or other controls.

The salons with a high s-index adopt software features earlier, use software features more intensely, enter more unique text descriptions of services into the software (column 2 of Table 6), and offer more unique discounts of physical products (column 3 of Table 6).[6]

4 Theory

In this section, I first design a model where firms with product market power choose labor demand, task assignments and prices. Second, I prove that equilibrium outcomes in this model are both tractable and consistent with the stylized facts presented earlier. Third, I show that despite its analytical tractability, the model allows for endogenous organizational complementarities between workers. As a result, firms in the same market can vary greatly in the types of workers they hire and the way they substitute across workers in response to wage shocks.

6. Because the data on physical products are limited, most of the paper focuses on the sale of services.

4.1 Model

For ease of exposition, the model is specified for a single market and a single period, with the additional subscripts kept implicit. There are three important groups of objects in the economy: firms, indexed by $j = 1, ..., J$; *individual* workers, indexed by $m = 1, ..., M$; and task types, indexed by $k = 1, ..., K$.

Firms and Tasks. The J firms differ in their organization cost $\gamma_j \in \mathbb{R}_+$, discussed below. Each firm produces a single differentiated product, also indexed by j. Producing a single unit of good j requires $a_j \in \mathbb{R}_+$ units of labor. The fraction of total labor that must be assigned to each type of task is called the task mix ($\alpha_j \in \mathbb{R}_+^K$). In addition to incurring an organization cost, a firm producing good j incurs a per unit cost $\alpha_j \cdot c + \omega_j$, where ω_j captures Hicks-neutral productivity differences across firms, and $\alpha_j \cdot c$ captures variable material costs.

Workers. The M workers are each described by an inelastic labor supply $l_m \in \mathbb{R}_+$ and skills, which I decomposed into a skill level $\bar{\theta}_m \in \mathbb{R}$ and a skill set vector $\theta_m \in \mathbb{R}^K$. Worker m performs task k with quality $\bar{\theta}_m + \theta_m(k)$. The skill level $\bar{\theta}$ captures vertical skills, while the skill set vector captures horizontal differences in worker skill portfolios. There are $N \in \mathbb{N}_+$ unique skill sets, which are indexed by i.[7] Worker characteristics are common knowledge to all actors in the model.

Firm Strategies. Each firm simultaneously chooses the price of its product $p_j \in \mathbb{R}_+$, a relative labor demand for each worker $E_j \in \mathbb{R}_+^M$, and a task assignment for each worker $b_j \in \mathbb{R}_+^{M \times K}$. The sum of a firm's relative labor demand across all workers is 1: $\sum_m^M E_j(m) = 1$. The sum of a firm-worker task assignment across all task types is 1: $\sum_k^K b_j(i,k) = 1$.[8]

A task assignment specifies how much labor the firm assigns to each worker-task pair. Therefore, the sum of a worker's task assignments across all tasks and all firms implies a labor demand for that worker. In equilibrium, I require the prices and task assignments of all firms be such that each firm can produce the amount of its product demanded by consumers given its task-based production technology.

Organization Costs. The firm must communicate each worker's task assignment. Prior to communication, each worker knows only their firm's overall task-based product

7. The total labor available from skill set i is denoted by L_i.

8. The firm's product price controls the overall amount of labor demanded, while relative labor demands and task assignments control the division of labor within the firm.

function. Therefore, the amount of information that must be communicated per unit of labor assigned to worker m is given by the Kullback–Leibler divergence between the task assignment of that worker $\{b_j(m,k)\}_{k=1}^K$ and the task mix $\{\alpha_j(k)\}_{k=1}^K$.[9] The total communication is this quantity multiplied by the total labor assigned summed across all workers at the firm. The cost of communication is borne by the firm. Specifically, γ_j is the cost for firm j to communicate 1 unit of information.

Labor Market. The labor market is competitive, with worker-specific wages $w \in \mathbb{R}_+^M$ per unit of labor. The labor market for worker m clears if total labor demanded from worker m across all firms is equal to worker m's labor supply (l_m).

Consumers and Demand. There is a mass of consumers interested in purchasing at most one of the J products. Consumers observe firm task assignments and prices prior to purchase. Consumer z's utility for good j is represented by the logit utility function

$$u_{z,j} = \xi_j + \nu_j - \rho p_j + \epsilon_{z,j},$$

where ξ_j is the average quality of the tasks performed to produce product j given the firm's task-assignment strategy, ν_j captures the components of quality that cannot be controlled by the firm, $\epsilon_{z,j}$ captures idiosyncratic consumer preferences over products, and ρ captures consumer price sensitivity. I assume $\epsilon_{z,j}$ is distributed i.i.d. Type 1 extreme value across consumers and products.[10] The outside option for consumers is assigned index $j = 0$, and its utility is normalized to $u_{z,0} = \epsilon_{z,0}$.

Equilibrium. The equilibrium concept I use consists of two conditions. First, firm labor demand, task assignments and pricing strategies must form best responses to all other firm strategies at wage vector w. Second, firm labor demands at wage vector w must clear the labor market for each of the M workers. Note that for any fixed wage vector w, the model is a well-defined game, and the first part of the equilibrium definition amounts to a Nash equilibrium. For any fixed wage w, I refer to this game as the fixed-wage subgame.

9. Throughout, the natural logarithm is used so that information is measured in natural units of information (nats). One nat is equal to around 1.44 bits.

10. Most theoretical results also extend to other demand systems, including multinomial logit, nested logit and mixed logit with a non-random price coefficient.

4.2 Model Comments

Workers in the model differ in their labor supply, skill level and skill set. In Section 4.3, I show that the dimension of heterogeneity which matters for most equilibrium outcomes is the skill set. The other dimensions are included so that the model can be mapped to the data. The skill set can be thought of in a general sense (cognitive skill vs. manual) or a very occupation specific sense (coloring or cutting hair). The empirical application in this paper is the later, however the underlying model is meant to be a general way to model specialization costs in equilibrium settings.

Firms in the model differ in their required labor, task mix, Hicks-neutral productivity, exogenous product quality and organization cost. In Section 4.3, I show that the task mix and organization costs will determine firm task assignments for any fixed wages. As a result, the other dimensions of firm heterogeneity will impact task assignments (both own and competitors') indirectly via wages. The role of the other dimensions is to allow for firm differences in prices and market share that are not related to internal task assignments.

Notably, most elements of a firm's costs are scaled by required labor (a_j), so most costs are per unit of labor, and labor requirements are allowed to vary across firms. This is in line with the choice to model task assignment per unit of time rather than per task. Variation in required labor allows firms to differ in their quantity-based productivity as well as their quality-based productivity.

The model specifies that task assignments must be communicated to each worker. The firm bares the cost of communication,[11] and firms differ in their cost of communication. These differences can reflect either greater management efficiency or a lower opportunity cost of time for the firm owner. This specification provides a micro-foundation for why the s-index is costly and varies significantly across firms. It is equally valid to discard the micro-foundation and directly view the organization cost of a firm as capturing a reduced-form specialization cost.

The model allows for strategic product market power (oligopoly) but requires firms to take wages as given. In the model's application to hair salons, firms will be small and thus close to monopolistic rather than oligopolistic. However, allowing for strategic market

11. This is a common assumption in models of within-firm communication (Garicano 2000).

power means the model can be used for merger analysis. Given that salons are small employers, the main threat to the wage-taking assumption is firm-specific job amenities that give firms monopsonistic (rather than oligopsonistic) labor market power.

The assumption that labor markets clear is not necessary for estimation. However, if one allows for unemployment and ants to conduct counterfactuals, it will be necessary to identify the skills of the unemployed. This task is challenging, and the identification strategy in this paper would not apply because by definition unemployed workers do not have current task assignment information. The challenge could be resolved with data about worker skills (schooling, licenses, etc.) or information about past task assignments combined with a skill depreciation assumption.

4.3 Theoretical Results

This section analyzes equilibrium task assignments. I derive model properties which I later use to prove identification and construct an estimation procedure.

In the model, firms demand labor from specific workers and assign specific workers to tasks. This allows the model to be directly mapped to the empirical application, where rich information on how specific workers use their time is available, but direct information about worker skills is not. The cost of this realism is that whenever there are many workers in a labor market (which is almost always the case), task-assignment strategies are high-dimensional objects. The firm must choose a portfolio of workers and assign tasks given this portfolio. The first step in the theoretical analysis is to understand how firms make this choice. The key insight is that while firms can in principle make task assignments depend on all three dimensions of worker heterogeneity (i.e., labor supply, skill level and skill set), in equilibrium, task assignments at a given firm are the same for all workers with the same skill set.

Proposition 1 *In any equilibrium task-assignment strategy, all workers at the same firm with the same skill set are assigned the same distribution of time across tasks.*

The proof is provided in Appendix Section A.6; however, the intuition is straightforward. The firm can engage in a very complicated division of labor by specifying different task assignments along all the dimensions of worker heterogeneity. However, among workers with the same skill set, tailoring task assignments holding fixed labor demand has no benefit but is costly because it requires more communication. This proposition also has implications for equilibrium wages.

Proposition 2 *In any equilibrium, all workers with the same skill set and skill level are paid the same wage per unit of labor, and wages are such that firms are indifferent between all workers with the same skill set but different skill level.*

The proof is provided in Appendix Section A.7. Workers with the same skill set but different skill level are ranked by absolute advantage. The proposition establishes that workers extract the value of their absolute advantage via wages in a way that perfectly offsets the productivity improvements. Specifically, for any two workers m, m' with the same skill set but different skill levels, wages are such that $w_m - w_{m'} = \rho^{-1}(\bar{\theta}_m - \bar{\theta}_{m'})$. As a result, firms are indifferent between all workers with the same skill set. Firms that happen to employ workers with a high skill level will produce higher-quality products but will pay out the revenue from quality improvements to workers. In this way, the model is able to accommodate vertical skill differences without loss of tractability. This is helpful in settings where there are wage differentials among workers with the same skill set.[12]

As these propositions make clear, there are many equilibria, but the actors in the model (i.e., consumers, firms, workers) are indifferent between them.[13] Further, all equilibria imply the same task assignments and the same labor productivity. For these reasons, and with some abuse of notation, I recast all firm strategies in terms of worker skill sets (rather than worker identities), which are indexed by i. Thus, $E_j(i)$ is the relative labor demand for skill set i at firm j, and $\{b_j(i,k)\}_{k=1}^{K}$ is the task assignment of skill set i at firm j.

I analyze firms as assigning tasks to a representative worker of each skill set, with the understanding that these tasks will be executed in equilibrium by potentially multiple actual workers with different labor supplies and skill levels. I recast the wage vector to be length N, with one wage for each skill set, with the understanding that this wage reflects both the wage of that skill set and the average

12. With wage data, the distribution of skill levels within skill sets could be estimated.
13. This is because the difference across these equilibria is which individual worker is employed at which individual firm, and the model does not include workplace amenities.

Given the Type-1 extreme value distribution of consumer taste shocks (McFadden 1973), I can write firm j's profit-maximization problem as

$$\max_{p_j, b_j, E_j} \frac{exp(\xi(b_j, E_j) - \rho p_j))}{1 + \xi(b_{j'}, E_{j'}) - \rho p_{j'})} \left(p_j - a_j \gamma_j \sum_i D_{KL}(b_j(i, \cdot) || \alpha_j) - a_j \sum_i w_i E_j(i) - \alpha_j c - \omega_j \right) \quad (1)$$

s.t.$\sum_{i,k} E_j(i) b_j(i,k) = \alpha_j(k) \forall k$.

Though pricing and internal organization may appear intertwined, the problems naturally separate in the following way.

Proposition 3 *A relative labor demand and task assignment are profit-maximizing if and only if they solve*

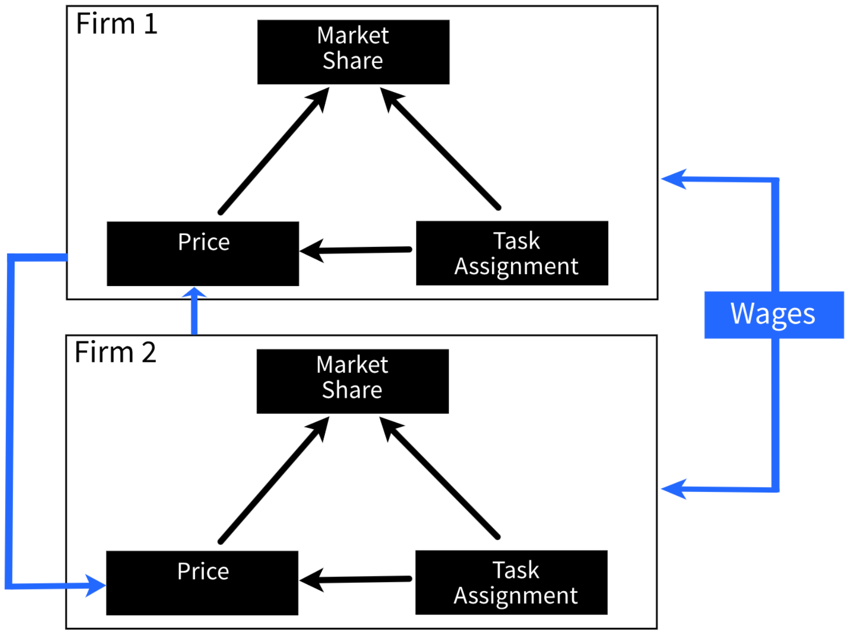
$$\min_{b_j, E_j} \gamma_j \sum_i E_j(i) D_{KL}(b_j(i,k) || \alpha_j) + \sum_i E_j(i) w_i - \rho^{-1} \sum_i E_j(i) \sum_k \theta_i(k) b_j(i,k) \quad (2)$$

s.t. $\sum_{i,k} E_j(i) b_j(i,k) = \alpha_j(k) \forall k$.

The proof is provided in Appendix Section A.8. The result implies that while task assignments impact prices, prices do not impact task assignments. Further, task assignments are impacted by competitors' choices only via wages. Prices, however, are fully strategic and depend only on own task assignments, competitor task assignments and competitor prices. These relationships are summarized in Figure 6. This separation makes the model computationally tractable, and it also allows the researcher to shut down internal organization and have a benchmark to understand how firms would behave if they could not reorganize.

I now relate the task-specialization index (s-index) to the model. Consider a firm j with an observed task-assignment matrix B_j, where task assignments are with respect to worker identities indexed by m rather than worker skill sets indexed by i. Denote the skill set of worker m by t_m. Because workers with the same skill set are assigned the same

Figure 6: Strategic Interactions in the Model



distribution of tasks, $\frac{B_j(m,k)}{\sum_{k'} B_j(m,k')} = b_j(t_m, k)$. The s-index can then be expressed as

$$
\begin{aligned}
I(B_j) &= \sum_{m,k} B_j(m,k) log\left(\frac{B_j(m,k)}{B_j^G(m,k)}\right) && (3)\\
&= \sum_{m,k} B_j(m,k) log\left(\frac{B_j(m,k)}{[\sum_{k'} B_j(m,k')]\alpha_j(k)}\right) && (4)\\
&= \sum_{m,k} \left[\sum_{k'} B_j(m,k')\right] b_j(t_m,k) log\left(\frac{b_j(t_m,k)}{\alpha_j(k)}\right) && (5)\\
&= \sum_{m} \left[\sum_{k'} B_j(m,k')\right] \sum_k b_j(t_m,k) log\left(\frac{b_j(t_m,k)}{\alpha_j(k)}\right) && (6)\\
&= \sum_{m} \left[\sum_{k'} B_j(m,k')\right] D_{KL}(b_j(t_m,k)||\alpha_j) && (7)\\
&= \sum_{i} E_j(i) D_{KL}(b_j(i,k)||\alpha_j). && (8)
\end{aligned}
$$

This final expression is the total amount of communication that must occur within the firm per unit of labor. Thus, the observed s-index over worker identities is exactly equal to the amount of communication that occurs within the firm. Thus, the s-index is the appropriate way to measure specialization within the firm under the model. This fact, combined with Proposition 3, provides a micro-foundation for the observed heterogeneity in the s-index.

Proposition 4 *The communication required to implement a profit-maximizing task assignment is equal to the **observed** s-index (I_j) and is strictly decreasing in γ_j for all values of firm-level heterogeneity ($a_j, \alpha_j, \nu_j, \omega_j$) until it reaches 0.*

The first part of the proof was provided earlier, and the second part follows from applying the envelope theorem to Equation 2.[14] Unobserved differences in the cost of communication across firms (γ_j) give rise to observed differences in task specialization across firms in equilibrium. This variation is not driven by firm size. Because firms can differ in their exogenous quality ν_j or exogenous marginal cost shifter ω_j, firms of the same size can have different levels of specialization.

Importantly, the s-index can be inverted in equilibrium to recover firm-level organization costs. This is similar in spirit to how Garicano and Hubbard (2016) invert the span of control to recover a manager's skill. This property plays an important role in the identification proof and estimation strategy. A corollary of this result is that the model replicates the positive correlation between the s-index and firm productivity documented in Section 3.

Corollary 0.1 *All else constant, firms with a lower organization cost (γ_j) have a higher s-index, a higher market share and higher profits.*

The proof is provided in Appendix Section A.9. Recall that γ_j represents the management technology, relationships, knowledge and practices specific to the firm which determine the cost of communicating task assignments to workers. Corollary 0.1 implies more organizationally efficient firms are larger and more profitable, and can produce better-quality goods at a lower cost. This is in line with the findings of Kuhn et al. (2023), who use surveys and administrative data to show that more coordinated or specialized firms are more profitable. In the model, firms can be large or small for reasons unrelated to organizational costs.[15] However, there are no synergies between firms expecting to be large

14. The formal proof is in Appendix Section A.9.
15. Due to heterogeneity in marginal cost (ω_j) or quality (ν_j).

and task specialization.[16]

Theorem 1 *Profit-maximizing task assignments for any worker with skill set i at firm j can be expressed as*

$$b_j(i,k) = \alpha_j(k) \frac{exp[\gamma_j^{-1}(\rho^{-1}\theta_i(k) - w(i))]}{\sum_{i'} E_j(i') exp[\gamma^{-1}(\rho^{-1}\theta_{i'}(k) - w(i')]}, \tag{9}$$

and they satisfy the following properties:

1. ***Relative Law of Demand:*** *As $w(i)$ increases, skill set i's share of labor at firm j ($E_j(i)$) decreases.*

2. ***Incomplete Specialization:*** *All workers employed by firm j ($E_j(i) > 0$) spend a strictly positive amount of time on all tasks performed at the firm ($\{k|\alpha_j(k) > 0\}$).*

3. ***Maximum Coworker Diversity:*** *Either the number of skill sets employed at a firm is less than or equal to the number of tasks, or there exists another profit-maximizing task assignment strategy where this is true.*

The proof, provided in Appendix Section A.10, involves manipulating the first-order conditions of the firm and relying on the equivalence between Equation 2 and a class of problems that are well-studied in computer science and information economics. Equation 9 reveals that the profit-maximizing task assignment balances the tasks that need to be done (α_j), the firm's organizational costs (γ_j), wages (w_i) and skill sets (θ_i).

The result regarding incomplete specialization is obtained because of the functional form of the organization cost, specifically, the Kullback–Leibler divergence. Because many of the workers in my context are in the same occupation (cosmetology) and perfect specialization at the firm level is uncommon, it is reasonable to assume incomplete specialization here. However, in other contexts (such as manufacturing which has assembly lines) it is not. In these cases, there are alternative cost functions that can be used which allow for complete specialization of some workers.

Because the relative labor demand of each skill set ($E_j(i)$) is endogenous, the expression given for optimal jobs is not a closed-form solution. Despite this, I will show that

16. This is a major difference between this paper and Adenbaum (2022), where firms specialize more precisely because they expect to be able to spread fixed costs.

the expression will enable identification and greatly simplify computation of equilibria. The last two properties in the theorem motivate a natural restriction on the skill sets in the economy.

Assumption 1 *There are $N = K$ skill sets that can be collected into a matrix Θ, which is positive definite.*[17][18]

Some aspects of this assumption are without loss. In particular, adding a new skill set that vertically dominates an existing skill set in the sense of Proposition 2 has no impact on either firm profit or consumer welfare or task assignments. Additionally, Theorem 1 proves that if there are more skill sets than tasks ($N > K$), every firm has a profit-maximizing task-assignment strategy that employs weakly fewer than K skill sets at a time. However, different firms may utilize different subsets of worker types, and in this sense I am restricting the space of workforce compositions.[19] This assumption is used to constructively identify the model later, and to derive the following uniqueness result.

Proposition 5 *There exists a unique Nash equilibrium in prices (p_j), task assignments (b_j) and relative labor demands (E_j) for every fixed-w subgame with strictly positive wages.*

When strategies are formulated in terms of worker skill sets, the issues with uniqueness are resolved. The proof is provided in Appendix Section A.11. I utilize a uniqueness result in the rational inattention literature (Matêjka and McKay 2015), Nash equilibrium uniqueness of Bertrand pricing games with multinomial logit demand (Caplin and Nalebuff 1991) and the Schur product theorem. The proposition means that equilibria are identified by wage vectors. Put another way, if wages are known, all firm strategies are uniquely determined. This property will be useful for the empirical application, where wages are parameters that are estimated. However, it is important to note that this property does not establish equilibrium uniqueness of the full model, as more than one wage vector may clear the market.

17. The worker and task types can be reordered because the original index was arbitrary. Additionally, the same constant vector can be added to each skill set to make all entries weakly positive without changing the theory.

18. I do not require symmetry. One way to check for positive definiteness of a non-symmetric matrix is to check that all eigenvalues of $(\Theta + t(\Theta))/2$ are positive.

19. Still, it is a weaker assumption than those in the literature, where it has been typically assumed that tasks can be ordered in a single dimension and that worker skills satisfy a supermodularity condition in that single dimension.

4.4 A Simple Example

The model assumes that workers are perfect substitutes in production, both in terms of quantity and quality. To see this, notice that absent organization costs ($\gamma_j = 0$), the firm minimizes a constrained linear objective with payoffs determined by wages and skill sets. The solution of such problems always lies at the extreme points of the set, that is assign all of a task to a single worker skill set.

I illustrate this with a simple version of the model with three worker skill sets. Suppose wages are fixed at $w = (20, 15, 21)$ and the price sensitivity is $\rho = 1$, and skill sets are given by:[20]

$$\begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} = \begin{bmatrix} 23 & 19 & 15 \\ 15 & 15 & 15 \\ 15 & 19 & 26 \end{bmatrix}$$

Proposition 3 implies that firms consider quality-adjusted wages (or wage-adjusted skills), which can be stacked into a matrix:

$$\begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix} - \rho w = \begin{bmatrix} 3 & -1 & -5 \\ 0 & 0 & 0 \\ -6 & -2 & 5 \end{bmatrix}$$

A firm without any organization costs ($\gamma_j = 0$) will simply choose the best worker for each task at prevailing wages, which corresponds to the row of the cell with the highest number in each column. In this case, they will assign all of task 1 to skill set 1, all of task 2 to skill set 2, and all of task 3 to skill set 3.

Firms with organization costs will make choices that depend on the task mix. A firm with a task mix of $(1/3, 1/3, 1/3)$ and a high enough organization cost ($\gamma_j > 5$) will choose a generalist structure, and assign all 3 tasks to skill set 2. In equilibrium, these choices will also change firm pricing strategies, which will impact labor demands and feed back into wages.

20. These parameter values are based on an example in Csaba (2021).

5 Empirical Application

This section proves point identification of all parameters of the model, including wages. Because the proof is constructive and the actual estimation procedure follows the most steps, it is presented in full. The actual estimation procedure is then outlined, with special attention paid to modifications made to improve power given the limited sample size. The section concludes with the estimation results.

5.1 Identification

For identification, I assume that the econometrician observes the following data. For workers, the econometrician observes only the task assignment distribution of each worker ($\{b_m(i,k)\}_{m=1}^{M}$).[21] For each firm, the econometrician observes required labor, task mix, price and market share $\{(a_j, \alpha_j, p_j, s_j)\}_{j=1}^{J}$. I consider wages, skill set parameters Θ, and the type of each firm (in terms of γ_j) and worker (in terms of θ_m) as parameters to be estimated — not data.

I collect all parameters of the model into three groups. First, there are market parameters, denoted by Ω: wages, material costs, consumer price sensitivity and skill set vectors. Second, there is the skill set group membership of each worker (i.e., which of the N skill sets they possess). Third, there are the organization cost parameters of all firms ($\{\gamma_j\}_{j=1}^{J}$). I make two assumptions. First, the wage-adjusted skill matrix $\Theta - \rho(we')$ is full rank. Conditional on Θ being positive definite (a maintained assumption), this assumption requires ruling out a measure 0 set of wages where $\rho(\Theta^{-1}w) \cdot e = 1$. Second, idiosyncratic product quality (ν_j) and cost shocks (ω_j) are mean zero and independent of organization cost and task mixtures.

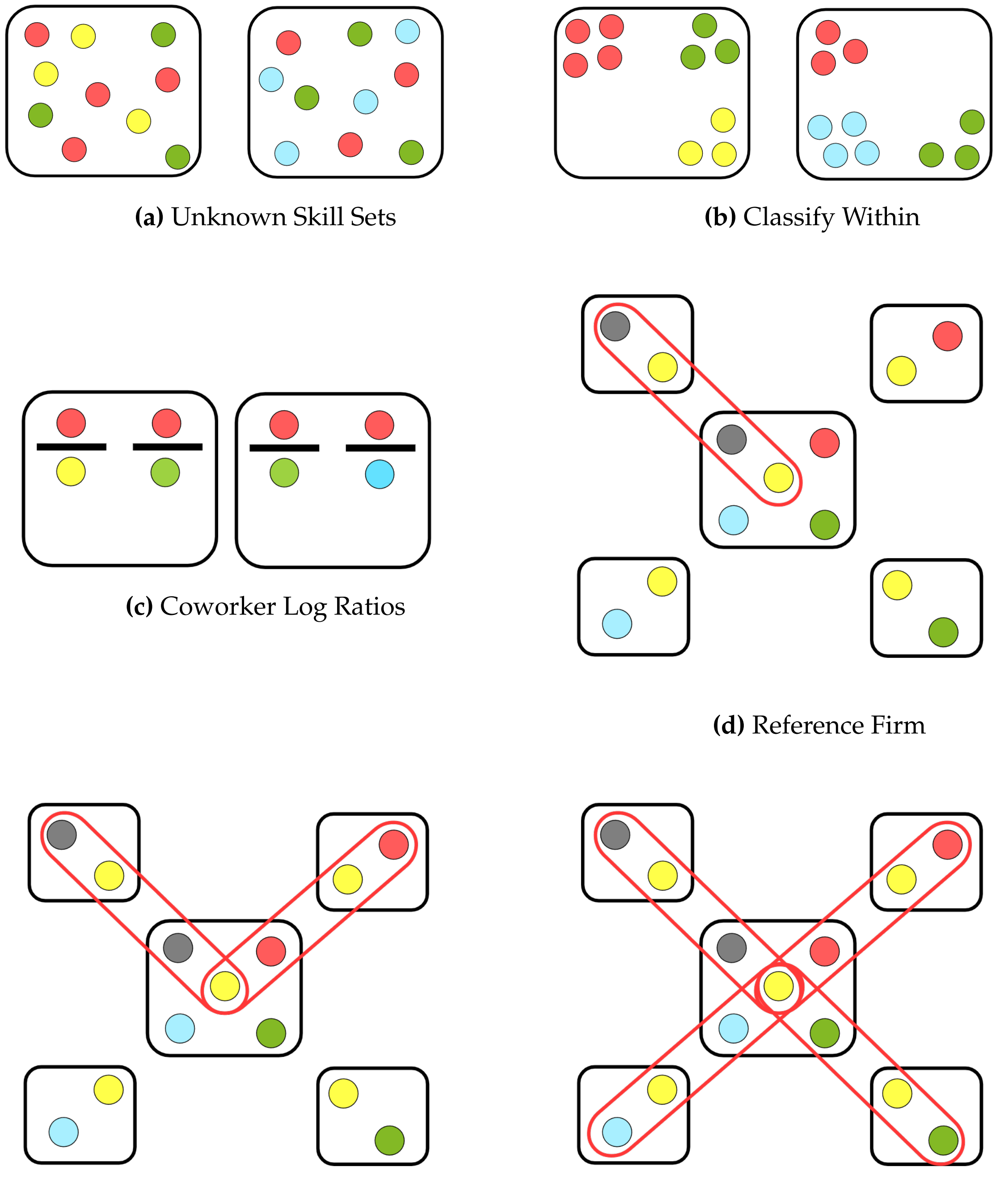
Theorem 2 *The market parameters (Ω) and the amount of labor of each skill set are identified. The organization cost parameters (γ_j) and the skill sets of all workers ($\{\theta_m\}_{m=1}^{M}$) at firms with a strictly positive s-index ($I_j > 0$) are identified.*[22]

The formal proof of the theorem is provided in Appendix Section A.12. The intuition for the proof is provided in Figure 7. The idea is that workers at different firms can

21. Importantly, the econometrician cannot access demographic or wage data about individual workers.
22. Lower bounds on the organization cost parameters of firms with an s-index of 0 are identified.

be classified into skill sets by comparing their task assignments to those of all of their coworkers. This allows the researcher to purge firm-specific task mixtures and organization costs, and recover the worker's skills. This approach can be used to classify workers among a set of firms with sufficient overlap in their workforces, and it can also be used to recover organization costs.

Figure 7: Identification of Worker Skill Sets



(a) Unknown Skill Sets **(b)** Classify Within

(c) Coworker Log Ratios

(d) Reference Firm

(e) Indirect Linkage

(f) Classified Across

Note: A graphical representation of how worker skill sets are identified via task assignments. Balls represent workers, and colors their skill sets. Boxes are firms.

Theorem 2 shows that the equilibrium wage vector is identified. Since multiple equilibria, if they exist, are distinguished by different wage vectors, the data identify which of the multiple equilibria is being played. Additionally, the identification proof is constructive and suggests the following estimation procedure:

1. Set aside firms that do not perform all tasks and firms with an s-index of 0.
2. Classify workers into groups with mutually exclusive skill sets within firm based on the similarity of task assignments.
3. Classify workers across firms by comparing the log ratio of their task assignments to that of coworkers of different skill sets.
4. Regress log relative market shares on prices and the task assignments, instrumenting for price using relative organization costs.
5. Obtain marginal costs by adjusting prices by the markup. Regress marginal costs on the relative labor demand of the worker skill sets and relative organization costs.
6. For the set-aside firms with a positive s-index, invert the s-index to obtain γ_j. For those with a zero s-index, invert the s-index to obtain the minimum threshold $\bar{\gamma}_j$.

This estimation procedure is computationally light. The steps are not nested and do not need to be repeated, because the implied classification procedure is hierarchical and deterministic and the regression coefficients all have a closed form. I provide a globally convergent contraction mapping that can be used to invert the s-index in the final step; this operation needs to be done only once after all market parameters are estimated.

5.2 Estimation

The model is estimated for three counties (Manhattan, NY, Cook County, IL, and Los Angeles, CA) and 12 quarters (2018Q1 through 2021Q2 excluding 2020Q2 and 2020Q3, which were impacted by the COVID-19 pandemic). This subset represents 997 data points of the full sample of 4,599 salon-quarters. The summary statistics in Table 7 reveal that this subset is positively selected on average relative to the rest of the data, with slightly more employees, revenue and task specialization.

Table 7: Estimation Sample Summary Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Revenue	997	313,110.80	364,023.00	81.00	2,559,703.00
Employees	997	16.68	13.45	1	76
Customers	997	1,296.46	1,183.14	2	7,420
S-Index	997	0.25	0.22	0.00	0.92
Share Haircut/Shave	997	0.40	0.23	0.00	1.00
Share Color/Highlight/Wash	997	0.40	0.20	0.00	0.93
Share Blowdry/Style/Treatment/Extensions	997	0.11	0.14	0.00	1.00
Share Admininstrative	997	0.05	0.15	0.00	1.00
Share Nail/Spa/Eye/Misc.	997	0.03	0.10	0.00	0.79

Note: Summary statistics for the subset of data used to estimate the structural model.

Model estimation requires taking a stand on the size of the potential market. I assume that the potential market is the resident population for each county-year obtained from the U.S. Census Bureau 2010 and U.S. Census Bureau 2024. I divide the number of unique customers at a salon-quarter by the number representing the potential market to obtain market shares. I assume the fraction of consumers choosing the outside option to be the fraction of consumers from the associated county-quarter who do not spend any money on salons, according to the Consumer Expenditure Survey (U.S. Bureau of Labor Statistics,U.S. Bureau of Labor Statistics).

I make an assumption on wages and exploit the panel nature of the data to improve power. I assume the 25 skill set parameters and price sensitivity are fixed across time within a county. Additionally, I include county-quarter fixed effects in both the marginal cost and quality equations. I also assume that the wages of each worker skill set evolve in parallel across time in each county. The number of wage parameters that need to be estimated is thus reduced. Because wages move in parallel, I can also group workers into skill sets across quarters as well as across firms. Moreover, I restrict wages to be at least the minimum wage for small employers in that county and quarter. I allow firms' organization costs (γ_j), task mix (α_j) and required labor ($\bar{a}_j$) to vary across quarters. Finally, I assume that the material cost parameters vary across quarters but are the same across counties.

The estimation procedure, detailed in Appendix Section A.4, follows the spirit of the identification proof except for a few key differences. First, I instrument for prices in the demand equation using a variable constructed from the producer price index of hair dye interacted with the share of labor assigned to the hair dye task. With the estimated skill matrix in hand, I then estimate wages in each market to match the average share of labor demanded from each worker type, inverting the s-index to recover each firm's organization cost parameter with each guess of wages. Finally, the other parameters in the pricing equation (material costs, wage level, etc.) are estimated as described in the proof, with the exception that wage levels are constrained so that implied wages of all workers are at least the minimum wage in that county-quarter.

This modified procedure improves power while also resolving issues with internal consistency. For example, this procedure imposes that the wages are such that the worker type shares obtained via the classification step are consistent with the skill set matrix. The procedure still brings significant computational savings, as the worker skill set matrix is obtained in closed form in a first step. The way in which the skill set matrix is estimated remains similar to how it is identified in the proof. Additionally, in Appendix Section A.5 I provide a full solution contraction mapping which further simplifies and speeds up the process of inverting the s-index.

Standard errors are obtained by the Bayesian bootstrap method described by Rubin (1981). This procedure re-weights the data instead of resampling them, which is important for obtaining standard errors for many quarter-county fixed effects. Weights are drawn at the establishment level to account for serial correlation. This procedure has some benefits and limitations relative to alternatives, which I discuss at the end of Appendix Section A.4.

6 Model Estimates

This section reports the results of estimating the model and provides tests of model fit and validation. It concludes by illustrating the partial equilibrium implications of the model estimates for internal organization.

6.1 Parameter Estimates

I begin by reporting the price sensitivity parameters (ρ) in Table 8, and the wages and skill set parameters (w, θ) in Tables 9, 10 and 11.

Table 8: Price Sensitivity Parameters

County	Price Sensitivity
Cook	0.027
	(.010)
Los Angeles	0.016
	(.004)
New York	0.018
	(.014)

Note: The price sensitivity parameters for the three counties analyzed in this paper. Standard errors are from 500 bootstrap replications.

Table 9: Worker Wages and Skill Parameters for Cook County Salons

Worker Skill Set	Wage	Administrative	Blowdry/Style/Etc.	Color/Highlight/Wash	Haircut/Shave	Nail/Misc.
1	-	-0.993	12.340	-0.421	0.955	-37.562
1	-	(6.184)	(8.431)	(1.434)	(3.241)	(14.675)
2	-127.629	-0.372	10.695	-5.088	1.100	56.239
2	(54.518)	(1.592)	(7.919)	(2.370)	(2.556)	(32.155)
3	-80.328	-1.533	33.242	-2.516	0.721	-1.909
3	(60.416)	(1.118)	(28.023)	(2.395)	(2.531)	(4.621)
4	537.723	-1.186	-14.376	14.264	-5.015	-9.803
4	(183.391)	(2.457)	(44.545)	(25.922)	(8.842)	(3.439)
5	-122.678	6.755	9.516	-4.148	0.751	-4.197
5	(55.150)	(4.877)	(5.038)	(4.186)	(3.327)	(4.548)

Note: The parameters associated with the skill sets. Standard errors are from 631 bootstrap replications.

Note that worker skill sets are numbered arbitrarily. It is exactly the skill parameters in these tables which make each skill set distinct. For example, Skill Set 2 in New York County is a high-wage color–blow-dry/style specialist because they have the highest value in the columns corresponding to those tasks. Wages are relative to Skill Set 1. For an extended discussion of how to interpret wages in this model given heterogeneity in skill levels, see Appendix Section A.3.

The material costs, demand levels, cost levels and wage levels parameters across all market-quarters are presented in Appendix Table A1. For an extended discussion of how to interpret the material cost coefficients, see Appendix Section A.3.

The organization cost parameters (γ_j) for each salon-quarter are visualized in Figure 8. To provide a sense of magnitude, I can divide each salon-quarter's organization cost ($a_j \gamma_j I(B_j)$) by the observed price. The interquartile range of this object is $(0.06, 0.16)$, implying organization costs account for a sizable fraction of the observed price.

Table 10: Worker Wages and Skill Parameters for Los Angeles County Salons

Worker Skill Set	Wage	Administrative	Blowdry/Style/Etc.	Color/Highlight/Wash	Haircut/Shave	Nail/Misc.
1	-	-0.028	-0.275	0.876	-5.248	-61.626
1	-	(4.874)	(2.737)	(1.175)	(1.509)	(29.540)
2	536.753	-5.466	13.326	2.332	-6.157	-9.492
2	(210.962)	(3.919)	(10.040)	(1.968)	(2.535)	(2.699)
3	-7.202	0.043	1.570	-0.439	-3.733	-6.118
3	(24.149)	(1.343)	(2.155)	(.965)	(.701)	(10.649)
4	20.981	-0.305	3.759	0.751	-5.383	-3.982
4	(33.875)	(.954)	(2.710)	(1.231)	(1.351)	(2.395)
5	59.820	0.946	-2.708	1.654	-3.703	-3.676
5	(33.640)	(1.662)	(1.189)	(1.108)	(1.232)	(1.419)

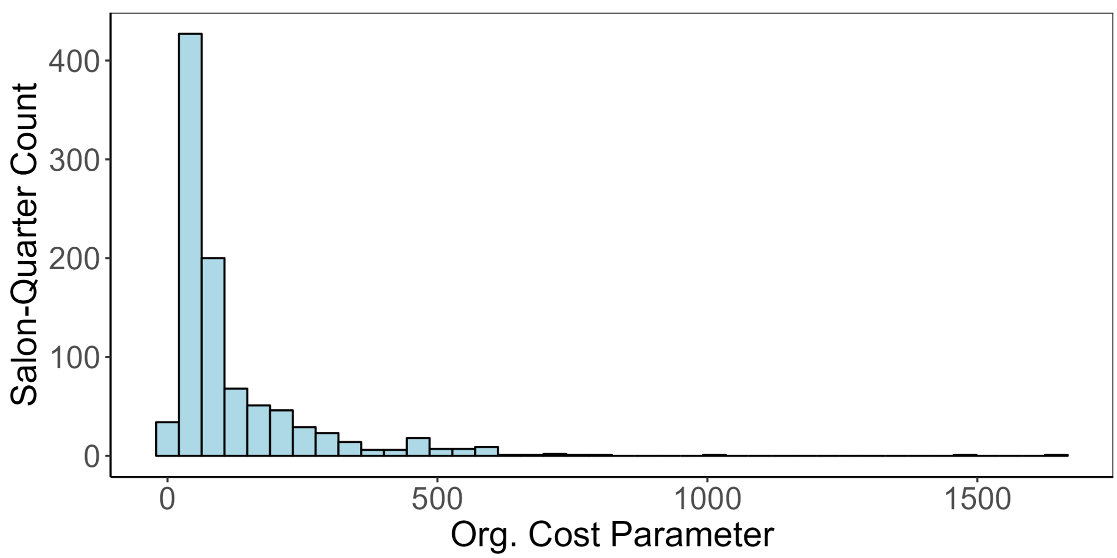
Note: The parameters associated with the skill sets. Standard errors are from 631 bootstrap replications.

Table 11: Worker Wages and Skill Parameters for New York County Salons

Worker Skill Set	Wage	Administrative	Blowdry/Style/Etc.	Color/Highlight/Wash	Haircut/Shave	Nail/Misc.
1	-	-29.238	2.254	1.103	1.647	0.206
1	-	(21.703)	(4.086)	(2.771)	(7.516)	(3.837)
2	-70.085	-0.795	2.752	1.991	-5.408	-3.038
2	(64.121)	(3.339)	(4.172)	(1.958)	(1.444)	(2.787)
3	-166.154	-4.001	-6.377	-0.745	-1.541	8.193
3	(69.972)	(11.974)	(5.772)	(3.531)	(2.934)	(8.411)
4	-141.734	11.461	-3.885	0.683	-3.853	9.979
4	(65.275)	(22.542)	(1.209)	(1.868)	(2.847)	(8.459)
5	660.399	47.273	16.775	-10.078	-4.238	22.728
5	(132.957)	(34.174)	(45.639)	(4.806)	(3.451)	(14.573)

Note: The parameters associated with the skill sets. Standard errors are from 631 bootstrap replications.

Figure 8: Estimated Salon-Specific Organization Costs



Note: The figure provides a a histogram of the organization costs recovered from inverting the s-index of all salons-quarters in the estimation sample.

6.2 Model Fit and Validation

I assess model fit by comparing data- and model-generated moments used in estimation in Table 12. Because there are many moments, I group them into four categories. The demand-side moments derived from the log market share equation are matched perfectly because they are estimated by two-stage least squares. The other moments do not come from a closed-form procedure, but they are still matched almost perfectly. The differences are mainly due to the constraint imposed that wages be at least the minimum wage and numerical precision in finding the wages which zero the labor demand moment conditions.

Table 12: Model Fit

Equation	Instrument	Count	Avg. Model	Avg. Data	R2
Log Market Share	County-Dye Instrument	3	-126.46	-126.46	1.000
Log Market Share	County-Quarter	33	-0.22	-0.22	1.000
Log Market Share	County-Task Assignments	75	-0.24	-0.24	1.000
Labor Demand	County-Skill Set	15	0.07	0.07	0.998
Price	County-Quarter	36	6.87	6.92	0.997
Price	Quarter-Task Mix	48	3.38	3.49	0.997
Price	County-Quarter-Labor	36	19.41	19.34	1.000

Note: This table summarizes how well the model matches the moments used in estimation.

In order to validate the model, I assess how well it can match the observed task content of jobs at the worker level. Although firm-level relative labor demands and task specialization were used in estimation, the task content of individual worker jobs was not. Despite this, I show in Table 13 that the model can match many of the patterns observed in the data. These include the correlation between different task dimensions, which represents the way in which tasks are packaged together.

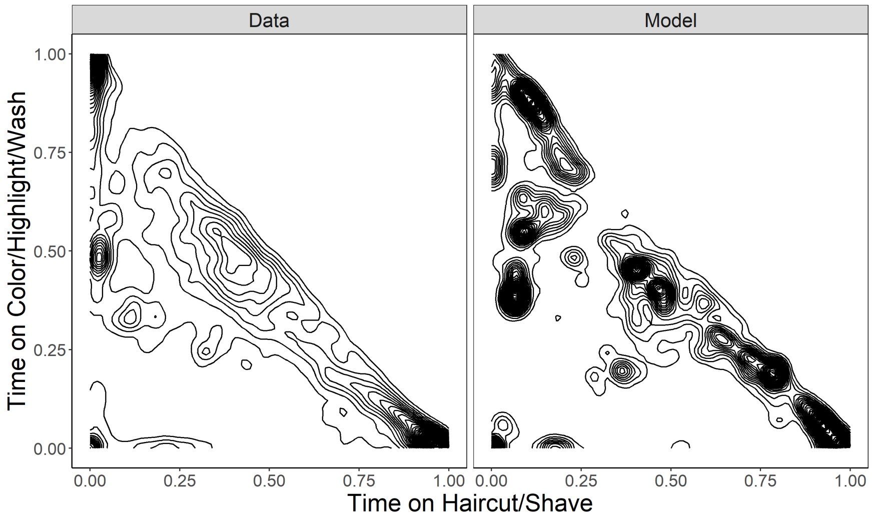
Table 13: Validating the Model Using Job Task Content

	Task	Variance	Cor. Task 1	Cor. Task 2	Cor. Task 3	Cor. Task 4	Cor. Task 5
Model	1	0.105	1.000	-0.678	-0.392	-0.259	-0.171
Data	1	0.107	1.000	-0.745	-0.260	-0.285	-0.184
Model	2	0.084		1.000	-0.154	-0.164	-0.156
Data	2	0.094		1.000	-0.080	-0.143	-0.234
Model	3	0.033			1.000	-0.013	-0.077
Data	3	0.014			1.000	0.013	-0.083
Model	4	0.019				1.000	-0.039
Data	4	0.019				1.000	-0.026
Model	5	0.014					1.000
Data	5	0.021					1.000

Note: This table summarizes how well the model matches moments not used in estimation. Variances and correlations are weighted by the labor time associated with each job.

Figure 9 illustrates the ability of the model to predict the task content of jobs. I plot the joint distribution of the Haircut/Shave task and the Color/Highlight/Wash task. The model replicates many of the important peaks of the distribution. Notably, the model has more peaks than the data because the model uses firm differences rather than worker differences to generate jobs. Examination of the marginal distributions reveals that in the data, the task content of several jobs is trimodal. The model is also able to match this trimodal structure, as seen in Appendix Figure A6.

Figure 9: Model and Data Joint Distributions of Job Task Content



Note: This model validation exercise compares jobs generated by the model to jobs in the data in terms of their task content. Jobs are weighted by the amount of labor assigned. The plot is a contour density plot, where steeper contours indicate greater probability mass.

6.3 Labor-Labor Substitution Patterns

I conduct two partial equilibrium exercises to illustrate how the model changes standard economic intuition. First, I explore labor-labor substitution patterns. I do this by increasing the wage of each worker skill set and examining how relative labor demand for each worker skill set responds across the heterogeneous firms in each market. The results are in Table 14.

Table 14: Labor-Labor Substitution Patterns

		Skill Set 1			Skill Set 2			Skill Set 3			Skill Set 4			Skill Set 5		
County	Skill Set	Max.	Med.	Min.	Max.	Med.	Min.	Max.	Med.	Min.	Max.	Med.	Min.	Max.	Med.	Min.
Cook	1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Cook	2	0.000	0.000	0.000	0.000	-0.003	-0.509	0.036	0.001	-0.133	0.014	0.000	-0.060	0.266	0.002	0.000
Cook	3	0.000	0.000	0.000	0.018	0.001	-0.094	0.000	-0.009	-0.199	0.493	0.000	0.000	0.071	0.000	0.000
Cook	4	0.000	0.000	0.000	0.005	0.000	-0.040	0.145	0.000	0.000	0.000	0.000	-0.475	0.052	0.000	0.000
Cook	5	0.000	0.000	0.000	0.603	0.002	0.000	0.231	0.000	0.000	0.083	0.000	0.000	0.000	-0.003	-0.337
New York	1	0.000	-0.013	-0.651	0.343	0.000	-0.459	0.644	0.000	0.000	0.005	0.000	-0.500	0.386	0.000	-0.079
New York	2	0.166	0.000	-0.105	0.000	-0.005	-0.429	0.076	0.000	-0.205	0.401	0.000	-0.010	0.153	0.000	-0.119
New York	3	0.707	0.008	0.000	0.566	0.000	-0.459	0.000	-0.004	-0.655	0.613	0.000	0.000	0.081	0.000	-0.471
New York	4	0.016	0.000	-0.280	0.474	0.000	-0.043	0.384	0.000	0.000	0.000	0.000	-0.401	0.429	0.000	-0.100
New York	5	0.023	0.000	-0.002	0.054	0.000	-0.012	0.011	0.000	-0.014	0.035	0.000	-0.015	0.000	-0.001	-0.109
Los Angeles	1	0.000	-0.002	-0.525	0.054	0.000	-0.345	0.311	0.000	-0.161	0.329	0.000	-0.212	0.377	0.000	0.000
Los Angeles	2	0.107	0.000	-0.176	0.000	-0.001	-0.266	0.152	0.000	-0.204	0.184	0.000	0.000	0.000	0.000	-0.132
Los Angeles	3	0.041	0.000	-0.152	0.097	0.000	-0.221	0.000	-0.001	-0.221	0.143	0.000	-0.096	0.112	0.000	0.000
Los Angeles	4	0.554	0.000	-0.328	0.521	0.000	0.000	0.424	0.000	-0.452	0.000	0.000	-0.521	0.414	0.000	-0.020
Los Angeles	5	0.802	0.000	0.000	0.000	0.000	-0.389	0.480	0.000	0.000	0.536	0.000	-0.039	0.000	0.000	-0.613

Note: This table depicts the labor-labor substitutions patterns across different worker-skill sets. For each skill set listed in column 2, I increase the wage by 1% holding all other wages fixed. I then measure the change in relative labor demand across all salons and all worker skill sets. I report the minimum, median and maximum changes across all salon-quarters in each county, for each skill set. Thus, row 1 (Cook County Skill Set 1) column 3 (Skill Set 1 Max.) represents the maximum change in relative labor demand of skill set 1 after a 1% change in the wage of workers with skill set 1, similar to an own-wage elasticity of labor demand. Row 1 column 6 (Skill Set 2 Max.) would represent a cross-wage elasticity.

Although the numbers are not exactly elasticities,[23] the table has a similar interpretation as demand substitution patterns tables common in industrial organization. Three patterns emerge.

First, an increase in own wage reduces relative labor demand across all firms, as expected. Second, in all markets there exist pairs of worker skill sets that are substitutes at one firm and complements at another. Third, substitution effects are small at some firms but large at others. For example, it is common for the median firm to experience close to 0 change in relative labor demand in response to a wage increase, while the most responsive firm sees a change of 20 percentage points or more. These patterns occur not just because of differences in task mixtures across firms but also because of differences in organization costs.

6.4 Productivity Spillovers

In this section, I consider the same wage shocks as in the last subsection. However, I now ask how these change the productivity of different worker skill sets at different salons. Theoretically in Theorem 1 and empirically in Table 14, I show that as the wage of a particular skill set rises, the law of demand holds and each salon demands relatively less of that particular worker. However, the tasks that were previously performed by that skill set still need to be completed, and the salon incorporates them into the task assignments of the workers that remain. Intuitively, coworkers must pick up the slack of those who are let go.

Table 15 shows how such task reassignment impacts the quality per unit of labor produced (labor productivity). I only include skill set combinations that are employed together in the initial equilibrium.

23. I do not estimate elasticities because many firms have an initial relative labor demand of 0 for one or more worker types.

Table 15: Effects of Wage Increases on Own and Coworker Productivity

		Skill Set 1			Skill Set 2			Skill Set 3			Skill Set 4			Skill Set 5		
County	Skill Set	Max.	Med.	Min.	Max.	Med.	Min.	Max.	Med.	Min.	Max.	Med.	Min.	Max.	Med.	Min.
Cook	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Cook	2	NA	NA	NA	0.069	0.001	0.000	0.002	0.000	-0.024	0.024	-0.002	-0.006	0.000	-0.001	-0.018
Cook	3	NA	NA	NA	0.029	-0.001	-0.008	0.065	0.007	0.000	-0.005	-0.122	-0.196	0.000	0.000	-0.006
Cook	4	NA	NA	NA	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Cook	5	NA	NA	NA	0.007	0.000	-0.009	0.000	0.000	-0.008	-0.001	-0.019	-0.030	0.016	0.001	-0.001
New York	1	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.001	0.000	0.000	0.000	0.000	0.000	0.000
New York	2	0.004	0.000	-0.001	0.003	0.000	0.000	0.000	0.000	-0.002	0.000	0.000	-0.001	0.001	0.000	0.000
New York	3	0.001	-0.001	-0.044	0.001	0.000	-0.003	0.010	0.000	-0.052	0.006	0.000	0.000	0.005	0.000	-0.002
New York	4	0.005	0.000	-0.052	0.000	-0.004	-0.020	0.000	0.000	-0.005	0.169	0.001	-0.001	0.015	0.000	-0.003
New York	5	0.005	-0.006	-0.176	0.001	-0.012	-0.163	0.056	0.001	-0.001	0.036	0.000	-0.015	0.167	0.009	-0.003
Los Angeles	1	0.001	0.000	0.000	0.000	0.000	-0.002	0.000	0.000	-0.001	0.001	0.000	0.000	0.000	0.000	-0.001
Los Angeles	2	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Los Angeles	3	0.001	0.000	0.000	0.002	0.000	-0.001	0.000	0.000	-0.001	0.001	0.000	-0.002	0.004	0.000	-0.001
Los Angeles	4	0.003	0.000	-0.001	0.005	0.001	0.000	0.000	0.000	-0.001	0.002	0.000	-0.004	0.004	0.000	-0.005
Los Angeles	5	0.005	0.000	0.000	0.003	0.000	-0.002	0.000	0.000	-0.002	0.000	0.000	-0.004	0.005	0.000	-0.004

Note: This table depicts the percent increase in labor productivity due to an increase in own and coworker wage at different salons. For each skill set listed in column 2, I increase the wage by 1% holding all other wages fixed. I then measure the change in quality per unit of labor across all salons that employ both skill sets and all worker skill sets. I report the minimum, median and maximum changes across salon-quarters in each county, for each skill set. Thus, row 1 (Cook County Skill Set 1) column 3 (Skill Set 1 Max.) represents the maximum change in productivity of workers in skill set 1 after a 1% change in the wage of workers with skill set 1. Salons are only included in these calculations if their relative labor demand for both worker skill sets was greater than 0.01 initially. Skill Set 1 in Cook County is very uncommon, as a result it never satisfies this criteria.

In general, a wage increase for one worker can move the productivity of a coworker in either direction. This is because individual salons are not maximizing productivity (quality per unit of labor) alone but are also balancing the total organizational cost as well as the wage bill. Taking on tasks that were previously performed by a coworker can raise (lower) productivity if the receiving worker is (not) particularly skilled in that task relative to what they were originally doing at the salon.

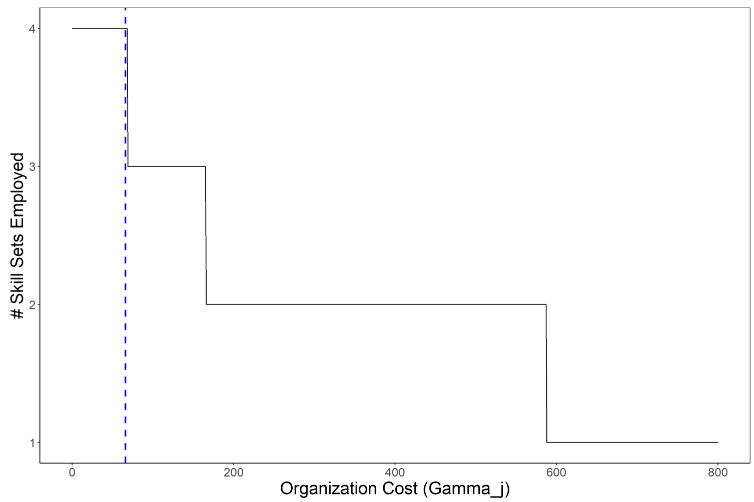
When positive productivity effects occur, they tend to be own-wage effects. That is, a worker's productivity tends to rise when their own wage increases. The intuition is that if a salon retains a worker after a wage increase, the worker is more likely to be doing tasks at which they are relatively skilled.

6.5 Workforce Expansion Paths

In this partial equilibrium counterfactual, I ask how the composition of a firm's workforce changes as organization costs decline (or, equivalently, as organizational productivity improves). I hold wages fixed and consider only how relative labor demand (not the total amount of labor) changes.

The conventional wisdom is well-articulated by Adam Smith: Specialization is limited by the extent of the market. The implication for firms is then that larger firms should employ a more specialized workforce, and in order to do so, they should hire a more diverse portfolio of worker skill sets. In the model, as organization costs fall, Corollary 0.1 states that firms increase in size, and Figure 10 illustrates that for at least one firm at the estimated parameters, the number of unique skill sets hired increases.

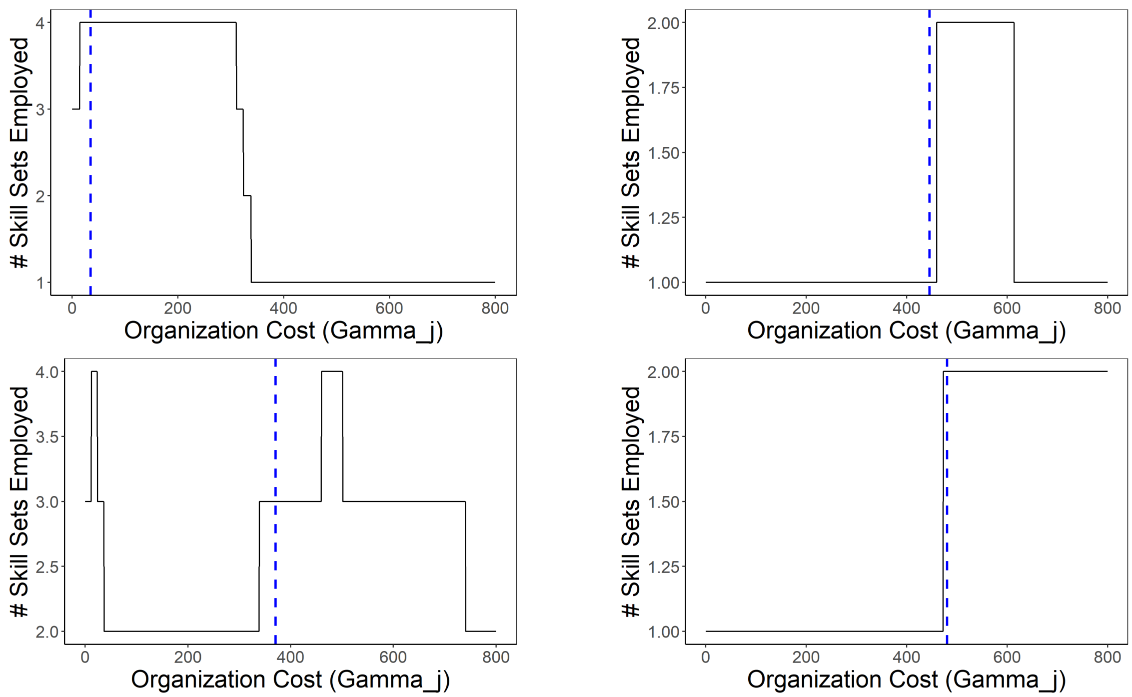
Figure 10: A Conventional Expansion Path of a Firm



Note: The y-axis plots the number of unique skill sets where relative labor demand is at least 0.01. The x-axis plots different values of the organization cost parameter, with the parameters estimated in the data via inversion of the s-index marked with the dashed vertical line.

However, Figure 10 is for a single firm, and it turns out that the conventional wisdom is more of a rule of thumb. Figure 11 shows that there exist many firms where expansion paths look qualitatively different. At some firms, the number of employed skill sets reaches a maximum at intermediate levels of organization costs. In some rare cases, the number of skill sets employed can increase as organization costs rise. These differences are driven by differences in task intensities and labor market environments.

Figure 11: A Conventional Expansion Path of a Firm



Note: Each plot represents the counterfactual composition of a specific salon-quarter's workforce as organization costs change but all other factors are held fixed. The y-axis plots the number of unique skill sets where relative labor demand is at least 0.01. The x-axis plots different values of the organization cost parameter, with the parameters estimated in the data via inversion of the s-index marked with the dashed vertical line.

7 Counterfactuals

This section uses the model to study the impact of several counterfactual economic shocks on aggregate labor productivity. These counterfactuals are chosen to represent important questions across economics, to illustrate how endogenous and heterogeneous task assignment within the firm determines the impact of the economic shocks on industry-wide productivity.

To do this, I first define the reallocation equilibrium. It is the outcome when firms are allowed to adjust prices (p_j) but organization structures (B_j) are fixed at the initial equilibrium choices. Because prices control quantities, this equilibrium allows firms to adjust the total labor they hire but not the division of labor within the firm. I define the reorganization equilibrium as the outcome when firms are allowed to fully adjust their task assignments. It is the full equilibrium described in Section 4.1.

Different economic shocks change the competitive position of firms within an industry, depending on their initial structure and task mix. The reallocation equilibrium quantifies how reallocation of labor across firms shifts outcomes. Task assignments differ across firms, even if the task assignments are held fixed, so reallocation can change aggregate productivity and task specialization. The spirit of the reallocation equilibrium is not novel, as in many models with heterogeneous firms, shocks reallocate labor inputs in ways that change aggregate productivity. However, the reorganization equilibrium is novel, because firms with different organization costs are internally changing in response to an external economic shock.

I solve for counterfactual equilibria for each county in 2021 Q2, as this is the final quarter in the estimation sample, and typically has the most salons due to greater adoption of the software over time. For additional details, see Appendix Section A.14.

I consider four counterfactuals.

- **Sales Tax Increase.** I impose a 4 percentage point increase of the tax on salon services. In Los Angeles and Cook Counties, salon services are exempt from sales tax, so in these two counties, this counterfactual changes the sales tax rate from 0% to 4%. In New York County, salon services are already taxed at a rate of 4%, so this change brings the tax rate to 8%.

- **Management Diffusion.** Each salon learns and then adopts the management practices of the next best salon. I implement this change by ordering the salons by their organization cost parameter and then changing each salon's organization cost to be that of the salon one rank above it. I leave the salon with the lowest organization cost unchanged.

- **Immigration.** There is a 10% increase in the total labor supply of the worker skill set with the lowest wage in each market.

- **Increase in Market Concentration.** Half of the salons in each market are removed. Because each salon in the data represents a number of actual salons in each market, this change is implemented by reducing the number of actual salons that each salon in the data represents.[24]

Table 16 presents the results from these counterfactuals. I solve for the reallocation and reorganization equilibrium for each counterfactual, and compute the percentage change in the average s-index (weighted by total labor) and labor productivity relative to baseline. Labor productivity is measured as the task-assignment component of quality produced by all workers in the market.[25] Because none of the counterfactuals generate unemployment, the set of labor being used remains fixed, total quality is a measure of labor productivity.

Table 16: Counterfactual Productivity and Specialization Changes

		Reallocation		Reorganization	
County	Counterfactual	S-Index Change	Prod. Change	S-Index Change	Prod. Change
Cook	Immigration	-0.017	0.006	0.017	0.018
New York	Immigration	-0.030	0.015	-0.018	0.015
Los Angeles	Immigration	-0.014	-0.002	0.004	0.022
Cook	Incr. Concentration	0.000	0.000	0.010	0.003
New York	Incr. Concentration	0.000	0.000	-0.013	0.005
Los Angeles	Incr. Concentration	0.002	0.001	-0.008	-0.019
Cook	Management Diffusion	0.000	0.000	0.010	0.000
New York	Management Diffusion	0.000	0.000	0.007	0.000
Los Angeles	Management Diffusion	0.001	0.001	0.045	0.011
Cook	Sales Tax	0.000	0.000	-0.010	-0.002
New York	Sales Tax	0.000	0.000	0.007	-0.006
Los Angeles	Sales Tax	0.000	0.001	-0.047	-0.007

Note: Effects are percent changes from the baseline equilibrium.

A first pattern across all shocks is that reallocation effects tend to be small, with the exception of the immigration shock. This finding is in line with the idea that most of the shocks impact all firms similarly, while the immigration counterfactual raises the supply and lowers the wage of a particular worker type, reducing the marginal cost of a particular subset of firms. Depending on the identity of these advantaged firms, the aggregate productivity effect can be positive (as in New York County) or negative (as in Los Angeles County).

A second pattern is that preventing internal reorganization tends to underestimate the productivity implications of economic shocks, in some cases flipping the sign of the

24. This is similar to merging salons with the same characteristics.

25. Formally, for a worker of skill set i at firm j, labor productivity is $\sum_k E_j(i)b_j(i,k)\theta_i(k)$ multiplied by total labor. Labor productivity is the sum of this quantity across all salons and workers.

change in aggregate labor productivity. As an example, imposing a sales tax on salon services in Los Angeles appears to have a small positive impact when firms are prevented from reorganizing. This implies that absent reorganization, the sales tax causes labor to flow to firms where on average workers are being better utilized. However, imposing a sales tax makes it more difficult for firms to pass on the costs of specialization to consumers, and as a result specialization declines by 5 percent. Thus, when reorganization is accounted for, the labor productivity declines by 0.7%.

Similarly, in Cook County, immigration reduces task specialization when firms are prevented from reorganizing, because it shifts labor towards salons with lower levels of specialization. When salons are allowed to reorganize, task specialization rises by an equal and opposite amount, tripling the aggregate productivity impact.

In general, I find that immigration and management diffusion improve productivity across all markets, and sales tax increases reduce productivity across all markets. However, the effects of increased concentration vary. In Los Angeles County, increased concentration leads to a 1.9% decline in productivity, while in New York and Cook Counties, it leads to small increases. These differences are due to the differential wage effects across different types of workers.

These aggregate measures mask significant heterogeneity across workers. Appendix Table A2 displays productivity and wage effects by worker skill set for the reorganization equilibrium. Although the diffusion of management practices in Los Angeles County improves productivity in aggregate by around 1%, it decreases the productivity of skill-set-3 workers by 1.2%, increases the productivity of skill-set-4 workers by 2.5%, increases the productivity of skill-set-5 workers by 1.8%, and is close to neutral for all other workers.

8 Conclusion

This paper provides evidence that task specialization within establishments is related to productivity differences across establishments. It also provides a structural model which can be used to understand how endogenous and heterogeneous task assignment reacts to economic shocks. The counterfactual exercises illustrate that allowing internal organization to be endogenous and heterogeneous qualitatively changes the economic forces at play.

In addition, this paper shows that task assignment data can be used to enrich our understanding of how firms determine aggregate productivity. Traditional employee-employer matched data, which often contain wage information, has already contributed to our understanding of labor markets. Future work should consider bringing these data sources together. This united perspective will help us understand new questions, including how labor market power and management hierarchies interact with task assignment.

References

Adenbaum, Jacob. 2022. "Endogenous Firm Structure and Worker Specialization."

Alchian, AA, and H Demsetz. 1972. "Production, information costs, and economic organization." *American Economic Review* 62 (5): 777–795.

Alvarez, Jorge, Felipe Benguria, Niklas Engbom, and Christian Moser. 2018. "Firms and the decline in earnings inequality in Brazil." *American Economic Journal: Macroeconomics* 10 (1): 149–189.

Aravindakshan, Ashwin, and Brian Ratchford. 2011. "Solving share equations in logit models using the LambertW function." *Review of Marketing Science* 9.

Autor, David H., Frank Levy, and Richard J. Murnane. 2003. "The Skill Content of Recent Technological Change: An Empirical Exploration." *The Quarterly Journal of Economics* 118, no. 4 (November): 1279–1333.

Baker, George, Robert Gibbons, and Kevin J. Murphy. 2002. "Relational Contracts and the Theory of the Firm." *The Quarterly Journal of Economics* 117, no. 1 (February): 39–84.

Baker, George P., and Thomas N. Hubbard. 2003. "Make Versus Buy in Trucking: Asset Ownership, Job Design, and Information." *American Economic Review* 93, no. 3 (June): 551–572.

Bassi, Vittorio, Jung Hyuk Lee, Alessandra Peter, Tommaso Porzio, Ritwika Sen, and Esau Tugume. 2023. *Self-employment within the firm.* Technical report. National Bureau of Economic Research.

Berge, Laurent, Sebastian Krantz, Grant McDermott, and Russell Lenth. 2021. *Package 'fixest'.*

Blahut, Richard E. 1972. "Computation of Channel Capacity and Rate-Distortion Functions." *IEEE Transactions on Information Theory,* no. 4.

Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and explaining management practices across firms and countries." *The quarterly journal of Economics* 122 (4): 1351–1408.

Bonhomme, Stéphane, and Elena Manresa. 2015. "Grouped patterns of heterogeneity in panel data." *Econometrica* 83 (3): 1147–1184.

Caliendo, Lorenzo, Giordano Mion, Luca David Opromolla, and Esteban Rossi-Hansberg. 2020. "Productivity and organization in portuguese firms." *Journal of Political Economy* 128, no. 11 (November): 4211–4257.

Caplin, Andrew, and Barry Nalebuff. 1991. "Aggregation and Imperfect Competition: On the Existence of Equilibrium." *Econometrica* 59, no. 1 (January): 25.

Card, David, Jörg Heining, and Patrick Kline. 2013. "Workplace heterogeneity and the rise of West German wage inequality." *The Quarterly journal of economics* 128 (3): 967–1015.

Coraggio, Luca, Marco Pagano, Annalisa Scognamiglio, and Joacim Tåg. 2023. *JAQ of All Trades: Job Mismatch, Firm Productivity and Managerial Quality.* Technical report. Research Institute of Industrial Economics.

Csaba, Dániel. 2021. "Attention Elasticities and Invariant Information Costs *."

Dessein, Wouter, and Tano Santos. 2006. "Adaptive organizations." *Journal of Political Economy* 114, no. 5 (October): 956–995.

Dowle, Matt, Arun Srinivasan, Jan Gorecki, Michael Chirico, Pasha Stetsenko, Tom Short, Steve Lianoglou, Eduard Antonyan, Markus Bonsch, Hugh Parsonage, et al. 2019. "Package 'data. table'." *Extension of 'data. frame* 596:952.

Du, Yu, and Ravi Varadhan. 2018. "SQUAREM: An R package for off-the-shelf acceleration of EM, MM and other EM-like monotone algorithms." *arXiv preprint arXiv:1810.11163.*

Freund, Lukas. 2022. "Superstar Teams: The Micro Origins and Macro Implications of Coworker Complementarities." *SSRN Electronic Journal* (December).

Garicano, L. 2000. "Hierarchies and the organization of knowledge in production." *Journal of Political Economy* 108 (5): 874–904.

Garicano, Luis, and Thomas N. Hubbard. 2016. "The Returns to Knowledge Hierarchies." *The Journal of Law, Economics, and Organization* 32, no. 4 (November): 653–684.

Garicano, Luis, and Yanhui Wu. 2012. "Knowledge, communication, and organizational capabilities." *Organization Science* 23, no. 5 (February): 1382–1397.

Haanwinckel, Daniel. 2023. *Supply, demand, institutions, and firms: A theory of labor market sorting and the wage distribution.* Technical report. National Bureau of Economic Research.

Haegele, Ingrid. 2022. "Talent hoarding in organizations." *arXiv preprint arXiv:2206.15098.*

Hlavac, Marek. 2022. "stargazer: beautiful LATEX, HTML and ASCII tables from R statistical output." *The Comprehensive R Archive Network.*

Horn, Roger A, and Charles R Johnson. 2012. *Matrix analysis.* Cambridge university press.

Johnson, Matthew S., and Michael Lipsitz. 2022. "Why Are Low-Wage Workers Signing Noncompete Agreements?" *Journal of Human Resources* 57, no. 3 (May): 689–724.

Jung, Junehyuk, Jeong Ho (John) Kim, Filip Matějka, and Christopher A. Sims. 2019. "Discrete Actions in Information-Constrained Decision Problems." *The Review of Economic Studies* 86, no. 6 (November): 2643–2667.

Kuhn, Moritz, Jinfeng Luo, Iourii Manovskii, and Xincheng Qiu. 2023. "Coordinated firm-level work processes and macroeconomic resilience." *Journal of Monetary Economics* 137:107–127.

Lazear, Edward P. 2009. "Firm-specific human capital: A skill-weights approach." *Journal of Political Economy* 117, no. 5 (October): 914–940.

Lindenlaub, Ilse. 2017. "Sorting Multidimensional Types: Theory and Application." *The Review of Economic Studies* 84, no. 2 (April): 718–789.

Martinez, Elizabeth A., Nancy Beaulieu, Robert Gibbons, Peter Pronovost, and Thomas Wang. 2015. "Organizational Culture and Performance." *American Economic Review* 105, no. 5 (May): 331–35.

Matêjka, Filip, and Alisdair McKay. 2015. "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model." *American Economic Review* 105, no. 1 (January): 272–98.

McFadden, D. 1973. "Conditional logit analysis of qualitative choice behavior."

Meier, Stephan, Matthew Stephenson, and Patryk Perkowski. 2019. "Culture of trust and division of labor in nonhierarchical teams." *Strategic Management Journal* 40, no. 8 (August): 1171–1193.

Minni, Virginia Magda Luisa. 2023. *Making the invisible hand visible: Managers and the allocation of workers to jobs.* Technical report. London School of Economics and Political Science Centre for Economic Performance.

Moser, Stefan M, and Po-Ning Chen. 2012. *A student's guide to coding and information theory.* Cambridge University Press.

Ocampo, Sergio. 2022. "A Task-Based Theory of Occupations with Multidimensional Heterogeneity *."

Rubin, Donald B. 1981. "The bayesian bootstrap." *The annals of statistics,* 130–134.

Sattinger, Michael. 1975. "Comparative Advantage and the Distributions of Earnings and Abilities." *Econometrica* 43, no. 3 (May): 455.

Shannon, C. E. 1948. "A Mathematical Theory of Communication." *Bell System Technical Journal* 27 (3): 379–423.

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter. 2019. "Firming up inequality." *The Quarterly journal of economics* 134 (1): 1–50.

Syverson, Chad. 2004. "Product substitutability and productivity dispersion." *Review of Economics and Statistics* 86 (2): 534–550.

Tian, Jianrong. 2019. "Attention Costs without Distraction and Entropy."

Tishby, Naftali, Fernando C Pereira, and William Bialek. 2000. "The information bottleneck method."

U.S. Bureau of Labor Statistics. *Consumer Expenditure Survey Public Use Microdata (PUMD) 2018-2020.* Accessed: 2022-07-11.

———. *Consumer Expenditure Survey Public Use Microdata (PUMD) 2021.* Accessed: 2023-10-23.

U.S. Census Bureau. 2010. *County Population Totals: 2010-2019.* Accessed: 2024-03-10.

———. 2024. *County Population Totals: 2020-2022.* Accessed: 2024-03-10.

Varadhan, Ravi, and Paul Gilbert. 2010. "BB: An R package for solving a large system of nonlinear equations and for optimizing a high-dimensional nonlinear objective function." *Journal of statistical software* 32:1–26.

A Online Appendix

A.1 Data Construction Details

The data throughout the paper is constructed using the following procedure. First, I limit the data to establishments that identify as hair salons, barbershops or blowouts and styling. I exclude two establishments with revenue information that appears to be incorrectly entered.

Service descriptions are classified as described in Appendix Section A.16. Importantly, some service descriptions are classified as multiple task categories. When this occurs, the service descriptions are broken into component tasks.

After this process, the amount of time spent on each task is constructed using a variable capturing the total time spent on an appointment and the times of individual services. The more detailed time variable is only available for 68% of the data. When it is not available, I use the appointment-level duration variable.

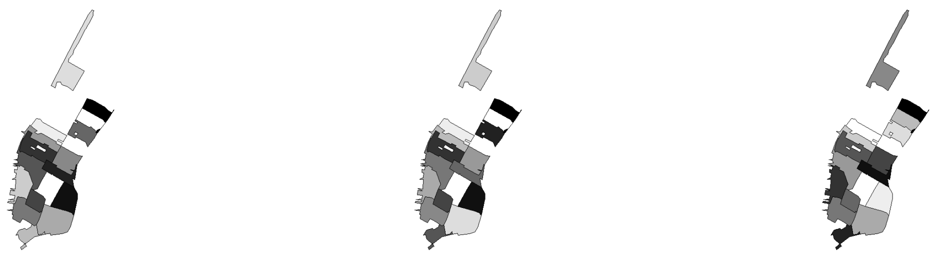
For services that consist of multiple task categories, and for appointments where the more detailed time variable is not available and multiple services are performed, the total time spent on the appointment is split across the tasks. First, I compute the average amount of time spent on each task category among only single-task appointments. Second, I compute the fraction of time to assign to each task as the corresponding task average divided by the sum of the averages of all other tasks in that appointment. Third, I distribute the total time spent on the appointment across the tasks using this imputed fraction.

I remove 5 task assignments (out of 13.7 million) with negative time spent. Slightly less than 2% of task assignments have a time spent of 0. For these I impute the time spent as the average duration of other tasks in that category. The data.table package (Dowle et al. 2019), fixest package (Berge et al. 2021), stargazer package (Hlavac 2022), and squarem package (Du and Varadhan 2018) were all used to build the data and analyze the results.

A.2 Spatial Correlation

In this section, I visualize the spatial distribution of return rates, productivity (revenue per minute), and task specialization. I plot the distribution by taking average by zip code across the three markets for which the model is estimated. Zip codes without activity are excluded from the maps.

Figure A1: New York County Spatial Distribution



(a) S-Index **(b)** Productivity **(c)** Return Customers

Figure A2: Los Angeles County Spatial Distribution

(a) S-Index **(b)** Productivity **(c)** Return Customers

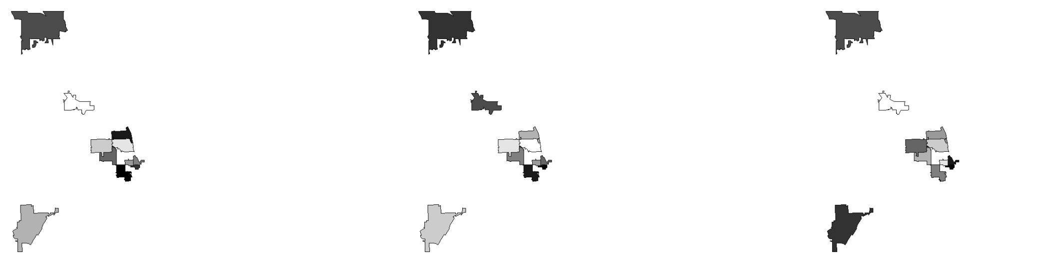
There are zip codes where productivity, specialization and the fraction of repeat customers are all high. There are also zip codes where repeat customers are relatively low but productivity and the s-index are high. Finally, there are zip codes where productivity is high

A.3 Interpreting Material Costs and Wage Levels

The magnitude of the estimated wages and material cost coefficients may appear unreasonably large at first glance. This section argues that they are reasonable once one considers how they impact model outcomes.

To understand what an administrative task material cost coefficient of -6098 means, note that the median firm in the estimation sample spends 0.56% of total time on this task.

Figure A3: Cook County Spatial Distribution



(a) S-Index **(b)** Productivity **(c)** Return Customers

Note also that the coefficient is relative to the haircut task. This means that at the median firm, the effect is $-6098 * 0.0056 = -34$. So if the median firm completely eliminated administrative tasks and all time was allocated to the haircut task, marginal costs would increase by only \$34.

To understand wages, first recall that proposition 2 implies that if there is vertical skill level differences across workers, these will manifest as wage differences that exactly make firms indifferent within skill set. Specifically, for any two workers m, m' with the same skill set but different skill levels, wages are such that $w_m - w_{m'} = \rho^{-1}(\bar{\theta}_m - \bar{\theta}_{m'})$.

Given that ρ is often on the order of 0.02 or less, workers with a skill level difference of 5 will have a wage differential of $5 \cdot 0.02^{-1} = 250$. Because I do not have wage data, I do explicitly incorporate this into estimation. This means that, assuming random sorting across firms, the wage estimates reflect the average skill level of workers in that skill set. Thus large wages may reflect the fact that some skill sets have a high average skill level within them.

A.4 Estimation Procedure

The estimation procedure consists of the following steps:

1. Classify workers within firm based on their job's task content.

2. Set aside salon-quarters where workers cannot be grouped across salon-quarters.[26]

26. This includes firms with an s-index of 0, firms with only 1 worker skill set, and firms that do not perform one or more tasks.

3. Classify workers across firms using their normalized coworker log ratio vectors.

4. Using recovered task assignments and relative labor demands, recover price sensitivity, skill set parameters, and county-quarter demand levels via two-stage least squares applied to (20). As an instrument for price, use the producer price index for synthetic organic dies multiplied by the task mix for the color task.

5. Guess relative wages.

6. Recover the organization cost parameter of each salon by inverting the s-index.

7. Solve each salon-quarter's internal organization problem. Compare the average relative labor demands to those implied via the classification procedure.

8. Return to 5 and repeat until convergence.

9. Since relative wages are now known, recover the organization costs and relative labor demands of the set-aside firms by inverting the s-index.

10. Estimate material costs, county-quarter wage levels and cost levels via linear GMM applied to (22) with all variables instrumenting for themselves. Constrain wage levels such that all wages are above county-quarter specific minimum wages. If not for the constraint, this would be exactly ordinary least squares.

I use inversion, as opposed to the connected set procedure from the identification proof to recover salon-quarter organization costs for two reasons. First, this imposes that each firm's s-index in the data match its model generated s-index at the market parameters. Second, it leads to more reasonable organization costs in practice. The main negative of inversion is that it greatly increases computational time, as it requires solving each firm's internal organization problem for each guess of the wage vector. I mitigate this cost by accelerating the contraction mapping proposed in Appendix Section A.5.

I classify workers within a firm using hierarchical clustering with complete linkage where the distance is Euclidean. I set the stopping threshold for a market to be the smallest value such that no firm has more than 5 worker types.

Classifying workers across firms is more involved. The proof of identification implies a clustering procedure, where workers are grouped across firms into skill sets based on the ratio of their time use with coworkers from other skill sets. To translate the theoretical procedure into a practical estimation procedure, I must choose an algorithm to implement it. Several factors constrain the procedures I can use. First, the ratio comparisons are not proper distances. That is, we can make statements of the form: workers 1 and 2 are close. But we cannot say that 1 and 2 are closer than 3 and 4, and I cannot say that 1 and 2 are close to some fixed point in euclidean space. We also cannot say that just because 1 and 2 are far, they are not the same type. This is because two workers may have different log ratios simply because they have different coworker types.

These issues imply I cannot use any clustering algorithms that rely on centroids or that aggregate over multiple coworker pairs when forming clusters. This rules out popular non-hierarchical algorithms like k-means and also many hierarchical algorithms that use linkage methods which involve some sort of aggregation (average, complete, Ward, median, etc.). This leaves the single-linkage algorithm as a natural candidate. I now outline the exact single-linkage procedure that is used to estimate the model.

1. Start with any firm that has all worker skill sets (5 distinct worker groups). Call this the reference firm. Fix the labels of the worker skill sets to be the labels of the 5 distinct worker groups at the reference firm.

2. Consider any firm that is not the reference firm. Take all the distinct worker groups at this firm and label them every possible way.

3. For each labeling, compute the sum of the Manhattan distance between each worker group at the firm and the worker group with the same label at the reference firm. The final classification of the worker groups is the labeling that yields the smallest value.

4. Repeat for all firms in the market.

This procedure is both internally consistent and computationally feasible. This is because under the assumption that a firm with 5 worker skill sets exist, all other firms should be able to be linked to this firm both theoretically (under optimal task assignment) and numerically (one labeling of workers at each firm will minimize the distance of normalized coworker log ratios from the reference firm normalized coworker log ratios). Additionally, the procedure never contradicts itself. For example, if workers A and C are at the same firm, and the first clustering procedure grouped them into different skill sets within the firm, the algorithm will never generate a grouping that places workers A and C in the same skill set.

There exist other procedures to classify workers across firms, however they vary both in their computational complexity and their internal consistency. For example, one can attempt to use classic single linkage clustering methods with constraints that specify forbidden linkages. However, this problem is known to be NP-hard and difficult to implement.

The proof of identification uses variation in relative organization costs across firms to identify price sensitivity. In practice, relative organization costs are estimated using small samples and are therefore noisy. Instead, I use the share of the color task multiplied by the producer price index of synthetic organic dyes in that quarter.

In practice this provides better power. Because the task mix is taken to be exogenous, this instrument satisfies an exclusion restriction under the theory. It is also relevant to firm pricing decisions both in theory and empirically.[27] Intuitively, by observing how variation across time in the price of dye and variation in the share of the color task across firms is passed on to consumers via prices identifies the price sensitivity parameter (ρ).

Standard errors are obtained via the Bayesian bootstrap (drawing random weights) instead of resampling for two reasons. First, is the common issue that if data is resampled some samples will leave out quarter-county fixed effects. Second, if the data is resampled, the labeling of the worker types from the classification procedure will not be consistent across bootstrap replications. Essentially, skill set 1 will not correspond to skill set 1 across replications. Ten programs were run simultaneously on a parallel computing cluster for 10 days. There are 631 bootstrap replications because this is the number that finished in the allocated 10 day period.

Importantly, the classification procedure of workers into skill sets depends only on the support of observed task assignments, not on relative frequencies. Thus, it is invariant to different weighting of the data and will not change when new bootstrap weights are

27. It is strongly positively correlated with prices.

drawn almost surely.[28] In this sense, the current standard errors do not account for the classification step. However, because there are a finite number of worker skill sets that are small relative to the number of establishments and workers, it is possible that an argument similar to that given in Bonhomme and Manresa 2015 holds. In that case, not accounting for classification in asymptotic inference is without loss.

For the classification step of estimation only, I Laplace smooth the task assignments within a firm. Specifically, I add 47.67 minutes (the average time spent on a single task) to the time spent by each employee on each task in a quarter. The idea is to capture the fact that due to the finiteness of the data, there is a positive probability that even if a worker is supposed to be assigned a task, they are not observed being assigned that task. The practical reason for doing this is in order to classify workers across firms, which requires taking the ratio of task assignments. Note this is only done at the worker level. Salon-quarters where some tasks are not performed are not smoothed. Their organization costs can be recovered by inverting their s-index after skills, wages and other market parameters have been estimated.

A.5 Full Solution Contraction Mapping

Recovering the organization costs of firms that only employ one skill set (the final step of estimation) and performing counterfactuals requires solving the for the firm's optimal task-assignment strategy. Although the theoretical results imply that the firm's task-assignment strategy is fully characterized by the division of time across worker skill sets, this is still a $K \times K$ matrix that must be found for all J firms. It turns out that there exists a globally convergent contraction mapping which delivers the firm's profit-maximizing organization structure given values for the market parameters (Ω).

Proposition 6 *Given market parameters (Ω), and firm j's task mix and organization cost parameter, the Blahut–Arimoto algorithm delivers relative labor demands and jobs for each skill set which maximize firm j's profit.*

The proposition follows directly from the fact that the firm's strategy solves an equivalent rate-distortion problem, which can be solved using the Blahut–Arimoto algorithm.[29] This

28. As long as a weight of 0 is not drawn, an event that occurs with probability 0.
29. See Tishby, Pereira, and Bialek 2000 or Blahut 1972.

equivalence is proven in the course of proving Theorem 1. The Blahut–Arimoto algorithm (Blahut 1972) is a fixed-point algorithm which iterates on two optimality conditions and can be described as follows:

0. Guess some relative labor demands E^0. Create matrix V: $V_{i,k} = exp[\gamma_j^{-1}(\rho^{-1}\Theta(i,k) - w_i)]$.
1. Compute interim organization structure $B_j(i,k)^t = \alpha_j(k) \frac{V_{i,k} \frac{E^t(i)}{\sum_{i'} E^t(i')}}{\sum_{i'} \frac{E_j^t(i')}{\sum_{i''} E_j^t(i')} V_{i,k}}$.
2. Compute interim relative labor demands $\frac{E_j^{t+1}(i)}{\sum_{i'} E_j^{t+1}(i')} = \sum_k B(i,k)^t$.
3. If converged, exit; else return to Step 1 and advance t.

This algorithm converges to a global optimum from any feasible starting point (Tishby, Pereira, and Bialek 2000). For fixed wages, there is one global optimum so the algorithm converges to the unique profit-maximizing strategy for each firm. Using the algorithm allows the researcher to estimate the model and perform counterfactuals without numerically searching for the firm's profit-maximizing strategy. Solving for a counterfactual equilibrium then consists of only two additional steps: solving for equilibrium product prices (a standard problem in industrial organization) and solving for the wages which clear the labor market.

A.6 Proof of Propositions 1

Proof. First consider any set of workers that differ only in their labor supply. Conditional on a task being assigned to this set of workers, who it is assigned to does not impact quality because all have the same skills. Take any two workers in the set and suppose they have different task assignment. There always exists a way to accomplish the same work but with each worker performing the same task assignment. This alternative task assignment is a less fine division of labor, and by the "distraction-free" property of mutual information (Tian 2019) it requires strictly less communication, contradicting optimality of the original task assignment.

Second, consider workers that have the same skill set but different skill level. If labor demand is held fixed, wage costs are sunk, and how a firm divides tasks among workers with the same skill set does not impact service quality. However, assigning all workers with the same skill set the same distribution of tasks strictly reduces communication and thus organization costs by the argument from the last paragraph. Therefore, workers with the same skill set but different skill level are assigned the same distribution of tasks.■

A.7 Proof of Propositions 2

Proof. First consider two workers that differ only in labor supply. Because they will be assigned the same distribution of time across tasks and they have the same skills, if they have different wages per unit of labor, firms will demand no labor from the one with the higher wage, and the labor market will not clear.

Second consider two workers with the same skill set but different skill level $\bar{\theta}_m > \bar{\theta}_{m'}$. The last result proves that conditional on being hired, a firm will assign both the same distribution of time across tasks. Therefore the impact of hiring m compared to m' on profit comes only through wage differences and skill level differences. Therefore firm j hires worker m over worker m' if: $w_m - w_{m'} \geq \rho^{-1}(\bar{\theta}_m - \bar{\theta}_{m'})$. Notice that this inequality is the same for all firms. This implies that labor markets do not clear unless wages are such that firms are indifferent: $w_m - w_{m'} = \rho^{-1}(\bar{\theta}_m - \bar{\theta}_{m'})$.■

A.8 Proof of Proposition 3

This result is proven for a general demand system under-which demand for product j is D_j. The only restriction is that demand for each product depend only on quality and price through and be strictly increase in a quality price index ($\xi_j - \rho - p_j$). Multinomial logit, nested logit and mixed logit with a non-random price coefficient all satisfy. Mixed logit with consumer price sensitivity heterogeneity would not. Variables denoted by $-j$ represent the vector of all competitor objects other than firm j.

For any given organization structure, the firm will choose prices only weakly above marginal cost; otherwise, it receives negative profit. Without loss, I therefore restrict the set of price-structure pairs considered to be those where price exceeds marginal cost. For this proof, I work in the space of organization structures, defined as: $B_j(i,k) := E_j(i)b_j(i,k)$. I also use the notation $\xi(B_j) := \sum_i \theta_i(k)B_j(i,k)$. The proof is performed without quality and marginal cost heterogeneity (ν_j, ω_j) for expositional convenience only.

First, I prove that if an organization structure B_j^* solves the simpler problem (Equation 2), then it is profit-maximizing ("only if" direction). I need to show that for any price-organization structure pair (p_j', B_j') there exists p_j such that profit under (p_j, B_j^*) is weakly higher than profit under (p_j', B_j'). I do this by construction. Denote B_j^* as a structure which solves Equation (2). Such a structure always exists because Equation (2) is a rate-distortion/rational inattention problem, as I prove in the following lemma.

Lemma 1 *Equation 2 is a rate-distortion or rational inattention problem.*

Proof of Lemma. For this proof, I work in the space of organization structures, defined as: $B_j(i,k) := E_j(i)b_j(i,k)$. Equation (9) from Theorem 1 can be rewritten as:

$$\gamma_j \min_{B_j \in \mathbb{B}} \left\{ I(B_j) + \gamma_j^{-1} \Big[W(B_j) - \rho^{-1}\xi(B_j) \Big] \right\}. \tag{10}$$

I can rewrite (10) as a maximization problem:

$$\max_{B_j \in \mathbb{B}} \left\{ \sum_{i,k} B_j(i,k)(\rho^{-1}\theta_{i,k} - W_i) \Big] - \gamma_j I(B_j) \right\}. \tag{11}$$

Comparing (11) to formulations in papers such as Jung et al. (2019) illustrates that this is a rational inattention problem with mutual information attention costs. I rewrite Equation 10 one last time:

$$\gamma_j \min_{B_j \in \mathbb{B}} \left\{ I(B_j) + \gamma_j^{-1} \sum_{i,k} B_j(i,k)(W_i - \rho^{-1}\theta_{i,k}) \Big] \right\}. \tag{12}$$

Comparing Equation (12) to formulations such as Equation 6 in Tishby, Pereira, and Bialek (2000) demonstrates this is a well-understood minimization problem from information theory called a rate-distortion problem.■

For any price p_j' and any structure B_j', I can construct $p_j = p_j' + \gamma_j I(B_j^*) + W(B_j^*) - \gamma_j I(B_j') - W(B_j')$. The price p_j is positive and therefore feasible. Recall that profit evaluated at (p_j, B_j^*) is

$$D_j(\xi(B_j^*) - \rho p_j, p_{-j}, \xi_{-j}) \Big[p_j - \gamma_j I(B_j^*) - W(B_j^*) \Big].$$

The second multiplicative term of profit is equal under (p_j, B_j^*) and (p_j', B_j'). The first term (demand) is strictly increasing in the quality-price index $\xi(B_j) - \rho p_j$; therefore, it is sufficient to show that this index is weakly higher for (p_j, B_j^*). I show this by rewriting $\xi(B_j^*) - \rho p_j$:

$$
\begin{aligned}
&= \xi(B_j^*) - \rho[p_j' + \gamma_j I(B_j^*) + W(B_j^*) - \gamma_j I(B_j') - W(B_j')] && (13)\\
&= \xi(B_j^*) - \rho[p_j' + \gamma_j I(B_j^*) + W(B_j^*) - \gamma_j I(B_j') - W(B_j')] + \xi(B_j') - \xi(B_j') && (14)\\
&= \xi(B_j') - \rho[p_j' + \gamma_j I(B_j^*) + W(B_j^*) - \gamma_j I(B_j') - W(B_j') - \rho^{-1}\xi(B_j^*) + \rho^{-1}\xi(B_j')] && (15)\\
&= \xi(B_j') - \rho p_j' - \rho[\gamma_j I(B_j^*) + W(B_j^*) - \rho^{-1}\xi(B_j^*) - \{\gamma_j I(B_j') + W(B_j') - \rho^{-1}\xi(B_j')\}] && (16)\\
&\geq \xi(B_j') - \rho p_j'. && (17)
\end{aligned}
$$

The second to last line occurs because B_j^* is a minimizer. This proves the "only if" direction. I now prove that if a structure B_j^* is profit-maximizing, it solves Equation (2) (the "if" direction). Suppose for sake of contradiction there exists B_j' which is profit maximizing but does not solve Equation (2). Then, as in the first part of the proof, there exists B_j^* which does solve Equation (2). Then I can construct p_j as before for any p_j' that is weakly higher than marginal cost under B_j'. However, because B_j' does not minimize Equation (2), $\xi(B_j^*) - \rho p_j > \xi(B_j') - \rho p_j'$, and thus profit is strictly higher under B_j^*, p_j. This contradicts optimality of B_j' and concludes the proof.

A.9 Proof of Proposition 4

Denote by Q the quality-adjusted wages. Denote by $I^*(Q)$ the optimal complexity as a function of quality-adjusted wages. Rate-distortion equivalence, proven in Lemma 1, implies $I^*(Q)$ is continuous, convex and decreasing. It is also strictly decreasing above some threshold $\bar{Q}$ (Moser and Chen 2012). The firm's choice of quality-adjusted wages solves

$$V := \min_Q \gamma I^*(Q) + Q.$$

The envelope theorem implies the index and thus profit are increasing in γ:

$$\frac{\partial V}{\partial \gamma} = I^*(Q) \geq 0.$$

taking the first-order condition:

$$\frac{dI^*(Q) + \gamma^{-1}Q}{dQ} = \frac{dI^*(Q)}{dQ} + \gamma^{-1} = 0 \implies \frac{dI^*(Q)}{dQ} = -\gamma^{-1}.$$

Because I^* is decreasing and convex, its derivative is negative and increasing. Therefore, Q which solves is increasing in γ. Thus profit and complexity will be positively correlated via γ.

A.10 Proof of Theorem 1

For this proof, I suppress the firm subscript j. I define $h_{i,k}$ as the fraction of task k performed by worker i. The first-order condition of (2) using this choice variable is givne by:

$$h_{i,k} = \frac{E_i}{Z(k,\lambda)} exp\bigg(- \lambda(\rho w_i - \theta_{i,k})\bigg).$$

Summing over i yields

$$\sum_i h_{i,k} = \frac{1}{Z(k,\lambda)} \sum_i E_i exp\bigg(- \lambda(\rho w_i - \theta_{i,k})\bigg) = 1.$$

Therefore,

$$Z(k,\lambda) = \sum_i E_i exp(-\lambda(\rho w_i - \theta_i \delta^{\mathbb{I}\{\kappa_i \neq k\}}))$$

and

$$h_{i,k} = \frac{e_i exp(\lambda(-\rho w_i + \theta_{i,k}))}{\sum_{i'} e_{i'} exp(\lambda(-\rho w_{i'} + \theta_{i,k}))}.$$

Substituting for λ yields

$$h_{i,k} = \frac{E_i exp(-\gamma^{-1} w_i + (\rho\gamma)^{-1}\theta_{i,k}))}{\sum_{i'} E_{i'} exp(-\gamma^{-1} w_{i'} + (\rho\gamma)^{-1}\theta_{i',k}))}.$$

By the definition of $h_{i,k}$,

$$b(i,k) = \alpha(k)h_{i,k}/E(i).$$

therefore:

$$b(i,k) = \alpha(k)h_{i,k}/E(i) = \frac{\alpha(k)exp(-\gamma^{-1}w_i + (\rho\gamma)^{-1}\theta_{i,k}))}{\sum_{i'} E_{i'}exp(-\gamma^{-1}w_{i'} + (\rho\gamma)^{-1}\theta_{i'}(k)))}.$$

This illustrates that optimal jobs take an almost-logit form. I can also derive this result by applying Theorem 1 from Matêjka and McKay (2015).

The fact that all hired worker types spend a positive amount of time on each task is a direct application of Lemma 1 from Jung et al. (2019). An increase in wage corresponds to a decrease in the "payoff" to the firm of using workers of type i in all tasks (i.e., states of the world in the rational inattention literature). This means I can apply Proposition 3 from Matêjka and McKay (2015) to say that an increase in w_i leads to a decrease in E_i all else constant. I can even say that E_i is strictly decreasing in w_i whenever the initial share of worker i is strictly interior, i.e., $0 < E_i < 1$.

A.11 Proposition 5

To recover the best responses of the firm's problem, I use the fact that the joint maximization of any function is equivalent to the sequential maximization. Thus I can write the firm's problem as

$$\max_{b_j,E_j} \max_{p_j}, \frac{exp(\xi(b_j,E_j)-p_j))}{1+\xi(b_{j'},E_{j'})-p_{j'})}\left(p_j - a_j\gamma_j\sum_i D_{KL}(b_j(i,\cdot)||\alpha_j) - a_j\sum_i w_iE_j(i) - \alpha_jc - \omega_j\right)$$

s.t.

$$\sum_{i,k} E_j(i)b_j(i,k) = \alpha_j(k)\forall k$$

I first study the inner pricing problem. Fixing an organization structure, the model reduces to a logit Bertrand game with heterogeneous costs and qualities. Proposition 7 of Caplin and Nalebuff (1991) proves that such a game has a unique pure-strategy Nash equilibrium in prices. Therefore, for any chosen organizational structure, there is a single best-response price.

I now move on to the choice of task assignments and relative labor demands (the outer maximization). In Proposition 3, I show that when prices are chosen to maximize profit, the internal organization choice separate from the pricing problem and solves:

$$\min_{b_j, E_j} \gamma_j \sum_i E_j(i) D_{KL}(b_j(i,k)||\alpha_j) + \sum_i E_j(i) w_i - \rho^{-1} \sum_i E_j(i) \sum_k \theta_i(k) b_j(i,k) \tag{18}$$

s.t. $\sum_{i,k} E_j(i) b_j(i,k) = \alpha_j(k) \forall k$

The best-response structure will therefore depend on other actions of the firm only through wages. I show in Appendix Section A.9 that this is a rational inattention problem with a mutual information cost function. I can appeal to Matêjka and McKay (2015) to say that there exists an organization structure which maximizes profit for each firm. This establishes equilibrium existence.

For uniqueness, the online Appendix of Matêjka and McKay (2015) contains a result which implies that whenever the following condition holds, the relative labor demands and task assignments will have a unique solution:

Assumption 2 *Define the wage-quality vector of a worker of type i at firm j as $v_{i,j} = \{exp(\gamma_j^{-1}(\rho^{-1}\theta_i(k) - w_i))\}_{k=1}^{K}$. The set of wage-quality vectors $\{v_{i,j}\}_{i \in \mathcal{I}}$ is affinely independent.*

Notice that the following can be re-written as:

$$exp(\gamma_j^{-1}(\rho^{-1}\theta_i(k) - w_i)) = (exp(\rho^{-1}\gamma_j^{-1}\theta_i(k)) exp(-w_i \gamma_j^{-1})) \tag{19}$$

To apply the result from Matêjka and McKay (2015), it is necessary to show that (19) is affinely independent. I have that the skill sets θ_i can be arranged into a positive definite matrix Θ. I begin with the following lemma.

I begin by showing that applying the function $f(x) = exp(\rho^{-1} \cdot x)$ element-wise to A is generates a positive definite.

Lemma 2 *Suppose a matrix A is positive definite. Given a scalar $a > 0$, the function $exp(a \cdot x)$ applied element-wise to A is also positive definite.*

Proof Suppose a matrix A is positive definite. We wish to show that given a scalar $a > 0$, the function $exp(a \cdot x)$ applied element-wise to A is also positive definite. To begin, note that A multiplied by a constant scalar is positive definite, because for any vector Z we have that $z^T a \cdot Az = az^T Az$ which since $z^T Az > 0$ is positive. Therefore we suppress a and consider whether $exp(x)$ applied elementwise to a positive definite matrix A is positive definite. Note that we can express $exp(x)$, by the power series, as:

$$exp(x) = \sum_{k=0}^{\infty} \frac{x^i}{k!}$$

Thus we can express the elementwise application of $exp(x)$ to A as the sum of elementwise products (Hadamard products) of A plus the addition of 1 to each element. That is, each element of the sum is a Hadamard product of A with itself k times multiplied by a positive scalar. By the Schur product theorem, the Hadamard product of two positive definite matrices is positive definite. Thus the Hadamard product of A with itself any number of times is positive definite. We have just shown that multiplying by a scalar preserves positive definiteness. Therefore every element of the series is positive definite, except of course the first element of the series (which is a matrix of all ones). Thus the sum of all the elements except the first is positive definite. Now take a positive definite matrix B and add one to every element. Then we have that for any vector v: $\sum_{i,k}(B_{i,k} + 1)v_i v_k = \sum_{i,k} B_{i,k} v_i v_k + (\sum_i v_i)^2$. Since the first term is strictly positive and the second is weakly positive then we have that a positive definite matrix with 1 added to each element is also positive definite. Thus $exp[a \cdot A]$ is positive definite.■

Part of this Lemma is proven in Horn and Johnson (2012) Theorem 7.5.9. Using this lemma, the function $exp(\rho^{-1}\gamma_j^{-1} \cdot x)$ applied element-wise to Θ generates a positive definite matrix, call it Θ'. Positive definite matrices are full rank (linearly independent rows). Multiplying by $exp(-w_i\gamma_j^{-1})$ in (19) is equivalent to multiplying each row i of Θ' by the scalar $exp(-\gamma_j^{-1} w_i)$. This is a scaling of each row by a non-zero constant, which is an elementary row operation, which preserves the rank of the matrices and therefore linear independence. Thus the collection of vectors is linearly independent, which implies affine independence, and Assumption 2 is satisfied.

Notice that this result is true for any wage vector w where no wages are 0. Thus there is a unique task assignment and relative labor demand for all strictly positive wage vectors w. Combined with the earlier result, I have the result that there is a unique Nash equilibrium

A.12 Proof of Theorem 2

Proof. I first set aside all firms where one of the K task types is not performed at all ($\exists k s.t. \alpha_j(k) = 0$) or where the s-index is 0 ($I_j = 0$). Until stated otherwise, I work with only a dataset of firms where all tasks are performed and the s-index is strictly positive. Propositions 1 and 2 establishes that all workers at the same firm are assigned the same distribution of time across tasks if and only if they have the same skill set. As a result, time use data allows workers to be grouped into mutually exclusive types within each firm. That is, we can group all workers with different skill sets at a firm into representative workers. However, these representative workers are not comparable across firms. If I observe 2 groups of workers at firm A and 2 at firm B, I do not know which if any of the groups are the same skill set. I know only that group 1 and group 2 at firm A have different skill sets than each other, and group 1 and group 2 at firm B have different skill sets than each other. Note that the firms currently being analyzed all have strictly positive s-index, which implies they will employ at least two worker types.

In order to partition all workers in the market into skill sets, I need to group workers across firms. The key challenge is that the task content of a worker's job depends on both the worker's skill and the firm at which the worker works. Given a worker l at firm j with unknown skill set t_l, I can capture this issue using the characterization from Theorem 1:

$$b_j(t_l, k) = \alpha_j(k) \frac{exp(-\gamma^{-1} w(t_l) + (\rho\gamma_j)^{-1}\theta_{t_l}(k)))}{\sum_{i'} \frac{E_j(i')}{\sum_{i''} E_j(i'')} exp(-\gamma_j^{-1} w(i') + (\rho\gamma)^{-1}\theta_{i'}(k))}$$

To identify t_l, recall that at this stage workers with different types are grouped into mutually exclusive groups within the firm. At any firm with at least two worker skill sets, I can divide the distribution of time across tasks of two workers who are known to be different skill sets and take the logarithm to obtain:

$$log\left(\frac{b_j(t_l, k)}{b_j(t_{l'}, k)}\right) = (\rho\gamma_j)^{-1}\Big([\theta_{t_l}(k) - \rho w(t_l)] - [\theta_{t_{l'}}(k) - \rho w(t_{l'})]\Big)$$

This expression does not depend on firm-specific factors except for the fact that it is scaled by the reciprocal of organization costs γ_j. It is also defined for all K tasks and for every coworker who is a different skill set. Thus every worker has several such log-ratio vectors, representing how similar the task content of their job is to all of their coworkers. The following lemma, which uses the fact that wage-adjusted skills are linearly independent, establishes that the normalized log ratio between two pairs of workers at a firm matches if and only if the skill sets of all workers match.

Lemma 3 *Suppose worker 1 and 2 are at firm j and workers 3 and 4 are at firm j'. Suppose firm j and j' are in the same market. The log ratio of relative time-use between workers 1 and 2 divided by its Euclidean norm is equal to the log ratio of relative time-use between workers 3 and 4 divided by its Euclidean norm if and only if worker 1 is the same type as worker 3 and worker 2 is the same type as worker 4.*

Proof. Normalizing the log ratio removes the firm-specific object γ_j:

$$\frac{log\left(\frac{\frac{\tilde{B}_j(t_l,k)}{\tilde{E}_i}}{\frac{\tilde{B}_j(t_{l'},k)}{\tilde{E}_{t_{l'}}}}\right)}{\left|\{log\left(\frac{\frac{\tilde{B}_j(t_l,k')}{\tilde{E}_i}}{\frac{\tilde{B}_j(t_{l'},k')}{\tilde{E}_{t_{l'}}}}\right)\}_{k'=1}^{K}\right|} = \frac{(\rho\gamma_j)^{-1}\left[\tilde{\theta}_{t_l}(k)-\tilde{\theta}_{t_{l'}}(k)\right]}{\left(\sum_{k'}\left((\rho\gamma_j)^{-1}\left[\tilde{\theta}_{t_l}(k')-\tilde{\theta}_{t_{l'}}(k')\right]\right)^2\right)^{1/2}}$$

$$\frac{log\left(\frac{\frac{\tilde{B}_j(t_l,k)}{\tilde{E}_i}}{\frac{\tilde{B}_j(t_{l'},k)}{\tilde{E}_{t_{l'}}}}\right)}{\left|\{log\left(\frac{\frac{\tilde{B}_j(t_l,k')}{\tilde{E}_i}}{\frac{\tilde{B}_j(t_{l'},k')}{\tilde{E}_{t_{l'}}}}\right)\}_{k'=1}^{K}\right|} = \frac{\left[\tilde{\theta}_{t_l}(k)-\tilde{\theta}_{t_{l'}}(k)\right]}{\left(\sum_{k'}\left[\tilde{\theta}_{t_l}(k')-\tilde{\theta}_{t_{l'}}(k')\right]^2\right)^{1/2}}$$

From this expression it is clear that for any fixed pair of types $t_l, t_{l'}$ the normalized log ratios will be the same at all firms. This proves one direction. I now prove that if the normalized log ratios are the same, the types correspond. Suppose for sake of contradiction that the normalized log ratios correspond but the types are different. Then:

$$\frac{\left[\tilde{\theta}_{t_1}(k)-\tilde{\theta}_{t_2}(k)\right]}{\left(\sum_{k'}\left[\tilde{\theta}_1(k')-\tilde{\theta}_{t_2}(k')\right]^2\right)^{1/2}} = \frac{\left[\tilde{\theta}_{t_3}(k)-\tilde{\theta}_{t_4}(k)\right]}{\left(\sum_{k'}\left[\tilde{\theta}_3(k')-\tilde{\theta}_{t_4}(k')\right]^2\right)^{1/2}}$$

Denoting the denominators as $\frac{1}{A}$ and $\frac{1}{B}$:

$$A\left[\tilde{\theta}_{t_1} - \tilde{\theta}_{t_2}\right] = B\left[\tilde{\theta}_{t_3} - \tilde{\theta}_{t_4}\right] \leftrightarrow A\tilde{\theta}_{t_1} - A\tilde{\theta}_{t_2} - B\tilde{\theta}_{t_3} + B\tilde{\theta}_{t_4} = 0$$

This can be re-arranged as saying that one wage-adjusted skill can be written as a combination of other wage-adjusted skills:

$$\tilde{\theta}_{t_1} = \frac{B}{A}\tilde{\theta}_{t_3} - \frac{B}{A}\tilde{\theta}_{t_4} - \tilde{\theta}_{t_2}$$

Note that $t_1 \neq t_2, t_3 \neq t_4$. This is a valid linear combination and therefore a contradiction unless $t_1 = t_3, t_2 = t_4$ in which case $A = B$ and we have that all coefficients are 0. Even in the case when $t_1 = t_4, t_2 = t_3$ this is a contradiction because then $B = A$ but the coefficients are $2, -2$. Thus the ratios coincide if and only if the types of the workers coincide.■

With this result in hand, I can classify two workers as the same skill set if one of their log ratio vectors matches. By repeating this process, I can partition all workers into the N mutually exclusive skill set groups except for unconnected workers. Unconnected workers are those workers that have coworkers not of a skill set that can be linked to the other firms. For example, if many firms employ skill set 1 and 2, and only one firm employs skill set 2 and 3 it is not possible to link skill set 2 workers at this final firm to the rest of the skill set 2 workers. This type of disconnectedness resolves in large samples and is not a threat to identification. One sufficient condition is that there exists a firm which employs all worker types.

At this point in the proof, workers are classified into their skill set group. Although the skill set parameters and wages are not yet identified, the labeling itself is arbitrary. As a result, I can treat the index of these workers' skill set as identified, and jobs $b_j(i,k)$ as data for the rest of the proof. I now identify the market parameters Ω. It is convenient to define $B_j(i,k) := \frac{E_j(i)b_j(i,k)}{\sum_{i'} E_j(i')}$, the share of total labor assigned to worker i and task k. This object is identified and can be considered data in what follows because $b_j(i,k), E_j(i)$ are identified. Under the demand system, market shares can be written as:

$$log\left(\frac{s_j}{s_0}\right) = \sum_{i,k} \theta_i(k) a_j B_j(i,k) - \rho p_j + \nu_j \tag{20}$$

The theoretical results imply the firm's task-assignment strategy ($\{b_j(i,k)\}_{i,k}$) and relative labor demands ($\frac{E_j(i)}{\sum_{i'} E_j(i')}$) do not depend on ν_j or prices p_j. They do depend on the task mix (α_j) and organization costs (γ_j) but I assumed that these objects are independent of ν_j. Therefore $B_j(i,k)$ is independent of ν_j and I have the moment conditions $E[\nu_j B_j(i,k)] = 0$. Prices are set strategically, and they depend directly on ν_j. However, the model provides a natural instrument: relative organization costs ($\frac{\gamma_j}{\gamma_1} I_j$). This object has already been identified and can therefore be considered data. It is independent of ν_j because organization cost parameters are independent of ν_j and the amount of communication is proven to not depend on ν_j. It also has a positive relationship with prices because it increases marginal costs. Thus I obtain the moment condition $E[\nu_j \frac{\gamma_j}{\gamma_1} I_j]$. Collecting these I obtain the following set of moment conditions:

$$E\left[\left(log\left(\frac{s_j}{s_0}\right) - \sum_{i,k} a_j \theta_i(k) B_j(i,k) - \rho p_j\right) \begin{pmatrix} \{B_j(i,k)\}_{i,k}^{N,K} \\ \frac{\gamma_j}{\gamma_1} \gamma_j I_j \end{pmatrix}\right] = 0 \tag{21}$$

This is a linear system of $N \times K + 1$ equations and $N \times K + 1$ unknowns, so there exists a single solution under a standard full rank condition on the moments.[30] Thus the skill set parameters (Θ) and price sensitivity (ρ) are identified. To identify the remaining parameters, consider the firm's first-order condition for price, which takes the standard Bertrand-logit form of a markup plus marginal cost:

$$p_j = \frac{1}{\rho(1-s_j)} + \gamma_1 \frac{\gamma_j}{\gamma_1} a_j I_j + \sum_{i,k} w_i a_j \frac{E_j(i)}{\sum_{i'} E_j(i')} + \omega_j$$

Price sensitivity has already been identified, so the markup $\frac{1}{\rho(1-s_j)}$ is identified. I can therefore adjust prices by the markup to obtain marginal costs:

$$p_j - \frac{1}{\rho(1-s_j)} = \gamma_1 \frac{\gamma_j}{\gamma_1} a_j I_j + \sum_{i,k} w_i a_j \frac{E_j(i)}{\sum_{i'} E_j(i')} + \omega_j \tag{22}$$

The theoretical results imply the firm's relative labor demands ($\frac{E_j(i)}{\sum_{i'} E_j(i')}$) and relative organization cost ($\frac{\gamma_j}{\gamma_1} I_j$) do not depend on ω_j or prices p_j. They do depend on the task mix

30. This is the condition: $rank\{E[\left(\{B_j(i,k)\}_{i,k}^{N,K} \quad \frac{\gamma_j}{\gamma_1} I_j\right) \left(\{B_j(i,k)\}_{i,k}^{N,K} \quad \frac{\gamma_j}{\gamma_1} I_j\right)']\} = N \times K + 1$

(α_j) and organization costs (γ_j) but I assumed that these objects are independent of ω_j. Therefore both are independent of ω_j and I have the moment conditions $E[\omega_j \frac{E_j(i)}{\sum_{i'} E_j(i')}] = 0$ and $E[\omega_j \frac{\gamma_j}{\gamma_1} I_j] = 0$. Collecting these I obtain the following set of moment conditions:

$$E\left[\left(\tilde{p}_j - \gamma_1 \frac{\gamma_j}{\gamma_1} a_j I_j - \sum_i w_i a_j \frac{E_j(i)}{\sum_{i'} E_j(i')}\right) \begin{pmatrix} \{a_j \frac{E_j(i)}{\sum_{i'} E_j(i')}\}_i^N \\ \frac{\gamma_j}{\gamma_1} I_j \end{pmatrix}'\right] = 0 \tag{23}$$

This is a linear system of $N+1$ equations and $N+1$ unknowns, so there exists a single solution under the standard full rank condition on the moments. Thus the skill-set specific wages ($\{w_i\}_i^N$ and the organization cost of the reference firm (γ_1) are identified under a standard rank condition on the moments.[31] Thus all of the parameters in Ω are identified.

All that remains to be identified are the organization cost parameters and skill sets of workers at firms that either (1) do not use all task types but have a strictly positive s-index or (2) have an s-index of 0.

I begin with the firms in group (1). By Proposition 4, the s-index I_j is strictly decreasing in γ_j until it reaches 0. The s-index chosen by the firm for each γ_j is known because all market parameters are now identified. Therefore, the organization cost parameter γ_j is identified for these firms as the value where the model predicted level of communication matches the observed s-index. Task-assignment strategies are unique by Proposition 5, therefore identification of γ_j implies identification of the skill sets of the workers employed at the firm.

Now I address (2). By Proposition 4, the organization cost parameter of these firms is only set identified: γ_j can be any number above some threshold $\bar{\gamma}_j$. Because the s-index is strictly decreasing in γ_j below this threshold, the threshold is identified. The firm's task-assignment strategy (and therefore the composition of its workforce) is the same for all $\gamma_j > \bar{\gamma}_j$, therefore since task-assignment strategies are unique by Proposition 5 identification of the threshold $\bar{\gamma}_j$ implies identification of the skill sets of the workers employed at the firm.■

31. This is the condition: $rank\left\{E\left[\left(\{\frac{E_j(i)}{\sum_{i'} E_j(i')}\}_i^N \quad \frac{\gamma_j}{\gamma_1} I_j\right)\left(\{\frac{E_j(i)}{\sum_{i'} E_j(i')}\}_i^N \quad \frac{\gamma_j}{\gamma_1} I_j\right)\right]\right\} = N \times K + 1$

A.13 Closed-Form Logit Price Expression

Demand for a product j is given by

$$s_j(p_j) = \frac{exp(-\rho p_j + \xi_j)}{\sum_{j'=0}^{J} exp(-\rho p_{j'} + \xi_{j'})}.$$

Optimal pricing in a Bertrand Nash equilibrium with single-product firms is then given by

$$p_j = MC_j + \frac{1}{\rho(1 - s_j(p_j))}.$$

I now follow the arguments laid out in Aravindakshan and Ratchford (2011). I rewrite this expression as

$$p_j = c_j + \frac{1}{\rho(1 - \frac{exp(-\rho p_j + \xi_j)}{exp(-\rho p_j + \xi_j) + \sum_{j' \neq j} exp(-\rho p_{j'} + \xi_{j'})})}.$$

I rewrite it again as

$$p_j = c_j + \frac{1}{\rho} + \frac{exp(-\rho p_j + \xi_j)}{\rho \sum_{j' \neq j} exp(-\rho p_{j'} + \xi_{j'})}.$$

Multiplying by ρ and subtracting ξ_j yields

$$\rho p_j - \xi_j = \rho c_j + 1 + \frac{exp(-\rho p_j + \xi_j)}{\sum_{j' \neq j} exp(-\rho p_{j'} + \xi_{j'})} - \xi_j.$$

Now denote

$$E_j = \sum_{j' \neq j} exp(-\rho p_{j'} + \xi_{j'})$$

$$\frac{exp(-\rho p_j + \xi_j)}{E_j} + \xi_j - \rho p_j = -1 - \rho c_j + \xi_j$$

$$exp\left(\frac{exp(\xi_j - \rho p_j)}{E_j}\right) exp\left(\xi_j - \rho p_j\right) E_j^{-1} = exp\left(-1 + \xi_j - \rho c_j\right) E_j^{-1}$$

and

$$\tilde{W} = exp\left(\xi_j - \rho p_j\right) E_j^{-1}.$$

Then the expression becomes

$$\tilde{W}e^{\tilde{W}} = exp\Big(-1 + \xi_j - \rho c_j \Big) E_j^{-1}.$$

The left-hand side expression is the form required by Lambert's W, so the $\tilde{W}$ which solves is given by Lambert's W function of the right-hand side by definition. Thus optimal price solves

$$W\Bigg(exp\Big(-1 + \xi_j - \rho c_j \Big) E_j^{-1} \Bigg) = exp\Big(\xi_j - \rho p_j \Big) E_j^{-1}.$$

A property of this function is that $log(W(x)) = log(x) - W(x)$. Using this fact yields

$$-1 + \xi_j - \rho c_j - log(E_j) - W\Bigg(exp\Big(-1 + \xi_j - \rho c_j \Big) E_j^{-1} \Bigg) = \xi_j - \rho p_j - log(E_j),$$

which can be solved for the optimal price:

$$\frac{1}{\rho} + c_j + \rho^{-1} W\Bigg(exp\Big(-1 + \xi_j - \rho c_j \Big) E_j^{-1} \Bigg) = p_j^*. \tag{24}$$

A.14 Counterfactual Procedure

In order to perform counterfactuals, it is necessary to impose additional assumptions. First, the data contain only a small fraction of the total set of salons, but to solve for a new equilibrium I must understand how all firms respond. I do this by assuming there are n copies of each salon in the data, where n is set to be the number such that the sum of market shares in the data multiplied by n equals 1 minus the number of consumers choosing the outside option. This has clear limitations, most notably that it overestimates the level of competition faced by the salons in the data, who are positively selected relative to their competitors.

Second, I assume the total labor supplied by each skill set is the sum of labor demands under the estimated wages across all firms scaled by the weight. This is line with the assumption of workers in-elastically supplying labor.

I set the initial benchmark to be the model fully solved given the total labor supplies just calculated. This is different than using the values from the estimated model and data, because it imposes the full structure of the model, in particular because it involves solving the Bertrand pricing game.

Each salon-quarter has exogenous quality heterogeneity (ν_j) and marginal cost heterogeneity (ω_j). I estimate ν_j as the log relative market share less the skill set multiplied by the observed task assignment plus price component (ρp_j). I estimate ω_j as the price observed in the data less organization costs and relative wages. I fix these exogenous components across counterfactuals, with the caveat that I truncate the estimate if it would cause a negative predicted price. By fixing quality and cost heterogeneity, I am implicitly fixing the skill level of workers employed at each salon.

For each counterfactual, I begin by guessing an initial wage for each skill set. For the reorganization equilibrium, I first solve each salon's task assignment problem via the contraction mapping given in Appendix Section A.5. For the reallocation equilibrium, I fix the salon's task assignments at their initial position. I then iterate on each firm's Bertrand pricing best response function until I achieve convergence. I use the formulation derived in Appendix Section A.13. Using the implied labor demand from each firm, I check whether the labor market clears. If it does not, I repeat with a new wage guess.

Solving for wages in this model is complicated by the rich set of possible labor-labor substitution patterns, and the two nested loops involved. I address these challenges using the BB package developed in Varadhan and Gilbert (2010). The package provides a root finding algorithm that works well for this model, specially given the frequent failure of traditional gradient based methods.

A.15 A Quantity-Based Model

In some contexts, such as manufacturing, one may wish to model organizational efficiency as allowing firms to produce greater quantity rather than greater quality. Indeed, this is the default definition of productivity in economics. The model can also be extended to accommodate this: one can simply interpret the skill sets as denoting the amount of time required by the worker to complete task k (therefore smaller $\theta_{i,k}$ are better). Then the production function becomes a function of organization structure:

$$F_{\alpha,B}(a_j) = \min\left\{\frac{a_1}{\alpha_1 \sum_i \theta_{i,1} B_j(i,1)}, \ldots \frac{a_k}{\alpha(k) \sum_i \theta_{i,k} B_j(i,k)}, \ldots, \frac{a_K}{\alpha(k) \sum_i \theta_{i,K} B_j(i,K)}\right\}.$$

Given any fixed organizational structure, the efficient way to produce a single unit of output is to set $a_k = \alpha(k)\sum_i \theta_{i,k}B_j(i,k)$. Thus the per-unit wage bill is given by

$$\sum_i W_i \sum_k \alpha(k) \sum_i \theta_{i,k}B_j(i,k).$$

Marginal costs are constant and consist of the per-unit wage bill and organization costs:

$$MC_j = \sum_i w_i \sum_k \alpha(k) \sum_i \theta_{i,k}B_j(i,k) + \gamma_j I(B_j).$$

All of the benefits of a more complex organization come through a reduction in the per-unit wage bill. In this way, the intuition from the original model extends directly to the quantity case: firms with greater organizational efficiency (lower γ_j) can produce more of the good with the same workforce. I did not use this as the main model because the following property is not compatible with the empirical application to hair salons:

Proposition 7 *Under a quantity model with multinomial logit demand, prices are decreasing with organizational complexity.*

The proof of this proposition is given in the next paragraph. Intuitively, under the quantity model with logit demand, all the benefits of a complex organization come from greater output rather than from greater revenue per unit. The reduction in marginal cost outpaces the increase in the markup, resulting in lower prices. This implies a negative correlation between prices and complexity, which is shown not to be true for hair salons. However, for manufacturing firms, it appears to be true. Caliendo et al. (2020) finds that prices (revenue-based productivity) decline when manufacturing firms reorganize.

Proof. Equation 24 from Appendix Section A.13 provides a closed-form expression for price in any Nash Equilibrium under logit demand:

$$\frac{1}{\rho} + c_j + \rho^{-1}W\Big(exp\Big(-1+\xi_j-\rho c_j\Big)E_j^{-1}\Big) = p_j^*.$$

Taking the derivative w.r.t. c_j yields

$$\frac{\partial p_j^*}{\partial c_j} = 1 - exp\Big(-1+\xi_j-\rho c_j\Big)E_j^{-1}W'\Big(exp\Big(-1+\xi_j-\rho c_j\Big)E_j^{-1}\Big).$$

A property of the Lambert W function is that

$$W'(x) = \frac{W(x)}{(1 + W(x))x}.$$

Thus, I can simplify the expression to

$$\frac{\partial p_j^*}{\partial c_j} = 1 - \frac{W(exp[-1 + \xi_j - \rho c_j]E_j^{-1})}{1 + W(exp[-1 + \xi_j - \rho c_j]E_j^{-1})}.$$

The Lambert W function is weakly positive for values which are weakly positive; therefore, the derivative is positive, and price is decreasing in cost. The firm minimizes cost:

$$\min_{B \in \mathbb{B}} \gamma I(B_j) + W(B_j).$$

This is again a rate-distortion problem. Denoting the optimal wage-bill as $D = W(B_j^*)$, I can reformulate the problem as before, with the firm choosing D given some optimal organization cost and wage bill:

$$\min_{D} \gamma I(D) + W(D),$$

where I and W are expressed as functions of D instead of B_j. Then, as before, there is a negative cross-partial derivative:

$$\frac{\partial \gamma I(D) + W(D)}{\partial D \partial \gamma} = I'(D) < 0$$

with strict inequality whenever $I(D)$ is strictly positive. This establishes strict decreasing differences of D in γ; thus D is strictly decreasing in γ, and since $I(D)$ is a strictly decreasing function, it is also strictly decreasing in γ. Therefore, prices should be decreasing as γ decreases, while complexity should be increasing.

A.16 Task Classification Process: Further Details

A licensed cosmetologist was paid to categorize 20,560 salon services performed according to their descriptions. As part of the agreement, the person provided a picture of their cosmetology license. The cosmetologist was provided with a blank spreadsheet with pre-defined subcategories and was instructed to mark all subcategories where the description matched with a 1. They were instructed that some subcategories may not be mutually exclusive, so they should mark all that applied. The initial job description was as follows:

> I have a list of approx. 20,560 short descriptions of salon services (mainly hair salons, but also some nail/spas). I would like someone with knowledge of the industry to mark whether each descriptions fits into one of several categories (male/female service, coloring, cutting, highlighting, washing, etc). This amounts to putting a 1 in each column that fits the description.

In a follow-up message I further clarified the instructions:

> Here are the descriptions. I did the first few to give you a sense of the task. Basically read the description and then put a 1 in all categories that fit. Sometimes a description may match many, sometimes 1, rarely none. If you start reading them and see that it may be worth adding a separate category let me know. The idea though is to capture the core "tasks" or services performed at hair salons, like cut, color, highlight, style, etc and also to get some info on gender and typos.

After the first draft was submitted, I checked the coding, looking for any mistakes or missed descriptions, and sent the document back to the cosmetologist several times for revision. A sample from the final spreadsheet is displayed in Figure A4.

	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	Service Description	Typo	Cut	Shave	Style	Extensions	Color/Bleach	Highlight/Balayage	Other Treatment	Blowdry	Wash/Shampoo	Eyebrow/Eye Lash Service	Admin/Consult	Spa Service	Nail service	Male	Female	Child	No Info
1319	'Add-on Camo								1				1						
1320	'Add-on Clipper cut (for booking with other F		1										1			1			
1321	'Add-on Color (1OZ)						1						1						
1322	'Add-on Color Melt						1						1						
1323	'Add-on Curling Iron				1								1						
1324	'Add-on Extra 20oz Lightener						1						1						
1325	'Add-on Haircut		1										1				1		
1326	'Add-on Haircut - (Thick hair)		1										1				1		
1327	'Add-on Haircut - with Color		1										1				1		
1328	'Add-on Haircut -(Fine Hair)		1										1				1		
1329	'Add-on Iron Service				1								1				1		
1330	'Add-on Kids Cut (when added to service of a		1										1					1	
1331	'Add-on Lightener						1						1						
1332	'Add-on Olaplex						1						1						
1333	'Add-on Pixie cut		1										1				1		
1334	'Add-on Salon Metal Detox												1						
1335	'Add-on Scalp Treatment								1				1						
1336	'Add-on Service Only												1						
1337	'Add-on Toner						1						1						
1338	'Add-on Treatment								1				1						
1339	'Added Base						1						1						
1340	'Added Biolage Color						1						1						
1341	'Added Color For Eyebrows											1	1						
1342	'Added Deep Conditioner								1				1						
1343	'Added Elumen Glaze								1				1						
1344	'Added Glaze								1				1						
1345	'Added Highlight							1					1						
1346	'Added K18								1				1						
1347	'Added Nectaya Single												1						

Figure A4: Final Task Subcategorization Spreadsheet from Cosmetologist

Since the subcategories were very detailed, I hired the same cosmetologist, at a rate of $100, to classify the subcategories into six task categories. The specific instructions given to the cosmetologist were as follows:

> Please categorize the 13 tasks from before into "groups." For the 6 group column, put the 13 tasks into 6 groups that are most similar in terms of who would do them/tasks they would require. for example, if color and highlight are similar, mark both as

> number 1. Number the groups 1 through 6. For the four group column, make 4 groups, etc. Underneath, please write a small note describing why you put the tasks together the way you did.

I use the six-category grouping provided by the cosmetologist with one modification: I combine the extension task with the blow-dry task to create five final task categories, because the extension task is very sparse–for Manhattan in 2021 Q2, fewer than 10 hours were dedicated to this task. This sparsity leads to estimation problems, as parameters tied to this task have a negligible effect on observable outcomes.

A.17 Measurement Error in the S-index

Complexity is estimated based on the observed task assignments within firm, yet the empirical part of this paper treats complexity as if it were observed or measured without error. One justification is that many assignments are observed per firm per quarter, so estimation error should be small. If estimation error at the quarter level is small, the correlation between complexity measures at the month level within quarter should be large. This section illustrates that this is indeed the case.

To do this, I recompute complexity for each month within a quarter so that I have three measurements of complexity per salon-quarter observation. In the full sample, the pairwise correlation between the first and second month is 0.945, the first and third is 0.98, and the second and third is 0.939. When 2020 (the onset of the coronavirus pandemic) is excluded, the pairwise correlations are 0.978, 0.962 and 0.976, respectively. The high correlation between complexity measurements within quarters suggests that complexity at the quarter level is measured precisely.

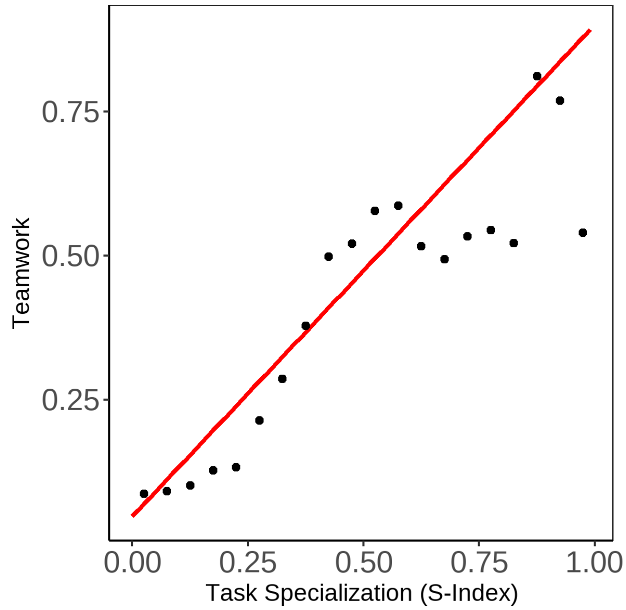
A.18 Supplementary Tables and Figures

Table A1: Auxiliary Parameter Estimates

County	Parameter	2018Q1	2018Q2	2018Q3	2018Q4	2019Q1	2019Q2	2019Q3	2019Q4	2020Q1	2020Q4	2021Q1	2021Q2
All	Material Cost Administrative	-6098.768	-4265.294	-2514.718	-3946.164	-1190.256	-1766.626	223.741	-13.654	-977.093	-472.657	65.660	323.883
All	Material Cost Administrative	(3663.985)	(2734.785)	(1371.918)	(1790.838)	(1043.558)	(1217.646)	(445.846)	(424.071)	(965.873)	(710.394)	(527.082)	(293.799)
All	Material Cost Blowdry/Style/Etc.	1483.836	1188.110	553.194	1868.320	-142.391	292.990	16.451	297.970	28.164	-401.204	550.294	373.979
All	Material Cost Blowdry/Style/Etc.	(3184.089)	(1389.424)	(859.982)	(1335.990)	(223.318)	(665.346)	(297.786)	(392.240)	(208.693)	(541.656)	(571.934)	(256.142)
All	Material Cost Color/Highlight/Wash	-1546.921	-207.271	-2117.150	-2739.052	-456.463	-880.497	-157.536	22.821	130.361	-502.550	26.707	465.137
All	Material Cost Color/Highlight/Wash	(1573.211)	(1215.911)	(1308.662)	(1508.346)	(363.753)	(878.046)	(321.387)	(460.486)	(273.499)	(485.579)	(426.057)	(252.031)
All	Material Cost Nail/Misc.	2790.696	2417.343	-2033.100	-4277.314	-1798.830	0.656	-2371.790	-2323.012	-1648.472	-2114.878	-1024.013	-144.601
All	Material Cost Nail/Misc.	(7160.424)	(1819.574)	(2914.020)	(3207.783)	(787.872)	(2691.813)	(1336.522)	(1649.135)	(1144.572)	(1021.066)	(574.367)	(367.303)
Cook	Cost Level	426.337	49.647	253.919	747.773	178.851	646.232	-61.117	40.854	145.273	422.205	-129.679	-866.647
Cook	Cost Level	(565.251)	(.552)	(.722)	(1.315)	(.843)	(1.056)	(.936)	(.836)	(.659)	(1.882)	(1.736)	(1.531)
Cook	Demand Level	-	-0.951	-0.352	-1.615	-1.359	-3.108	-3.035	-1.882	-3.455	-1.424	-1.122	-1.178
Cook	Demand Level	-	(.552)	(.722)	(1.315)	(.843)	(1.056)	(.936)	(.836)	(.659)	(1.882)	(1.736)	(1.531)
Cook	Wage Level	403.538	314.855	305.628	387.967	232.118	173.868	207.819	181.394	163.145	135.729	187.753	377.364
Cook	Wage Level	(174.290)	(139.554)	(116.638)	(124.465)	(73.895)	(70.763)	(65.860)	(62.460)	(67.737)	(57.116)	(193.166)	(141.559)
Los Angeles	Cost Level	830.870	136.964	1202.044	1416.270	414.613	420.455	16.107	77.849	73.637	-152.968	-247.498	-158.846
Los Angeles	Cost Level	(811.759)	(.581)	(.576)	(.643)	(.617)	(.626)	(.683)	(.773)	(.883)	(.868)	(.826)	(.857)
Los Angeles	Demand Level	-	-0.719	-1.031	-1.191	-0.269	-0.417	0.044	-0.276	-0.734	-1.569	-0.925	0.008
Los Angeles	Demand Level	-	(.581)	(.576)	(.643)	(.617)	(.626)	(.683)	(.773)	(.883)	(.868)	(.826)	(.857)
Los Angeles	Wage Level	17.702	17.702	19.202	19.202	19.202	19.202	125.198	20.452	20.452	269.782	166.130	37.436
Los Angeles	Wage Level	(21.145)	(21.425)	(93.164)	(33.585)	(38.293)	(28.355)	(104.740)	(107.051)	(34.292)	(181.314)	(111.114)	(58.224)
New York	Cost Level	290.341	24.061	1070.865	1149.234	345.135	353.060	130.273	-601.465	-86.848	387.032	-38.754	-354.297
New York	Cost Level	(1796.466)	(.819)	(1.200)	(1.249)	(.955)	(1.407)	(1.041)	(1.591)	(.962)	(.961)	(1.256)	(1.056)
New York	Demand Level	-	0.257	0.928	1.410	0.835	1.812	0.316	1.324	-0.447	-0.404	-0.806	0.576
New York	Demand Level	-	(.819)	(1.200)	(1.249)	(.955)	(1.407)	(1.041)	(1.591)	(.962)	(.961)	(1.256)	(1.056)
New York	Wage Level	371.681	178.154	178.154	178.154	179.654	179.654	179.654	564.807	226.201	181.154	181.154	214.664
New York	Wage Level	(778.278)	(147.109)	(190.426)	(325.077)	(86.678)	(240.757)	(109.651)	(217.530)	(93.957)	(65.736)	(82.948)	(127.463)

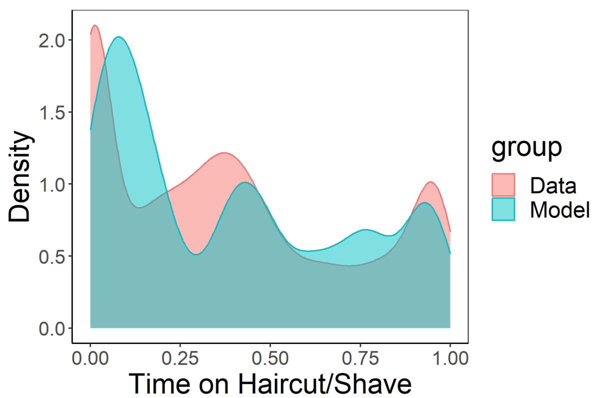
Note: The parameters associated with material costs, demand levels, wage levels and cost levels. Standard errors are from 631 bootstrap replications.

Figure A5: Binned Scatter Plot of Teamwork and Task Specialization

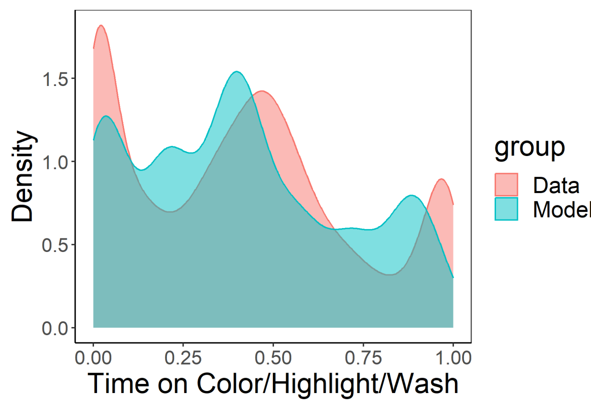


Note: The plot displays the average of the teamwork variable across bins of size 0.05 of the s-index. The red line is from a linear regression of teamwork on the s-index. There is a positive association between the two variables, with a slope close to 1.

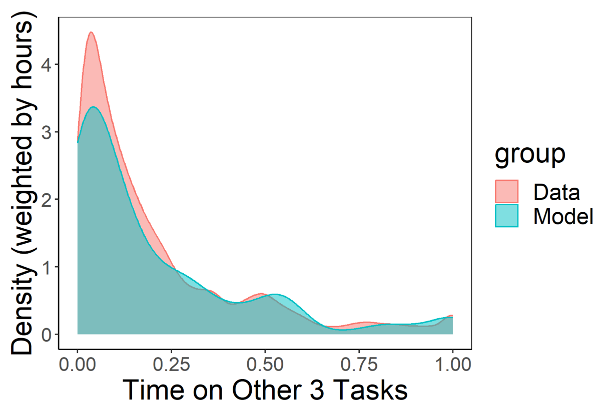
Figure A6: Model Validation: Marginal task Distributions



(a) Haircut/Shave



(b) Color/Wash



(c) Other 3 Categories

Note: The distribution of tie spent on each task, where density is based on the amount of time assigned to the particular worker-quarter in the data and firm-skill set-quarter in the model.

Table A2: Counterfactual Productivity and Wage Effects by Worker Skill Set

		Skill Set 1		Skill Set 2		Skill Set 3		Skill Set 4		Skill Set 5	
Counterfactual	County	Prod.	Wage	Prod.	Wage	Prod.	Wage	Prod.	Wage	Prod.	Wage
Immigration	Cook	0.003	-0.001	0.018	0.001	0.015	-0.001	0.000	0.000	0.085	-0.002
Immigration	Los Angeles	0.105	-0.028	0.000	0.001	0.007	-0.009	0.019	-0.008	0.030	-0.013
Immigration	New York	0.001	-0.004	-0.001	-0.003	0.105	-0.006	0.002	-0.004	0.002	0.002
Incr. Concentration	Cook	0.002	-0.017	0.003	-0.012	0.008	-0.012	0.000	-0.007	-0.028	-0.013
Incr. Concentration	Los Angeles	-0.009	-0.159	0.000	-0.031	-0.024	-0.164	-0.021	-0.130	-0.027	-0.102
Incr. Concentration	New York	0.001	-0.037	0.000	-0.037	0.006	-0.052	0.017	-0.044	0.002	0.001
Management Diffusion	Cook	-0.014	-0.068	0.000	0.000	0.000	0.001	0.000	0.003	0.008	0.003
Management Diffusion	Los Angeles	0.006	0.007	0.000	0.007	-0.012	0.003	0.025	0.003	0.018	0.002
Management Diffusion	New York	-0.002	0.000	0.000	0.001	0.002	0.000	-0.004	0.001	0.003	0.007
Sales Tax	Cook	0.018	0.027	-0.001	-0.026	-0.004	-0.027	0.000	-0.033	0.009	-0.027
Sales Tax	Los Angeles	-0.004	-0.024	0.000	-0.043	0.002	-0.018	-0.013	-0.028	-0.013	-0.026
Sales Tax	New York	0.001	-0.012	0.000	-0.015	-0.011	-0.002	-0.015	-0.011	-0.008	-0.059

Note: Effects are percent changes from the baseline equilibrium. The table presents effects for each worker skill set.