The Inner Beauty of Firms: Internal Organization in Industry Equilibrium

Jacob Kohlhepp*

[Click Here for Latest Version]

March 28, 2023

Abstract

This paper studies how the internal organization of firms interacts with labor and product markets using millions of task assignments within hundreds of hair salons. I develop a measure of organization complexity and provide evidence of firm-specific organization costs, which grant complex salons a comparative advantage in producing high-quality products. Based on these facts, I develop a model where oligopolistic firms with different organization costs choose their internal structure. Complexity is costly, but it allows firms to improve product quality by better matching workers with multidimensional skills to tasks. I characterize the profit-maximizing organization, and identify and estimate the model for Manhattan hair salons. Counterfactuals reveal that allowing internal organization to be heterogeneous and endogenous changes the equilibrium effects of policy. A sales tax cut increases specialization and therefore the productivity of all workers, while a minimum wage increase generates new types of wage spillovers.

*Department of Economics, University of California, Los Angeles. Contact: jkohlhepp@ucla.edu
I am grateful for helpful suggestions from Tara Sinclair, Maurizio Mazzocco, Simon Board, Moritz Meyer-ter-Vehn, Daniel Haanwinckel, Alex Bloedel, Ichiro Obara, Jay Lu, Felipe Goncalves, Till von Wachter, Coen Teulings, Tyler Ransom, Rodrigo Pinto, Jesse Shapiro, John Kennan, Stepan Aleksanko, 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

Greater specialization allows markets to better use the unique talents of individuals. As early as Adam Smith's pin factory, economists have recognized that much of this division of labor occurs within the firm, a process often referred to as internal organization. In practice, firms differ in their ability to organize people and use a wide variety of organization structures. How do firms choose their internal organization, and how does this choice interact with product markets, labor markets and government policy?

To answer this question, I propose a framework to study firms' equilibrium choice of internal organization. Using a set of stylized facts from management software data, I model firms as deciding which workers to hire and how to assign them to tasks. More complex assignments are costly, but they improve product quality through a better match of skills to tasks. Because firms differ in their organizational capabilities, they choose different internal structures. Additionally, because firms share a product and a labor market, the internal organization structures of competing firms are intertwined in equilibrium. I estimate the model for Manhattan hair salons, and show that allowing internal organization to be heterogeneous and endogenous qualitatively changes the effect of counterfactual policies. For example, a minimum wage raises equilibrium specialization for minimum wage workers, reduces specialization for non-minimum wage workers, and causes wage spillovers which are not monotone in initial wage.

In the first part of this paper, I use novel data to establish empirical patterns in firm internal organization. The data, which come from a management software company, allow me to observe the assignment of millions of tasks to individual workers across hundreds of hair salons. I view firms as choosing organization structures, which are matrices where rows represent workers, columns represent tasks, and each element is the fraction of total time assigned to each worker-task pair. I create a measure called organization complexity, which quantifies the amount of information that must be communicated within a firm in order to implement a given organization structure.

I document three facts about salon internal organization. First, complexity varies significantly across salons but very little across time, with few salons engaging in complete specialization. This is evidence of firm-specific and time-invariant organization costs which prevent full specialization. Second, complex firms have higher revenue and employment. This indicates firms with lower organizational costs have a competitive advantage in the product market. Third, complex firms have higher prices and more repeat customers. This is evidence the organizational competitive advantage operates through quality rather than quantity, meaning organizationally efficient salons have a comparative advantage at producing higher-quality products.

In the second part of the paper, I build a model consistent with these facts. In this model, firms with product market power choose product prices, the composition of their workforce, and worker task assignments. Workers differ in their skill at each task. Assigning tasks to the most skilled worker raises product quality but also increases organization complexity. Firms differ in the cost of complexity and their task-based production function, which causes them to choose different internal structures. Firms compete in a common product and labor market, so their choices of internal structures both shape and are shaped by wages, prices and qualities.

The main theoretical result in this paper is a characterization of the firm's optimal organization structure enabling analysis, identification and estimation. My model differs from past task-assignment models along three dimensions: firms face heterogeneous organization costs which prevent full specialization, firms have market power, and workers have horizontal skill heterogeneity.[1] Because of these differences, I cannot use existing approaches to make the firm's problem tractable. Instead, I show that the profit-maximizing organization also solves an equivalent rational inattention problem with mutual information attention costs. This equivalence allows me to weave together existing results to prove the other propositions in the paper.

Using this model, I analyze the theoretical forces that shape a firm's choice of internal structure. I show that firms navigate a trade-off between organization complexity, wage and quality, where they attempt to produce the highest-quality product using the simplest task assignment and the lowest wage bill. I prove that even though firms are choosing the task assignment of each individual worker, at a high level, the firm is choosing a point along a convex frontier that divides two dimensions: organization complexity and wages adjusted for product quality. The firm chooses the point along the frontier that is tangent to its isoprofit curves, which I show are straight lines with a slope determined by the firm's organization cost parameter.

In the third part of the paper, I identify and estimate a structural version of the model for hair salons in New York City. The distribution of organization costs is identified by the complexity of a firm's task assignments. Further, organization costs and structures are

1. In the beauty industry horizontal skill heterogeneity is important: one worker may be skilled at cutting and unskilled at coloring while another may be skilled at coloring and unskilled at cutting.

known functions of the data and the other parameters, and do not need to be estimated.[2] Variation in the interaction of task intensity and organization complexity across firms in the same market allows the identification of the other parameters. Intuitively, firms intense in task k and organizationally complex hire a large share of task k specialists and assign a large amount of task k to these specialists. The quality of these firms identifies the skill of task k specialists, while the cost of these firms identifies the wage of task k specialists. I provide a computationally light, nested fixed-point estimation procedure which implements this identification strategy.

The estimated model reveals that even within a single industry (hair salons) and occupation (cosmetologists), variation in task specialization is large and depends on unobserved worker skills and unobserved firm organizational differences. Firms in the bottom quartile of organization costs (efficient salons) on average assign 90% of tasks to the associated specialist, while firms in the top quartile (inefficient salons), assign only 67%. Haircut specialists spend most of their time cutting, but blow-dry specialists spend less than half of their time blow-drying. I also show that internal organization is a large source of productivity differences across firms, accounting for 40% of variation in marginal costs.

In the fourth part of the paper, I study two counterfactual policy changes, one in the product market and one in the labor market. In both cases, the fact that internal organization is heterogeneous and endogenous introduces new economic forces and qualitatively changes the total economic impact of each policy. The structure of the model allows any policy to be cleanly decomposed into a reallocation effect, where labor shifts across firms but internal organization remains fixed, and a reorganization effect, where task assignments within the firm are allowed to adjust. The reallocation effect is driven by the heterogeneity of internal organization, while the reorganization effect is driven by the endogeneity of internal organization.

In the first counterfactual, I eliminate the 4.5% New York City sales tax on services. The reallocation effect improves the competitive position of complex salons who were initially providing high-quality services. The reorganization effect induces almost all salons to reorganize in order to increase quality. Both effects increase equilibrium task specialization across all workers and increase equilibrium labor productivity. Workers capture most of the productivity gains through higher wages.

In the second counterfactual, I increase the minimum wage from \$15 to \$20. The reallocation effect reduces the competitive position of firms with internal structures that rely on minimum wage workers. Thus, non-minimum wage workers initially employed alongside minimum wage workers see a reduction in labor demand. The reorganization

2. This is because complexity reveals organization costs, by the logic in the last paragraph.

effect causes firms to lay off more minimum wage workers and shift their tasks onto other workers. This increases task specialization for minimum wage workers but reduces it for other workers. Although the labor market is competitive, organizational heterogeneity and endogeneity allow the model to generate aggregate labor-labor substitution patterns that are not possible with standard models. For Manhattan hair salons, reallocation and reorganization together produce wage spillovers that are non-monotonic in initial wage, with high- and low-wage workers seeing wage increases and workers in the middle seeing wage decreases.

In this paper I draw insights from organizational economics and the task-based literature in labor economics in order to understand how internal organization decisions shape economic outcomes. The primary contribution of the paper is to build and estimate a model where organizationally unique firms make task assignment decisions which have labor and product market consequences.

The literature in organizational economics provides many ways in which firms can allocate talent better than markets do. 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). Just as Holmstrom and Milgrom (1994) view the firm as an incentive system, in this paper I view the firm as a system of organizational practices. Once one adopts this view, firms should have heterogeneous organizational capabilities depending on their particular mix of practices (Argyres et al. 2012). Using firm-specific costs, I capture this heterogeneity in order to study its impact on market outcomes. I find that organizational heterogeneity is important both for determining the division of labor across the economy and for understanding the distributional impact of policy changes.

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), firms with multiple worker types (Haanwinckel 2020), organization costs (Adenbaum 2021, Garicano 2000), and firm-specific task demands (Lazear 2009). I also incorporate product market power. This combination of features allows for flexible labor-labor substitution patterns that are determined by the distribution of skills, organization costs and task demands in the economy. This flexibility is why I find that a minimum wage generates non-monotonic wage spillovers even in a competitive labor market.[3] Additionally, my

3. This is similar to how Teulings (2000) showed that imperfect substitution along a single dimension

model generates jobs which are bundles of tasks and which vary from firm to firm even for the same type of worker. This makes my model more realistic than past models, which typically generate fully specialized jobs that are homogeneous within industry.

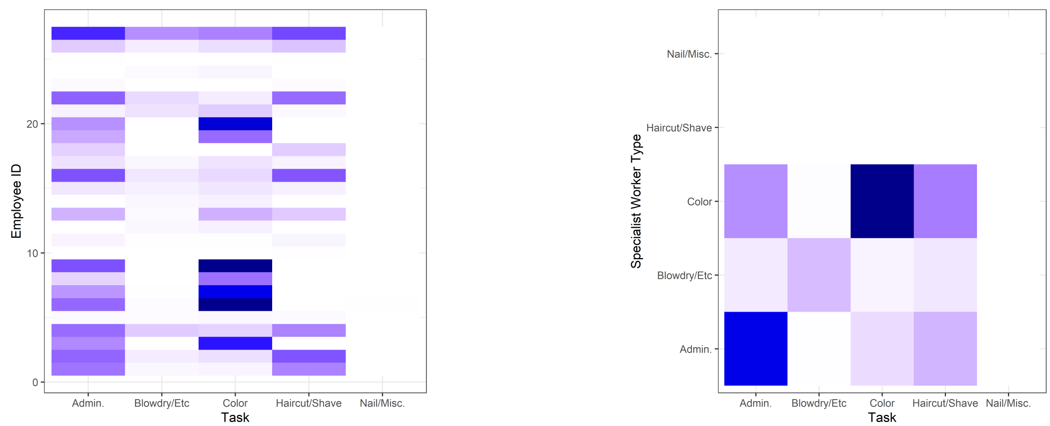
Finally, this paper makes a methodological contribution. In the majority of task-based models and hierarchy models (Garicano and Rossi-Hansberg (2006), Caliendo et al. (2012), Garicano and Hubbard 2016), workers are matched to tasks according to a single dimension, typically involving education. Since it is known prior to estimation that wages are increasing in this observed dimension, information on the wage distribution identifies features of the task-based production function. Direct information on tasks is then typically not used for estimation. I consider the opposite case, when information on wages is limited but information on tasks is rich. I show that the parameters of the task-based production function can be inferred based on differences in qualities and costs across firms intense in different tasks but operating in the same market. Further, task information allows the incorporation of workers that have unobserved horizontal differences in skill. This makes the model useful for incorporating skill differences which cannot be inferred from observed characteristics of workers.

I am able to maintain computational and analytical tractability by establishing that the profit-maximizing organization is the solution to a class of well-studied problems from the information theory and rational inattention literature. In a similar fashion, Ocampo (2022) and Adenbaum (2021) show the firm's task assignment problem is an optimal transport problem, while Freund (2022) uses the Fréchet distribution combined with results from the trade literature. Despite this theoretical similarity, this paper is distinct in that it uses data on task assignments within firms directly for both the design and estimation of the model. An illustration of the methodological contribution is given in Figure 1.

The remainder of the paper is organized as follows. Section 2 describes the management software data. Section 3 describes three stylized facts about hair salons that are used to build the model. Section 4 specifies the theoretical model. Section 5 theoretically analyzes the model. Section 6 discusses the identification and estimation of a structural version of the model. Section 7 presents the parameter estimates and assesses the model fit. Section 8 performs two counterfactual policy experiments. Sections 9 discusses implications and Section 10 concludes.

changes how we should analyze minimum wages.

Figure 1: Utilization of Task Assignment Data



(a) Observed **(b)** Estimated

Note: Darker colors indicate a higher fraction of total work at the salon. The model in this paper takes in establishment-level data about the task assignments of employees with unknown skills (Panel A) and returns the task assignments of worker types with known skills (Panel B). Even though the displayed salon in New York City employs 26 people, the model infers these represent only three of the five worker types available in the market and that most specialization is within the color and administrative tasks.

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, unique IDs allow a researcher to track employees and clients across time within a salon.[4]

A sample from the data is provided in Table 3, 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 separate employees, Tyler and Ben. Both of these salons are in the same zip code.

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 discussions with the company and analysis of internal data revealed.

2.2 Mapping Descriptions to Tasks

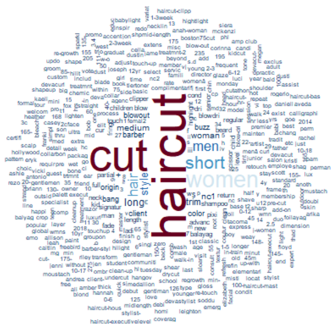
The data contain 20,560 unique text descriptions of services. This section describes the process used to categorize these descriptions into five tasks.

A licensed cosmetologist was hired to group the tasks into a manageable number of categories. Appendix Section B.1 describes the instructions sent to the cosmetologist and displays part of the final spreadsheet. 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.

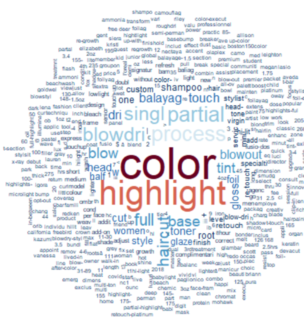
If a service is marked as multiple task categories, I divide the service into unique tasks in the following way. First, I compute the average amount of time spent on each task among services that are marked only as one task. Second, I compute the fraction of time to assign to each task as the corresponding task average divided by the sum of the averages of all other tasks marked for that service. Third, I distribute the total time spent on the service across the tasks using this imputed fraction. This process generates task categories that are mutually exclusive.

4. IDs are salon specific, so I cannot track employees or clients if they move across salons.

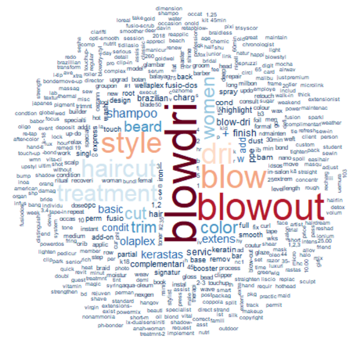
Figure 2: Task Categories



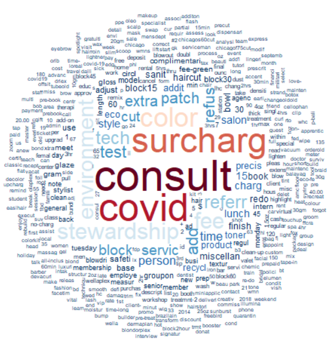
(a) "Haircut/Shave:" 38%



(b) "Color/Etc.:" 41%



(c) "Blowdry/Etc.:" 8.9%



(d) "Administrative:" 6.1%

(e) "Nail/Misc.:" 6.3%

Note: The 20,560 service descriptions grouped into task categories by a cosmetologist. Word size is proportional to the number of tasks which include the word.

2.3 Descriptive Statistics

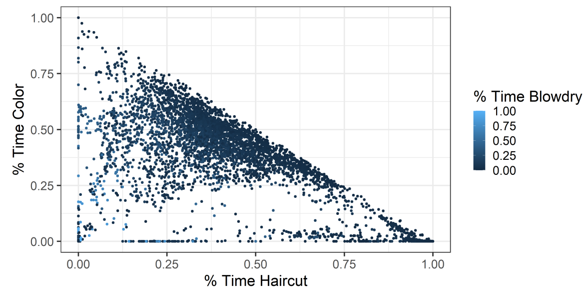
The data used in this section and the stylized facts include all observed firm-quarters where revenue per customer is positive. I exclude 2021 Q3, because I observe only part of the quarter. I also exclude an establishment in Kentucky with revenue that is implausibly high. The data contain information on 445 hair salon establishments, which represent 316 unique businesses, 9,179 hair stylists, 1,654,233 customers and 10.8 million services performed. Establishments first appear in the data when they adopt the management software. The last complete month with available data is August 2021. Although the software is used by salons across the country, users are concentrated in New York and California.

I aggregate the data to the firm-quarter level, for analysis. Descriptive statistics at this level are provided in Table 4. Throughout the paper, I refer to the price as the average revenue per customer per quarter. The salons have an average price of $200. Even though there is significant 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. Firm-quarter heterogeneity in the task mix is illustrated in Figure 3. Firms vary in their intensity in each task.[5]

Figure 3: Variation in Firm-Quarter Task Mix

Share of Labor	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Haircut/Shave	4,558	0.41	0.23	0.00	0.26	0.52	1.00
Color/Highlight/Wash	4,558	0.38	0.20	0.00	0.29	0.52	1.00
Blowdry/Etc	4,558	0.09	0.12	0.00	0.03	0.11	1.00
Administrative	4,558	0.05	0.11	0.00	0.002	0.04	1.00
Nail/Etc	4,558	0.06	0.16	0.00	0.00	0.05	1.00

(a) Summary Statistics



(b) Variation in 3 Main Tasks

Note: Panel A provides summary statistics about the share of time spent on each task across all firm-quarters. Panel B illustrates this variation for the three most common tasks. Each point is a firm-quarter.

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. It is important to be cautious when comparing self-reported survey estimates from other sources with management data (like this source), but given the subscription fee of the software, it is reasonable to conclude that the salons in my sample skew toward larger and higher-end salons. This suggests the heterogeneity found in this paper underestimates the heterogeneity in the universe of U.S. salons.

3 Stylized Facts

The model I use to study the effect of internal organization on product and labor markets is inspired by three stylized facts. These facts require the definition of two concepts which will be used throughout the paper. To begin, denote workers by the index i, firms by the index j, and tasks by the index k.

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

5. To see variation across jobs, refer to Table B3 and Figure B12.

An example of two different organizational structures is given in Figure 4. The left structure is staffed by specialists while the right structure is staffed by generalists.

Figure 4: Two Organization Structures

Specialist Salon

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 Salon

Employee	Tasks 1	2	3	
A	1/6	1/12	1/12	1/3
B	1/6	1/12	1/12	1/3
C	1/6	1/12	1/12	1/3
Tot.	1/2	1/4	1/4	1

Note: Two organizational structures a firm with a task mix of $1/2, 1/4, 1/4$ could choose. Column sums represent the task mix, and row sums represent the fraction of work performed by each employee.

The second concept, complexity, measures the minimum amount of information that must flow through the firm in order for it to implement a given structure, and it is based on a literature in information theory starting with Shannon (1948).[6]

Definition 2 *The complexity of an organization structure B_j is*[7]

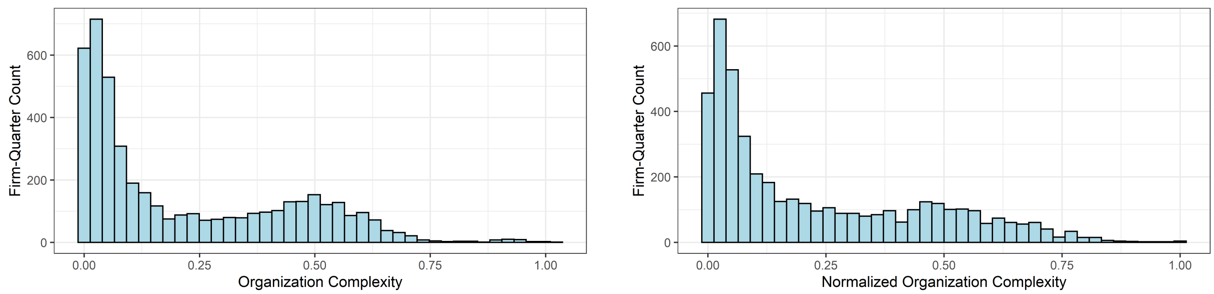
$$I(B_j) = \sum_{i,k} B_j(i,k) log\left(\frac{B_j(i,k)}{\underbrace{\sum_{k'} B_j(i,k')}_{\text{Worker } i \text{ Labor Share}} \underbrace{\sum_{i'} B_j(i',k)}_{\text{Task } k \text{ Labor Share}}} \right)$$

Consider the two structures in Figure 4. The firm can implement the chair-renter structure (right) by randomly assigning workers to tasks. This implementation does not require information about tasks or worker identities, so the complexity is 0 in this case. To implement the employee structure (left), the firm must tell each worker exactly which task to perform. The firm can write an employee manual, stating "assign the task to employee A if you observe '0', assign to B if you observe '01', and assign to C if you observe '10'." The expected number of bits (or amount of information) is $1 \times 1/2 + 2 \times 1/2 = 1.5$. This is the minimum information required to communicate this assignment, so the complexity in this case is 1.5.

6. It is the mutual information of the joint distribution over workers and tasks. Mutual information is a well-understood way to measure information costs. It has many desirable properties and many microfoundations.

7. When computing this measure, I assume that $0log(0) = 0$.

Figure 5: Histogram of Normalized Complexity



(a) Raw **(b)** Normalized

Note: Includes all firm-quarter observations. Both normalized and raw complexity vary significantly. Normalized complexity is observed achieving both its lower bound (0) and upper bound (1).

I now present three stylized facts about internal organization. Throughout the rest of the paper, complexity is assumed to be measured without error. Appendix Section B.4 provides evidence that measurement error is small.

Fact 1 *Complexity varies significantly across firms and little across time.*

To establish this fact, I first compute I_j^{max}, which is the maximum value of complexity given a firm's task mix in a given quarter. I construct normalized complexity $\bar{I}_j$ as raw complexity divided by I_j^{max}. Normalized complexity $\bar{I}_j$ has a minimum of 0 (like raw complexity) and a maximum of 1 (unlike raw complexity). I plot a histogram of normalized complexity in Figure 5 and observe that complexity varies significantly across firm-quarters and has a long right tail. In particular, I observe that while some firms have very complex organizations (close to the upper bound), others have very simple organizations (complexity of 0). To understand whether the variation is across time or across salons, I decompose complexity into a salon-specific component, a time-specific component and a residual component:

$$\bar{I}_{j,t} = \bar{I}_j + \bar{I}_t + e_{j,t}$$

I estimate the firm and year components by regressing normalized complexity on time and salon fixed effects. This allows me to decompose the total variance of complexity into the three components:

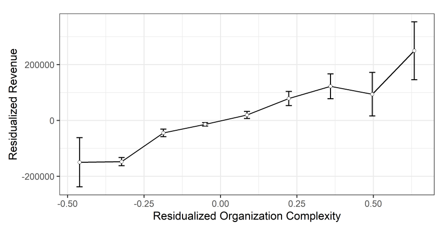
$$\underset{.0516}{Var(I_{j,t})} = \underset{.0464}{Var(\bar{I}_j)} + \underset{.0002}{Var(\bar{I}_t)} + \underset{-.0009}{2Cov(\bar{I}_j, \bar{I}_t)} + \underset{0.0059}{Var(e_{j,t})}$$

These results demonstrate that 90 percent of the variance in normalized complexity is attributable to the firm component and only 0.4 percent to the time component. Therefore, complexity varies significantly across firms but little across time. This is evidence the choice of complexity is driven by a time-invariant, firm-specific organization cost.

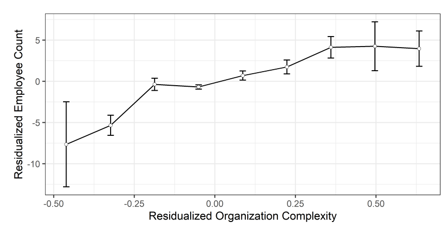
Fact 2 *Complex firms have higher revenue and employment.*

Complexity is positively correlated with revenue and employment, as well as several other measures of firm size. This correlation is depicted in Figure 6, which shows binned scatter plots of residualized complexity against residualized revenue employment, customers and visits. The plots control for the task mix, county, and quarter fixed effects. Table 2 demonstrates via a series of regressions that the correlation is positive for all firm size variables and statistically significant at the 5 percent level for revenue and employment. The positive relationship between revenue and complexity is robust; it remains when only Manhattan hair salons are analyzed and when employee count is interacted with complexity.[8]

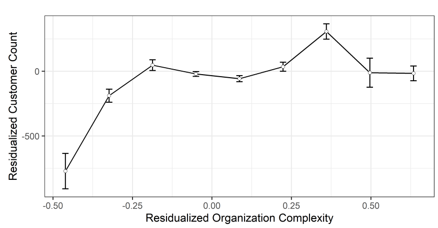
Figure 6: Organization Complexity and Firm Size



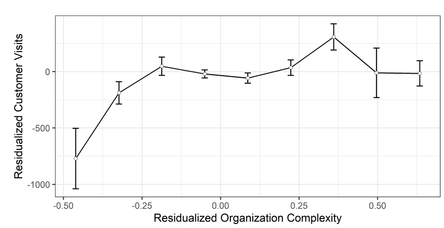
(a) Revenue ($)



(b) Employees



(c) Unique Customers



(d) Customer Visits

Note: Each panel illustrates the positive relationship between organization complexity and a different measure of firm size. All variables are residualized for quarter, county and task mix. Firm-quarters are grouped into equally spaced bins based on complexity.

The positive correlation between firm size and complexity suggests some salons have an organizational competitive advantage in that they find it easier than competitors to adopt productive organizational practices. This allows them to implement more complex task assignments at a lower cost.

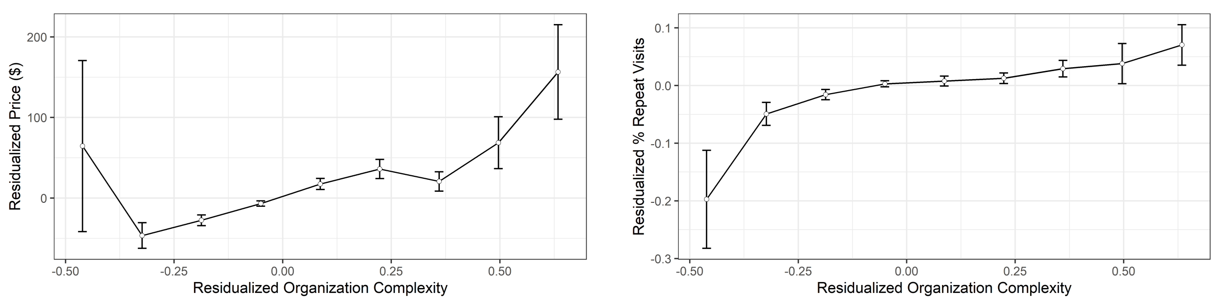
8. See Appendix Section B.5 for details on these additional results.

Fact 3 *Complex firms have higher prices and more repeat customers.*

Complexity is positively correlated with price, as shown in Panel A of Figure 7. Appendix Section A.12.5 proves that this pattern in the data is inconsistent with a model where organizational competitive advantages operate only through marginal cost reductions. In such a model, prices should be decreasing in complexity. The fact that the opposite is true suggests salons with higher internal complexity are producing services with higher unobserved quality and thus higher costs.[9]

To test this quality channel, I use the share of repeat visits as a proxy for quality in Panel B of Figure 7. It is reasonable to assume that a customer who returns was satisfied with the quality of the original service. The fraction of visits by return customers rises with complexity, evidence of a link between quality and organization. This suggests that the organizational advantage described in Fact 2 operates through unobserved quality rather than quantity. In the next section, I build a model inspired by this and the other two facts. Appendix Section B.3 discusses the robustness of the stylized facts.

Figure 7: Organization Complexity, Prices and Repeat Customers



(a) Price

(b) Repeat Customers

Note: The positive relationship between organization complexity and price (panel A), and the relationship between organization complexity and the fraction of customers that return (panel B). All variables are residualized for quarter, county and task mix. Firm-quarters are grouped into equally spaced bins based on complexity.

4 Model

This section specifies a model where firms choose prices and organizational structures simultaneously in order to compete for consumers. Consistent with the stylized facts, firms choose their organization structure subject to heterogeneous organization costs. The

9. Kugler and Verhoogen (2012) use a similar argument to conclude that endogenous product quality is important.

main benefit of a complex organization is the ability to produce a higher-quality product. There are three important groups of objects in the economy: firms, indexed by $j = 1, ..., J$; worker types, indexed by $i = 1, ..., N$; and tasks, 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 good[10] using a Leontief task-based production function described by $\alpha \in \mathbb{R}_+^K$, which I refer to throughout the paper as the task mix. The task mix is homogeneous in the theoretical section only for exposition: all results are obtained when it varies by firm. To produce one unit of the good, the firm must allocate α_k labor to task k, where I normalize $\sum_k \alpha_k = 1$. The firm can choose how these tasks are assigned to workers in a process that is described shortly.

Workers and Labor Markets. Each of the N worker types is characterized by inelastic total labor supply L_i and skill set vector θ_i. Element $\theta_i(k)$ is the quality with which worker i performs task k. The labor market is competitive with type-specific wages w_i, which I collect into a wage vector w.

Firm Strategies. Firms choose the price of their product $p_j \in \mathbb{R}_+$ and their organizational structure $B_j \in \Delta^{N\times K}$, where $\Delta^{N\times K}$ is a $N \times K$-dimensional unit simplex. Element $B_j(i, k)$ of an organization structure specifies the fraction of total labor allocated to worker type i and task k.[11] An organizational structure B_j is feasible if it is consistent with the task-mix vector: $\sum_i B_j(i, k) = \alpha_k \,\forall\, k$. The workforce composition, $E_j(i) = \{E_j(1), ..., E_j(N)\}$, is the fraction of total labor demanded that is from each worker type. By definition,

$$E_j(i) = \sum_k B_j(i, k). \tag{1}$$

10. Considering only a single good allows me to focus on internal organization. Nocke and Schutz (2018) shows we can represent a pricing game with multi-product firms as one with single-product firms by adjusting qualities and costs when demand takes the multinomial logit form.

11. This definition of task assignment treats all workers with a given set of skills symmetrically. In the model brought to the data, workers with the same skills may have different labor supplies. I show in Appendix Section A.6 that, due to an invariance property of the organization cost function, even if a firm could treat different workers with the same skills differently, it would not choose to do so in equilibrium. Thus this abstraction is without loss of generality under mutual information organization costs.

The cost of a firm's organization structure is the firm-specific parameter γ_j multiplied by the complexity of the organization structure $I(B_j)$. Recall complexity is defined as[12]

$$I(B_j) = \sum_{i,k} B_j(i,k) log\Bigg(\frac{B_j(i,k)}{\underbrace{\sum_{k'} B_j(i,k')}_{\text{Type } i \text{ Labor Share, } E_i} \underbrace{\sum_{i'} B_j(i',k)}_{\text{Task-mix, } \alpha_k}}\Bigg).$$

A firm's organizational structure determines the match between worker skills and tasks. As a result, it determines product quality ($\xi(B_j)$). I specify that product quality is a weighted average of task quality:

$$\xi(B_j) = \sum_{i,k} B_j(i,k)\theta_i(k).$$

Since quality is valued by consumers, increased quality is the main benefit of carefully assigning workers to tasks. A firm's organization structure also determines its per-unit wage bill:

$$W(B_j) = \sum_{i,k} w_i B_j(i,k)$$

Demand. Total market demand for good j is given by a function D_j which maps the prices and qualities of all firms into a quantity demanded for firm j. I assume that demand for good j depends on own-price and own-quality only through the quality-price index $\xi(B_j) - \rho p_j$, where ρ is a consumer price sensitivity parameter. I also assume demand for good j is strictly increasing in good j's quality-price index. This implies the demand can be written as $D_j(\xi(B_j) - \rho p_j, p_{-j}, \xi_{-j})$.[13] I place further parametric assumptions on consumer utility only when the model is estimated.

The Firm's Problem. Per-unit organization costs and competitive labor markets imply marginal costs are constant. I denote the feasible set of organization structures $\mathbb{B} = \{B \in$

12. This mutual information-based functional form is used because it is the only cost function in a certain class where complexity over types will be equal to complexity over worker identities under a general matching process (Bloedel and Zhong 2021).

13. There are several random utility models and representative consumer models consistent with this assumption, including multinomial logit, nested logit and mixed logit with a non-random price coefficient. A mixed logit model with consumer price sensitivity heterogeneity would violate the assumption.

$\Delta^{N\times K} | \sum_i B(i,k) = \alpha_k \, \forall \, k\}$. The firm's problem can now be defined:

$$\max_{p_j \in \mathbb{R}_+, B_j \in \mathbb{B}} D_j(\underbrace{\xi(B_j)}_{\text{quality}} - \rho p_j, p_{-j}, \xi_{-j}) \left[p_j - \underbrace{\left(\overbrace{\gamma_j I(B_j)}^{\text{org. cost}} + \overbrace{W(B_j)}^{\text{avg. wage}} \right)}_{\text{constant marginal cost, } MC_j} \right] \tag{2}$$

Equilibrium. An equilibrium consists of firm strategies $\{p_j, B_j\}_{j=1}^J$ and wages w such that:

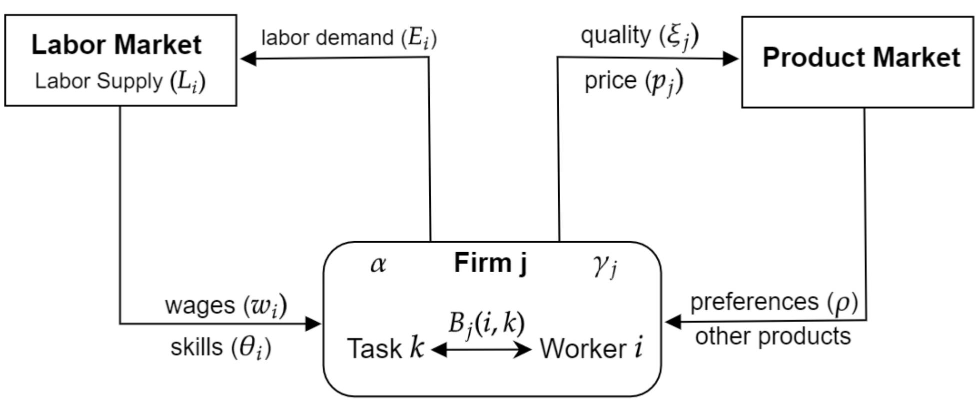
1. Firms choose prices p_j and organizational structures B_j to maximize (2).

2. Labor markets for each worker type clear:

$$\sum_j D_j(\xi(B_j) - \rho p_j, p_{-j}, \xi_{-j}) \sum_k B_j(i,k) = L_i \, \forall \, i = 1, ..., N.$$

Model Summary. Figure 8 illustrates the model from the perspective of a single firm. The firm chooses B_j (i.e., determining who it hires and how hired workers are assigned to tasks) and prices taking into account internal factors (i.e., the task mix and organization costs), labor market factors (i.e., wages and skills), and product market factors (i.e., consumer price sensitivity, the prices and qualities of other products). The choice of B_j feeds back into the product market by determining product quality and prices, and into the labor market by determining labor demand across worker types.

Figure 8: Illustration of the Model



4.1 Discussion of Organization Costs

This section describes different ways of interpreting γ_j, the cost of increasing the complexity of firm j's organizational structure by 1 unit. The multifaceted nature of γ_j can account for several dimensions of organizational heterogeneity all at once. Unfortunately, this also means the mechanisms driving the value of γ_j at any particular firm can not be separated.

Coordination Costs. Under this interpretation, γ represents the fact that firms are "second-best solutions to transactions plagued by various forms of contractual incompleteness" (Gibbons 2020) and that "firms can never succeed in conquering the nonrational dimensions of organizational behavior" (Williamson 1984). As γ approaches 0, coordination costs disappear and a firm can design any organizational structure it chooses at 0 cost. When γ becomes sufficiently large, firms will resort to assigning every worker the same job. In the latter case, workers are essentially firms, since they perform all of the tasks the firm performs and do not rely on coworkers. In this way the distribution of γ traces out the value of firms. A firm with low γ is greater than the sum of its workers, producing a product superior to that which any of its workers could produce alone.[14]

Rational Inattention. The mutual information form of organization costs gives it a rational inattention micro-foundation. We can interpret γ as the level of "managerial talent" (Lucas 1978) which determines the attention cost needed to allocate tasks to workers. Similarly, organization costs also capture contractual inattention, such as those described by Tirole (2009). Different firms may find it more or less costly to write down complex contracts in order to support complex organizational structures.

Incentives for Teams. Under this interpretation, organization costs reflect losses due to free-riding. Dai and Toikka (2022) show that the profit of a firm managing a team of multiple workers increases with the productivity of the known technology. Thus, heterogeneity in γ_j reflects the fact that some firms know a larger number of technologies, or ways to combine tasks and workers.

Costly Specialization. Because complexity is a measure of distance from the random assignment of workers to tasks,[15] γ_j can be interpreted directly as a firm-specific specialization cost.

An example from the hair salon industry makes these ideas concrete. There are two main ways to organize a salon. In the chair-renter arrangement, stylists pay a fixed fee to

14. Total organization cost increases linearly with the size of the firm for a fixed organization, and per-unit organization costs increase non-linearly with specialization. This is different than the coordination costs found in papers such as Becker and Murphy (1992).

15. See Appendix Section A.8 for a proof.

the salon owner and keep all revenue. Chair renters set their own hours and develop their own client lists. In the employee arrangement, stylists are paid by the salon owner and do not run their own business. The chair renters are independent contractors and have little need for coordination; in the language of the model, there is little organizational complexity in this arrangement. In contrast, the employee arrangement exhibits complex contracts (including non-competes, commission-based compensation, etc.) and requires coordination. In the language of the model, there is significant organizational complexity in this arrangement.[16]

5 Theoretical Results

This section analyzes the theoretical model. I first show the profit-maximizing organization structure is also the solution to a simpler problem that is well studied in information theory and behavioral economics. I use this equivalence to understand the economic forces which determine each firm's internal structure.

5.1 Main Characterization

The firm's problem as written in Equation (2) appears complicated at first glance; there are $1 + N \times K$ choice variables and the objective is highly non-linear. The following theorem reveals the firm's problem can be greatly simplified.

Theorem 1 *An organizational structure (B_j^*) is profit-maximizing if and only if it solves*

$$\min_{B_j \in \mathbb{B}} \gamma_j I(B_j) + W(B_j) - \rho^{-1}\xi(B_j), \tag{3}$$

which is a rate-distortion problem and a rational inattention problem.

The proof of the result is provided in Appendix Section A.2. The main idea of the proof is that if an organization structure does not solve Equation (3), the firm can switch to a structure that does and adjust its prices to strictly improve its profit. In this way, even though price and organization structure appear entangled in the firm's problem, they can be separated during analysis. The result relies on the fact that the quality-price index $\xi(B_j) - \rho p_j$ is sufficient for price and organization structure in demand, and demand is

16. Johnson and Lipsitz (2022) survey salon owners and find that 48% are employee salons and 52% are independent contractor salons.

strictly increasing in the quality-price index. The result does not rely on the functional form of organization costs.

Theorem 1 is useful for three reasons. First, it allows the model to be taken to the data. Because (3) is a rate distortion and rational inattention problem, and these problems are well studied in information theory and behavioral economics, I can weave together results across the two strands of literature to identify firm-specific organization costs, prove a form of equilibrium existence and uniqueness, construct an estimation algorithm, and solve for counterfactual equilibria.

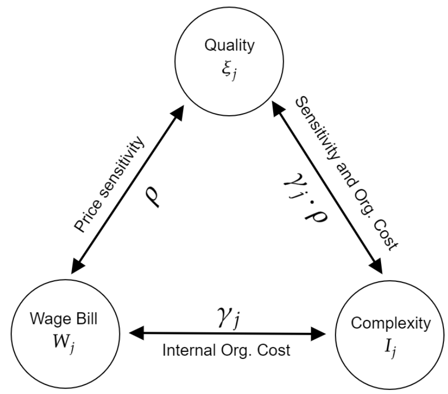
Second, the pricing and organization decisions can be separated when solving for an equilibrium for fixed wages. Specifically, a firm's internal organization *directly* affects own and competitor prices, but these do not directly affect internal organization. Additionally, one firm's internal organization does not *directly* impact a competitor's internal organization; however, each firm's internal organization is *indirectly* impacted by all prices and all competitor internal organizations via wages.

The separation implied by Theorem 1 has practical implications: it means equilibria are robust to timing. Although I assume firms choose organizations and prices simultaneously, if firms chose organizations first and then competed in prices, the outcomes would be the same. The separation implied by Theorem 1 also greatly improves tractability because for fixed wages, the organization problem can be solved first and then used to derive equilibrium prices. This simplifies counterfactual analyses and allows for policies to be decomposed in useful ways.

One question is whether the separation implied by the model is reasonable, or, equivalently, is it the case that wages are the main connection between different firms' internal organizations? The answer appears to come down to whether the labor market is well approximated by perfect competition and whether demand satisfies the index restriction. These assumptions seem reasonable in the case of hair salons, because they sell a horizontally differentiated product and are small in terms of employment, but they may not be in industries where differentiation is largely vertical (e.g., supermarkets) or where individual firms employ a large share of the labor market (e.g., manufacturing company towns). Appendix Section A.12.5 discusses ways in which the model can be extended to accommodate these other contexts.

Third, Theorem 1 reveals the forces that shape a firm's internal organization. Examining Equation (3) shows firms face a triple trade-off, as depicted in Figure 9. Each firm wishes to achieve the lowest complexity and wages while achieving the highest quality. How it navigates this trade-off depends on its internal organization cost γ_j, consumer price sensitivity ρ and their interaction $\gamma_j \cdot \rho$.

Figure 9: The Complexity-Wage-Quality Trade-Off



If a firm wishes to increase quality, it has two options: (1) hire better workers and incur a wage cost or (2) rearrange its current workforce to better leverage existing worker skills and incur an organization cost. Intuitively, when consumers are price sensitive (ρ is high), the firm cannot pass on costs to consumers via prices. Thus, firms prioritize minimizing costs over maximizing quality by choosing less complex organizations.

To analyze how the firm navigates the complexity-wage-quality trade-off, I define the *organization frontier* as the set of all organization structures which minimize complexity for some quality-adjusted wages (Q). The frontier consists of the simplest organization that achieves some quality-adjusted wages. I wish to study the relationship between quality-adjusted wages and complexity along the frontier:

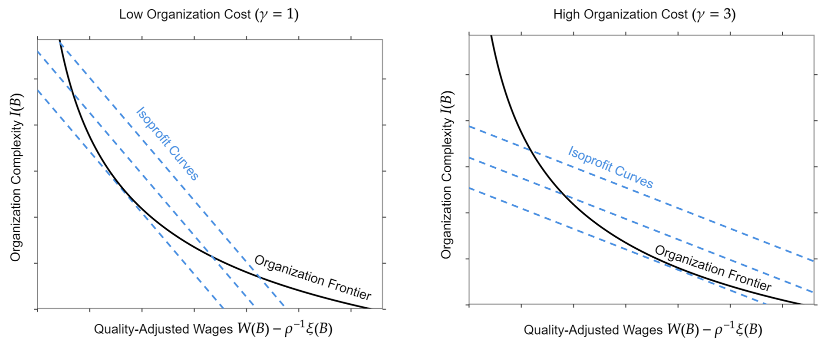
$$I^*(Q) = \min_{B_j \in \mathbb{B}} I(B_j) \text{ s.t. } W(B_j) - \rho^{-1}\xi(B_j) \leq Q.$$

The characterization provided in Theorem 1 allows me to apply existing results from information theory to understand the general shape of this relationship.

Proposition 1 *Organization complexity along the organization frontier ($I^*(Q)$) is continuous, convex and decreasing in quality-adjusted wages.*

The proof is provided in Appendix Section A.3. The proposition implies the choice of a high-dimensional organization structure can be thought of as a two-dimensional choice, similar to a classic expenditure minimization problem from consumer theory, as illustrated in Figure 10. Although B_j (i.e., how a firm chooses its workers and how they are assigned to tasks) is a high-dimensional object, the firm essentially solves a two-dimensional trade-off between complexity and quality-adjusted wages. The firm's optimal structure will be the point of tangency between the organization frontier and the

Figure 10: Choosing an Organizational Structure



Note: Although B_j (i.e., how a firm chooses its workers and how they are assigned to tasks) is a high-dimensional object, the firm essentially solves a two-dimensional trade-off between complexity and quality-adjusted wages. The firm's optimal structure will be the point of tangency between the organization frontier and the best possible isoprofit curve.

best possible (leftmost) isoprofit curve. The firm's isoprofit curves have a slope equal to $-\gamma_j^{-1}$. As the organization cost parameter (γ_j) rises, the curves become flatter, causing the tangent point to shift right and reducing organizational complexity while increasing quality-adjusted wages. A more complex organization allows a firm to produce a higher-quality good at a lower wage, but it requires a greater organization cost. An immediate consequence is that a lower organization cost parameter grants the firm an organizational competitive advantage in the product market.

Proposition 2 *In equilibrium, firms with a lower organization cost (γ_j) have higher organization complexity, market share and profits.*

The proof is provided in Appendix Section A.3. Recall that γ_j represents the management technology, relationships, knowledge and practices specific to the firm which make it easier or harder for the firm to implement a given organizational structure. Proposition 2 implies more organizationally efficient firms are larger and more profitable, and can produce better-quality goods at a lower cost. Importantly, this proposition confirms that the model is consistent with Fact 2: complexity should be positively correlated with measures of firm size. This is in line with the findings of Kuhn et al. (2022), who use surveys and administrative data to show that more coordinated or specialized firms are more profitable.

5.2 Workforce Heterogeneity

The model assumes that workers are perfect substitutes in production, both in terms of quantity and quality. To see this, set $\gamma_j = 0$ and examine Equation (3). Without organization costs, the firm minimizes a constrained linear objective with weights determined by wages and skill sets. All complementarities between workers arise endogenously via organization costs. Because these costs are firm specific, this allows for rich heterogeneity within a product and labor market, both in terms of labor-labor substitution patterns and workforce composition.

I illustrate this with a simple version of the model with three worker types. Suppose wages are fixed at $w = (21, 20, 15)$, the task mix is $\alpha = (1/3, 1/3, 1/3)$, price sensitivity is $\rho = 1$, and worker skill sets are given by Θ (defined shortly). Under this worker-type space, there are two worker types that are specialists in task 1 and 3 relative to each other, but that have higher absolute skill in all tasks compared to a third type. When I "adjust" skills for wages, it can be seen that in relative terms, there are two workers who are optimal to hire for task 1 and task 3, and one jack-of-all-trades who is a safe option for all tasks:[17]

$$\Theta = \begin{bmatrix} 15 & 19 & 26 \\ 23 & 19 & 15 \\ 15 & 15 & 15 \end{bmatrix} \implies \theta - \rho w = \begin{bmatrix} -6 & -2 & 5 \\ 3 & -1 & -5 \\ 0 & 0 & 0 \end{bmatrix}.$$

Firms facing the same market conditions and task mix can have heterogeneous workforce compositions, as illustrated in Figure 11 panel A. Organizationally efficient firms employ an equal share of each worker because they can fully utilize the specific skills of each worker type. Firms with intermediate organization costs hire only two worker types. Organizationally inefficient salons employ only type 3 workers (jacks-of-all-trades), because these firms cannot utilize the specific skills of the specialist types.

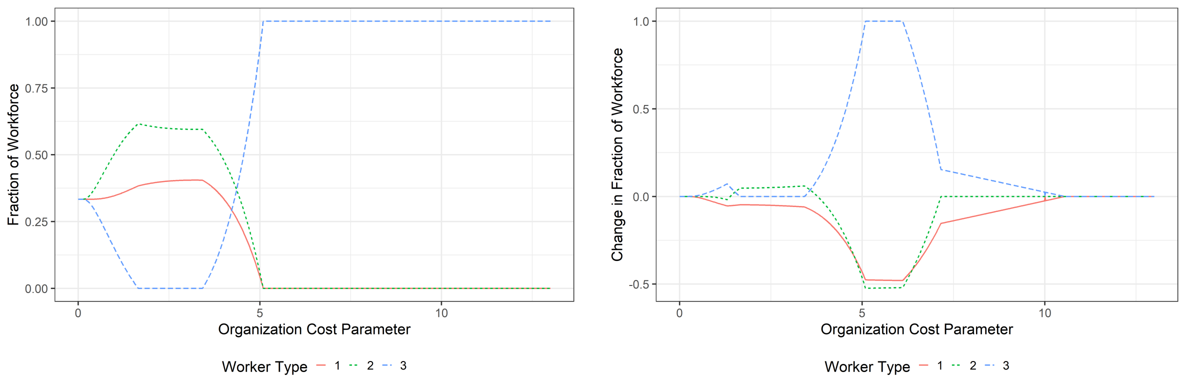
Additionally, firms facing the same market conditions and task mix can exhibit very different labor-labor substitution patterns, as demonstrated in Figure 11 panel B. When the wage of type 1 workers is increased by 1, firms with different organization costs react differently. Firms with very high or very low organization costs reduce the share of type 1 workers and increase the share of the other two types. Firms with intermediate costs reduce the share of both type 1 and type 2 workers. Thus, these two types are substitutes at extreme firms, but are complements at intermediate firms.

6 A Structural Model of Internal Organization

Understanding the quantitative relationship between internal organization and the labor and product market requires a structural model that can be taken to the data. This section

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

Figure 11: Organizational Heterogeneity



(a) Workforce Composition

(b) Labor-Labor Substitution

Note: Panel A illustrates that as organization costs change, the composition of a firm's workforce changes in a non-monotone fashion. Panel B illustrates the change in the share of each worker type due to a 1-unit increase in the wage of type 1 workers. Type 1 and type 2 workers are substitutes at extreme firms (i.e., those with very high or very low organization costs) and complements at intermediate firms.

describes such a model, which preserves the spirit of the theory developed in Section 5 while allowing for additional firm and worker heterogeneity. I prove the identification of the distribution of organization costs and provide a computationally light, nested-fixed point generalized method of moments estimation procedure.

6.1 Econometric Model

I define labor markets and product markets as counties, and time periods as quarters. I estimate the model for New York County (Manhattan) 2021 Quarter 2, the last full quarter with available data in my sample. I add several types of heterogeneity to the theoretical model introduced in Section 4 to better fit the data. The theoretical results in Section 5 continue to apply to the econometric model.

Consumers. I assume a parametric form for demand. There is a mass M of consumers interested in purchasing at most one of the J final products, where M is set to be the population of Manhattan. Consumer z's utility for good j is represented by the logit utility function

$$u_{z,j} = \xi(B_j) - \rho p_j + \epsilon_{z,j},$$

where $\epsilon_{z,j}$ is distributed i.i.d. Type 1 extreme value across consumers and products. The outside option for consumers is assigned index $j = 0$, and its utility is normalized to

$u_{z,0} = \epsilon_{z,0}$. As in McFadden (1973), market demand for good j can be written as

$$D_j(\xi(B_j) - \rho p_j, p_{-j}, \xi_{-j}) = \frac{exp(\xi(B_j) - \rho p_j)}{\sum_{j'} exp(\xi(B_{j'}) - \rho p_{j'})}. \tag{4}$$

Expression (4) reveals the multinomial logit demand system satisfies the quality-price index restriction assumed during theoretical analysis.

Production Function Heterogeneity. The task mix is firm-specific and therefore indexed by j (α_j). This allows firms in the same product and labor market to have different organizational frontiers. Since tasks are observed, the distribution of time across task categories can be computed.[18]

Marginal Cost Heterogeneity. Marginal cost may depend on the firm-specific task mix (α_j) to capture the costs of materials relating to specific tasks (e.g., dyes) as well as an idiosyncratic marginal cost shifter ϕ_j, which has expectation 0 and is independent of γ_j and α_j. I measure $\bar{a}_j$ as the average number of hours salon j spends on a customer in a quarter. I specify that organization costs and wages are per hour of labor. This allows each firm to have a different required labor per unit ($\bar{a}_j$) so that I can capture traditional productivity differences across firms.[19] With these modifications, marginal cost can be expressed as

$$MC_j = \bar{a}_j \left[\gamma_j I(B_j) + W(B_j) \right] + \sum_k m_k \alpha_j(k) + \phi_j.$$

Quality Heterogeneity. In addition to endogenously chosen quality $\xi(B_j)$, each firm also has exogenous unobserved quality ν_j, which has expectation 0 and is independent of γ_j, α_j and $\bar{a}_j$. Exogenous unobserved quality ν_j represents reputation and other attributes that impact quality but are fixed in a given period and unrelated to labor. Inclusion of ν_j ensures only quality differences correlated with observed organization complexity (I_j) will be attributed to internal organization. Quality is now $\xi(B_j) + \nu_j$.

Worker Labor Supply Heterogeneity. Workers with the same skill set may differ in their labor supply. This clarifies the relationship between worker identities (observed in the data) and worker types in the model (unobserved). Specifically, in addition to being char-

18. It is assumed that enough tasks are observed so that the computed shares are exactly equal to the underlying shares.

19. This is similar to specifying a production function of the form $\bar{a}_j \min \left\{ \frac{a_1}{\alpha_1}, ... \frac{a_k}{\alpha_k}, ..., \frac{a_K}{\alpha_K} \right\}$.

acterized by their skill set, workers are also characterized by an inelastic person-specific labor supply.[20]

Worker-Firm Matching. I augment the game by specifying that firms first demand an amount of labor of each skill set, and then an unspecified process matches workers to firms. The only assumption I place on this process is that the firm's labor demand from the first stage is exactly met. Thus if a firm demands 10 hours of a skill set, this amount may be met by any combination and number of workers, but no more or less than 10 hours is supplied in total. Following the matching process, firms then select an organization structure $\tilde{B}_j$, which is an assignment of worker identities to tasks.[21]

Worker Skill Sets. I assume there is one specialist worker type for each of the five tasks. Tasks are performed with a base skill level β_k when assigned to a non-specialist, and are performed with an additive skill gap θ_k when assigned to a task k specialist. The matrix of skill sets, where each row denotes a worker type and each column a task, can be written as[22]

$$\Theta = \begin{bmatrix} \theta_1 + \beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_5 \\ \beta_1 & \theta_2 + \beta_2 & \beta_3 & \beta_4 & \beta_5 \\ \beta_1 & \beta_2 & \theta_3 + \beta_3 & \beta_4 & \beta_5 \\ \beta_1 & \beta_2 & \beta_3 & \theta_4 + \beta_4 & \beta_5 \\ \beta_1 & \beta_2 & \beta_3 & \beta_4 & \theta_5 + \beta_5 \end{bmatrix}.$$

Sales Tax. The state of New York does not tax hair services. However, New York City levies a 4.5 percent tax on beauty services. Therefore, I denote the sales tax τ and assume it is 4.5% initially.

Outside Option. I define consumers' outside option as not buying services from a salon. I use Consumer Expenditure Survey micro-data to compute the share of individuals from a county in a quarter who spend $0 at salons, and take this to be the share of people who choose the outside option. Based on this methodology, the share of New York County residents selecting the outside option in 2021 Q2 is 40%.

20. If the set of labor supplies is Λ, the worker type space is now $\Theta \times \Lambda$.

21. This means that in principle a firm may employ multiple workers of the same skill set and assign them different distributions of tasks. This allows me to bring the model to the data, where I observe worker identities but not worker types.

22. In the absence of matched employer-employee data, which would allow the researcher to infer worker types, this restriction of the worker type space is likely necessary.

Profit. Under the econometric model, a firm's profit can be written as

$$\frac{exp(\xi(B_j) - \rho(1+\tau)p_j + \beta\alpha_j + \nu_j)}{\sum_{j'} exp(\xi(B_{j'}) + -\rho(1+\tau)p_{j'} + \beta\alpha_{j'} + \nu_{j'})}\left[p_j - \bar{a}_j\left(\gamma_j I(B_j) + W(B_j) + m\alpha\right) - \phi_j\right],$$

where the features added to the theoretical model are written in blue. Fixing an equilibrium, the parameters of the model can be divided into two groups. The first group is the firm-specific organization cost coefficients $\{\gamma_j\}_{j=1}^{J}$. The second group is the parameters that define worker skills (10 parameters), wages (5 parameters), material costs (5 parameters) and price sensitivity (1 parameter). I call these *market parameters* and denote them by Ω.

6.2 Equilibrium Existence and Uniqueness with Fixed Wages

In the empirical application of the model, I treat wages as fixed parameters to be estimated. Prior to identification and estimation, I establish that for fixed wages, there almost always exists a unique equilibrium.

Proposition 3 *Suppose wages are fixed parameters. A pure strategy equilibrium always exists, and it is unique except over a set of parameters with measure 0.*

The proof of this result is provided in Appendix Section A.5, and it relies on the equivalence to a rational inattention problem established in Theorem 1. This result means that multiplicity arises only in knife-edge cases. Proposition 3 does not establish equilibrium uniqueness or existence in the full model with wages determined endogenously by labor market clearing. Nevertheless, Proposition 3 is crucial for proving Proposition 4, the main identification result in this paper.

Several aspects of the model make Proposition 3 surprising. First, each firm has 26 choice variables, and quality and marginal cost are endogenous. Many models where product positioning is endogenous (including the canonical two-stage Hotelling model) suffer from equilibrium existence and uniqueness problems.[23] This idea is captured well by Shaked and Sutton (1987): "It is notorious in models of product differentiation that equilibria may fail to exist for many reasonable specifications of the standard models."

The result is also useful for counterfactual analysis, because it means the model almost always delivers one and only one internal organization structure for each firm. The model

23. In a two-firm Hotelling model with product positioning, it is known that a pure strategy equilibrium does not exist for linear transportation costs. When transportation costs are quadratic, there are two equilibria (d'Aspremont, Gabszewicz, and Thisse 1979)

will almost never suffer from the inverse identification problems, at least when wages are held fixed.

6.3 Identification of Firm-Specific Organization Costs

The organization costs $\{\gamma_j\}_{j=1}^{J}$ are important parameters, determining product quality, organization complexity and marginal costs for each firm. However, the fact that there is one parameter per firm and that I place no restrictions on the economy-wide distribution raises concerns for identification and estimation. I alleviate this concern with the following result.

Proposition 4 *Organization costs (γ_j) and organization structures (B_j) are a known function of firm task mixtures (α_j), complexities (I_j) and market parameters (Ω) for all firms with positive complexity, except for a set of market parameters with measure 0.*

The proof is fully described in Appendix Section A.6, and it makes use of the essential equilibrium uniqueness result given in Proposition 3. A key hurdle is that I do not observe worker types, but worker identities within firms. Because I allow for a flexible matching process, a given firm may in principle employ multiple workers of the same skill set and assign these workers different tasks. However, a property of the mutual-information based organization cost ensures that if firms do employ multiple workers with the same skill set, they assign these workers the same tasks. This implies that the observed organization complexity based on worker identities is equal to the true organization complexity based on worker skill sets.

The intuition for the identification of γ_j is demonstrated in Figure 10. Suppose two firms with the same task mix (α_j) are observed in the same product and labor market. This means they have the same organization frontier. If firm A has a higher complexity, it must be that the slope of firm A's isoprofit curve is steeper than B's, which can only be because A has a lower organization cost. Therefore, I can order the firms by organization cost without knowing the market parameters. Once these parameters are known, I can find the cardinal values of each firm's organization cost.

The proposition implies that organization costs do not need to be estimated in the statistical sense. For any market parameters, there are unique organization cost parameters which rationalize the observed organization complexities and task mixtures. This is similar to the way unobserved product quality is a known function of market shares in Berry, Levinsohn, and Pakes (1995). Another similarity is the lack of a closed form for the known function. Instead, I provide a fixed-point algorithm as part of my suggested estimation routine.

Beyond estimation, Proposition 4 also implies that observing the task mix (a vector of length K) and organizational complexity (a scalar) is enough to estimate the model. It is not necessary to observe the individual assignments of workers to tasks; the researcher need only observe complexity. This means the model can be estimated in settings where rich assignment data are not available.[24] It also means that a researcher who has assignment data can estimate the model using only complexity and the task mix and use the rest of the data to conduct validation exercises. This is precisely what I do in Section 7.2.

6.4 Estimation of Market Parameters

I have established that organization costs are a known function of the data and market parameters. This section derives a set of moments and assumptions under which the market parameters can be estimated via the generalized method of moments. I note that material costs are not separately identified from $\mathbb{E}[\phi_j]$ and the skill base is not separately identified from $\mathbb{E}[\nu_j]$. I therefore estimate $\mathbb{E}[\nu_j] + \beta_{cut}$ and $\mathbb{E}[\phi_j] + m_{cut}$ and call them the demand and cost intercepts, respectively. This means all skills parameters and material costs parameters are relative to the haircut/shave task.

To construct moment conditions, I follow a common approach in the literature and use one demand-side and one supply-side equation. Starting with the supply side, the equilibrium pricing equation can be written as

$$p_j = \frac{1}{\rho(1+\tau)(1-s_j)} + \bar{a}_j\Big[\gamma(\Omega, I_j, \alpha_j)I_j + W(\Omega, I_j, \alpha_j)\Big] + m\alpha_j + \phi_j. \tag{5}$$

Because the demand system takes a multinomial logit form, market shares can be expressed as

$$log(s_j) - log(s_0) = \xi(\Omega, I_j, \alpha_j) - \rho(1+\tau)p_j + \beta\alpha_j + \nu_j. \tag{6}$$

I interact firm-level covariates with Equations 5 and 6. I use covariates that are relevant to the determination of prices and market shares but also independent of ν_j, ϕ_j. The firm organizational complexity (I_j) and task-mix vector (α_j) fit these requirements, because they change organization costs but are not impacted by ν_j, ϕ_j. Additionally, I include their interaction, $\alpha_j \cdot I_j$. A discussion of how this variation identifies specific parameters is provided in Section 6.6.

I add one additional wage moment. For each county and quarter, I compute the average wage of hair stylists using the Quarterly Census of Employment and Wages. I take the

24. For example, privacy concerns may often prevent the disclosure of employee-client assignments.

total quarterly wages of establishments with NAICS code 812112 and divide them by the number of establishments. This generates the average total wage bill, which corresponds to $W(\Omega, I_j, \alpha_j)$ multiplied by the number of customers. I allow for classical measurement error (e_j) which yields

$$W_j = Ms_j a_j W(\Omega, I_j, \alpha_j) + e_j.$$

Taken together, the moment conditions used for estimation are

$$\mathbb{E}\left[\begin{pmatrix}\phi_j(\Omega, I_j, \alpha_j, p_j, s_j)\\ \nu_j(\Omega, I_j, \alpha_j, p_j, s_j)\end{pmatrix}\begin{pmatrix}\alpha_j & \alpha_j I_j\end{pmatrix}\right] = 0 \qquad \mathbb{E}[e_j(\Omega, I_j, \alpha_j)] = 0.$$

For a single market and quarter, I obtain a total of 21 moments to estimate 21 market parameters. The model is globally identified if I assume that Ω is the unique vector of parameter values which satisfies the moment conditions. With this assumption, I estimate the model using the generalized method of moments (GMM). Denote the sample analogue of the moments as $G(\Omega)$. Then, to estimate Ω, I solve

$$\arg\min_{\hat{\Omega}} G(\hat{\Omega})' W G(\hat{\Omega}), \tag{7}$$

where W is a weighting matrix. Note that to evaluate this GMM objective, I must recover the vector of organization costs implied by the data and each guess of the market parameters. This requires solving each firm's internal organization problem many times per evaluation of the GMM objective. The next section provides a result which greatly reduces the computational resources needed to do this.

I take the weighting matrix W to be a diagonal matrix, where each diagonal element is the sample variance of the independent variable involved in the moment. This standardizes the moments in the objective function, preventing variables with large nominal values (i.e., average hours per unit) from dominating during estimation. I constrain wages to be between \$15 (the minimum wage) and \$200 per hour. I require that the algorithm search only over parameters values yielding a positive demand share for each type of labor.

6.5 A Computationally Light Estimation Procedure

Although organization costs are a known function of the data, there does not exist a closed-form expression for this function. This is a problem for estimation, because it is necessary to numerically solve each firm's internal organization problem many times until the model-produced complexities match the complexities in the data. This is computationally intensive when there are many tasks and many firms, because each firm's problem is a high-dimensional non-linear minimization problem. This section solves this problem by providing a nested-fixed point algorithm which efficiently solves the firm's problem and is proven to globally converge.

To derive the algorithm, I use the equivalence to a rate-distortion problem proved in Theorem 1.

Lemma 1 *Given market parameter values, the Blahut–Arimoto algorithm with Lagrangian multiplier* $(\bar{a}_j\gamma_j)^{-1}$ *delivers an organizational structure* B_j *which maximizes firm profit.*

The lemma follows directly from Theorem 1 and well-known results in information theory.[25] The Lagrangian multiplier involves $\bar{a}_j$ because marginal costs are $\bar{a}_j(\sum_i w_i E_i + \gamma_j I(B_j)) - \rho^{-1}\xi(B_j)$. The Blahut–Arimoto algorithm (Blahut 1972) is a fixed-point algorithm which iterates on two optimality conditions and can be described as follows (suppressing firm subscripts):

0. Guess some labor demand E^0. Create matrix V:

$$V_{i,k} = exp[(\bar{a}\gamma)^{-1}(\rho^{-1}\theta_{i,k} - w_i)].$$

1. Compute B^t as

$$B(i,k)^t = \alpha_k \frac{V_{i,k}E^t(k)}{\sum_i E^t(i)V_{i,k}}.$$

2. Compute E^{t+1} as

$$E^{t+1}(i) = \sum_k B(i,k)^t.$$

3. If converged, exit; else return to Step 1 and advance t.

The Blahut–Arimoto algorithm is proven to converge to a global optimum from any feasible starting point (Tishby, Pereira, and Bialek 2000). It avoids the need to repeatedly use nonlinear optimization routines, and because it is a fixed-point method, acceleration routines such as Du and Varadhan (2020) can be used to increase the speed of convergence. The algorithm also delivers the entire internal organization of the firm, B_j. Practically, a researcher can use the algorithm to search for the γ_j which makes the model-generated complexities match the complexities observed in the data. Because complexity is monotone in γ_j, a researcher can use the bisection method for this task. Thus, the full estimation procedure is as follows:

25. See Tishby, Pereira, and Bialek 2000 or Blahut 1972.

1. Given a guess of the market parameters $\hat{\Omega}$, use the Blahut–Arimoto algorithm to find the organization costs $\gamma_j(\hat{\Omega})$ which rationalize each firm's observed organizational complexity I_j.

2. Using $B_j(\hat{\Omega}), \gamma_j(\hat{\Omega})$, compute firm-specific wage bills and endogenous quality.

3. Evaluate the GMM objective given by Equation 7.[26] If the objective is minimized, stop; otherwise, return to step 1 with a new market parameter guess, $\hat{\Omega}$.

This estimation algorithm is similar in spirit to the demand estimation procedures that have become popular in industrial organization since Berry, Levinsohn, and Pakes (1995). Just as those procedures invert market shares using a contraction mapping to derive unobserved product qualities, my procedure inverts organization complexities using a contraction mapping to obtain unobserved organization costs. Implicit in this inversion procedure is that complexity is measured without error. Appendix Section B.4 provides evidence that measurement error is small.

6.6 Identifying Variation

Proposition 4 establishes that given fixed values for the market parameters, organization costs (γ_j) are identified by variation in complexity and the task mix across firms. The purpose of this section is to discuss the sources of identifying variation for the market parameters.

Consumer price sensitivity is identified by the pass-through of average wages to consumers. If wage costs are passed through to consumers via higher prices, consumers are not price sensitive and ρ is low. Once price sensitivity (ρ) is known, the marginal cost of each firm can be obtained by subtracting the markup from prices. Similarly, service quality can be obtained from observed prices and market shares. For this reason, I discuss identification of the other parameters as if quality and cost were observed.

The base skill parameters (β) and the material costs (m) are identified by variation in the task mixtures (α_j) across firms. When I observe a firm that is intense in task k, I its cost is informative about m_k and its quality is informative about β_k. This is why α is interacted with the demand supply side residuals to obtain a first set of moments.

Recall that a complex firm generally has a specialized workforce. When I observe a firm that is complex and is intense in a task k, this implies two things. First, the firm likely uses a large share of task k specialists. Therefore, the observed price (and thus cost)

26. I use the R package "gmm" described in Chausse (2021) to perform estimation.

of that firm largely reflects the wage of task k specialists (w_k). Second, the firm assigns a large amount of task k to those specialists. Therefore, the observed market share (and thus quality) largely reflects the skill gap of task k specialists (θ_k). This is why $\alpha_j \cdot I_j$ is interacted with the demand supply side residuals to obtain a second set of moments.

7 Empirical Results

This section summarizes parameter estimates and uses the model to analyze the sources of variation in task content for hair stylists in Manhattan.

7.1 Parameter Estimates

Estimated wages, skill parameters and material costs are organized by task in Table 5. Price sensitivity and the intercepts are provided in Table 6. Standard errors are computed as the sample standard deviation of the parameter estimates from 500 bootstrap replications, with the procedure described in Appendix Section B.10.

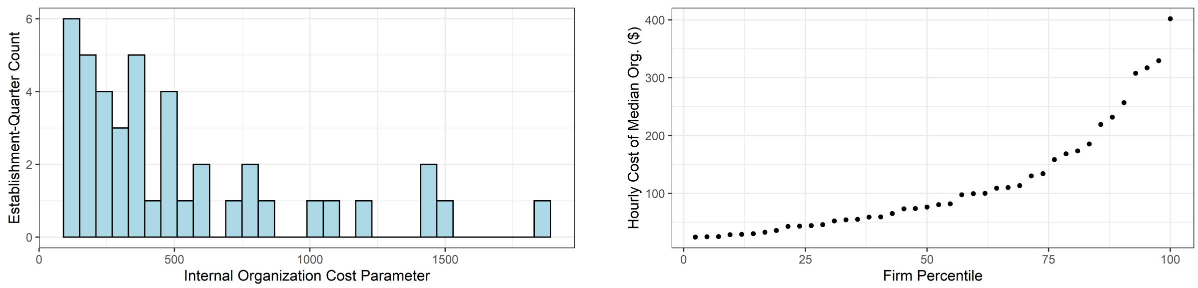
The coefficients associated with the color and haircut tasks are the most precisely estimated. This is not surprising, as these tasks are the most common and, as a consequence, their associated parameters will have the most statistical power. Across all tasks, the skill-gap parameters are positive, indicating that assigning the task to the associated specialist increases quality. The skill-gap parameters can be interpreted as the dollar value to a consumer of increasing task specialization in that task by 4 percentage points. Wages are in 2021 dollars per hour. Material costs are in terms of 2021 dollars per service.

Wages for color specialists are more than double the wages for haircut specialists, and the skill gap for color specialists is nearly double the skill gap for haircut specialists. This is in line with folk wisdom in the industry that it is hard to master coloring. The material costs are largest for the color task, in line with the fact that coloring is intense in expensive non-labor supplies, such as hair dye.

The organization costs (γ_j) for each firm are shown in Figure 12a. To provide a magnitude for these estimates, I plot the cost of implementing the median-complexity organization structure across all firms in Figure 12b. There are large differences in organization costs. Firms in the bottom quartile of organization costs can implement the structure for less than \$50 an hour. It would cost firms in the top quartile over \$150 an hour. The estimates imply that variation in organization costs explains 40% of total variation in prices across firms.[27]

27. This is obtained as the R-squared of regressing price on $\bar{a}_j \gamma_j I(B_j)$.

Figure 12: Estimated Organization Costs



(a) Organization Costs (γ_j)

(b) Cost of Median Org. Structure

Note: Panel A displays the estimated organization cost (γ_j) parameters for Manhattan. These can be interpreted as a measure of organization frictions at each firm, with lower values indicating less friction. Panel B displays the magnitude of these differences, by plotting the cost (in dollars) to each firm of implementing the median-complexity organization structure.

For each firm, I recover the unobserved, equilibrium organization structure B_j. Four examples are visualized in Figure 13. These matrices represent the amount of time allocated to each task and each worker type.[28]

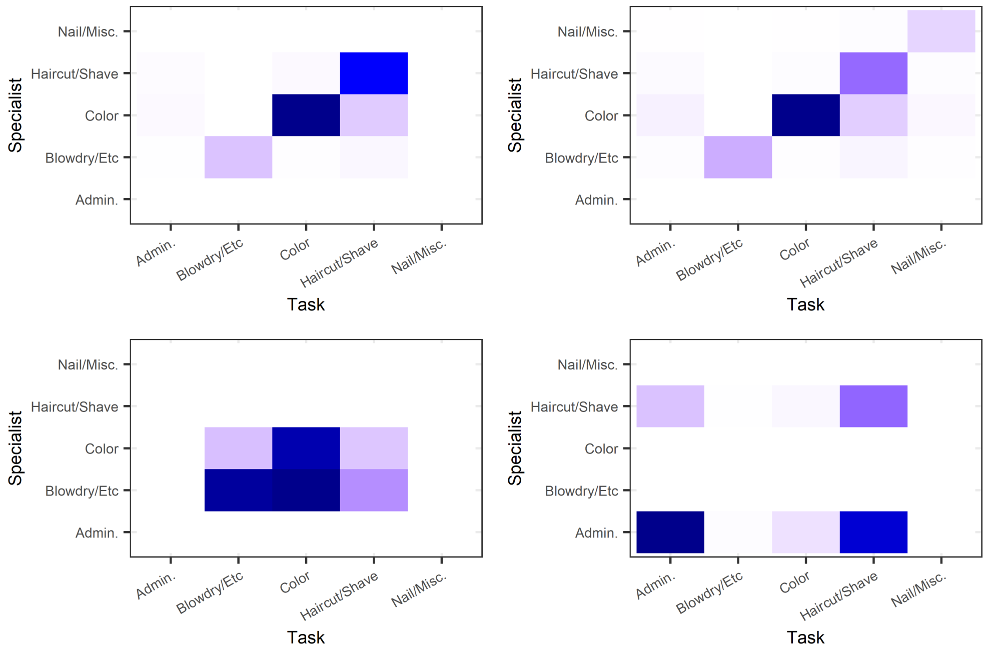
7.2 Model Fit and Validation

I assess model fit by comparing the predicted and actual relationship between prices and various organizational variables in Figure 14. The model captures the shape of the relationships.

Although the model delivers an entire predicted organization structure (B_j) for each firm, the estimation procedure uses only some of this information. I use the additional predicted information to validate the model. In particular, I compare the model-generated distribution of task content to the observed distribution of task content. Recall that the jobs within firm j are denoted by b_j, which is a matrix where element i, k denotes the time worker type i spends on task k. Using the model, I can compute b_j for each firm among worker types that it hires. In the data, I can compute $\tilde{b}_j$, which are jobs within firm j, where element i, k denotes the time worker i spends on task k. The main difference between $\tilde{b}_j$ and b_j is that the first is with respect to worker identity, and the second is with respect to worker types. To make them comparable, I can weight each job by the total amount of labor it represents. Combining all J firms yields an unobserved and model-based distribution of job task content for each of the five tasks, where jobs are weighted by their effective labor.

28. I present two estimated structures in tabular form in Appendix Table B5.

Figure 13: Estimated Organization Structures



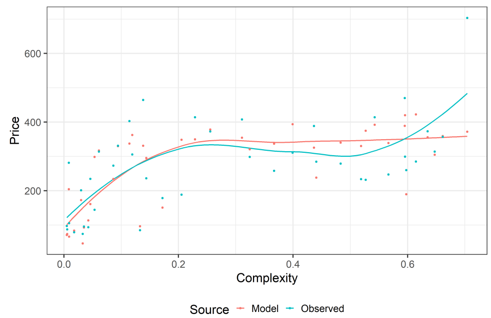
Note: Darker colors indicate a higher fraction of total labor was allocated to that worker-task pair. The first two structures are for organizationally efficient salons, while the last two are for organizationally inefficient salons.

Tables 7a, 7b and 7c compare the model and observed mean, median, variance, 25th percentile and 75th percentile of job task content. The estimated results exactly match the mean and between-firm variance of job task content because the model imposes that organization structures must be consistent with the task mix α_j, which is exactly the average amount of time spent on each task at each firm. The estimates are also reasonable approximations of the total variances of task content and the 75th percentile of the job task-content distribution. The model is not able to match the median and the 25th percentile.

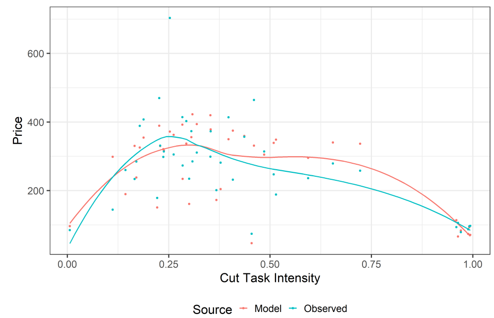
The statistics related to the color/highlight/wash task are the hardest for the model to replicate, because the empirical distribution for this task is triple peaked. The empirical and model-generated task-content distributions are presented in full in Appendix Figure B11.[29]

29. Comparing the entire distribution of actual and model-predicted task content is a demanding test of the model.

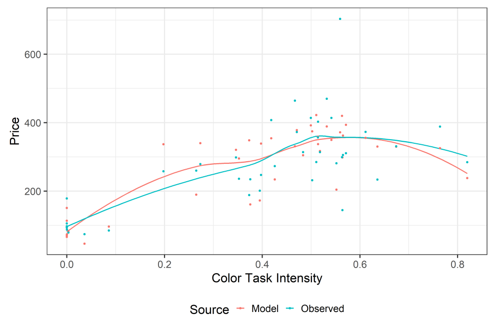
Figure 14: Model Fit



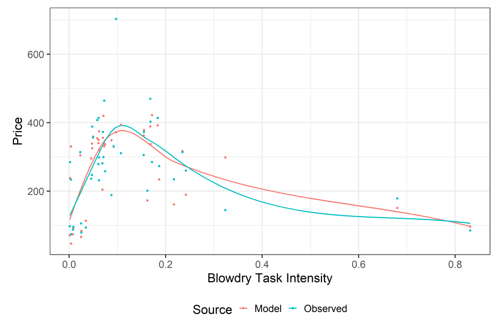
(a) Complexity



(b) Task Mix 1 (Haircut/Shave Task)



(c) Task Mix 2 (Haircut/Shave Task)



(d) Task Mix 3 (Haircut/Shave Task)

Note: Each panel plots the model and observed relationship between price and different firm variables. Dots represent individual firms, while lines are Loess smoothed fitted curves.

7.3 The Determinants of Task Specialization

The estimated model allows the researcher to understand how worker skills and firm internal organization determine the task specialization of jobs. Measuring task specialization as the amount of time a worker spends on their specialty task, I find that 45% of the variation in task specialization is attributable to firms, while 55% is attributable to worker skills. [30]

I calculate the variation in task specialization due to the firm component by computing the fraction of total firm labor spent on any worker's specialty. Firms with higher organization costs exhibit less task specialization. The magnitude of this effect is large: firms in the bottom quartile of organization costs (efficient firms) assign on average 90% of tasks to the associated specialist, while firms in the top quartile (inefficient firms), only 67%.

30. In the decomposition, I separate task-specialization variance into a within-worker type and across-worker type component. Since the only difference across worker types is firms, I can call the across-worker type component the firm component of variance.

There is also significant variation in specialization across worker types. Haircut/shave specialists work the most specialized jobs, spending 95% of their time on their specialty task. Blow dry/extension/style specialists work the most generalized jobs, spending only 48% of their time on their specialty task.

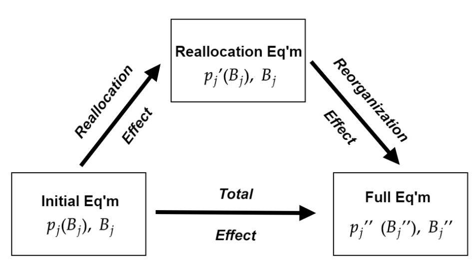
8 Counterfactuals

This section uses the estimated model to study two counterfactual policy changes, one impacting the product market and one impacting the labor market. Internal organization qualitatively alters the responses to these well-studied policies. The procedure used to solve for equilibria and conduct the analyses is described in Appendix Section B.12. The model allows me to distinguish between two effects of any policy: a reallocation effect and a reorganization effect.

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. The *reallocation effect* of any policy change is the change in outcomes between the reallocation equilibrium and the initial equilibrium. It captures changes due to the reallocation of labor across firms. Because firms differ in their organization costs and task mixtures, reallocation will change the task content of jobs, relative wages, and other outcomes.

The *reorganization effect* of any policy is the change in outcomes between the full equilibrium and the reallocation equilibrium. It captures changes due to reorganization of labor within firms. I define the *total effect* of any policy change as the change in outcomes from the initial to the full equilibrium. These relationships are summarized in Figure 15.

Figure 15: Reallocation, Reorganization and Total Effect



In the reallocation equilibrium, firms are acting as if they employ a composite worker. The worker's skills and wage are determined exogenously by the initial internal organization of the firm, B_j. The firm has the option of adjusting the total amount of labor it demands from this composite worker, but cannot adjust the worker's skills and wages. In the full equilibrium, the firm is free to fully adjust its internal structure.

8.1 Minimum Wage Increase

I study a counterfactual increase in the minimum wage in Manhattan from \$15 (the minimum in 2021) to \$20. An increase to \$20 is similar in magnitude to the increases that would occur if the minimum wage were pegged to inflation, as proposed in several pending pieces of legislation.[31]

To implement the counterfactual, I require that all equilibrium wages be at least \$20, and that markets clear for all worker types for which the wage is not binding. I allow there to be excess labor supply (unemployment) for those worker types facing a binding minimum wage. The model is well suited for studying large increases in the minimum wage because it allows salons to reorganize as well as raise prices. There are technical details that must be addressed when implementing the minimum wage counterfactuals, including the possibility of multiple equilibria and numeraire goods. I address these in Appendix Section B.12.2.

I find that the minimum wage binds for the haircut/shave specialist only. The new wages and employment levels across worker types are given in Table B7 (including values for the reallocation equilibrium). I first discuss the reallocation and reorganization effects of this policy change. I then analyze the overall impact of the new policy, using the reallocation and reorganization effects to understand the underlying forces.

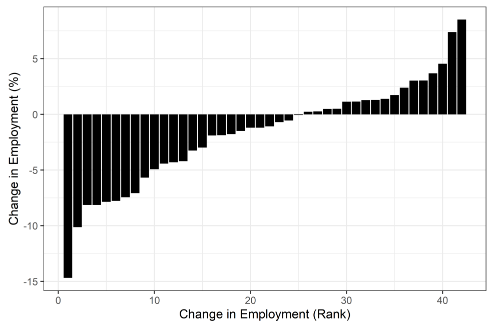
8.1.1 The Reallocation Effect

The impact of the minimum wage on individual salons depends partly on their initial internal structure. As a result, the minimum wage changes the competitive positions of salons and reallocates labor. By comparing the initial and reallocation equilibrium, I can hold each firm's internal structure fixed but allow firms to adjust prices. This captures the extensive margin adjustment of salons but prevents internal reorganization. Figure 16 presents the reallocation effect of the minimum wage in a series of three panels.

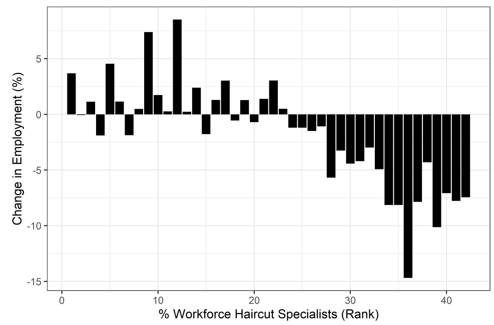
The minimum wage has a disproportionately negative impact on salons whose internal organization relies heavily on minimum wage workers. These salons see the largest increases in marginal costs and thus the largest decreases in output and employment. Because the minimum wage increases some salons' costs more than others', it changes the

31. Senate Bill S3062C and Assembly Bill A7503B.

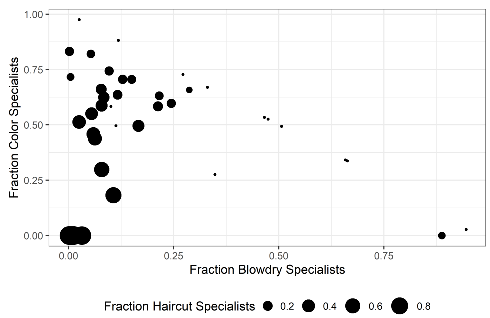
Figure 16: The Minimum Wage Reallocation Effect



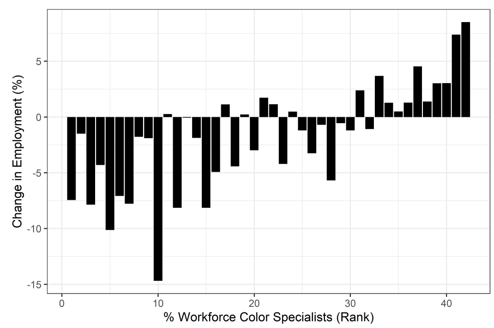
(a) Ordered by Employment Loss



(b) Ordered by Share Binding Workers



(c) Initial Coworkers



(d) Ordered by Share Color Specialists

Note: In panels A, B and D each bar is a firm, and employment changes are comparing the reallocation equilibrium to the initial equilibrium, holding fixed internal organization. Panel A orders firms by employment losses. Panel B reorders firms by the fraction of the workers that are haircut specialists in the initial equilibrium. Panel C plots firms by their initial workforce composition. Panel D orders firms by their share of color specialists.

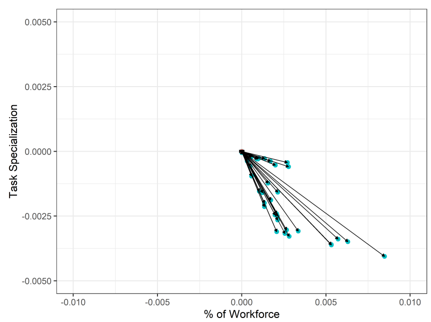
competitive position of firms in the product market. As can be seen in the figure, this effect is heterogeneous enough that some salons see employment increases.

Workers that are often employed alongside minimum wage workers initially see negative wage spillovers because the minimum wage erodes the competitive position of these firms. In the opposite way, workers employed at salons with few minimum wage workers initially see positive wage spillovers, because the minimum wage improves the competitive position of these firms. The effects of the minimum wage is contagious, and are spread across workers based on firm internal organization. In equilibrium, the minimum wage reallocates labor towards high-complexity, task-specialized salons and away from low-complexity, task-generalized salons, raising industry task specialization and average worker productivity.

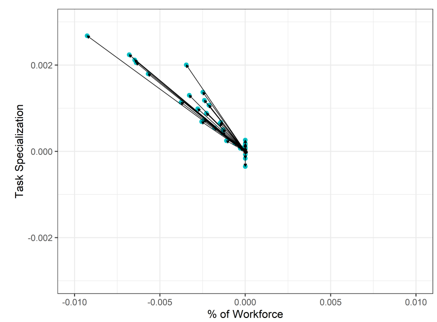
8.1.2 The Reorganization Effect

By comparing the full equilibrium and the equilibrium where firms can adjust only prices, I can study the effect of internal reorganization. In Figure 17, I plot the vectors representing firms, where the length and direction of the vector represents the change in the firm's relative labor demand for that worker type and the change in task specialization of that worker type at the firm.

Figure 17: Reorganization Effect Under a Minimum Wage Increase



(a) Non-Binding (Color/Highlight Specialists)



(b) Binding (Haircut/Shave Specialists)

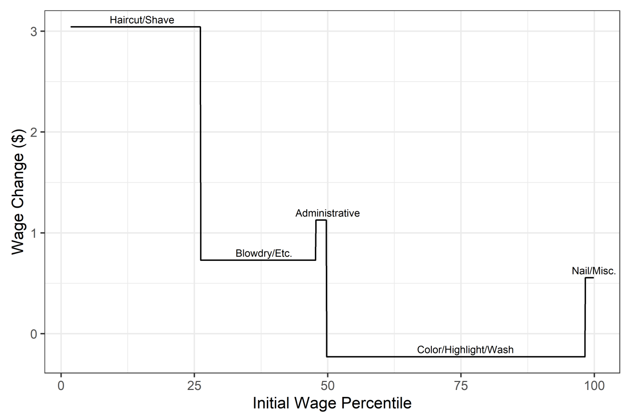
Note: Each arrow in both panels is a firm, with the blue dot at the end of the arrow representing the firm after the reorganization effect (the final position). Panel A displays the change in task specialization and relative employment for color/highlight specialists, a type for which the minimum wage is not binding. Relative employment increases, while task specialization decreases. Panel B displays the change in task specialization and relative employment for haircut/shave specialists, a type for which the minimum wage is binding. Relative employment decreases, while task specialization increases. This illustrates how firms are asking surviving workers to "pick up the slack."

The figure illustrates a general pattern. Salons reduce relative employment and increase task specialization of minimum wage workers. Salons reduce task specialization and increase relative employment for workers above the minimum wage. I call this a "pick-up-the-slack" effect. Intuitively, the minimum wage reduces the comparative advantage of workers for which the minimum wage binds in all tasks relative to other (non-binding) workers. Firms respond by laying off minimum wage workers and shifting tasks performed by them onto the relatively less expensive non-binding workers. Only minimum wage workers that are sufficiently productive survive, which are those who are task specialized. This implies that the minimum wage increases the absolute productivity of binding workers, but decreases the absolute productivity of non-binding workers.

8.1.3 Total Impact

Although the minimum wage is binding for only one worker type, all workers see wage changes. Table 8a shows that there are both positive and negative wage spillovers. The largest positive spillover is for administrative specialists, who see a wage increase of 4.2% (+$1.13). Color/highlight/wash specialists see a small wage decrease, of 0.7% (-$0.23). What is notable about these spillovers is that they are non-monotonic in initial wage. To see this, I plot the wage change experienced by different workers ordered by initial wage in Figure 18. Non-monotone spillovers occur because substitution patterns in the model are determined endogenously based on the distribution of firm organization costs and task mixtures.

Figure 18: Minimum Wage Spillovers Across the Initial Wage Distribution



Note: This figure plots the wage change experienced by different workers ordered by the initial wage of the worker. Haircut/shave specialists are the only binding type, so their wage increase is due to the direct effect of the minimum wage. All other wage changes are spillovers. While the majority of workers see wage increases, some see decreases. Spillovers are not monotone in initial wage.

These non-monotonic wage spillovers illustrate that internal organization can link workers that are very far apart in the initial wage distribution. Workers that differ horizontally in their specialty may be quite likely to work alongside each other, and they may have quite different wages depending on other factors. In this way the reallocation effect can cause large wage increases or decreases even for high-wage workers: indeed, this is exactly what I observe with haircut and nail specialists. Similarly, because the initial wage distribution is not determined by vertical skill differences, the reorganization effect will induce firms to shift tasks from minimum wage workers to workers across the wage distribution.

In Table 9 I decompose wage spillovers into those arising from the reallocation and the reorganization effect. As discussed in the prior two subsections, spillovers for each worker type are a combination of forces, with the reorganization and reallocation effects sometimes moving in opposite directions. For example, color specialists see negative wage spillovers because they are employed alongside minimum wage workers and the minimum wage increase disadvantages the salon where they work. But they also see positive wage spillovers because firms shift tasks from minimum workers to them during reorganization. The total wage spillover for color specialists is negative, as the reallocation effect is about double the reorganization effect. For binding workers (haircut/shave specialists) the two effects work in the same direction, increasing unemployment. In this sense, internal reorganization amplifies unemployment losses.

Table 9b shows that reorganization wage spillovers follow a pattern. Workers that see an increase in task specialization see a wage decrease (or unemployment increase), while workers that see a decrease in task specialization see a wage increase. Because task specialization determines worker productivity, this implies that internal reorganization causes absolute productivity and wages to move in opposite directions.

8.2 Sales Taxes

New York City is unique in that it levies a 4.5% sales tax on certain services, including those performed at hair salons. This section studies the effect of eliminating this sales tax. Formally, I estimate a new equilibrium with $\tau^{NEW} = 0$. The wages in this new equilibrium are provided in Table B8.[32] I first discuss the reallocation and reorganization effects of the policy. I then analyze the overall impact of the policy, using the reallocation and reorganization effects to understand their driving mechanisms.

8.2.1 Reallocation Effect

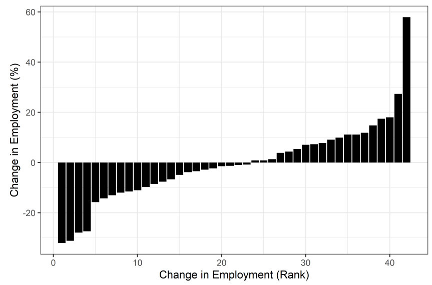
Eliminating the sales tax confers a competitive advantage on firms producing high-quality services in the initial equilibrium. Since salons with low organization costs tend to produce high-quality services, eliminating the sales tax reallocates labor towards organizationally efficient firms, as seen in Figure 19. These firms produce high-quality services using a more task-specialized internal structure. Thus the reallocation effect increases market-wide task specialization because more workers are working at task-specialized firms.

8.2.2 Reorganization Effect

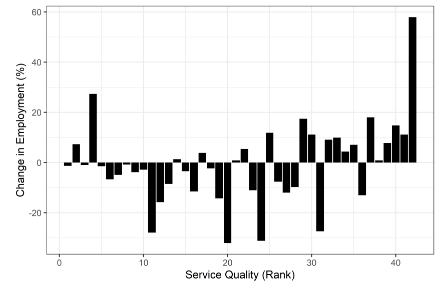
Eliminating the sales tax makes producing higher-quality products more attractive. In order to produce higher-quality products, firms choose internal organizations which are

32. Employment remains the same before and after the policy because labor supply is inelastic and the minimum wage is assumed not to be binding.

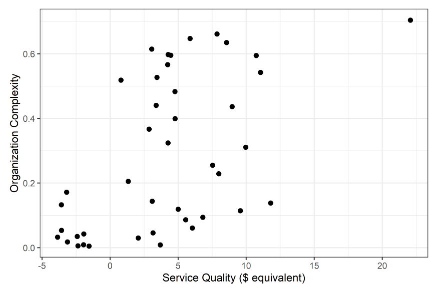
Figure 19: Sales Tax Reallocation Effect



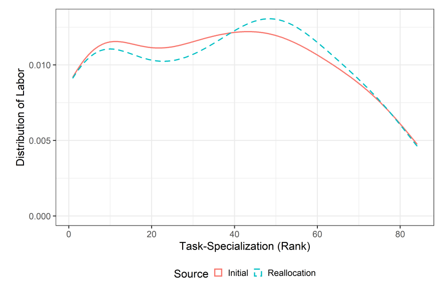
(a) Heterogeneous Employment Effects



(b) Ordered by Quality



(c) High-Quality Salons Are Complex



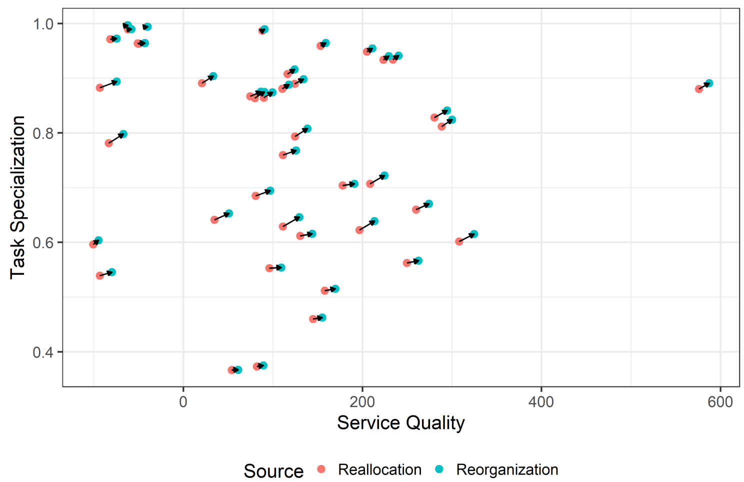
(d) Shift in Distribution of Labor

Note: Each bar is a salon. The sales tax elimination decreases employment at some salons and increases it at others (Panel A). Salons with high-quality service see an improvement in their competitive position (Panel B). These salons have a complex, task-specialized internal structure (Panel C). As a result, labor is reallocated to task-specialized firms, and workers become more specialized in equilibrium (Panel D).

on average 5.5% more complex, increasing average labor market task specialization by 0.9%. In terms of the three-way trade-off introduced in Figure 9, eliminating the sales tax has the same effect as reducing consumer price sensitivity (ρ). Average firm service quality rises by 10%. This is consistent with the quality-complexity-wage three-way trade-off discussion in the theoretical section.

Figure 20 illustrates that these market-wide patterns also happen at the firm level. However, the extent to which salons increase quality and increase task specialization depends on the firm's internal organization costs and its particular task mix. Thus, the slopes and lengths of the arrows in Figure 20 differ. Changing sales tax, a product market policy, influences what workers do and what workers are paid in the labor market.

Figure 20: Reorganization Effect Under a Sales Tax



Note: Each pair of dots connected by an arrow represents a firm, with red representing the firm before the sales tax and blue representing the firm after the sales tax. The direction of the arrows indicates that most salons increase quality by raising task specialization internally. The magnitude of this change (given by the length and angle of the arrow) depends on the firm's particular organization costs and task mixture.

8.2.3 Total Impact

Table 12a summarizes the effect of the policy on wages and task specialization. All worker types see wage increases and task-specialization increases. Wage increases are not proportional to task-specialization increases: even though blow-dry specialists see the largest increase in specialization, they see the lowest increase in wages. This is because the size of wage increases is partly driven by how the policy impacts the competitive position of firms.

The welfare effects of the policy are summarized in Table 12b. Overall, eliminating the sales tax leads to a small welfare increase, of 0.19%. However, the effects are quite different for different actors in the model. Firms respond to the sales tax elimination by increasing quality by 10%. Firms capture the surplus from improved quality and reduced taxes from consumers by raising prices by 8.7%. Firm profit increases by a modest 0.58% because workers capture most of the surplus from firms through higher wages, which rise by a dollar amount that is comparable to the total lost tax revenue. This is consistent with workers capturing almost all of the productivity improvements from increased task specialization.

9 Conclusion

This paper studies how internal organization decisions within firms interact with markets outside firms. I develop a structural model, grounded in a set of stylized facts, which allows firms to differ in their internal organization and to change it in response to market conditions. Workers have multidimensional skills and different wages. Internal organization matters because the match between workers and tasks determines wage costs and product qualities. Firms in the same market choose different internal organizations because firms vary both in their ability to internally organize and in their task-based production functions.

The counterfactual exercises illustrate that allowing internal organization to be endogenous and heterogeneous qualitatively changes the impact of policy. Minimum wage increases generate new types of wage spillovers that cannot occur in many other models of the labor market. Sales tax cuts induce firms to reorganize their workforce, changing the task composition of jobs. Although these effects are specific to the salon industry, they indicate that internal organization is an important force that deserves careful study in a variety of contexts.

The framework provided in this paper provides a starting point for researchers to do exactly this. The approach in this paper can be extended to accommodate quantity-based (rather than quality-based) productivity, continuous task spaces, labor market power and more sophisticated demand systems. These extensions, combined with traditional employer-employee matched data will be important to answer future questions, including the effect of internal organization on human capital accumulation and the welfare implications of endogenous task assignment for workers.

References

Adenbaum, Jacob. 2021. "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.

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

Argyres, Nicholas S., Teppo Felin, Nicolai Foss, and Todd Zenger. 2012. "Organizational Economics of Capability and Heterogeneity." *https://doi.org/10.1287/orsc.1120.0746* 23, no. 5 (October): 1213–1226.

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.

Becker, Gary S., and Kevin M. Murphy. 1992. “The division of labor, coordination costs, and knowledge.” *Quarterly Journal of Economics* 107 (4): 1137–1160.

Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. “Automobile Prices in Market Equilibrium.” *Econometrica* 63 (4): 841–890.

Blahut, Richard E. 1972. “Computation of Channel Capacity and Rate-Distortion Functions.” *IEEE TRANSACTIONS ON INFORMATION THEORY,* no. 4.

Bloedel, Alexander, and Weijie Zhong. 2021. “The Cost of Optimally Acquired Information*.” June.

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.

Caliendo, Lorenzo, Ferdinando Monte, and Esteban Rossi-Hansberg. 2015. “The Anatomy of French Production Hierarchies.” *The Journal of Political Economy* 123, no. 4 (July).

Caliendo, Lorenzo, Esteban Rossi-Hansberg, Lorenzo Caliendo, and Esteban Rossi-Hansberg. 2012. “The Impact of Trade on Organization and Productivity.” *The Quarterly Journal of Economics* 127, no. 3 (August): 1393–1467.

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

Card, David, Ana Rute Cardoso, Joerg Heining, and Patrick Kline. 2018. “Firms and labor market inequality: Evidence and some theory.” *Journal of Labor Economics* 36, no. S1 (January): S13–S70.

Chausse, Pierre. 2021. *Package 'gmm'.*

Chen, Po-Ning. n.d. “Chapter 6 Lossy Data Compression and Transmission.”

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

d’Aspremont, C., J. Jaskold Gabszewicz, and J.-F. Thisse. 1979. “On Hotelling’s ”Stability in Competition”.” *Econometrica* 47, no. 5 (September): 1145.

Dai, Tianjiao, and Juuso Toikka. 2022. “Robust Incentives for Teams.” *Econometrica* 90, no. 4 (July): 1583–1613.

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

Du, Yu, and Ravi Varadhan. 2020. “SQUAREM: An R package for off-the-shelf acceleration of EM, MM and other EM-like monotone algorithms.” *Journal of Statistical Software* 92:1–41.

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 Esteban Rossi-Hansberg. 2006. “ORGANIZATION AND INEQUALITY IN A KNOWLEDGE ECONOMY*.”

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

Gibbons, Robert. 2020. “March-ing toward organizational economics.” *Industrial and Corporate Change* 29, no. 1 (February): 89–94.

Haanwinckel, Daniel. 2020. “Supply, Demand, Institutions, and Firms: A Theory of Labor Market Sorting and the Wage Distribution.”

Holmstrom, Bengt, and Paul Milgrom. 1994. “The Firm as an Incentive System.” *American Economci Review* 84 (4): 972–991.

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.

Kremer, Michael. 1993. "The o-ring theory of economic development." *Quarterly Journal of Economics* 108 (3): 551–575.

Kugler, Maurice, and Eric Verhoogen. 2012. "Prices, Plant Size, and Product Quality." *The Review of Economic Studies* 79, no. 1 (January): 307–339.

Kuhn, Moritz, Jinfeng Luo, Iourii Manovskii, and Xincheng Qiu. 2022. "Coordinated Firm-Level Work Processes and Macroeconomic Resilience *."

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

Li, Fei, and Can Tian. 2013. "Directed search and job rotation." *Journal of Economic Theory* 148, no. 3 (May): 1268–1281.

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

Lipnowski, Elliot, and Doron Ravid. 2022. "Predicting Choice from Information Costs" (May).

Lucas, Robert E. 1978. "On the Size Distribution of Business Firms." *The Bell Journal of Economics* 9 (2): 508.

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.

Nocke, Volker, and Nicolas Schutz. 2018. "Multiproduct-Firm Oligopoly: An Aggregative Games Approach." *Econometrica* 86, no. 2 (March): 523–557.

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

Pomatto, Luciano, Philipp Strack, and Omer Tamuz. 2022. "The Cost of Information."

Rosen, Sherwin. 1982. "Authority, Control, and the Distribution of Earnings." *Bell Journal of Economics* 13 (2): 311–323.

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

Shaked, Avner, and John Sutton. 1987. "Product Differentiation and Industrial Structure." *The Journal of Industrial Economics* 36, no. 2 (December): 131.

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

Shkel, Yanina, and Sergio Verdú. 2018. "A Coding Theorem for f-Separable Distortion Measures." *Entropy* 20, no. 2 (February).

Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter. 2019. "Firming Up Inequality." *The Quarterly Journal of Economics* 134, no. 1 (February): 1–50.

Teulings, Coen N. 2000. "Aggregation Bias in Elasticities of Substitution and the Minimum Wage Paradox." *International Economic Review* 41, no. 2 (May): 359–398.

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

Tirole, Jean. 2009. "Cognition and Incomplete Contracts." *American Economic Review* 99, no. 1 (March): 265–94.

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

Williamson, Oliver. 1984. "The Incentive Limits of Firms: A Comparative Institutional Assessment of Bureaucracy," 736–763.

Tables

Table 1: Regressions of Worker Specialization on Organization Complexity

Dependent Variable:	Worker Task Specialization		
Model:	(1)	(2)	(3)
Variables			
Organization Complexity	0.2853***	0.2862***	0.2922***
	(0.0313)	(0.0310)	(0.0392)
Fixed-effects			
Quarter-Year		Yes	Yes
County			Yes
Observations	62,452	62,452	62,452
R^2	0.10184	0.10901	0.21483

Standard errors clustered at the salon level.
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Note: Task specialization is measured as the maximum fraction of time spent on a single task by a worker. Complexity is measured at the salon level. Across all specifications, complexity (a salon-level measure) can account for 10% of the variation in worker specialization.

Table 2: Regressions of Salon Size on Organization Complexity

Dependent Variables:	Revenue	Employees	Utilized Labor	Customers	Visits
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Org. Complexity	347549.2***	9.75**	26481	334.6	731.7
	(79546.2)	(3.016)	(35653.2)	(259.6)	(450.1)
Fixed-effects					
Quarter-Year	Yes	Yes	Yes	Yes	Yes
County	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	4,558	4,558	4,558	4,558	4,558
R^2	0.32465	0.34319	0.28918	0.34901	0.35004

Standard-errors clustered at the salon level.
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Note: Observations are salon-quarters. Regressions illustrate a positive correlation between complexity and several measures of salon size after controlling for county and quarter fixed effects and the composition of tasks performed at the salon in the quarter.

Table 3: Salon Activity Data Sample

Firm	Salon	App.	Cust.	Service	Staff	Time Stamp	Price	Duration
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 IDs are replaced by pseudonyms.

Table 4: Summary Statistics for All Salon-Quarters

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Revenue	4,558	213,201.30	248,359.90	5	58,912.5	271,236.5	2,559,703
Price	4,558	199.73	135.16	0.20	111.71	261.88	3,180.44
Employees	4,558	13.38	10.79	1	6	17	92
Customers	4,558	1,159.23	1,098.45	1	397	1,619	16,768
Task Categories	4,558	4.45	0.86	1	4	5	5
Labor per. Customer	4,558	2.15	1.63	0.10	1.52	2.57	61.33

Note: The table displays summary statistics for the main variables of interest with data aggregated at the salon-quarter level. There is significant variation across salons in complexity, number of employees, revenue and many other dimensions.

Table 5: Parameter Estimates, Tasks

Task	Associated Specialist Skill Gap	Associated Specialist Wage	Skill Base	Material Cost
Administrative	43.29*	26.99	-16.16	-147.60*
	(21.66)	(63.75)	(14.58)	(13.47)
Blowdry/Etc.	141.69*	20.91	-70.56*	12.39
	(36.67)	(40.22)	(13.57)	(16.65)
Color/Highlight/Wash	60.03*	37.75*	-9.69	56.49*
	(21.24)	(7.00)	(11.97)	(15.79)
Haircut/Shave	32.45*	16.96*	.	.
	(13.07)	(8.32)	.	.
Nail/Spa/Eye/Misc.	66.48	81.16	-252.58*	-1061.12*
	(37.72)	(53.52)	(11.47)	(10.73)

Note: Standard errors from 500 bootstrap replications in parentheses; * indicates significance at the 0.05 level. For each task, the table lists the skill gap and wage of the associated specialist in 2021 dollars. The skill gap is the change in quality when a task is assigned to a specialist. Also listed are the skill base, the quality when the task is performed by a non-specialist, the material cost, and the non-wage costs associated with the task (e.g., dye for coloring). Material costs and skill base are relative to the haircut task. Wages are per hour, while material costs and skills are per unit.

Table 6: Parameter Estimates, Other

Parameter	Estimate
Price Sensitivity	0.04*
	(0.01)
Cost Intercept	27.95
	(15.21)
Utility Intercept	-24.77*
	(8.36)

Note: Standard errors from 500 bootstrap replications in parentheses; * indicates significance at the 0.05 level. Consumer price sensitivity (ρ) is the main determinant of demand elasticities.

Table 7: Model Validation: Estimated vs. Observed Job Task Content

(a) Mean and Median

Task	Mean		Median	
	Model	Observed	Model	Observed
Haircut/Shave	0.4094	0.4094	0.2816	0.3357
Color/Highlight/Wash	0.4058	0.4058	0.3067	0.4042
Blowdry/Style/Treatment/Extension	0.1179	0.1179	0.0162	0.0704
Administrative	0.0278	0.0278	0.0050	0.0040
Nail/Spa/Eye/Misc.	0.0391	0.0391	0.0049	0.0000

(b) Variance

Task	Total Variance		Between Firm Variance	
	Model	Observed	Model	Observed
Haircut/Shave	0.1110	0.1268	0.0597	0.0597
Color/Highlight/Wash	0.1127	0.1105	0.0365	0.0365
Blowdry/Style/Treatment/Extension	0.0472	0.0194	0.0111	0.0111
Administrative	0.0098	0.0080	0.0063	0.0063
Nail/Spa/Eye/Misc.	0.0120	0.0171	0.0050	0.0050

(c) Interquartile Range

Task	p25		p75	
	Model	Observed	Model	Observed
Haircut/Shave	0.1583	0.0469	0.8013	0.7577
Color/Highlight/Wash	0.0417	0.0388	0.7020	0.6383
Blowdry/Style/Treatment/Extension	0.0004	0.0110	0.0726	0.1892
Administrative	0.0027	0.0000	0.0166	0.0108
Nail/Spa/Eye/Misc.	0.0000	0.0000	0.0329	0.0106

Note: The table compares model-generated and observed job task content along several dimensions. The model is designed to exactly match the average market-wide amount of time spent on each task and the between-firm variance. The other moments were not targeted, and assessing their match serves as a validation exercise.

Table 8: Total Effects of Increasing the Minimum Wage

(a) Wage Changes by Worker Type

Type	Wage Change	Total Wages Gained/Lost
Haircut/Shave - UNEMPLOYED	-100.00%	-$600,240
Haircut/Shave - EMPLOYED	17.95%	$1,528,205
Color/Highlight/Wash	-0.61%	-$228,453
Blowdry/Style/Treatment/Extension	3.48%	$323,374
Administrative	4.17%	$47,154
Nail/Spa/Eye/Misc.	0.68%	$19,319

(b) Welfare Breakdown

Source	Change	Percent Change
Salon Profit	-$714,413	-0.472%
Consumer Welfare	-$2,528,784	-1.671%
Employed Wages	$1,689,600	1.116%
Unemployed Wages	-$600,240	-0.397%
Total Welfare	-$2,153,838	-1.423%

Note: Increasing the minimum wage generates both positive and negative wage spillovers for workers on whom it is not binding. Positive spillovers are larger and occur for most worker types. Overall, wage increases for employed workers are more than salon profit losses and wage losses of unemployed workers combined. Total welfare declines, as consumers see higher prices and slightly lower quality.

Table 9: Spillovers from an Increase in the Minimum Wage

(a) Reallocation Effect

Type	Reallocation Change		
	Employment	Task-Spec.	Wage
Haircut/Shave	-5.85%	-0.04%	17.95%
Color/Highlight/Wash	0%	-0.17%	-1.13%
Blowdry/Style/Treatment/Extension	0%	-0.40%	4.63%
Administrative	0%	0.09%	5.22%
Nail/Spa/Eye/Misc.	0%	-0.03%	0.58%

(b) Reorganization Effect

Type	Reorganization Change		
	Employment	Task-Spec.	Wage
Haircut/Shave	-0.73%	0.12%	0%
Color/Highlight/Wash	0%	-0.33%	0.52%
Blowdry/Style/Treatment/Extension	0%	0.03%	-1.15%
Administrative	0%	0.03%	-1.05%
Nail/Spa/Eye/Misc.	0%	-0.00%	0.10%

Note: The minimum wage increase has positive spillovers for some workers and negative spillovers for others. These spillovers can be further decomposed to those resulting from reorganization and those resulting from reallocation. Most spillovers come from the fact that the policy favors salons that have internal organizations intense in binding workers initially (reallocation). Some spillovers occur because the policy induces firms to shift tasks from binding to non-binding workers (reorganization).

Table 10: Summary of All Minimum Wage Increase Effects

Statistic	Reallocation	Reorganization	Total
Avg. Price	1.96%	-0.29%	1.67%
Avg. Complexity	0.00%	-0.46%	-0.46%
Avg. Quality	0.00%	-0.54%	-0.54%
Avg. Hourly Wage	3.40%	0.20%	3.60%
Std. Dev. Wage	-8.91%	1.03%	-7.88%
Task Specialization	-0.61%	-0.18%	-0.79%
Employment	-1.53%	-0.19%	-1.72%
Market Served	-2.69%	-0.12%	-2.81%
Total Profit	-2.69%	-0.12%	-2.81%
Consumer Welfare	-2.64%	-1.19%	-3.83%
Total Wages	1.81%	0.00%	1.82%
Total Welfare	-0.88%	-0.54%	-1.42%

Note: This table summarizes the impact of increasing the minimum wage from \$15 to \$20 on different actions and market outcomes in the Manhattan hair-salon market.

Table 11: Summary of All Sales-Tax-Elimination Effects

Statistic	Reallocation	Reorganization	Total
Avg. Price	4.70%	3.99%	8.68%
Avg. Complexity	0.00%	5.53%	5.53%
Avg. Quality	0.00%	10.03%	10.03%
Avg. Hourly Wage	18.32%	1.02%	19.34%
Std. Dev. Wage	22.67%	-6.32%	16.35%
Task Specialization	0.90%	0.93%	1.83%
Total Profit	4.32%	-0.60%	3.71%
Consumer Welfare	-0.18%	-0.57%	-0.75%
Total Wages	18.32%	1.02%	19.34%
Total Welfare	0.14%	0.05%	0.19%

Note: This table summarizes the impact of eliminating the service sales tax on different actions and market outcomes in the Manhattan hair-salon market.

Table 12: Total Effects of a Sales-Tax Elimination

(a) Wage Changes by Worker Type

Type	Wage Change	Task-Spec. Change
Haircut/Shave	31.99%	0.29%
Color/Highlight/Wash	20.09%	2.57%
Blowdry/Style/Treatment/Extension	6.06%	3.01%
Administrative	17.99%	1.03%
Nail/Spa/Eye/Misc.	12.74%	2.39%

(b) Welfare Breakdown

Source	Change	Percent Change
Salon Profit	$942,740	0.58%
Consumer Welfare	-$494,199	-0.30%
Wages	$11,603,777	7.12%
Tax Revenue	-$11,739,300	-7.20%
Total Welfare	$313,017	0.19%

Note: Eliminating the sales tax raises wages most in percentage terms for haircut specialists. Workers gain the most from the elimination of the sales tax: wage increases are almost equal to the lost revenue to the government.

A Theoretical Appendix

A.1 Rate Distortion and Rational Inattention Equivalence

Equation (3) from Theorem 1 can be rewritten as

$$\gamma_j \min_{B_j \in \mathbb{B}} \left\{ I(B_j) + \gamma_j^{-1} \Big[W(B_j) - \rho^{-1} \xi(B_j) \Big] \right\}. \tag{8}$$

I can rewrite (8) as a maximization problem:

$$\max_{B_j \in \mathbb{B}} \left\{ \sum_{i,k} B_j(i,k)(\rho^{-1}\theta_{i,k} - W_i) \Big] - \gamma_j I(B_j) \right\}. \tag{9}$$

Comparing (9) 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 8 one last time:

$$\gamma_j \min_{B_j \in \mathbb{B}} \left\{ I(B_j) + \gamma_j^{-1} \sum_{i,k} B_j(i,k)(W_i - \rho^{-1}\theta_{i,k}) \Big] \Big] \right\}. \tag{10}$$

Comparing Equation (10) 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.

A.2 Proof of Theorem 1

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.

First, I prove that if an organization structure B_j^* solves the simpler problem (Equation 3), 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 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 (3). Such a structure always exists because Equation (3) is a rate-distortion/rational inattention problem, as shown in Appendix Section A.1.

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 $\xi(B_j^*) - \rho p_j$:

$$
\begin{aligned}
&= \xi(B_j^*) - \rho[p_j' + \gamma_j I(B_j^*) + W(B_j^*) - \gamma_j I(B_j') - W(B_j')] && (11)\\
&= \xi(B_j^*) - \rho[p_j' + \gamma_j I(B_j^*) + W(B_j^*) - \gamma_j I(B_j') - W(B_j')] + \xi(B_j') - \xi(B_j') && (12)\\
&= \xi(B_j') - \rho[p_j' + \gamma_j I(B_j^*) + W(B_j^*) - \gamma_j I(B_j') - W(B_j') - \rho^{-1}\xi(B_j^*) + \rho^{-1}\xi(B_j')] && (13)\\
&= \xi(B_j') - \rho p_j' - \rho[\underbrace{\gamma_j I(B_j^*) + W(B_j^*) - \rho^{-1}\xi(B_j^*) - \{\gamma_j I(B_j') + W(B_j') - \rho^{-1}\xi(B_j')\}}_{\leq 0 \text{ because } B_j^* \text{ minimizes}}] && (14)\\
&\geq \xi(B_j') - \rho p_j'. && (15)
\end{aligned}
$$

This proves the "only if" direction. I now prove that if a structure B_j^* is profit maximizing, it solves Equation (3) (the "if" direction). Suppose for sake of contradiction there exists B_j' which is profit maximizing but does not solve Equation (3). Then, as in the first part of the proof, there exists B_j^* which does solve Equation (3). 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 (3), $\xi(B_j^*) - \rho p_j > \xi(B_j') - \rho p_j'$, and thus profit is strictly higher under B_j^*, p_j. This contradicts optimality of B_j' and concludes the proof.

A.3 Proof of Proposition 1 and 2

I have already shown in Theorem 1 that optimal B solves a rate-distortion problem.

- Denote by Q the quality-adjusted wages. Denote by $I^*(Q)$ the optimal complexity as a function of quality-adjusted wages.
- RD equivalence $\implies$ $I^*(Q)$ is continuous, convex and decreasing. It is also strictly decreasing above some threshold $\bar{Q}$ (Chen, n.d.).
- 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:

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

- Examine the FOC:

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

- 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.4 Optimal Jobs Within the Firm

The last result shows the originally high-dimensional problem of the firm can be reduced to a tractable two-dimensional trade-off. However, one of the goals of the model is to understand how firms assign workers to tasks. This section describes the properties of task assignments within the firm and shows that the firm customizes the bundles of tasks it assigns individual worker types. For this, I define the job of worker type i at firm j as a vector $(b_j(i,\cdot))$, where element k denotes the amount of i's time spent on task k. The jobs at a firm are the rows of the organization structure divided by the total labor of worker type i:

$$b_j(i,k) = \frac{B_j(i,k)}{\sum_{k'} B_j(i,k')}.$$

Proposition 5 *The profit-maximizing organizational structure satisfies the following properties.*

1. ***Law of Demand:*** *The share of workers of type i ($E_j(i)$) decreases as their wage increases.*
2. ***Incomplete Specialization:*** *All hired worker types spend a positive amount of time on each task whenever $\gamma_j > 0$.*
3. ***Optimal Jobs:*** *Jobs take the following logit-like form:*

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

I prove this result by appealing to the rational inattention literature. I derive the expression for optimal jobs by manipulating the first-order conditions and the constraints. The proof is provided in Appendix Section A.4.1. Even though at a high level the firms are trading off complexity and quality-adjusted wages, under the surface, they customize jobs for individual workers and tasks. The proposition illustrates that task assignments depend on skills through $\theta_{i,k}$, wages through w_i, consumer price sensitivity through ρ, the task mix through α_k, and organization costs through γ_j. This proposition highlights two important features of the model. First, whenever there are some organizational frictions within a firm, complete specialization will not occur. Every "job" will be a bundle of multiple tasks. Second, because jobs depend on organization costs, where someone works matters for what they do. That is, two identical workers will not perform the same tasks even in the same product and labor market. The tasks included in any job will depend on the firm where a worker is employed.

A.4.1 Proof of Proposition 5

For the purposes of this proof only, I define $h_{i,k}$ as the fraction of task k performed by worker i. Then the optimal job of worker i is given by

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

Summing over i yields

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

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 $h_{i,k}$,

$$B_{i,k} = \alpha_k h_{i,k}.$$

To get to jobs, I divide by E_i:

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

This illustrates that optimal jobs take a multinomial 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.5 Proof of Proposition 3

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

$$\max_{B_j \in \mathbb{B}} \max_{p_j \in \mathbb{R}_+} \underbrace{\frac{exp(\overbrace{\xi(B_j)}^{\text{quality}} - \rho p_j)}{\sum_{j'} exp(\xi(B_{j'}) - \rho p_{j'})}}_{\text{market share, } s_j} \left[p_j - \underbrace{\left(\overbrace{\gamma_j I(B_j)}^{\text{org.}} + \overbrace{\sum_{i,k} w_i B_j(i,k)}^{\text{avg. wage}} \right)}_{\text{constant marginal cost, } MC_j} \right].$$

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. In the proof of Theorem 1, I substituted the equation characterizing the optimal price into profit, and showed that the best response B_j also solves

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

The best-response structure will therefore depend on other actions of the firm only through wages. The theorem also establishes that this is equivalent to a rational inattention problem with a mutual information cost function. With the equivalence to a rational inattention problem, I can establish existence. I can then 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 all firms satisfying the following condition have a unique organization structure which maximizes profits:

Assumption 1 *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.*

Whenever this holds for all firms, there is a unique cost and quality for each firm, which, according to Caplin and Nalebuff (1991), implies there are unique equilibrium prices. Thus this condition is sufficient (but not necessary) to guarantee uniqueness. This condition is testable, but it requires many parameters, including wages, to be known. Since I wish to estimate the parameters, it is not satisfactory. To get a more general result, I now appeal to Lipnowski and Ravid (2022).

Note that a rational inattention problem with mutual information costs is a special case of the problems considered by Lipnowski and Ravid (2022). A stochastic choice rule in their language is an organization structure in mine. Proposition 1 of their paper (translated to the language of my model) states that if γ_j is known, the set of quality-adjusted wages which generate multiple organization structures is "meager and shy." Since I consider the case of finite tasks (finite Ω in their language), "meager and shy" implies a null set. This is only for one firm with a specific γ_j. The set of quality-adjusted wages which generate multiplicity for at least one firm will be the union of all sets which generate multiplicity for each individual firm. The union of countable null sets is also null; therefore, the set of quality-adjusted wages that generate multiplicity is null.

Denote the set of quality-adjusted wages which generate multiplicity as $\mathbb{M}$. The mapping from market parameters Ω to quality-adjusted wages is defined by a multivariate, vector-valued function $F : \mathbb{R}_+^{N\times K+N+1} \to \mathbb{R}_+^{N\times K}$. It can be shown that if F is smooth and the rank of the Jacobian of F is at least $N \times K$, then the measure of the pre-image of any measure 0 set is 0.

I now prove that F satisfies the rank condition. Recall that the quality-adjusted wage of worker i and task k has the form $w_i - \rho^{-1}\theta_{i,k}$. Collapse i, k into a single index, $y = 1, ..., N \times K$, where $\mathcal{I}(\cdot)$ and $\mathcal{K}(\cdot)$ return the task and worker type associated with the index y. Then that element y of F is

$$F(\Omega) = w_{\mathcal{I}(y)} - \rho^{-1}\theta_{\mathcal{I}(y),\mathcal{K}(y)}.$$

The Jacobian of this function is a rank of at least $N \times K$ because each skill parameter $\theta_{i,k}$

impacts only one quality-adjusted wage. Formally, there exist at least $N \times K$ columns of the Jacobian which are linearly independent of each other. Thus the pre-image of the null set $\mathbb{M}$ on F will be measure 0. Since the pre-image is the set of parameters which generate multiplicity, the set of parameters which generate multiplicity is measure 0.

An implication of this result is that a pure-strategy Nash equilibrium exists for any fixed wages. I conjecture that this proof could be extended to show equilibrium existence in the full model, that is, when wages are determined by market clearing. One approach would be to prove that excess labor demands satisfy Kakutani's fixed point theorem. Extending the uniqueness result to say that the equilibrium is unique for almost any total labor supplies may not be possible. This is because, in general, worker types may be complements or substitutes depending on their skill sets. If firms are homogeneous with respect to task mixtures and organization costs, the wages that clear the market may very well be the wages which induce indifference across multiple organization structures and multiple equilibria.

A.6 Proof of Proposition 4

As stated earlier in the main text, the proof of this proposition will consist of three parts. For simplicity, firm index j is suppressed throughout this section. I denote by $I(\tilde{B})$ the organization complexity based on worker identities. This is observed in the task assignment data. I denote by $I(B)$ the organization complexity based on worker skill sets. This is unobserved. I denote by $I^*(\gamma)$ the firm's complexity predicted by the model, where market parameters Ω and the task mix α are assumed to be known and thus are incorporated into the function and not left as arguments.

First, I prove that observed organization complexity based on worker identities ($I(B)$) is equal to unobserved true complexity based on worker skill sets ($I(\tilde{B})$). Consider the augmented model proposed in Section 6.1. In particular, recall that workers with different labor supplies match to firms by some unspecified matching process. I then prove the following:

Lemma 2 *All workers with the same skill set are assigned the same distribution of tasks regardless of their labor supply.*

Proof. A well-known property of mutual information attention costs is that they satisfy compression monotonicity or are "distraction-free " (Tian 2019). I will use this in the proof.

Suppose for the sake of contradiction the firm assigned two workers of the same skill set different distributions of tasks. Consider a different assignment of work such that the same amount of each task is accomplished, and both workers still are assigned the same total amount of work. Such an assignment always exists: I can just take the total time spent on each task by both workers and split it based on effective units of labor. By the strict distraction-free property of mutual information, this new assignment reduces organization costs. This does not impact the wage bill, since both workers have the same wage. Also, it does not impact quality, because the total amount of each task accomplished remains the same, and both workers have the same skill set. Thus quality-adjusted cost strictly decreases, so profit strictly decreases, contradicting the optimality of the original assignment. Therefore, all workers with the same skill set are assigned the same distribution of tasks regardless of their effective units of labor.

This lemma means that the firm treats workers with different labor supplies but the same skill sets as if they were a single, aggregate worker. Denote worker identities as indexed by n, and worker skill sets by i. Denote the organizational structure over worker identities as $\tilde{B}$. Then

$$\frac{\tilde{B}_{n,k}}{\sum_{k'} \tilde{B}_{n,k'}} = \frac{B_{i,k}}{\sum_{k'} B_{i,k'}} \forall i, k \text{ s.t. } \theta_n = \theta_i.$$

Because the total amount of each task is fixed at α_k,

$$\sum_{n'} \tilde{B}_{n',k} = \alpha_k = \sum_{i'} B_{i',k}.$$

Plugging these results into organization complexity yields

$$I(\tilde{B}) = \sum_{n,k} \tilde{B}_{n,k} log\left(\frac{\tilde{B}_{n,k}}{\sum_{k'} \tilde{B}_{n,k'} \sum_{n'} \tilde{B}_{n',k}}\right) = \sum_{n,k} \sum_{i} \frac{B_{i,k} \sum_{k'} \tilde{B}_{n,k'}}{\sum_{k'} B_{i,k'}} \mathbb{I}\{\theta_n = \theta_i\} log\left(\frac{B_{i,k}}{\sum_{k'} B_{i,k'} \sum_{i'} B_{i',k}}\right).$$

And rearranging terms yields

$$= \sum_{i,k} \frac{B_{i,k}}{\sum_{k'} B_{i,k'}} log\left(\frac{B_{i,k}}{\sum_{k'} B_{i,k'} \sum_{i'} B_{i',k}}\right) \sum_{n,k'} \tilde{B}_{n,k'} \mathbb{I}\{\theta_n = \theta_i\}.$$

The sum of all $\tilde{B}_{n,k}$ of workers with the same skill set but different labor supply is E_i, which is exactly equal to $\sum_{k'} B_{i,k'}$. Therefore, I can write

$$I(\tilde{B}) = \sum_{i,k} \frac{B_{i,k}}{\sum_{k'} B_{i,k'}} log\left(\frac{B_{i,k}}{\sum_{k'} B_{i,k'} \sum_{i'} B_{i',k}}\right) \sum_{k'} B_{i,k'} = \sum_{i,k} B_{i,k} log\left(\frac{B_{i,k}}{\sum_{k'} B_{i,k'} \sum_{i'} B_{i',k}}\right) = I(B).$$

Therefore, organization complexity based on worker identities is equal to organization complexity based on worker skill sets. Since I observe identities, this implies that I can compute organization complexity as the mutual information between worker identities and tasks.[33]

I next show that γ is identified. This requires that there be a unique γ such that $I^*(\gamma) = I(\tilde{B})$. Define $Q_j := W(B_j) - \rho^{-1}\xi(B_j)$. Applying Theorem 1, I can write the firm's problem in the following way:

$$V := \min_{B \in \mathbb{B}} \gamma I(B) + W(B) - \rho^{-1}\xi(B) = \min_{Q \in \mathbb{Q}} \gamma \tilde{I}(Q) + Q,$$

where $\tilde{I}$ is a continuous, decreasing and convex function. Further, it is strictly decreasing whenever $\tilde{I}(Q) > 0$ (Chen, n.d.). Consider only the case when $\tilde{I}(Q) > 0$. Then the FOC $\frac{dV}{dQ} = \gamma \frac{d\tilde{I}(Q)}{dQ} + 1 = 0$ and convexity imply the optimally chosen Q is strictly increasing in γ. This implies $\tilde{I}(B)$ is strictly decreasing in γ. Since $\tilde{I}(B) = I^*(\gamma)$ for optimal B, $I^*(\gamma)$ is strictly decreasing and identification is achieved whenever $I(\tilde{B}) > 0$.[34]

Theorem 1 established that the firm's problem is a rate-distortion problem. As a result, Blahut (1972) provides an algorithm that can be used to arbitrarily approximate $I^*(\gamma)$. Thus, because I^* is strictly decreasing, I can use this algorithm to invert complexity to retrieve γ as a known function of complexity and all other parameters.

To identify organization structures (B_j), I appeal to Proposition 3. Since wages are parameters during estimation, the proposition can be applied exactly, and I have that all organization structures are identified except over a set of market parameters with measure 0. Further, the algorithm given in Blahut (1972) constructs optimal B_j for each firm in the process of computing I^*. In the knife-edge cases where more than one structure is optimal for a firm, the algorithm will return one of them. Thus, organization structures are also a known function of the data and market parameters, except for a set of market parameters with measure 0.

A.7 Welfare

Preferences take a random utility form with Type 1 extreme value distribution for the horizontal taste heterogeneity $\epsilon_{i,j}$ in the population. I assumed throughout that this heterogeneity is distributed i.i.d. across consumers and alternatives. Therefore, expected

33. One can also appeal to the data-processing inequality (which holds with equality) to avoid much of this algebra.

34. Whenever complexity is 0 (it cannot be negative), any sufficiently large γ is consistent with the data.

utility of consumer i has the well-known closed form

$$V_i = \mathbb{E}[\max_j\{\xi_j - \rho p_j + \epsilon_{i,j}\}] = ln\bigg[\sum_{j=1}^{J} exp\bigg(\xi_j - \rho p_j\bigg)\bigg] + C,$$

where C is Euler's constant. There is a mass M of consumers; therefore, total consumer expected utility is $M \cdot V_i$.[35] I can then denominate this in dollar terms by dividing by the coefficient on price, ρ. My measure of total consumer welfare in dollar terms is

$$CS = \frac{M}{\rho}\bigg\{ln\bigg[\sum_{j=1}^{J} exp\bigg(\xi_j - \rho p_j\bigg)\bigg] + C\bigg\}.$$

With a sales tax τ, it is

$$CS = \frac{M}{\rho}\bigg\{ln\bigg[\sum_{j=1}^{J} exp\bigg(\xi_j - \rho(1+\tau) p_j\bigg)\bigg] + C\bigg\}.$$

Total welfare is measured as the sum of consumer surplus, firm profits and worker wages. This assumes an additive welfare function which weights all consumers, firms and workers equally.

A.8 Organization Complexity as Task Specialization

This section illustrates that complexity is a measure of average task specialization. To see this, first define a job as a vector, where component k is the fraction of a worker's total labor spent performing task k:

$$b_i(k) = \frac{B(i,k)}{E_i}.$$

I can measure the specialization of any job by comparing it to a benchmark "generalist job." I define the generalist job as the job where all workers are assigned exactly the task mix:

$$b_j^G(k) = \alpha_k.$$

Notice that when the firm gives all workers the generalist job, each worker is working as a miniature version of the firm itself. There is no sense in which a worker needs a coworker in order to produce output. With these two concepts in hand, I obtain the following result.

Proposition 6 *Complexity ($I(B_j)$) is the weighted-average Kullback-Leibler divergence between*

35. This assumes an additive welfare function which gives equal weight to all consumers.

the jobs at a firm and the firm's generalist job $b_j^G(k)$, where the weights are the share of each worker type.

Proof. Using the definition of mutual information, I can write complexity as

$$\begin{aligned}
I(B_j) &= \sum_{i,k} B(i,k) log\left(\frac{B(i,k)}{\sum_{k'} B(i,k') \sum_{i'} B(i',k)}\right) \\
&= \sum_{i,k} E_i \frac{B(i,k)}{E_i} log\left(\frac{B(i,k)}{E_i \alpha_k}\right) \\
&= \sum_{i} E_i \sum_{k} b_i(k) log\left(\frac{b_i(k)}{\alpha_k}\right) \\
&= \sum_{i} E_i \sum_{k} b_i(k) log\left(\frac{b_i(k)}{b_j^G(k)}\right) \\
&= \sum_{i} E_i D_{KL}(b_i || b_j^G).
\end{aligned}$$

This can also be proved more quickly using the well-known fact that mutual information is the expected Kullback-Leibler divergence of the conditionals with respect to a marginal distribution. An implication of this result is that assuming mutual information organization costs is then isomorphic to assuming the cost of an organization structure is proportional to its distance from the generalist structure.

A.9 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

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

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

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

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

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

which can be solved for the optimal price:

$$\frac{1}{\rho} + c_j + \rho^{-1} W\left(exp\left(-1 + \xi_j - \rho c_j\right) E_j^{-1}\right) = p_j^*. \tag{16}$$

A.10 A Microfoundation for the Organization Complexity Measure

This section provides a full microfoundation for thinking of mutual information as complexity when there are many tasks to be assigned.

Suppose the process of producing each unit of the final product consists of T steps. Each step is a type of task $k = 1, ..., K$ that is randomly drawn i.i.d. from the task mix α. Denote the random vector of task types X^T. A firm must choose a business plan, which consists of a task categorization system, f_T, and an employee handbook, g_T, which assigns task categories to workers. Here f_T is a vector-valued function which maps every possible combination of T tasks to an index, where without loss, the index is of the form $\{1, 2, ..., 2^{T \cdot I}\}$. Moreover, g_T is a function which maps this index back into a vector of length T, where element t specifies which worker type ($i = 1, ..., N$) does task t. The firm pays the wage cost as before. Average quality is similar to before, but with more notation:

$$\sum_{t=1}^{T} Pr(X^T = x^T) T^{-1} \sum_{i} \theta_{x_i^T, g_T(f_t(x^T))}.$$

The complexity of the business plan is the number of contingencies that need to be written in the employee handbook per step:

$$log_2(2^{T \cdot I})/T = I.$$

Organization cost is then complexity times the firm-specific organization cost coefficient γ_j. Equivalence to a rate-distortion problem allows me to say that as $T \to \infty$, optimal B_j from the original problem approximates the organization complexity and quality-adjusted wages which maximize profit in this more general problem. Put another way, each optimal B_j approximates the optimal business plan. This is quite interesting because the business plan is an even more complicated object than B_j, as it provides an assignment of tasks to workers for every possible realized set of tasks. Additionally, one can think of $I(B_j)$ as measuring the length of the business plan needed to implement a certain organization structure. This is why I call $I(B_j)$ complexity.

A.11 Other Organization Costs

Theorem 1 does not rely on organization costs taking the mutual information form. However, identification of the structural model relies heavily on organization costs taking this form: it allows me to equate the observed complexity over worker identities to the true complexity over types. This is one of the main reasons why the mutual information functional form is used for organization costs in this paper.

However, imposing mutual information costs ex ante imposes behavioral assumptions on firms. In particular, it makes assumptions about how firms trade off complexity with other concerns, and it imposes symmetry conditions on worker types. Working out the implications of these assumptions for substitution patterns is a matter of consulting the rapidly growing literature on information costs and mapping these results to the labor context. This is a non-trivial task that is beyond the scope of this paper but an area for future work.

Empirically, it would be interesting to identify the correct organization cost function. As Pomatto, Strack, and Tamuz (2022) note, the mutual information cost and the log-likelihood ratio cost imply quite different behavior. However, as Lipnowski and Ravid (2022) suggest, it will be necessary to observe more information about firm choice probabilities to distinguish between different costs. In my setting, this amounts to better information about worker skills (education, demographics, prior experience, etc.). Ideally, such a project would use matched employer-employee data with detailed demographic information alongside task information.

A.12 Extensions

Many of the modeling assumptions are made solely to achieve tractability or to match the hair-salon application. The core idea behind the model is general, and this section outlines several extensions which accommodate other contexts and additional economic forces. I also provide ways to use recent work in labor economics, rational inattention, and rate-distortion theory to implement these extensions in future work.

A.12.1 Labor Market Power

The model presented in this paper focuses on situations where firms have product market power but not labor market power. These assumptions are realistic when the product market is small relative to the labor market, either geographically or because of workers' ability to work in multiple industries. In many situations, such assumptions are not realistic, and one might expect firms to hold labor market power as well.

Introducing labor market power raises an interesting theoretical question which could make it the most important area for future work. Firms with labor market power have an incentive to reduce the number of workers they hire in order to mark down wages. How does this incentive interact with internal organization, and how does it change competition? Unlike firms in a competitive labor market, firms with labor market power will realize that demanding more of a certain type of worker increases wages.

Such an extension has the potential to help us understand two features of modern labor markets. First, we can measure the amenity value of task specialization to workers. In some industries, workers may find a specialized job unfulfilling or limiting, restricting their long-term career goals by pigeonholing them. However, in others, workers may find specialized jobs valuable because they deepen expertise. Second, in highly concentrated industries, we can study how internal organization choices are driven by a desire to make workers scarce for competitors. Anecdotal accounts in the technology sector suggest such talent wars occur. My model provides a way to study the trade-off between over-hiring a certain type of worker and trying to operate the firm.

A model with labor market power could be made tractable by assuming monopolistic competition in the product market and monopsonistic competition in the labor market. The labor market could be modeled using the framework introduced in Card et al. (2018). The novel internal organization cost introduced in this paper extends to such a model. However, because output would impact marginal costs (through wage markdowns), the characterization in Theorem 1 will no longer hold. New tools to solve and estimate the model would be needed. Because the new problem will be a non-linear rational inattention problem, results from Jung et al. (2019) may be helpful in this regard.

A.12.2 Large Firms

The model can be extended to the case where firms are "large," with a continuum of tasks and worker types.

Consider a firm which must complete a continuum of tasks to produce the final good. The task mix is now a distribution, which I assume to be normal: $k \sim N(0, \sigma^2)$. Suppose workers have a single specialty task, and that they are indexed by i in the order of their specialty task. An organization structure B is now a continuous bivariate joint distribution.

Suppose the quality of a performed task is given by the squared distance between the specialty of the worker and the task assigned, that is, $\xi = -\int (i-k)^2 dB(i,k)$, and denote $D = -\xi$. For simplicity, assume all workers have the same wages (skills are not priced by the market) and $\rho = 1$.

It can be shown that the organizational frontier in this special case has a closed form, and an organization structure B which maximizes profit is

$$\begin{pmatrix} i \\ k \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2 - ln(4)\gamma & \sigma^2 \\ \sigma^2 & \sigma^2 \end{bmatrix} \right).$$

To interpret this result, note that as γ approaches 0, the correlation between tasks and workers approaches 1 and the marginal distribution of hired worker types widens and approaches the distribution of tasks. In other words, the firm assigns each task entirely to the appropriate specialist. Whenever there are positive organization costs, the task content of a worker of type i is a normal distribution centered on that worker's specialty with variance $\sigma^2 - ln(4)\gamma$. Greater organization costs reduce task specialization. This illustrates two things. First, it is easy to extend the model to accommodate the large-firm case, where the task space is uncountable (as is the worker-type space). Second, the key role of organizational frictions is a deep property of the model that does not go away when the researcher makes organizations large and less "lumpy."

A.12.3 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:

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

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 16 from Appendix Section A.9 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.12.4 Non-Additive Quality

The model developed in this paper required the effect of the quality of each individual task to have an additive impact on overall quality. This assumption is natural in some settings, but unnatural in others. An excellent example is the launching of space shuttles. A single task performed poorly can be catastrophic, as illustrated by the Challenger explosion. In these contexts, nonlinear quality aggregation is necessary to model the production process. I can accommodate this within the model using multiplicative quality, similar in spirit to Kremer (1993):

$$\xi_j = \prod_{i,k} \theta_{i,k}^{B_{j,i,k}}.$$

Rewriting it using logarithms yields

$$\xi_j = exp\bigg(\sum_{i,k} B_{j,i,k} log(\theta_{i,k}) \bigg).$$

This is now an f-separable distortion measure, meaning I can apply recent work in information theory (Shkel and Verdú 2018) to adapt the Blahut–Arimoto algorithm and other tools to work with this extended model.

A.12.5 Quality Positioning and Richer Demand Systems

One surprising result from the theoretical section is that the choice of organization structure depends only on other firms' choices via wages. This derives from the quality-price index assumption placed on demand systems in the model. In some contexts this may be unrealistic, and one may believe that there is a "positioning effect," where the return to higher quality depends in part on how many other firms are also producing high quality. This section illustrates that this effect can be incorporated using mixed logit demand systems if a researcher is willing to sacrifice analytical and computational tractability.

Suppose consumers differ in their taste for quality. The utility of consumer z for product j is now given by

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

where q_z is distributed i.i.d. across consumers according to some distribution G. This utility specification now nests both pure vertical and pure horizontal differentiation models. When I specify that $\epsilon_{z,j}$ is a Type 1 extreme value and q_z is normally distributed, the result is a random coefficients logit model. Market share for product i among consumer segment z is given by

$$s_{j,z} = \frac{exp(q_z \xi_j - \rho p_j)}{\sum_{j'} exp(q_z \xi_{j'} - p_{j'})}.$$

To understand the effect of quality on market share, I can compute the derivative:

$$\frac{\partial s_j}{\partial \xi_j} = \int q_z s_{j,z}(1 - s_{j,z}) dG(z).$$

Two facts are apparent from this expression. First, the marginal revenue from increasing quality now depends on the quality position of other firms. Firms will find it more beneficial to raise quality when high-quality segments are relatively untapped by other firms. Second, the optimal organizational structure B_j will depend on the equilibrium quality choices of other firms.

The cost of this more-flexible demand system is tractability. Because of the dependence on the quality choices of other firms, the characterization in Theorem 1 no longer holds. In particular, the firm's problem is not a rate-distortion problem. Estimation requires solving the model for each firm using nonlinear convex optimization. Additionally, demand no longer takes a multinomial logit form, so there does not exist a closed-form solution relating market shares, prices and unobserved qualities. Estimation now also requires numerical integration and a BLP-style contraction mapping to invert market shares.

A.12.6 Richer Attention Elasticities

As noted by Csaba (2021), the mutual information cost restriction "attention elasticities" should be constant. To translate this to the context in this paper, consider a firm which is deciding how to split some amount of task k between two workers. Constant attention elasticities means that regardless of the task, and regardless of the initial skills and wages of the two workers, a 1-percent increase in the relative quality-adjusted cost of one of the workers relative to the other increases the probability task k is performed by that worker relative to the other by a constant percentage. Thus the mutual information cost function is building in symmetry and constant percentage changes when counterfactuals change the quality-adjusted costs of different worker types in equilibrium.

Csaba (2021) show that we can use α-mutual information cost functions to allow for more varied elasticities. Exploring such cost functions may be interesting, because they allow for a form of organizational inertia, where how elastic an organization structure is depends on the initial structure.

A.13 Knowledge Hierarchies

Rosen (1982) and Garicano (2000) envision firms as characterized by knowledge hierarchies. This idea is incorporated into a quantitative equilibrium model in Caliendo et al. (2012). In this conception, workers differ in their knowledge, tasks differ in their complexity or frequency, demand to each firm is exogenous, and firms choose the number of levels of their organization. There are also communication costs so that sending a problem up to a higher level of the organization has a cost.

As noted by Haanwinckel (2020), there are similarities between task-based models and knowledge hierarchy models. In particular, both models generally result in full specialization, where workers perform non-overlapping sets of tasks. Additionally, tractability in both settings is often maintained by ordering tasks and workers along a common dimension. Finally, estimation and worker types are often inferred from some combination of demographic information and wages. For example, following Caliendo, Monte, and Rossi-Hansberg (2015) it is common practice in knowledge hierarchy papers to group workers into management layers based on occupation and wages, where occupations with similar wages form a layer. In this sense, the difference between my model and knowledge hierarchy models is similar to the differences between my model and task-based models.

In another sense, knowledge hierarchy models seek to explain the hierarchical structure of firms, while my model tries to explain the task assignments within a firm. Because the models have different goals, they are designed differently. One way of bridging the gap is by introducing management worker types and introducing a management task into my framework. A key ingredient is that the management task generally must impact the other tasks multiplicatively. This can be accommodated, but requires new derivations. Refer to Section A.12.5, where I discuss multiplicative quality.

B Empirical Appendix

B.1 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 predefined 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 B1.

	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 Fa		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 B1: 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.

The final categories in their original spreadsheet are provided in Figure B2. The cosmetologist's reasons for grouping subcategories together are provided underneath the appropriate column.

	A	B	C	D
1	Task Category	Task Group (6 Groups)	Task Group (5 Groups)	Task Group (4 Groups)
2	Cut	4	1	1
3	Shave	4	1	1
4	Style	3	1	1
5	Extensions	2	2	1
6	Color/Bleach	1	1	1
7	Highlight/Balayage	1	1	1
8	Other Treatment	3	3	2
9	Blowdry	3	1	1
10	Wash/Shampoo	1	1	1
11	Eyebrow/Eye Lash Service	6	5	3
12	Admin/Consult	5	4	4
13	Spa Service	6	5	3
14	Nail service	6	5	3
15				
16				
17				
18				
19	Notes:	1 - Hair colouring services to the hair performed by a hairstylist. 2- Hair service that could be completed by a full cosmotologist or someone that is trained only in this service. 3. Hair service that could be performed by licenced cosmotologist but also an assistant / trainer or a service specific individual such as an aesthetics practictioner. 4. Hair cutting service typically performed by a hairstylist or barber. 5.A sevice that could be perfomed by either a hairstylist OR a Non licenced manager/ front of house staff member. 6. A beauty service to the body. Typically performed by an esthetican.	1 - Hair service only done by a licenced hairstyist, assistant hairstylist or barber. 2- Hair service that could be completed by a full cosmotologist or someone that is trained only in this service 3. Service that could be performed by a licenced professional or an assistant / trainee. 4. A sevice that could be perfomed by either a hairstylist OR a Non licenced manager/ front of house staff member. 5.A beauty service to the body. Typically performed by an esthetican.	1- Under most circumstances thiss service would be performed by a Hairsylist, as they are services to the hair/ scalp. 2. This treatment/ service could be performed by varied amount of staff members from esthetics practitioner to hairstylists. 3. These services would typically be performed by an esthetican as they are services to the face / body 4. This service could be performed by licenced professionals or non licenced team members such as front of house staff.

Figure B2: Final Task Subcategorization Spreadsheet from Cosmetologist

B.2 An Alternative Task-Category-Generation Process

The analysis in this paper relies on professionally categorized descriptions. This section describes an alternative natural language processing approach (NLP) used to generate task categories. This process has the benefit of being less labor intensive and not reliant on human judgment, but it has the disadvantage of making certain types of classification errors. An earlier version of this paper used the NLP method. Many of the stylized facts continue to hold using this alternative method of classification.

In order to minimize issues related to spelling and grammar, I cleaned the descriptions using standard natural language pre-processing techniques. This reduces issues related to spelling errors. Some descriptions contain only punctuation, symbols or "stop words" and as a result were removed from the analysis. To group descriptions into tasks, I used a hierarchical clustering method where the only choices are the number of categories and the measure of distance. I set the number of categories to be four and the distance measure to be the Euclidean distance.

B.3 Robustness of Stylized Facts

The first concern is one of reverse causality. Perhaps firm size allows firms to be organizationally complex and thus have a product market advantage.[36] Appendix Section B.6 shows that while this cannot be ruled out, it is not generating all of the observed relationships. Even among firm-quarters with the same number of employees, there is significant variation in complexity, and there is a positive association between complexity and the main market outcomes (i.e., revenue, prices and repeat customers).

A second concern is that the correlations are driven by demand-side factors, such as consumer preferences for particular stylists rather than firm choices. The software records when customers request a particular staff member. It would be concerning if there was a strong positive correlation between the request rate at a salon and complexity. Appendix Section B.7 shows that while many customers request specific staff, the rate of requests across salons is not correlated with organization complexity. Further, the correlation between the request rate and firm size is either zero or negative.

A third concern is that the correlations are driven by the specific functional form chosen for complexity. Appendix Section B.8 shows that the main patterns persist when complexity is replaced by within-visit specialization. Within-visit specialization is measured as the fraction of multi-service visits which are performed by a team (i.e., more than one employee).

B.4 Measurement Error in Organization Complexity

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

36. Li and Tian (2013) provide a theoretical mechanism for such an effect.

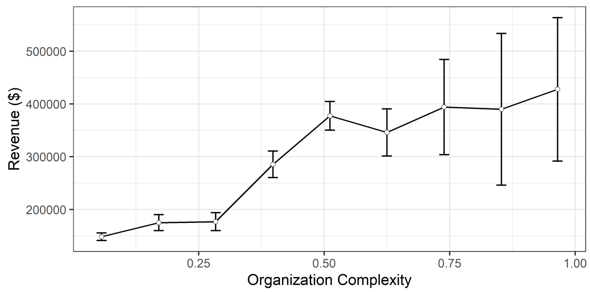
correlation between complexity measures at the month level within quarter should be large. This section illustrates that this is indeed the case.

To do this, I recompute complexity for each month within a quarter so that I have three measurements of complexity per firm-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.

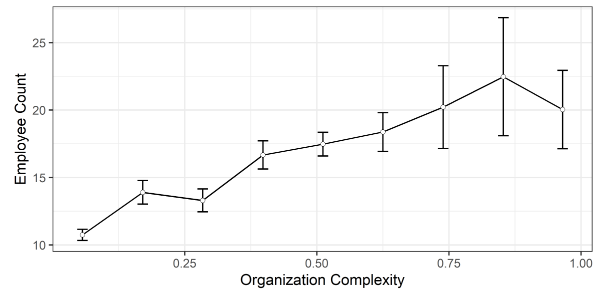
The assumption that complexity is measured without error allows me to "invert" complexity to obtain the underlying organization cost for each firm in a market. In a similar way, researchers in industrial organization often assume market shares are measured without error in order to invert them to obtain mean utilities for each firm. It is possible to relax this assumption and use the panel nature of the data to estimate each firm's organization cost parameter.

B.5 Firm Size and Complexity Associations

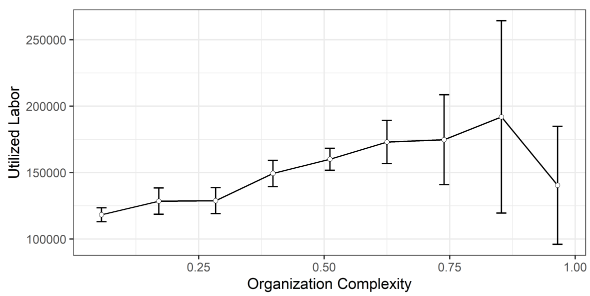
Figure B3: Organization Complexity and Firm Size



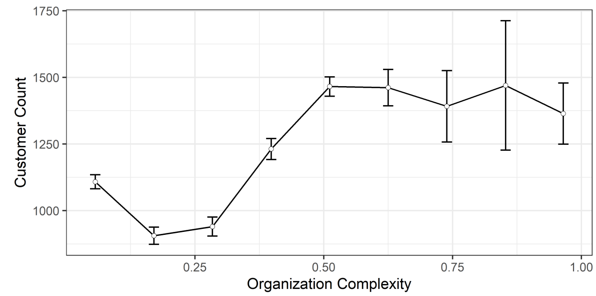
(a) Revenue ($)



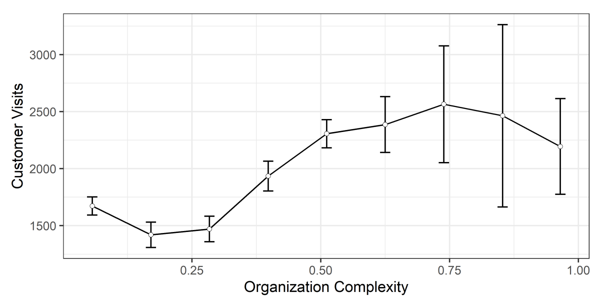
(b) Employees



(c) Utilized Labor (minutes)



(d) Unique Customers



(e) Customer Visits

Table B1: Regressions of Firm Size on Complexity, Manhattan Only

Dependent Variables:	Revenue	Employees	Utilized Labor	Customers	Visits
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Org. Complexity	430406.6*	12.55	-17733.9	277.2	876.9
	(179977.4)	(6.531)	(70765.2)	(600)	(907.1)
Fixed-effects					
Quarter-Year	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	595	595	595	595	595
R^2	0.33485	0.21039	0.20359	0.44164	0.48831

Clustered standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Note: This table repeats the regressions of revenue and other measures of firm size on complexity, but only for New York County (Manhattan). The positive relationship between revenue and complexity remains statistically significant.

Table B2: Regressions of Revenue on Complexity and Employee Count Interacted

Dependent Variable:	Revenue		
Model:	(1)	(2)	(3)
Variables			
(Intercept)	79487.9***		
	(19103.3)		
Complexity	-226181.8*	-242961.5*	-320973.6**
	(111684)	(110939.4)	(117545.2)
Employee Count	5652.8*	4871.6*	3878.9
	(2315.3)	(2257)	(2192.2)
Complexity × Employee Count	29487.9***	30187.8***	35052.8***
	(8587.8)	(8507.4)	(8528.5)
Fixed-effects			
Quarter-Year		Yes	Yes
County			Yes
Fit statistics			
Observations	4,558	4,558	4,558
R^2	0.4913	0.52042	0.61654

Clustered standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Note: This table presents regressions of revenue on complexity interacted with employee count. The mean number of employees is 13.38, so the marginal effects in all specifications evaluated at the mean are positive.

B.6 Complexity Relationships Among Similar-Size Firms

The main text of the paper established that complexity is correlated with the number of employees as well as other outcomes. This raises concerns about the direction of causality: are firms larger because they are more internally complex, or are larger firms naturally able to design more internally complex structures? The model in this paper specifies a common organization cost, which generates jointly both larger and more complex firms. In this sense, complexity does not cause a firm to be larger; rather a common, unobserved productivity heterogeneity generates both.

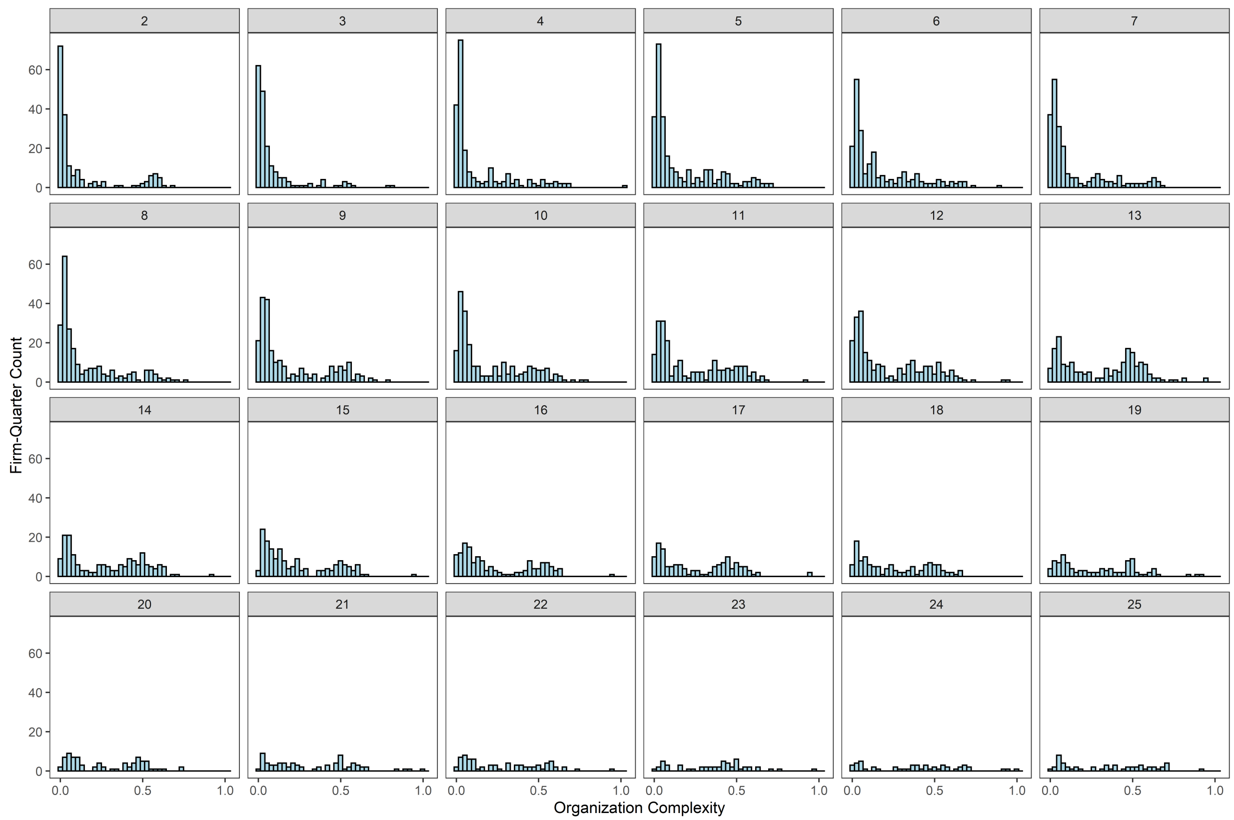
This answer still leaves several questions open. In particular, perhaps organization costs are more like fixed costs, so larger firms are better able to afford more complex organizations. Additionally, maybe larger firms have more organizational possibilities, and thus the relationships discussed are mechanical. I alleviate this concern by analyzing many of the outcomes among firms with the same number of employees. I begin by illustrating that there is significant heterogeneity in complexity across firm-quarters with the same number of employees by plotting histograms in Figure B4.

The positive correlation between complexity and revenue, prices and repeat customers persists among firm-quarters with the same number of employees. This can be seen in Figure B5, which shows scatter plots with linear best fit lines for firm-quarters with the same number of employees. There is a positive correlation within almost all firm sizes and for almost all variables. The exception is repeat customers among firms with 2–5 employees. In general, the positive correlation is larger in magnitude for firm-quarters with 13 or more employees.

Essentially, while complexity is correlated with both employee count and other market outcomes, and employee count is correlated with the other market outcomes, there seems to be a large, direct effect of complexity on market outcomes. Another way to see this is that when firm-size fixed effects are added to a regression of revenue on complexity, the point estimate for complexity decreases by around 60 percent, but remains economically and statistically significant. So while much of the effect of complexity on other outcomes seems to come through size, a sizable amount does not.

I do not interpret these correlations as causal. Rather, I take them as evidence that there is an organizational advantage that operates through channels beyond just firm size. For this reason, the model is built such that a common firm characteristic (γ_j) generates both larger firms and more complex organizations.

Figure B4: Organization Complexity for Similarly Sized Firms



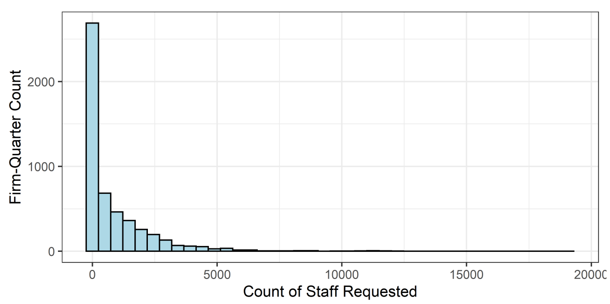
Note: Each plot is a histogram of complexity among firm-quarters with the shown number of employees. I perform the analysis for salons with fewer than 25 employees, as these represent the bulk of the data. Salons with 1 employee are excluded because mechanically complexity is 0 for these salons.

B.7 Consumers Requesting Particular Staff

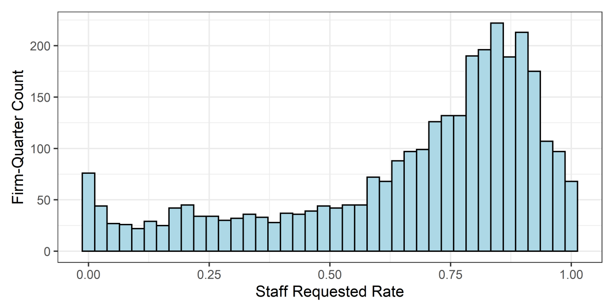
The stylized facts and the model treat the assignment of workers to tasks as a choice of the firm. In practice, some customers directly request particular stylists. The software allows salons to record when a staff member is requested for a task, and this information is captured in a variable titled ”Was Staff Requested.” This section establishes that although there is heterogeneity in how often staff are requested at different salons, this heterogeneity is not correlated with organization complexity.

I start by examining the variation in requests across salon-quarters in Figure B6. A large number of salon-quarters have no requests observed in a quarter (Panel A). Among those salon-quarters with at least one request, the request rate varies significantly, spanning close to 0 all the way to 1 with a mode around 0.8 (Panel B). Much of this heterogeneity comes form an aggregate increase in the request rate over time (Panel C). Therefore, I also run analyses excluding quarters before the first observed request for a salon. I call this sample “after adoption.”[37]

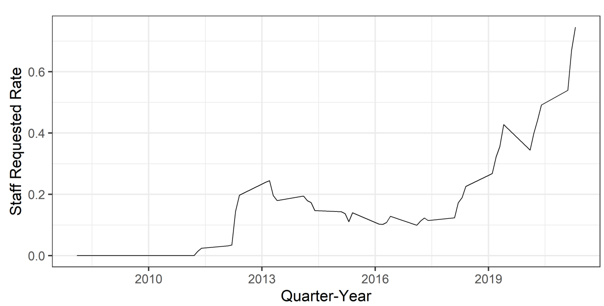
Figure B6: Was Staff Requested?



(a) Request Count Histogram



(b) Rate (After Adoption)



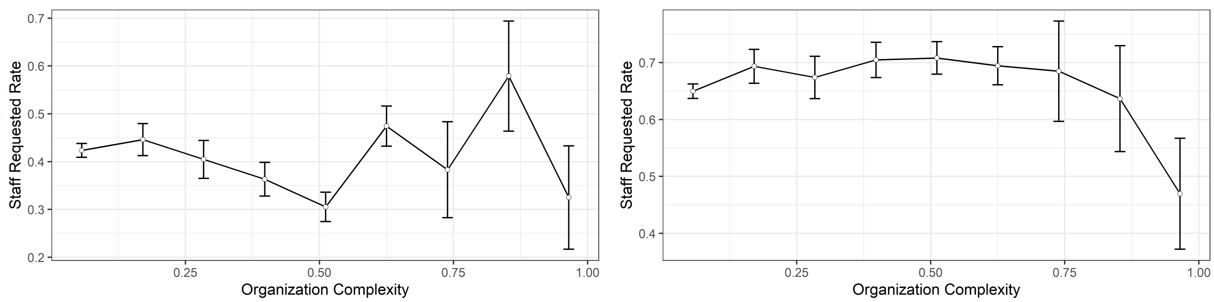
(c) Aggregate Rate Across Time

The primary question is whether consumer requests are driving observed organiza-

37. The idea is that salons differ in when they start using this feature, and I want to analyze variation only among salons that have chosen to use it.

tion complexity. I test this using binned scatter plots in Figure B7. Both unconditionally (Panel A) and among salon-quarters with one request (Panel B), complexity does not appear to have a systematic relationship with the request rate.

Figure B7: Request Rate and Organization Complexity

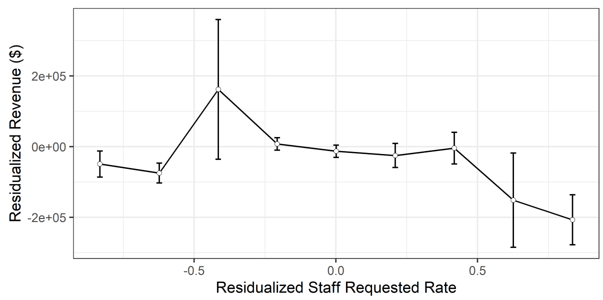


(a) All Salon-Quarters **(b)** After Adoption

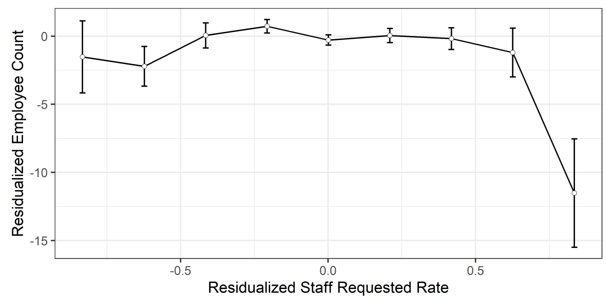
Regressions with standard errors clustered at the salon level also reveal mixed results. In the full sample, the coefficient on the request rate is statistically insignificant and negative. In the after-adoption sample, the coefficient is statistically insignificant and positive. In both cases, the coefficients are economically insignificant: they imply that a 1-standard-deviation increase in the request rate is associated with less than a 0.08-standard-deviation change in complexity.

Further, Figure B8 shows there is no evidence of a positive relationship between firm size and the request rate (if anything, there may be a negative relationship), which suggests the positive relationship between complexity and firm size documented in the stylized facts is not driven by customer request.

Figure B8: Request Rate and Firm Size

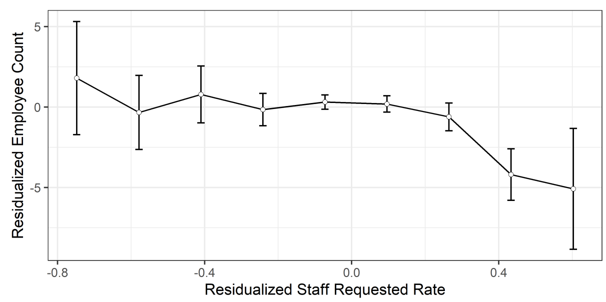


(a) All, Revenue



(b) All, Employees

(c) After Adoption, Revenue



(d) After Adoption, Employees

In summary, the data suggest that customers often request specific stylists (many salons have request rates around 0.8) but that this varies significantly across salon-quarters. This is in line with the intuition that requests are common. However, the correlation between the request rate and complexity at salons is statistically and economically weak, evidence that while consumers do request staff, these requests are not first-order determinants of the complexity differences across salons.

At a theoretical level, there is a question of whether a strong positive correlation would matter. Whether a consumer requests a stylist or a firm assigns a stylist, the firm still must bare the associated organization cost. Further, if consumers prefer a particular stylist for a task, this likely reflects the stylist's quality at that task and a match effect. Since quality differences across tasks are already built into the model, if the match effect is small, this phenomenon is captured by the existing model.

B.8 Within-Visit Specialization

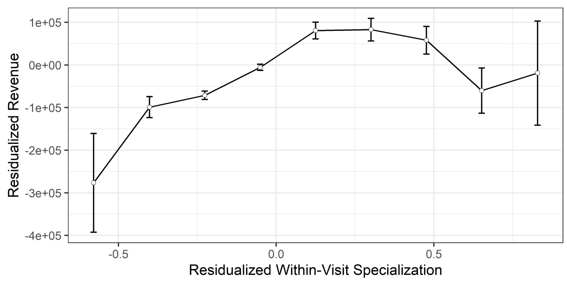
Section 3 documents robust, positive correlations between product market outcomes (revenue, customers, price, customer return rates) and complexity. One concern is that these correlations may be driven by the particular functional form of complexity.

This section alleviates this concern by showing that many of the patterns persist when I use a measure of within-visit specialization. I compute within-visit specialization as the number of customer visits[38] with two or more employees assigned divided by the number of customer visits with two or more services performed. Because a visit needs to have multiple services in order for more than one employee to be assigned in the system, this number is a fraction which captures the amount of within-visit specialization. It connects directly to the examples provided in Table 3.

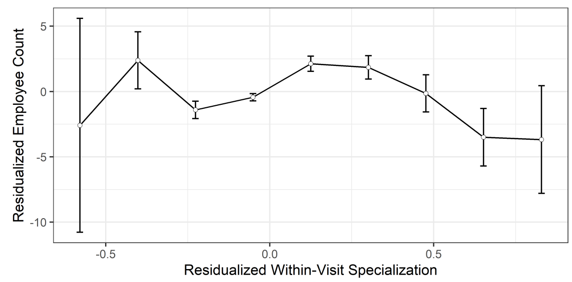
A histogram of this measure shows that it follows a similar power-law distribution as organization complexity, with observed values spanning the support and a long right tail. Like organization complexity, within-visit specialization is positively correlated with revenue, price and the share of repeat visits. However, unlike organization complexity, it has a non-monotone relationship with the number of employees.

38. Visits are the number of unique customer-date pairs in a quarter.

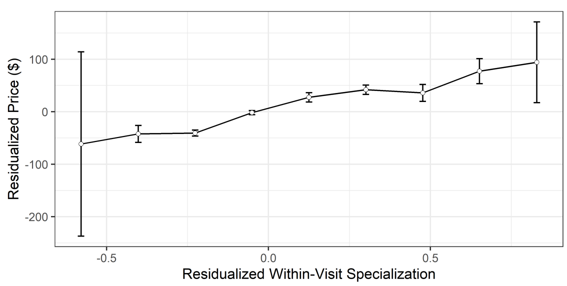
Figure B9: Within-Visit Specialization



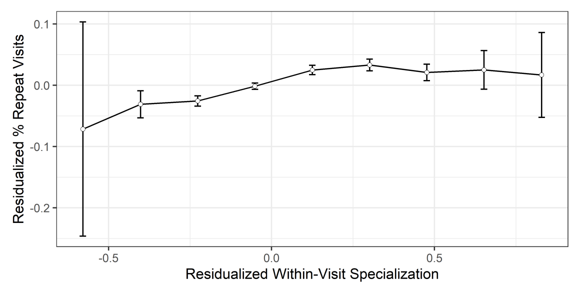
(a) Revenue



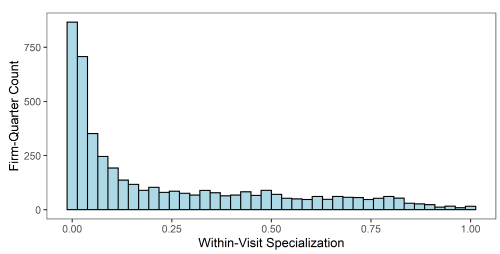
(b) Employees



(c) Price



(d) Repeat Customers

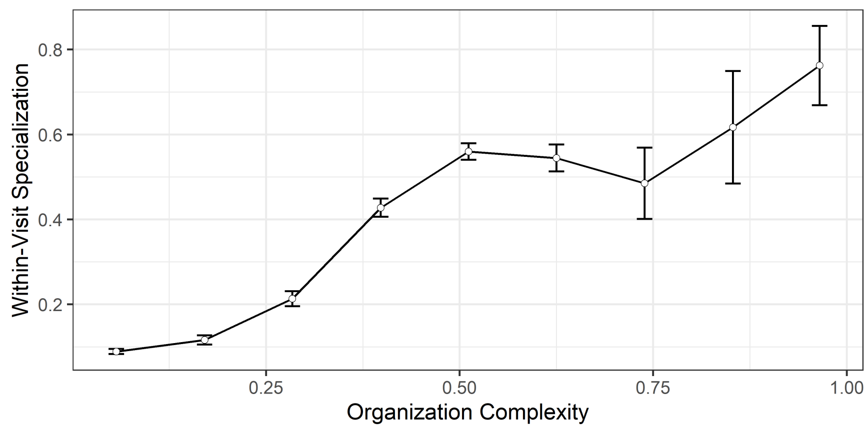


(e) Histogram

Note: Within-visit specialization is the share of visits with multiple services that are assigned to multiple employees.

These findings are further support that more internally specialized firms command a competitive advantage. To finish this section, I study the connection between complexity and within-visit specialization. A simple regression of complexity on within-visit specialization yields an R-squared of 0.50, suggesting that nearly half of the variation in complexity can be accounted for by specialization within-visit. Kendall's rank correlation coefficient between the two variables is 0.488. This can be interpreted as 25.6 percent of pairs of firm-quarters being discordant. Roughly, if two firm-quarters are drawn randomly, their ordering according to complexity and within-visit specialization will be the same 74.4 percent of the time. The strong connection between the two variables is visualized in a binned scatter plot in Figure B10.

Figure B10: Organization Complexity and Within-Visit Specialization



Note: There is a strong, positive correlation between complexity and within-visit specialization.

B.9 Task Content Variance Decomposition

Using the estimated model, I can study the determinants of the task content of hair-salon jobs in Manhattan. As a first step, I decompose the variation in task content into a worker and firm component. Using the distribution of model generated jobs, I can write

$$b_j(i,k) = \bar{b}(i,k) + (b_j(i,k) - \bar{b}(i,k)).$$

I then adapt the method used by Song et al. (2019) to my setting. Fixing k and taking the distribution to be weighted by effective units of labor, I then decompose the variance into a worker-type component and a within-worker-type component, where recall ω_i is the share of total labor represented by workers of skill set i:

$$Var_{i,j}(b_j(i,k)) = Var_i(\bar{b}(i,k)) + \sum_i \omega_i Var_j(b_j(i,k)|i).$$

However, because $b_j(i,k)$ is generated by the structural model, and I am considering a single labor and product market, the within-type component comes entirely from variation in firm attributes. Therefore, I have decomposed the total variance in task content into a worker and firm component. Dividing through by $Var_{i,j}(b_j(i,k))$ gives the share of variance due to each component, as shown in Table B9. For the main tasks (cut, color, blow-dry), between 8 and 22 percent of variation in job task content is attributable to firms.

B.10 Bootstrap Procedure

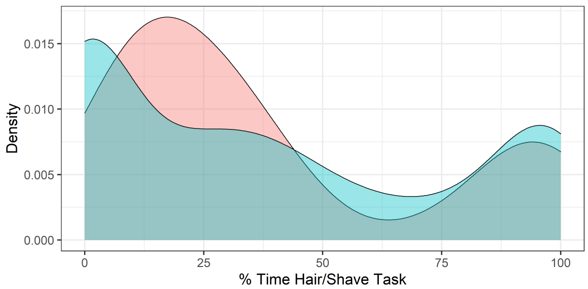
During each bootstrap replication, the model is fully re-estimated. The estimation procedure has three loops, which are run with slightly looser tolerances than the primary estimation algorithm. The innermost loop, which is the Blahut–Arimoto algorithm, is run with a convergence tolerance of 10^{-10}. The middle loop, for which I choose $gamma_j$ to match each firm's complexity, is run with a tolerance of 10^{-8}. The outer loop, which finds the market parameters, uses the Nelder-Mead method. Relative and absolute tolerance are set to 10^{-8} with a maximum number of 4000 iterations.

Standard errors are computed as the sample standard deviation of the bootstrap distribution of each parameter. To check the stability of standard errors, I ran an additional 23 replications. The standard errors with these additional replications are within 4% of the reported standard errors.

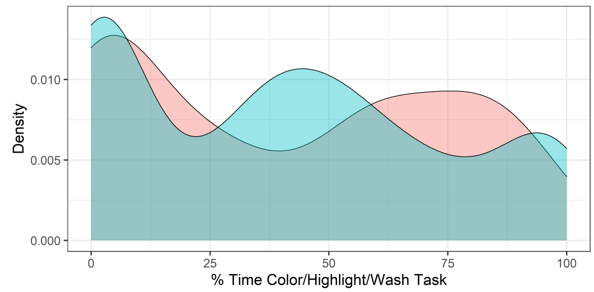
B.11 The Full Distribution of Task Content

I can also go beyond the variance and compare the entire distribution of model-generated and observed job task content. This is a strong test of the model, because I observe 509 stylists in Manhattan during 2021Q2 (the estimation period), and I am asking the model to match their jobs using only firm-based moments. Figure B11 plots the two distributions for each of the six tasks. Although the match is not perfect, the model is able to replicate important features of the data. As an example, panel B shows that the fraction of time stylists spend on the coloring task is tri-modal in the data, with peaks at 0%, 40% and 90%. The model is able to approximate this pattern with a bimodal distribution, with a wide first peak that merges the 0% and 40% observed in the data.

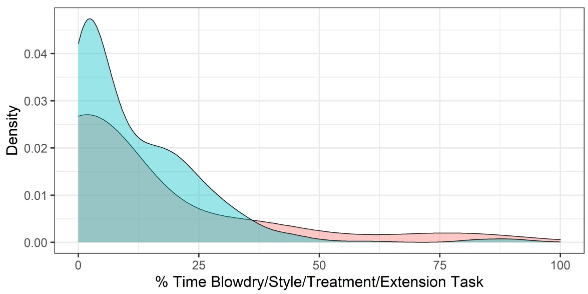
Figure B11: Model (Red) vs. Observed (Blue) Job Task Content in Manhattan



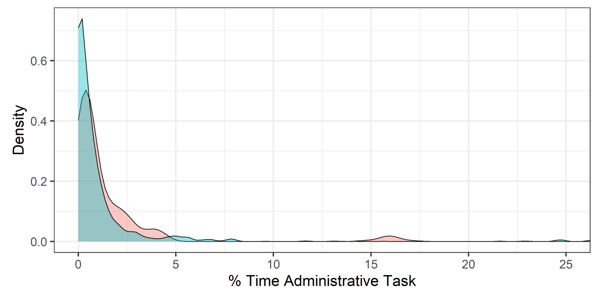
(a) Hair/Shave Task



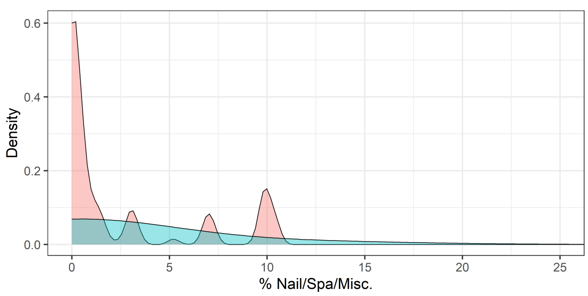
(b) Color/Highlight/Wash Task



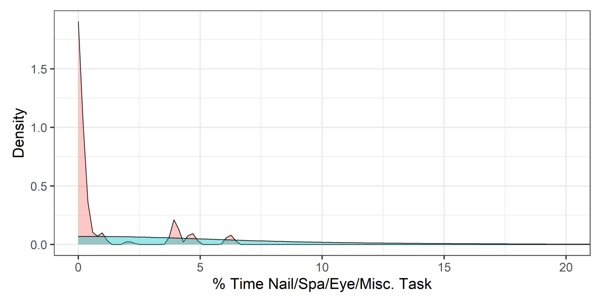
(c) Extension Task



(d) Blowdry/Style/Treatment Task



(e) Administrative Task



(f) Nail/Spa/Eye/Misc. Task

B.12 Counterfactual Procedures

B.12.1 General Procedure

The general procedure used in all counterfactuals is as follows.

1. Weight each firm such that the observed total market share matches the share of people purchasing some amount of hair salon services in the Consumer Expenditure Survey. This means that each Manhattan salon in the data is assumed to represent 23 salons.[39]

39. It is necessary to weight firms for counterfactual analysis but not estimation. This is because during estimation I fix an equilibrium, but in counterfactuals I must find a new equilibrium, and firm pricing

2. Compute the implied total labor supply of each worker skill set by summing all labor demands at initial wages over all firms.
3. Make the relevant parameter changes that correspond to the counterfactual.
4. Guess wages.
5. Solve for organization structure: If allowing internal reorganization, use the Blahut–Arimoto algorithm described in the estimation procedure to solve for each firm's organization structure. Otherwise, maintain the same organization structure for each firm.
6. Compute optimal pricing: Given that the organization structures, qualities and costs of all firms are now known, optimal pricing is computed by iterating on each firm's best response pricing function until convergence.
7. Check labor market clearing by comparing the new labor demands to the total labor supply computed in step 3. If supply and demand for each type match, exit. Otherwise, return to step 4 and guess a new wage vector.

I assume that the exogenous quality (ν_j) and exogenous marginal cost (ϕ_j) remain the same in the counterfactual analyses. To solve for market clearing wages, I minimize the sum of squared excess labor demand. I use the L-BFGS-B routine and stop only when the objective is less than 0.1. This corresponds to a very close match between labor supply and demand. I found it more efficient to use a minimization routine because the labor demands of each worker type depend in a complex manner on the entire vector of wages. That is, while each labor demand is monotone decreasing in own wage, firms can use other worker types as substitutes, so I cannot simply find each of the six market clearing wages sequentially.

Total welfare is defined as the sum of total wages, consumer welfare and total profit. Task specialization is defined as the total amount of labor spent on a worker's specialty tasks. Reported average prices, qualities and complexities are at the firm level, and are not weighted by market share. Wage statistics are weighted by the labor supply of each worker type.

B.12.2 Minimum Wage Technical Details

For the minimum wage counterfactuals, it is necessary to discuss two additional technical details: the numeraire good and multiple equilibria.

strategies depend on the number of other firms in the market.

The numeraire good in my model is the outside option, which is not buying services from a salon. I normalized its utility to be 0 in the random utility framework developed earlier. Wages are estimated using observed prices and hours, so they can be interpreted directly as nominal wages, or the wages we discuss in everyday language. When considering a counterfactual minimum wage increase to $30, I require equilibrium wages to be at or above $30, without any transformations. This is valid under the partial equilibrium assumption that the value of not buying services from a salon remains the same before and after the minimum wage increase. In general equilibrium models, counterfactual minimum wage changes must be implemented more carefully (for example, Haanwinckel (2020)).

In general, a minimum wage can result in multiple equilibria. To ensure that there are not multiple equilibria, I solved the model under every possible permutation of binding minimum wages. That is, I assumed the minimum wage binds for worker types 1 and 2 only, 1, 2 and 3 only, etc. With six worker types, this amounts to solving the model $2^6 = 64$ times. Each time I solved the model, I fixed the wages of the binding types at $30, and then solved for the wages of the other types which clear the labor market for only those other types. Afterwards, I checked three things:

1. Worker types with non-binding wages have wages greater than $30.

2. Worker types with binding minimum wages have excess labor supply.

Any solution which passed this check was considered a valid equilibrium. For example, for the case when the minimum wage is binding only for type 1, I set type 1's wage to $30 up-front, then solved for the other five wages which clear the market for the other five types. I then checked that types 1 through 6 have wages above $30 and type 1's excess labor supply is positive.

This process indicates there is a unique equilibrium in both counterfactuals. For both the full adjustment and no adjustment counterfactuals, only one of the 64 cases satisfied the checks as a valid equilibrium. One additional case in the full adjustment counterfactual never converged, meaning I could not find wages that cleared the labor market for the non-binding worker types. The wages, employment and task specialization in the initial, reallocation and full equilibrium are provided in Table B7.

B.13 Job-Level Heterogeneity

Table B3: Job Task Mix

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
Share Haircut/Shave	62,671	0.387	0.344	0.000	0.008	0.669	1.000
Share Color/Highlight/Wash/Extensions	62,586	0.371	0.322	0.000	0.025	0.599	1.000
Share Blowdry/Style/Treatment	62,564	0.102	0.162	0.000	0.008	0.124	1.000
Share Administrative	62,702	0.061	0.168	0.000	0.000	0.027	1.000
Share Nail/Spa/Eye/Misc.	63,012	0.076	0.227	0.000	0.000	0.010	1.000

Note: This table displays summary statistics about the time spent on each task at the worker level. While worker averages correspond roughly to firm averages, there is greater heterogeneity across workers, supporting the idea that within firms there are distinct roles.

Figure B12: The Job Task-Mix Distribution

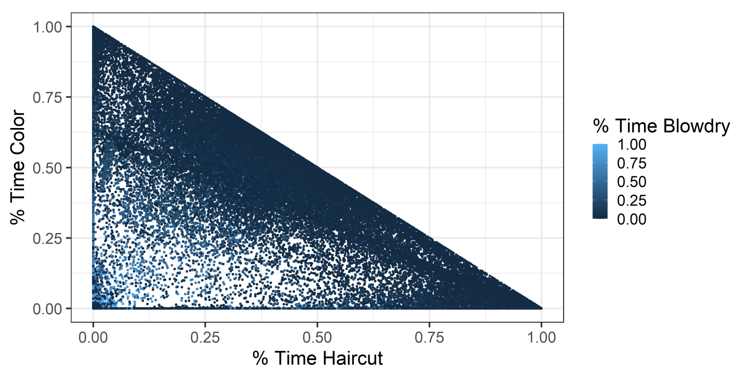


Table B4: Regressions of Revenue on Complexity

Dependent Variable:	Revenue					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
(Intercept)	121578.1***					
	(14835.8)					
Organization Complexity	456571.3***	440904.1***	485026.4***	486995.5***	271694.6**	261697**
	(100394.8)	(108427.1)	(116918.9)	(125004.8)	(87031.1)	(80920.6)
Task Mix 2				-19070.7	-7609.7	14482.9
				(93817.4)	(78597)	(67354.5)
Task Mix 3				-8011.8	116011.4	98022
				(81014.1)	(106735)	(98077.1)
Task Mix 4				-24893.1	76296.2	67131.1
				(113959)	(96547)	(95768.9)
Task Mix 5				43954.8	14593.5	33562.4
				(50238.8)	(47813)	(56691.1)
Staff Request Rate						-94370.7
						(89112.9)
Fixed-effects						
Quarter-Year		Yes	Yes	Yes	Yes	Yes
County			Yes	Yes	Yes	Yes
Firm Size					Yes	Yes
Fit statistics						
Observations	5,116	5,116	5,116	5,116	5,116	5,116
R^2	0.01475	0.01915	0.3104	0.31047	0.34273	0.34365

Clustered standard-errors in parentheses
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05*

Note: This table reports the regressions of revenue on complexity under various specifications, including controlling for the rate of staff requested.

Table B5: Two Estimated Organization Structures

Specialist	Cut	Color	Blow Dry	Admin.	Nail/Misc.	Total
Cut	0.15	0.01	0.001	0.06	0	0.22
Color	0	0	0	0	0	0
Blow Dry	0	0	0	0	0	0
Admin.	0.31	0.03	0.003	0.45	0	0.784
Nail/Misc.	0	0	0	0	0	0
Tot.	0.455	0.036	0.004	0.505	0	1

(a) Salon 1, $I_j = 0.03$

Specialist	Cut	Color	Blow Dry	Admin.	Nail/Misc.	Total
Cut	0.180	0.003	0	0.006	0.003	0.193
Color	0.057	0.553	0	0.016	0.009	0.116
Blow Dry	0.012	0.002	0.097	0.003	0.002	0.636
Admin.	0	0	0	0	0	0
Nail/Misc.	0.004	0.001	0	0.001	0.050	0.055
Tot.	0.253	0.559	0.097	0.026	0.064	1

(b) Salon 2, $I_j = 0.70$

Note: These are estimated organization structures (B_j) for a high- and a low-complexity salon in New York in Quarter 2, 2021.

Table B6: Variance Decomposition: Without a Model

Across Firms

Task	Share of Labor	Share of Variance: Firm	Within-Firm
Haircut/Shave	0.4049	0.3744	0.6256
Color/Highlight/Wash	0.3902	0.2899	0.7101
Blowdry/Style/Treatment/Extension	0.0850	0.5056	0.4944
Administrative	0.0590	0.4900	0.5100
Nail/Spa/Eye/Misc.	0.0610	0.4124	0.5876

Across Quarters

Task	Share of Labor	Share of Variance: Quarter	Within-Quarter
Haircut/Shave	0.4049	0.0057	0.9943
Color/Highlight/Wash	0.3902	0.0062	0.9938
Blowdry/Style/Treatment/Extension	0.0850	0.0111	0.9889
Administrative	0.0590	0.0193	0.9807
Nail/Spa/Eye/Misc.	0.0610	0.0118	0.9882

B.14 Supplementary Tables

Table B9: Model-Based Decomposition of Job Task-Content Variance

Task	Share of Task-Content Variance: Firm	Worker
Haircut/Shave	0.0761	0.9239
Color/Highlight/Wash	0.1194	0.8806
Blowdry/Style/Treatment/Extension	0.2180	0.7820
Administrative	0.0965	0.9035
Nail/Spa/Eye/Misc.	0.0865	0.9135

Note: The table displays a variance decomposition which uses the model to separate the variance of job task content into a worker and firm component.

Table B7: Minimum Wage Counterfactual Type-Specific Wages, Employment and Specialization

	Initial			Reallocation			Reorganization		
Worker Type	Hours	Wage	Task-Spec.	Hours	Wage	Task-Spec.	Hours	Wage	Task-Spec.
Haircut/Shave	537550	$16.96	0.9463	506090	$20.00	0.9459	502152	$20.00	0.947
Color/Highlight/Wash	997053	$37.75	0.7245	997053	$37.33	0.7233	997053	$37.52	0.7209
Blowdry/Style/Treatment/Extension	444040	$20.91	0.4837	444040	$21.88	0.4817	444040	$21.64	0.4819
Administrative	41860	$26.99	0.6801	41860	$28.40	0.6807	41860	$28.12	0.6809
Nail/Spa/Eye/Misc.	34844	$81.16	0.8262	34844	$81.63	0.826	34844	$81.71	0.826

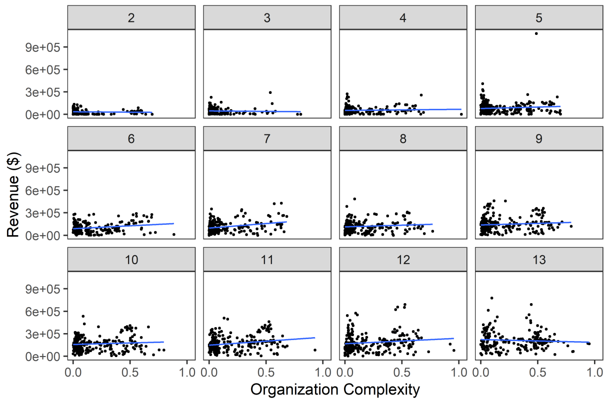
Note: This table displays employment and wage levels across the initial, reallocation and full equilibrium under a $20 minimum wage. It provides context for the main counterfactual results, which are reported in percentages. In both counterfactual equilibria, the minimum wage is binding only for haircut specialists.

Table B8: Sales Tax Counterfactual Type-Specific Wages, Employment and Specialization

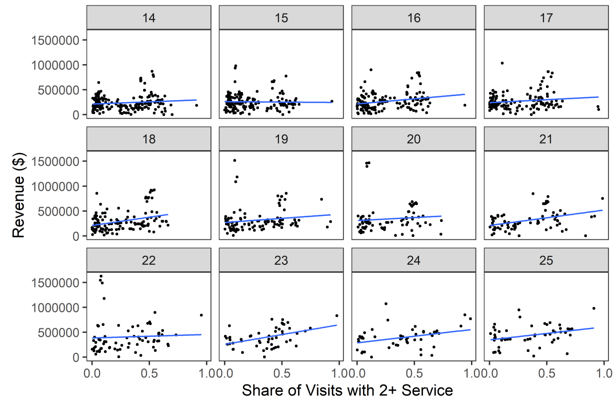
	Initial			Reallocation			Reorganization		
Worker Type	Hours	Wage	Task-Spec.	Hours	Wage	Task-Spec.	Hours	Wage	Task-Spec.
Haircut/Shave	537550	$16.96	0.9463	537550	$21.18	0.9471	537550	$22.38	0.9491
Color/Highlight/Wash	997053	$37.75	0.7245	997053	$45.99	0.7326	997053	$45.34	0.7432
Blowdry/Style/Treatment/Extension	444040	$20.91	0.4837	444040	$21.01	0.4946	444040	$22.18	0.4982
Administrative	41860	$26.99	0.6801	41860	$30.15	0.6786	41860	$31.85	0.6872
Nail/Spa/Eye/Misc.	34844	$81.16	0.8262	34844	$90.75	0.8351	34844	$91.49	0.846

Note: This table displays employment and wage levels across the initial, reallocation and full equilibrium under the elimination of the service sales tax. It provides context for the main counterfactual results, which are reported in percentages.

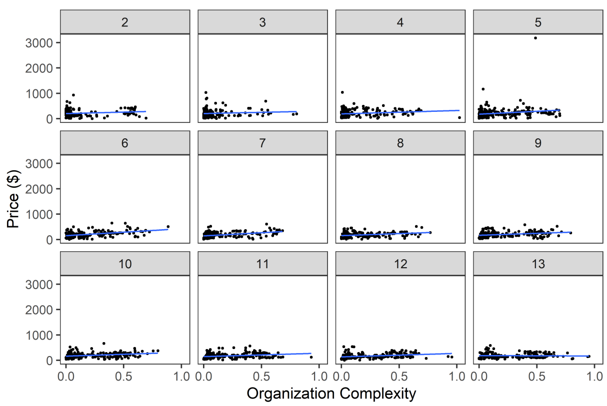
Figure B5: Complexity Relationships Among Similar-Size Firms: 2–13 Employees



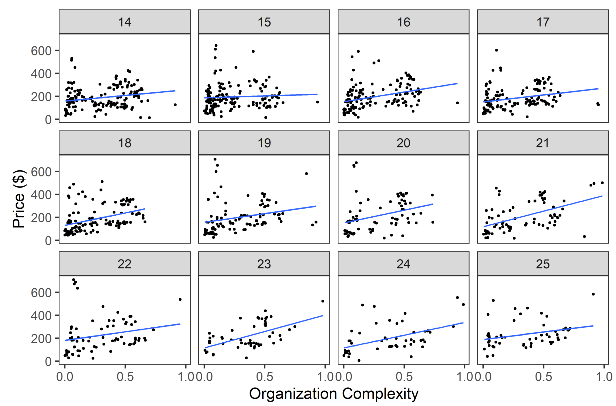
(a) Revenue (2–13 Employees)



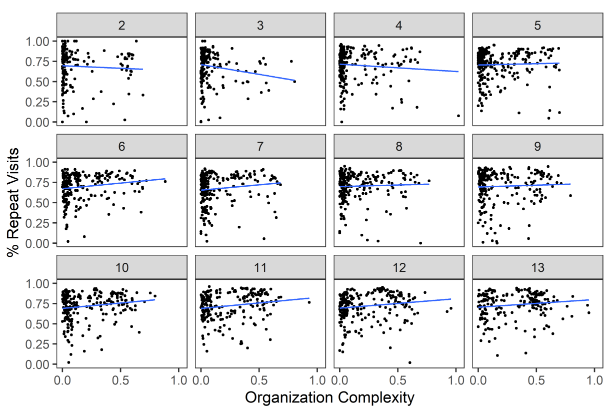
(b) Revenue (14–25 Employees)



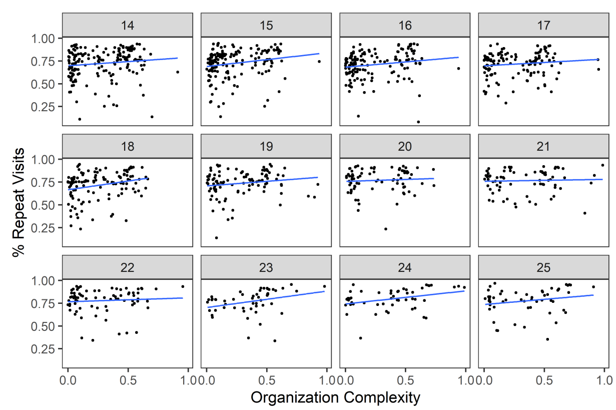
(c) Price (2–13 Employees)



(d) Price (14–25 Employees)



(e) Repeat Customers (2–13 Employees)



(f) Repeat Customers (14–25 Employees)