Competition and Internal Organization: Evidence from the Beauty Industry

Jacob Kohlhepp*

JOB MARKET PAPER

September 9, 2022

Abstract

This paper studies how internal organization decisions within the firm interact with products and labor markets outside the firm. Using data documenting the minute-by-minute assignment of millions of tasks to workers across competing salons, I show firms account for a large share of task variation, and the complexity of a firm's internal organization increases with revenue, employment, and prices. Based on these facts, I develop a model featuring multi-worker firms with different organization costs and workers with multidimensional skills. Firms choose their internal structure, including who to hire and the task-content of each worker's job, in order to compete in a differentiated product market. I prove an equivalence result, which enables theoretical analysis, identification and estimation. I estimate the model for Manhattan. A minimum wage increase from \$15 to \$21.25 generates wage increases which dominate salon, consumer and unemployed worker losses. This occurs for two reasons. First, labor is reallocated: employment losses are concentrated among firms intense in low-wage tasks with high organization costs, while some low organization cost firms not intense in low-wage tasks see employment gains. Second, firms internally reorganize by shifting tasks from laid-off workers to retained workers, amplifying unemployment. In equilibrium, employed workers for which the minimum wage binds (does not bind) see an increase (decrease) in task-specialization.

*Department of Economics, University of California, Los Angeles. Conant: jkohlhepp@ucla.edu
I am grateful for helpful feedback from Simon Board, Maurizio Mazzocco, Moritz Meyer-ter-Vehn, Daniel Haanwinckel, Till von Wachter, Alex Bloedel, Stepan Aleksanko, Jonathan Kowarski, Victoria Barone, and Coen Teulings. I am indebted to Tara Sinclair for helping me obtain the data used in this project.

"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

The way workers are assigned to tasks is an important determinant of many economic outcomes. Since at least Adam Smith's pin factory, economists have recognized that much of the division of labor occurs within firms. Most existing work on tasks treats this internal organization as homogeneous across firms or exogenous to product market conditions. While useful in some applications, these assumptions are inconsistent with a modern image of firms as strategic, organizationally unique entities.[1][2] The purpose of this paper is to combine insights from the task-based perspective, industrial organization and organizational economics to study how heterogeneous internal organization decisions within the firm interact with products and labor markets outside the firm.

When firm internal organization is heterogeneous and strategic, policies like minimum wages and sales taxes may have unforeseen effects not captured by traditional theory. Despite this importance, studying the interaction between internal organization and markets poses two challenges. Empirically, we need detailed data on the internal organization of many competing firms. Theoretically, we need a model where (1) firms differ in their organization costs (2) firms compete in a product and labor market (3) firms are allowed to choose their organization structure in response to their organization costs and market conditions.

This paper overcomes these challenges. I leverage novel management software data to design and estimate a structural model of firm internal organization. The management data document the minute-by-minute assignment of millions of tasks to workers across competing salons. I show 29 to 51% of variation in job task content occurs across firms, and the complexity of a firm's internal organization is positively correlated with revenue, employees, prices and repeat customers. Based on these facts, I develop a model where multi-worker firms with different organization costs choose their internal structure, including who to hire and the task-content of each worker's job. I prove an equivalence result, which allows me to study the model analytically, establish equilibrium existence and uniqueness for almost any parameter values, identify each firm's internal organization cost, and develop a computationally light estimation routine. The model illustrates that firms face a quality-wage-complexity trilemma when choosing their organizational

1. Gibbons 2020, Holmstrom and Milgrom 1994, Aghion, Bloom, and Van Reenen 2014
2. Tacit collusion among gas stations to bargaining leverage among coffee suppliers.

structure.

I estimate the model for Manhattan. Organization costs account for 40% of the variation in marginal costs. Job task content in the same occupation and the same product market varies based on the skill set of the worker and the firm at which they work. Workers at firms in the bottom quartile of organization costs (efficient firms) on average spend 90% of their time on their specialty task while workers in the top quartile (inefficient firms) spend 67%.

A minimum wage increase from $15 to $21.25 reduces consumer welfare by 0.9%, salon profit by 3.9%, and employment by 2.4%. Although the minimum is binding only for 1 of the 5 worker types, it results in higher wages for 4 of the 5 worker types. The wage increases for employed workers dominate the losses to firms, consumers and unemployed workers resulting in a small increase in total welfare of 1.3%. The minimum wage impacts firms in an unequal way. Employment losses and cost increases are concentrated among firms intense in low-wage tasks with high organization costs. Some firms with low organization costs and that are less intense in low wage tasks increase employment and profit, because the minimum wage increases their competitive advantage in the product market.

These welfare changes are driven by two effects. First, the minimum wage reallocates labor towards more organizationally efficient firms. Second, the minimum wage induces firms to internally reorganize, shifting tasks from laid-off workers to retained workers. This internal reorganization increases equilibrium unemployment, but it also changes the equilibrium division of labor. Surviving workers for which the minimum wage is binding experience an increase in task-specialization. Workers for which the minimum wage is not binding generally experience a decrease in task-specialization, as they are asked to "pick up the slack."

Eliminating the 4.5% New York City sales tax on services slightly reduces consumer welfare while increasing firm profit by 3.7%. Workers capture the majority of gains, with wages rising by 19.3%. This occurs because labor is reallocated from less to more organizationally efficient firms and because most firms increase service quality. They do this by increasing the complexity of their internal structure by 5.5%, which allows them to increase the task-specialization of their workforce. In equilibrium, all worker types are more productive, and these gains pass from consumers to firms via price increases and firms to workers through wage increases.

2 Related Literature

This paper draws from and contributes to three strands of literature.

Organizational theories of the firm. This literature is inspired by two questions: why do firms exist when we have markets (Coase 1937), and why do firms may differ in their "organizational capabilities" (Argyres et al. 2012). A few of the answers proposed include monitoring (Alchian and Demsetz 1972), relational contracts (Baker, Gibbons, and Murphy 2002), knowledge (Garicano and Wu 2012), coordination (Dessein et al. 2003) and culture (Martinez et al. 2015). The contribution of this paper is to aggregate these forces into firm specific organization costs so that they can be estimated and incorporated into a structural model. This contribution is also a limitation. By aggregating the forces, I am only able to say that firm A behaves as if they face greater organizational frictions. I cannot say what those frictions are.

Task-based perspectives of the labor market. Starting with Autor, Levy, and Murnane (2003), a growing literature has recast the labor market in terms of the tasks required for production. My model continues in this tradition, incorporating workers with different skill sets that impact how well they can perform different tasks. I contribute to this literature by studying the role of firms in the process.

Specifically, I study the strategic incentives firms face when the internal matching workers to tasks impacts marginal cost and product quality. The majority of work either assumes product markets are competitive or places restrictions on internal organization, often to study national labor markets or to study labor market power (Haanwinckel 2020a, Adenbaum 2021, (Ocampo 2018, Yamaguchi (2015),Lazear 2009). I also quantify and study the effect of organizational heterogeneity across firms. I allow firms to differ in their organization costs, meaning that the task-content of a worker's job will differ based on where they work and firms will adjust differently to policy changes.

Finally, I contribute methodologically to this literature by demonstrating how to use high-frequency management software data to study the intersection of labor and product markets. The majority of existing work uses O*NET or survey data. This data is more representative, but it is not well suited for studying task differences across firms in the same industry employing workers in the same occupation.

Management practices and technology. Work in this area links productivity to specific aspects of organizations, like the use of information or communication technology (Garicano and Heaton 2010, Garicano and Wu 2012), the level of trust (Bloom and Van Reenen 2007), individual incentive pay and human resource practices (Bloom and Van Reenen 2011), and management training (Giorcelli 2019). I contribute to this literature by providing a way to quantify management differences across firms via the firm-specific organizational cost parameter estimates. I also show, consistent with this literature, that even among similar firms competing in the same product market organizational differ-

ences are large. I am also able to analyze the impact of organizational convergence within an industry on consumers, workers and firms.

3 Stylized Facts

This section presents three facts about internal organization using data from a salon management software company. Details about the data are provided in Section 6. These facts inform the modeling choices made later.

Before presenting the facts, I define an organization structure, organizational complexity, the task content of a worker's job. These three concepts that will be used throughout the paper. I 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 B_j, is a matrix where element $B_j(i,k)$ is the fraction of labor assigned to worker i and task k.*

An example of two different organizational structures is given in Figure 1. The column totals represent the firm's task-mix: the amount of each task needed to produce one unit of output. This paper treats this margin as fixed, determined either by technology constraints or demand. The row margins represent the composition of a firm's workforce: how much work is assigned to each worker. The figure represents two different choices of organization structure, which both obey the column totals (the task-mix) but that assign work in very different ways.

Figure 1: Two Organization Structures

"Employee" 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

"Chair Renter" Salon

	Tasks 1	2	3	
A	1/6	1/8	1/8	1/3
B	1/6	1/8	1/8	1/3
C	1/6	1/8	1/8	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 left example is called an employee salon because each person performs a distinct task, and all workers are required to produce one unit. The right panel is called a chair renter salon because each employee is exchangeable and is not relying on the skills of others. It is as if there are three miniature salons.

Definition 2 *The complexity of an organization structure B_j is:*

$$I(B_j) = \sum_{i,k} B_j(i,k) log\left(\frac{B_j(i,k)}{\sum_{k'} B_j(i,k') \sum_{i'} B_j(i',k)}\right)$$

I refer to this quantity as complexity because it represents the minimum amount of information the firm must communicate to workers in order to implement the structure. This idea builds off decades of work in information theory, starting with Shannon (1948). To make this concrete, consider the two structures in Figure 1. The chair renter structure requires 0 information: the firm can have employees randomly decide who does which task. Therefore the complexity of this organization structure is 0.

For the employee arrangement, the firm must distinguish between a minimum of two cases. If the task is a cut, assign to A. The firm can communicate this to employees by sending a single bit of information. If not a cut, assign to B or C randomly. The firm can distinguish this by sending two bits. On average this is $1 \times 1/2 + 2 \times 1/2 = 1.5$ bits, which is why the complexity of the employee arrangement is 1.5.[3] For the task distribution $1/2, 1/4, 1/4$ it can be shown that the chair renter arrangement achieves the lowest possible complexity, while the employee arrangement achieves the highest possible complexity.

Definition 3 *The task-content of worker i's job at firm j is a vector where element $b_j(i,k)$ is the fraction of worker i's time spent on task k.*

A worker's job is related to the organizational structure of the firm at which they work. It is the ith row of the organizational structure normalized by the total amount of work performed by worker i:

$$b_j(i,k) = \frac{B_j(i,k)}{\sum_k B_j(i,k)}$$

3.1 29 to 53% of variation in job task content occurs across firms

To establish this fact, I decompose the variance into a within-firm and between firm component.[4] Similarly, I decompose variance into a between quarter-year and within-quarter component. The results are presented in Table 1. Between 28 to 53% of variation in task content occurs across firms. The decomposition reveals that difference across firms as well as difference within firms likely play a role in determining the task content of a given worker's job. In comparison, variation across time seems to play a minor role, explaining less than 2% of the variance in task content across all tasks.

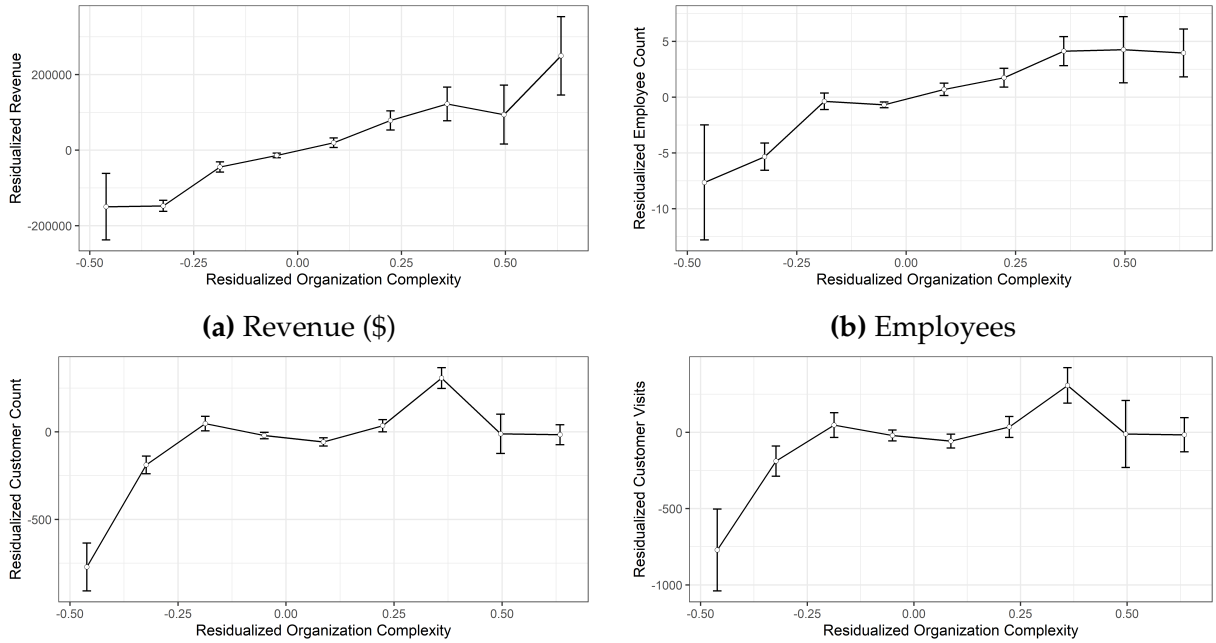
3. Notice the firm could also communicate this in a more complicated, less efficient way, by sending a separate signal for each type of task.

4. Formally: $Var(b_j(i,k)) = Var_j(\bar{b}_j(i,k)) + \sum_j \omega_j Var_i(b_j(i,k)|j)$

3.2 Complex firms have higher revenue and employment

Complexity is positively correlated with revenue and employment as well as several other measures of firm size. I depict this correlation visually in Figure 2 and via regression in Table 2. In both the graphs and regressions, I control for the task-mix, county and year-quarter. The relationship is strongest for revenue and employment, and is statistically significant at the 5 percent level.

Figure 2: Organization Complexity and Firm Size



(a) Revenue ($)

(b) Employees

(c) Unique Customers

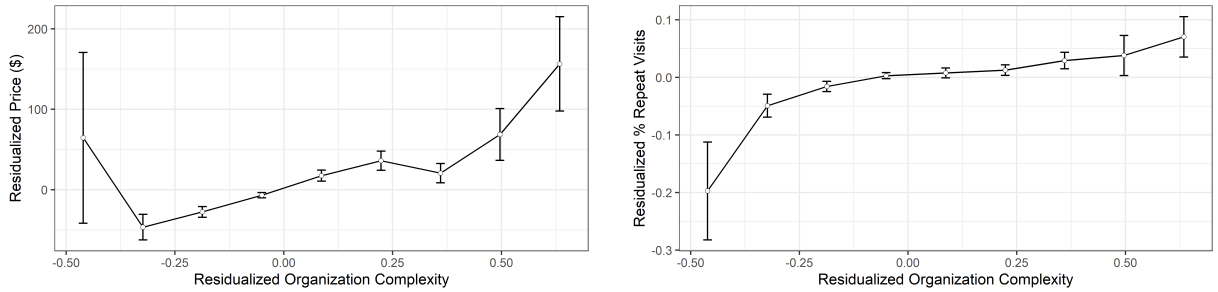
(d) Customer Visits

Note: Each panel illustrates the positive relationship between organization complexity and a different measure of firm size. Regressions indicate that only the revenue and employee relationships are statistically significant. The relationships support the idea that some firms have an organizational advantage. All variables residualized for quarter, county and task-mix. Firm-quarters are grouped into equally spaced bins by complexity.

3.3 Complex firms have higher prices and more repeat customers

I show in Section 10.5 that this pattern makes the data inconsistent with a model where organizational competitive advantages operate only through marginal cost reductions. In such a model, prices should be decreasing in complexity. I interpret this as evidence of endogenous quality: salons with organizational advantage do not just produce more, they produce higher quality services which they sell at a higher price. To further test the quality channel, I proxy for quality with the share of visits made by repeat customers in Figure 3.

Figure 3: Organization Complexity and Repeat Customers



(a) Price **(b)** Repeat Customers

Note: Panel A illustrates the positive relationship between organization complexity and price, while panel B illustrates the relationship between organization complexity and the fraction customers that return. These are inconsistent with a model of quantity-based productivity and supportive of a model of quality-based productivity. All variables residualized for quarter, county and task-mix. Firm-quarters are grouped into equally spaced bins by complexity.

Although repeat customers are not defined within the model, 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. In summary, the data suggest that variation in task content comes both across firms and within firms, suggesting both worker skills and firm organizational differences play important roles. Firms with complex internal structures earn higher revenue, suggesting an organizational competitive advantage in the product market. Finally, prices and return visitors rise with complexity, suggesting increased quality is an important mechanism through which the organizational competitive advantage operates.

4 A Model of Internal Organization with Competing Firms

This section specifies a model where firms choose prices and organizational structures simultaneously in order to compete for consumers. In order to match the facts established earlier, the model features endogenous vertical quality and heterogeneous organization costs. There are three important groups of objects in the economy. Firms, indexed by $j = 1, ..., J$. Worker types, indexed by $i = 1, ..., N$. Tasks, indexed by $k = 1, ..., K$.

Firms and Tasks. The J firms differ in their organizational cost γ_j discussed below. Each firm produces a single product by assigning labor to K tasks according to the task-mix vector, α. Specifically, α_k denotes the amount of labor the firm must devote to task k to

produce one unit of the final good.[5] I normalize $\sum_k \alpha_k = 1$.

Workers and Labor Markets. Each of the N workers types are 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. Let $\Delta^{N\times K}$ be a $N \times K$-dimensional unit simplex. Firms choose the price of their product $p_j \in \mathbb{R}_+$ and their organizational structure $B_j \in \Delta^{N\times K}$. An organizational structure B_j is a matrix, where element $B_j(i,k)$ specifies the fraction of total labor allocated to worker i and task k. An organizational structure B_j is *feasible* if they it is consistent with the task-mix vector, that is if:

$$\sum_i B_j(i,k) = \alpha_k \forall k$$

Sometimes it will be helpful to refer to the workforce composition, that is the fraction of total labor demanded that is from each worker type. I denote this E_j, where element $E_j(i)$ denotes the fraction of total labor performed by a worker of type i. Note that E_j is implied by B_j because $\sum_k B_j(i,k) = E_j(i)$.

The cost of an organizational structure is determined by its complexity. Specifically, the per-unit organization cost of firm j with organizational structure B_j is complexity multiplied by the firm-specific organization cost parameter γ_j. I use an entropy-based measure of complexity of an organizational structure:[6][7]

$$I(B_j) = \sum_{i,k} B_j(i,k) log\left(\frac{B_j(i,k)}{\sum_{k'} B_j(i,k') \sum_{i'} B_j(i',k)}\right)$$

A firm's organizational structure determines the match between worker skills and tasks. As a result, it determines product quality (ξ_j). I specify that product quality is a weighted-

5. This is similar to assuming all firms share a common Leontief task-based production function. Specifically, if a is a vector denoting the amount of labor devoted to each task, a firm with task-mix vector α has a task-based production function given by: $F_\alpha(a) = \min\left\{\frac{a_1}{\alpha_1}, ... \frac{a_k}{\alpha_k}, ..., \frac{a_K}{\alpha_K}\right\}$

6. The measure 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 gives the model many micro-foundations.

7. When computing this measure, I assume that whenever $B_{j,i,k} = 0$, that element of the sum is 0.

average of task quality:

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

Since quality is valued by consumers, this is the main benefit of carefully assigning workers to tasks.

Consumers and the the Firm's Problem. There are a mass 1 of consumers interested in purchasing at most one of the J final products. Consumer z's utility for good j is represented by the following utility function:

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

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 $u_{z,0} = \epsilon_{z,0}$. I assume throughout that utility is strictly decreasing in price, $\rho > 0$. Market shares (s_j) can be written as:

$$s_j = \frac{exp(\overbrace{\xi(B_j)}^{\text{quality}} - \rho p_j)}{\sum_{j'} exp(\xi(B_{j'}) - \rho p_{j'})}$$

Per-unit organization costs and competitive labor markets imply marginal costs are constant. This along with the expression for market share imply the firm's problem can be written as:

$$\max_{p_j \in \mathbb{R}_+, B_j \in \Delta^{N\times K}} \underbrace{\frac{exp(\overbrace{\xi(B_j)}^{\text{quality}} - \rho p_j)}{\sum_{j'} exp(\xi(B_{j'}) - \rho p_{j'})}}_{\text{market share, } s_j} \Bigg[p_j - \underbrace{\left(\overbrace{\gamma_j I(B_j)}^{\text{org.}} + \overbrace{\sum_{i,k} w_i B_j(i,k)}^{\text{avg. wage}} \right)}_{\text{constant marginal cost, } MC_j} \Bigg] \tag{1}$$

s.t.

$$\sum_i B_j(i,k) = \alpha_k \forall k$$

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

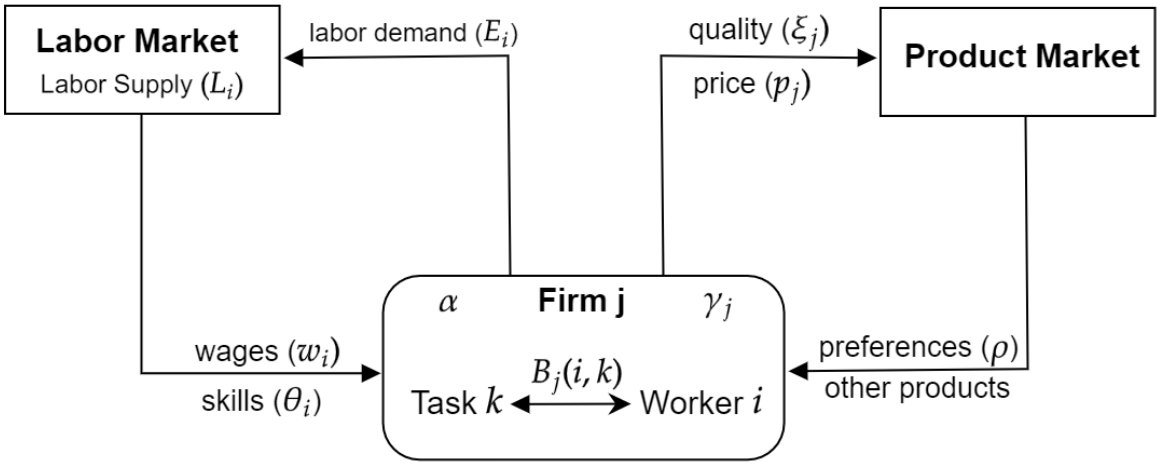
1. Firms choose prices p_j and organizational structures B_j to maximize (1).

2. Labor markets for each worker type clear:

$$\sum_j \frac{exp(\xi(B_j) - \rho p_j)}{\sum_{j'} exp(\xi(B_{j'}) - \rho p_{j'})} E_j(i) = L_i \, \forall \, i = 1, ..., N$$

Model Summary. The model for two firms is illustrated graphically in Figure 4. Firms choose B_j, which determines both who they hire and how hired workers are assigned to tasks. This choice impacts the product market directly by determining product quality and indirectly through prices. This choice impacts the labor market because it determines labor demand across worker types.

Figure 4: Illustration of the Model



Note: The diagram summarizes the different parts of the model. Each firm chooses B_j (who to hire and the assignment of workers to tasks) and prices taking into account internal factors like the task-mix (α) and organization costs (γ_j), labor market factors, like wages and skills, and product market factors, like price sensitivities (ρ) and the prices and qualities of other products. In turn, firm chooses feedback into both the product and labor market in equilibrium, because labor markets must clear and prices must form a Bertrand Nash equilibrium.

4.1 The Choice of Organizational Structure

An example from the hair salon industry makes the model concrete. There are two main ways to organize a salon. In the chair renter arrangement, stylists pay a fixed fee to 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 renter arrangement is characterized by a simple contracts and little need for coordination. In the language of the model, there is little organizational complexity. 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.

Although the employee arrangement is complex, it allows the firm to leverage the individual talents of workers to produce a higher quality product than any individual could produce alone. In contrast, the chair renter arrangement is simple but makes the salon more like a collection of individual businesses. The two organizational structures implied by this discussion are illustrated in the following figure.

4.2 Interpreting Organization Costs

This section describes different ways of interpreting γ_j. To summarize, γ_j is the cost of increasing the complexity of firm j's organizational structure by 1 unit. It represents the collection of relationships, management practices, contract knowledge, human resource technology and other factors which are inherent to firm j. Many of these interpretations can be formalized as micro-foundations. This both a strength and a limitation of the model. On the one hand, we can account for several organizational frictions together using one cost function and one parameter per firm. On the other hand, we cannot understand the mechanisms driving the value of γ_j at any particular firm.

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 and 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 this 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 people, producing a product superior to that which any could produce alone. A firm with γ very large is essentially a group of people each working independently.

Rational Inattention. The mutual information form of organizational costs gives it an 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. There is also a sense in which we can think of organizational costs as capturing contractual inattention, in the sense of Tirole (2009). That is different firms may find it more or less costly to write down complex contracts in order to support complex organizational structures.

Specialization is Costly. A less abstract interpretation is to view organizational efficiency as an assumption that specialization is costly, and this cost differs across firms according to γ. This interpretation arises from the fact that optimal $I(b, \alpha)$ is a measure of task-specialization.[8]

5 Theoretical Results

This section derives the main characterization of the firm's problem and provides theoretical intuition for the trade-offs firms face when choosing an organizational structure while operating in a shared product and labor market.

5.1 Main Characterization

The main theoretical result characterizes the firm's choice of organizational structure in a convenient way. All other results in the paper are applications of this characterization, including identification and estimation of the structural model developed later.

I define the feasible set of organization structures:

$$\mathbb{B} = \{B \in \Delta^{N \times K} | \sum_i B(i,k) = \alpha_k \forall k\}$$

We can now write the firms problem (1) in a simpler form:

$$\max_{p_j \in \mathbb{R}_+, B_j \in \mathbb{B}} \underbrace{\frac{exp(\overbrace{\xi(B_j)}^{\text{quality}} - \rho p_j)}{\sum_{j'} exp(\xi(B_{j'}) - \rho p_{j'})}}_{\text{market share, } s_j} \Bigg[p_j - \underbrace{\left(\overbrace{\gamma_j I(B_j)}^{\text{org.}} + \overbrace{\sum_{i,k} w_i B_j(i,k)}^{\text{avg. wage}} \right)}_{\text{constant marginal cost, } MC_j} \Bigg] \tag{2}$$

Taking the first-order condition with respect to price, we obtain the well-known equation which states that prices are marginal cost plus a mark-up.

$$p_j^*(B_j) = \frac{1}{\rho(1 - s_j(B_j))} + MC_j(B_j) \tag{3}$$

Notice the mark-up is a function of market-share, which is a function of the prices and qualities of other firms. Thus this equation represents a price best-response function. Also note that market-shares and marginal cost both depend on organizational structure, B_j.

8. The proof is in Section A.7.

At this point, it is useful to define quality-adjusted cost $MC_j(B_j) - \rho^{-1}\xi_j$. Given 3 we can relate profit and quality-adjusted marginal cost.

Lemma 1 *Given optimally chosen prices, profit is strictly decreasing in quality-adjusted marginal cost.*

The proof is mainly algebraic, and is provided in Appendix Section A.1. The lemma implies firms wish to choose an organization structure which achieves the lowest quality-adjusted marginal cost. In other words, the firm's optimal organization structure also solves the much simpler problem:

$$\min_{B_j \in \mathbb{B}_j} MC_j(B_j) - \rho^{-1}\xi_j(B_j) \tag{4}$$

Denote the firm-specific wage bill, complexity and quality as $W(B_j)$, $I(B_j)$, $\xi(B_j)$ respectively We can unpack and rearrange Equation 4 in the following way:

$$\begin{aligned}
\min_{B_j \in \mathbb{B}_j} MC_j(B_j) - \rho^{-1}\xi(B_j) &= \min_{B_j \in \mathbb{B}} W(B_j) + \gamma_j I(B_j) - \rho^{-1}\xi_j(B_j) \\
&= \gamma_j \left\{ \min_{B_j \in \mathbb{B}} I(B_j) + \gamma_j^{-1} \left[W(B_j) - \rho^{-1}\xi(B_j) \right] \right\} \\
&= \gamma_j \left\{ \min_{B_j \in \mathbb{B}} I(B_j) + \gamma_j^{-1} \sum_{i,k} B_j(i,k)(W_i - \rho^{-1}\theta_{i,k}) \right]
\end{aligned}$$

If we examine the final expression and compare it to formulations like Equation 6 in Tishby, Pereira, and Bialek (2000), we see that the choice of organizational structure also solves a rate-distortion problem. With this realization, the proof of the following theorem is complete.

Theorem 1 *The profit-maximizing organizational structure B^* also solves:*

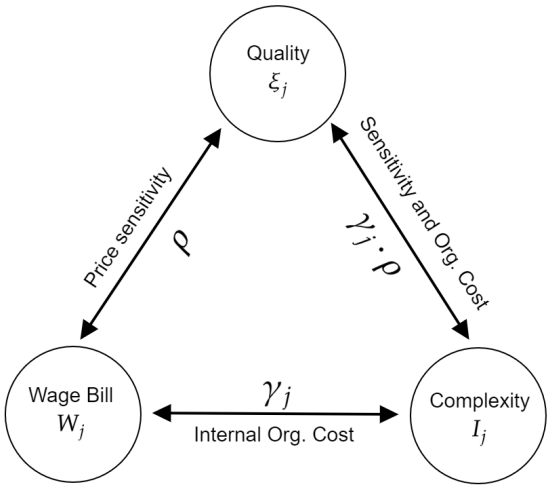
$$\min_{B_j \in \mathbb{B}} I(B_j) + \gamma^{-1} \left[W(B_j) - \rho^{-1}\xi(B_j) \right] \tag{5}$$

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

This characterization of the firm's problem is the main technical result of this paper, from which all other results follow. We can begin using it immediately to think about the choice of organizational structure. If we examine (5), we see that it depends on complexity, wages and quality. The firms faces a trilemma: it 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 5: The Complexity-Wage-Quality Trilemma



Note: The characterization derived in Theorem 1 reveals that firms face a three-way trade-off when designing their internal structure. This trade-off does not depend on other firms actions except through wages in equilibrium. The firm must choose between higher quality, higher organization costs, or a higher wage bill. The exact trade-off depends on consumer price sensitivity and the firm-specific organization cost parameter γ_j.

The intuition for why price sensitivity impacts organization structure is as follows. When ρ is high, consumers are very price sensitive and the firm cannot pass on any type of cost through prices. This means the firm will prioritize minimizing costs over maximizing quality, and will choose less complex organizations. When ρ is low, the firm can pass on most costs (including organization costs and wages) to consumers through prices. Because of this, it will prioritize improving quality. Since greater quality generally requires greater complexity, this causes the firm to choose a more complex organization. In the limiting case of $\rho = 0$, the firm can pass on all costs to consumers via prices. In this case any organizational advantage (granted to firms with low γ) is eliminated because all firms find it free to design complex organizations or hire high quality workers.

I now continue analyzing how the firm navigates the complexity-wage-quality trade-off. To do so, I define the organization frontier.

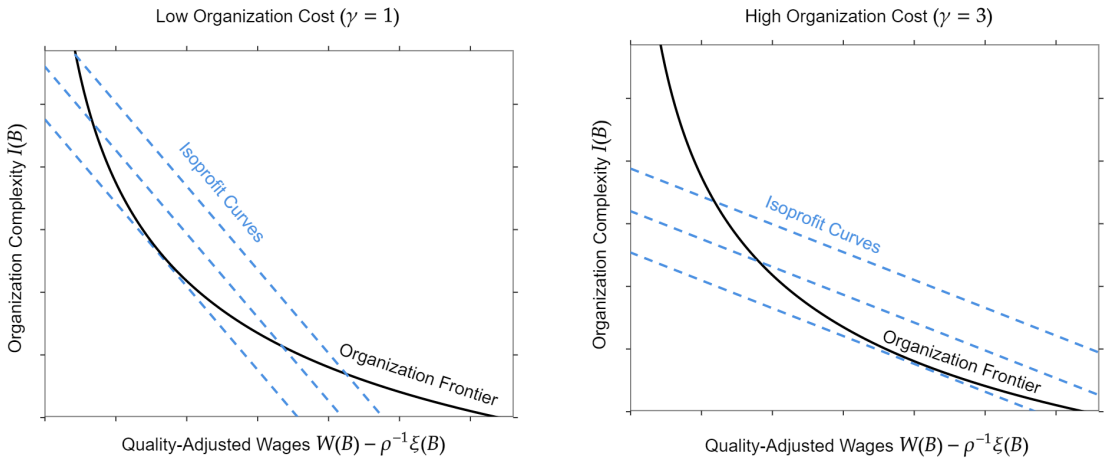
Definition 4 *The organization frontier is the set of all organization structures which minimize complexity for some quality-adjusted wages.*

The frontier consists of the simplest organizations which achieve some quality-adjusted wages. The characterization provided in Theorem 1 allows us to apply existing results from information theory to understand the general shape of the organizational frontier.

Proposition 1 *Organization complexity along the organization frontier is continuous, convex and decreasing in quality-adjusted wages.*

The proof is provided in the Appendix. This proposition implies that the firm's choice of organization structure boils down to a two-dimensional choice similar to a classic expenditure minimization problem from consumer theory. The firm trades-off organizational complexity with quality-adjusted wages. How it navigates this trade-off is determined by the organizational cost parameter γ. This is illustrated in Figure 6.

Figure 6: Choosing an Organizational Structure



The firm chooses the point along the frontier that achieves the left-most isoprofit curve. The firm's isoprofit curves have a slope equal to $-\gamma_j^{-1}$. As γ_j rises, the curves become flatter, which causes the tangent point to shift right, reducing organizational complexity and increasing quality-adjusted wages. A more complex organization allows a firm to produce a higher quality good at lower wage but it requires a greater organization cost. γ determines how each specific firm navigates this trade-off. 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 lower organization cost (γ_j) have higher organization complexity, market share, and profits.*

Recall that γ_j represents the management technology, relationships, knowledge, and practices specific to the firm. These factors make it easier or harder for a firm to implement a given organizational structure. Proposition 2 means that more organizationally efficient firms are larger and more profitable. They can produce better quality goods at a lower

cost. Importantly, this proposition is a testable implication of the model: firms with a more complex organization structure should be larger and more profitable than firms with a less complex structure.

One implication of the above results is that while firms interact strategically in their price decisions, it is as if firms have a dominant organizational strategy. That is, firms choose B_j to minimize quality-adjusted cost, which depends on other firm decisions only indirectly through wages. This greatly improves tractability because we can solve the organization problem first and then use the solution to solve for optimal prices. This is critical to the estimation procedure outlined later in the paper.

In many contexts this may be a reasonable assumption. The main way in which two firm's internal organization is often connected is via the labor market. However, there are situations where product market choices of other firms may directly impact internal organization. For example, in a model where consumers are more likely to substitute to products with similar qualities, the quality choice of the other firm may directly impact the returns to quality, and thus the returns to more complex organizations. Section 10.5 describes such a model in more detail and outlines some challenges involved.

5.2 Optimal Jobs Within the Firm

The last few results show the originally high-dimensional problem of the firm can thought of tractably as a 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.

I first denote the jobs of workers at firm j as b_j and define it as a matrix where element i, k is the fraction of worker i's total time spent on task k. Under the model, $b_j(i,k)$ corresponds to:

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

With this in hand, the following result describes several properties

Proposition 3 *The profit-maximizing organizational structures satisfies the following properties.*

1. ***Law of Demand:*** *The share of workers of type i ($E_j(i)$) hired 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:*

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

The proof is provided in Appendix Section A.3. If we examine the expression, we see that worker 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 organizational costs through γ_j.

This proposition highlights two important feature of the model. First, whenever there are some organizational frictions within a firm, complete specialization will not occur. Put another way, jobs are always a bundle of tasks. Second, because jobs depend on organizational 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. Their job will depend on the type of firm where they work.

6 Data

This section describes the salon management software data.

6.1 Context and Institutional Details

The data comes from a salon management software company. The software helps with many aspects of running a beauty business, including scheduling, pricing, payments, inventory, staffing, business reporting, client profiling, and marketing. As of July 19, 2022 the low-end subscription for the software is $175 per month, with an upgraded version available for $265. Although the company markets its products to spas, tanning salons and massage parlors, hair salons and barbers make up the majority of 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 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 documents the assignment of individual stylists to particular services with particular clients. It describes the duration of the appointment, the price paid, and a custom text description of the services performed. 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 is de-identified, unique ID allows a researcher to track employees and clients across time within salon.[9] A sample from the activity is provided in Table 3, with pseudonyms added.

From this data, we gain insight about how the different ways 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 is rich in terms of task content and worker assignments, information about employee compensation is sparse. The software is capable of tracking some compensation information, like tips, commissions and employment relationship. However, discussion with the software company and examining of the data reveal that these functions are not consistently used by the client firms.

I am permitted to use the data as part of a data sharing agreement. I designed the agreement as part of a consulting arrangement with the software company. During the first part of the consulting arrangement, the data was used to conduct internal analyses for the company. After, the agreement stipulated that I own the de-identified data, and the firm only had the right to review publications for privacy concerns.

6.2 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 six tasks.

First, I created 13 subcategories along with several other categories related to the type of service. The descriptions were sent to a certified cosmetologist hired through UpWork.[10] The cosmetologist was paid $1,500 to read all descriptions and mark all subcategories which fit the description. The cosmetologist was instructed to mark multiple columns if a service matched multiple subcategories. This process took approximately one week. The same cosmetologist was hired to group the subcategories into task categories based on the types of people who would typically do each. Appendix Section B.1 describes the instructions sent to the cosmetologist and displays part of the final spreadsheet. I use the 6-category grouping provided by the cosmetologist with one modification:

9. IDs are salon-specific, so we cannot track employees or clients across salons.
10. I received a copy of the person's cosmetology license.

I combine the extension task with the blowdry task. The reason is that the extension task is very sparse. For Manhattan in 2021 Q2, it has less than 10 hours. This sparsity causes estimation problems, as parameters tied to this task have a negligible effect on observable outcomes.

A description marked as multiple tasks is divided across multiple tasks in the following way. First, I compute the average amount of time spent on a task in the category among only descriptions with a single task marked. I then compute the fraction of time to assign to each task as the corresponding average divided by the sum of the averages of all marked tasks. This process generates a data set with worker-client-service observations grouped into six mutually exclusive tasks.

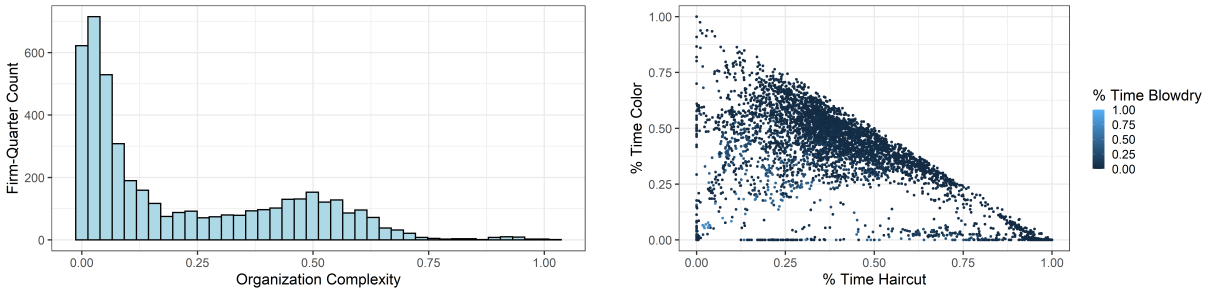
6.3 Descriptive Statistics

The data used in this section and the model tests includes all observed firm-quarters where revenue per customer is positive. I exclude 2021 Q3 because I only observe 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 of data is August of 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. The salons in the sample are higher end, bringing on average $200 in revenue per customer in a given quarter. Throughout the paper, I refer to the price as the average revenue per customer per quarter. Even though there is significant variation in the relative importance of services offered, most salons offer at least 4 of the 5 task categories. Organizational complexity using the natural log scale with five tasks is bounded above by approximately 1.61 and below by 0. Some firms have an organizational complexity of 0, while others have a complexity as high as 1.02. I visually represent firm-quarter heterogeneity in the task-mix and complexity in Figure 7. Firms vary in their intensity in each task, and there is a long right tail of high-complexity firms.[11]

11. To see variation across jobs, see Table 14 and Figure 23.

Figure 7: Heterogeneity in Organization Structures



(a) Organizational Complexity **(b)** Task-Mix

Note: Panel A is the distribution of organization complexity across all firm quarters in the data. Panel B is the distribution of firm-quarter task-mixtures for the first three tasks. These are the observable variables used to identify γ_j and each firm's organization structure B_j in the structural model.

6.4 The Strengths and Limitations of Salon Management Data

Hair salons are a good setting for understanding the connection between organization and competitive advantage for two reasons. First, labor is the primary input to production. Second, in the beauty industry tasks, which are typically not observable, map closely to services, which are observable. This means that transaction-level data can be used to measure tasks.

There is typically a breadth-depth trade-off when it comes to understanding the internal organizations of firms. Some papers are able to deeply understand the workings of a few establishments using site visits and case studies. Others are able to collect summary measurements from many firms. By using management software, I am able to do both. I am able to see the minute-by-minute assignment of hair stylists to tasks across hundreds of independent businesses.

Another unique strength of my approach is that I do not rely on survey information. Because hair salons use the software to coordinate economic activity in real time, the data do not suffer from recall bias. Since all firms are using the same software, there is also a sense in which my results are not confounded by differences in information technology.

There are two major issues when it comes to using salon management data. First, because salons must use the software to be in the data, my sample is selected. This is not as big a problem because my estimation approach effectively conditions on organizational efficiency. Second, firms code services themselves. This means the data is quite rich, but it also means that I must process descriptions into task categories which are meaningful across firms.

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

7.1 Econometric Model

I define labor markets and product markets as counties. I define time periods as quarters. I estimate the model for New York County (Manhattan) 2021 Quarter 2, the last full quarter of data in my sample.

I allow the task-mix (α_j) to be firm-specific. This is similar to allowing firms to have different task-based production functions. Since α will be equal to the distribution of time across tasks, it is fully observable in the data. This allows firms in the same product and labor market to have different feasible organizational structures.

I add additional quality heterogeneity. I specify that in addition to endogenously chosen quality ξ, each firm also has unobserved vertical quality ν which it cannot change. This represents reputation and other attributes that are not related to labor. When the model is estimated, only quality differences correlated with measured organization complexity (I_j) will be attributed to internal organization. I also allow demand to vary based on the task-mix, by explicitly including α in utility with a vector of coefficients β.

I add additional cost heterogeneity. I specify that in addition to the endogenous average wage bill, each firm also has a marginal cost shifter ϕ. This represents productivity differences for reasons other than internal organization. When the model is estimated, only variation in prices correlated with organization complexity are attributed to internal organization. I also specify that all firms share common material costs of production (m), which depend on the task-mix.

I allow workers to differ in their labor supply. This clarifies the relationship between worker identities (observed in the data) and worker types in the model. Formally, in addition to their skill set, individual workers also have an inelastic labor supply l_i. I augment the game by specifying that firms first demand an amount of labor of each skill set. Then an unspecified process matches workers to firms. The only assumption I place on this process is that firm's labor demand from the first stage is exactly met. Thus if a firm demands 10 hours of a skill set, it may be met by 10 hours from one worker or 5

hours from two workers, but no more or less than 10 hours is supplied. Following the matching process, firms then select an organization structure $\tilde{B}$, which is over worker identities rather than just worker skill sets.[12]

I allow firms to differ in the labor required to produce a unit of the final good. I specify that each firm requires $\bar{a}_j$ hours of labor to produce one unit of its good.[13] I measure $\bar{a}_j$ as the average number of hours salon j spends on a customer in a quarter. I continue to assume that organization costs and wages are per hour of labor.

I add a product market sales tax τ. Since the effective price faced by the consumer is $(1+\tau)p_j$, a sales tax is equivalent from the firm's perspective to making the consumer more price sensitive (increasing ρ). In most states, salon services are considered tangible personal property and are thus often exempt from sales tax. This is the case in California (*Tax Guide for Barbers and Beauty Shops* 2022). The state of New York also does not tax hair services. However, New York City (which includes New York County/Manhattan) levies a 4.5 percent tax on beauty services (New York City Department of Finance 2022).

Finally, I place restrictions on the worker type space. Specifically, I assume there are four worker types each with a different specialty task. Thus worker types can be indexed by their specialty task. Workers perform their specialty task i with skill $\bar{\theta}_i$. They perform all other tasks with a "typical" quality $\underline{\theta_i}$, which we normalize to 0. It is important to note that this restriction is done due to data limitations and concerns about identification. Given sufficient data on workers, the model can be identified and estimated for a richer worker type space. Since I have very little information about workers, this is not possible.[14]

Profit under the empirical model 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, we can divide the parameters of the model into two groups. The first are the firm-specific organizational cost coefficients $\{\gamma_j\}_{j=1}^{J}$ which are the main parameters of interest. The second are the market parameters that define the demand system (7 param-

12. This means that in principle a firm may employ multiple workers of the same skill set and assign them different distributions of tasks.

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

14. Typical quality because it is not separately identified from the β coefficients: all work on task k not assigned to the specialist in task k is assigned to a non-specialist. Thus we could raise $\underline{\theta_k}$ and decrease m_k and the model is observationally equivalent.

eters), worker skills (5 parameters), wages (5 parameters) and material costs (5 parameters). I denote these market parameters Ω. The following sections discuss identification and estimation of these parameters.

7.2 Equilibrium Existence and Uniqueness

Proposition 4 *Suppose wages are fixed. A pure strategy equilibrium exists, and is unique for almost any parameter values.*

By "almost any" I mean that the set of parameter values which generate a unique equilibrium is full measure. Put another way, there is a sense in which multiplicity arises only in knife-edge cases. The proof of this result is provided in the Appendix, and it relies on the equivalence to a rational inattention problem established in Theorem 1. Since wages are determined by market clearing, Proposition 4 does not establish equilibrium existence and uniqueness for the full theoretical model. Nevertheless, the result is still useful for three reasons.

First, even with fixed wages, the model in this paper features firms of heterogeneous productivity endogenously choosing their quality and marginal cost. Many models where product positioning is endogenous (including the canonical two-stage Hotelling model) suffer from equilibrium existence and uniqueness problems.[15] Second, the estimation procedure takes wages as parameters to be estimated. The result implies that if we have a uniform prior over the set of parameters, the probability the true parameters induce multiple equilibria is 0. Third, the result means that the model makes sharp internal organization predictions. That is, for almost any set of parameters, the model delivers one and only one internal organization structure for each firm. The model does not suffer from the inverse identification problem, which makes counterfactual analysis straightforward.[16] Additionally, the model almost always delivers one and only one distribution of task content across workers in a labor market. Because the distribution of task content across workers is observed in my data, I use this fact to validate the model.

7.3 Identification of Firm-Specific Organization Costs

The main parameters of interest are the organization costs, γ_j. They determine product quality, organization complexity and marginal costs for each firm. The fact that organiza-

15. In a 2 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)

16. Put another way, we will not need to select among different equilibria when studying a policy change.

tion cost parameters grow with the number of firms is concerning for identification and estimation. I alleviate this concern by proving that each firm's organization cost is identified given knowledge of the market parameters Ω. Further, I show organization costs are not statistical parameters to be estimated: they are a deterministic function of the data.

Proposition 5 *Suppose the market parameters Ω are known and the equilibrium is unique. If a firm's task-mix (α_j) is observed, and complexity (I_j) is observed and not 0, then the following are identified and a deterministic function of the data:*

1. *organization cost (γ_j)*
2. *organization structure (B_j)*

The formal proof consists of three parts: first we show that the observed organization complexity based on worker identities is equal to the unobserved true organizational complexity, which is based on worker types. Second, Proposition 2 implies that the firm's profit maximizing organization complexity is strictly decreasing in the organizational cost parameter γ whenever complexity is non-zero. This establishes a one-to-one mapping between an observed quantity (the mutual information between worker identities and tasks) and the unobserved firm-specific deep parameter, γ_j. Once γ_j is known, from equilibrium uniqueness we have that firm's organizational structure is uniquely pinned down. The proof is fully described in Appendix Section A.5.

A key hurdle in establishing identification is the fact that I do not observe worker types. Instead, I observe worker identities within firms. Because I allow for a flexible matching process, a given firm may employ multiple workers of the same skill set. The distraction free property of mutual information implies that firms assign workers with the same skill set the same distribution of tasks regardless of their labor supply. This is critical for identification, because it means that the observed organization complexity based on worker identities is equal to the true organization complexity based on worker skill sets.[17]

For intuition, recall Figure 6. Suppose we see two firms with the same task-mix (α_j) operating in the same product and labor market. This means they have the same organization frontier. If one firm has a higher complexity, it must be that the slope of their isoprofit line is steeper, which can only be because they have a lower organization cost.

17. Absent this property, we would need to find a way to map worker identities to types, perhaps using observed performance or demographic data.

Therefore we can actually order the firms by organization cost without knowing the market parameters. Once we know the market parameters, we 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 statisical sense. For any market parameters, there are unique organization cost parameters which rationalize the observed organization complexities and task-mixtures. This is similar to how unobserved product quality in Berry, Levinsohn, and Pakes (1995) is a deterministic function of market shares. Also similar to Berry, Levinsohn, and Pakes (1995), there is no closed-form for the function that maps the data into organization costs. Instead I provide a contraction mapping as part of my suggested estimation routine.

Beyond the main implication for estimation, Proposition 5 has two additional implications; one economic and one statistical. Statistically, it means that observing the task-mix (a length K vector) and organizational complexity (a scalar) is enough to estimate the model. It is not necessary to observe the individual assignments of workers to tasks, only an average measure of the complexity of the assignment. This means the model can be estimated in settings where rich assignment data is not available.[18] It also means that if a researcher has assignment data, they can estimate the model using only summary measures from this data and use the full data to conduct a validation exercise. This is precisely what I do in Section 8.1.

Economically, the result means that the task-mix (a length K vector) and organizational complexity (a scalar) are sufficient statistics for a firm's internal organization structure, B_j (an $N \times K$ matrix). When all firms share the same task-mix, this means that a firm's internal structure can be fully inferred by competitors through complexity alone. On the one hand, this illustrates that logit demand and mutual information-based organization costs are imposing a large amount of structure. On the other hand, to the extent logit demand and mutual information-based organization costs are reasonable, this framework has the potential to maintain tractability in complicated settings.

7.4 Identification of Market Parameters

Organization costs are a deterministic function of the data provided market parameters are known. This section derives a set of moments and assumptions under which the market parameters can be estimated via the generalized method of moments.

To construct moment conditions, I follow the industrial organization literature and use one demand side and one supply side equation. Starting with the supply side, we

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

can write the equilibrium pricing equation as:

$$p_j = \frac{1}{\rho(1+\tau)(1-s_j)} + \bar{a}_j \Big[\gamma(\Omega, I_j, \alpha_j) I_j + wE(\Omega, I_j, \alpha_j) \Big] + m\alpha_j + \phi_j \tag{6}$$

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

$$log(s_j) - log(s_0) = \xi(\Omega, I_j, \alpha_j) - \rho(1+\tau)p_j + \beta\alpha_j + \nu_j \tag{7}$$

I interact firm-level covariates with Equations 6 and 7. The covariates we interact must be relevant to the determination of prices and market shares but also must be independent of ν, ϕ. Firm organizational complexity (I) and task-mix vector (α) fit these requirements. Under the model, they change organizational costs but are not impacted by ν, ϕ. Additionally, I use their interaction, $\alpha \cdot I$. Average labor, $\bar{a}$, is also observable and relevant to prices and quantities. A discussion of how this variation identifies specific parameters is provided in Section 7.6.

To these 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. For each county and quarter, I take the total quarterly wages of establishments with NAICS code 812112 and divide 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. Assuming some measurement error in wages denoted e_j, we have that:

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

The identifying moment conditions can then be written as:

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

For a single market and quarter, this process yields a total of 27 moments to estimate 25 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. Denote the sample analogue of the moments as $G(\Omega)$. Then to estimate Ω, we solve:

$$\arg\min_{\hat{\Omega}} G(\hat{\Omega})' W G(\hat{\Omega}) \tag{8}$$

where W is a weighting matrix. Note that to evaluate this GMM objective, we must recover the vector of organization cost simplified 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.

For the hair salon application 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 effectively standardizes the moments in the objective, preventing random variables with large nominal values (like 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 search algorithm only search over parameters values where a positive share of each type of labor is demanded.

7.5 A Computationally Light Estimation Procedure

Because there is no closed-form for the mapping from organization complexities and market parameters to organization costs, 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 provides a nested-fixed point algorithm which efficiently solves the firm's problem and is proven to globally converge.

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

Lemma 2 *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.[19] 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 is a fixed-point algorithm which iterates on two optimality conditions (Blahut 1972) and can be described as follows, where I suppress firm subscripts:

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

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

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

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.

This algorithm is proven to globally 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 contraction mapping, a researcher can use acceleration algorithms, like Du and Varadhan (2020) to increase the speed of convergence. It 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 to do this. The full estimation procedure consists of the following steps.

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 8. If the objective is minimized stop, otherwise return to step 1 with a new market parameter guess, $\hat{\Omega}$.

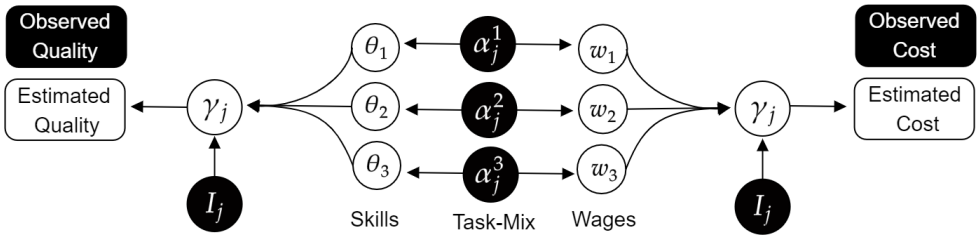
The logic of this estimation algorithm is similar in spirit to the demand estimation procedures that have become popular since Berry, Levinsohn, and Pakes (1995). Just as those prcoedures invert market shares using a contraction mapping to derive unobserved product qualities, my procedure inverts organization complexities using a contraction mapping to obtain unobserved organization costs.

7.6 Identifying Variation

Proposition 5 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 correlation of cost differences across firms with prices: high cost pass through implies a low price sensitivity. For the rest of this informal discussion, I will talk about firm costs and qualities as if they are observed. This is because once ρ is known, the marginal cost of each firm can be obtained by subtracting the markup from prices. Similarly, unobserved quality can be obtained from subtracting the disutility of prices from a firm's log relative market share. To understand the variation driving identification of the other market parameters, it is helpful to understand how these parameters are linked to observed variables and each other. The task-mix is linked to cost via wages and to qualities via skills. The relationships are mediated by complexity (I_j). I visualize this in Figure 8.

Figure 8: Identifying Variation for Skills and Wages



Note: The diagram visualizes the way market parameters (skill sets and wages) are identified from observable data. Black objects are observed, while white objects are parameters to be estimated or unobserved variables like organization costs. Organization costs mediate the relationship between particular skill and wage parameters and costs and qualities. Since organization costs are essentially deterministic functions of complexity, we can use variation in the interaction of complexity and the task-mix to identify the different parameters.

Recall that firms expend complexity to increase task-specialization. When a firm has high complexity and is intense in task k, the firm is employing a large share of task k specialists. A large share of its quality is driven by the skill parameter θ_k and a large share of its cost is driven by specialist k wages w_k. In this way, variation across firms in task-mixtures holding fixed complexity identifies the skills and wages of different specialists.

8 Empirical Results

The model is estimated for Manhattan during Quarter 2 of 2021. This section summarizes parameter estimates and uses the model to analyze the sources of variation in task content for hair stylists.

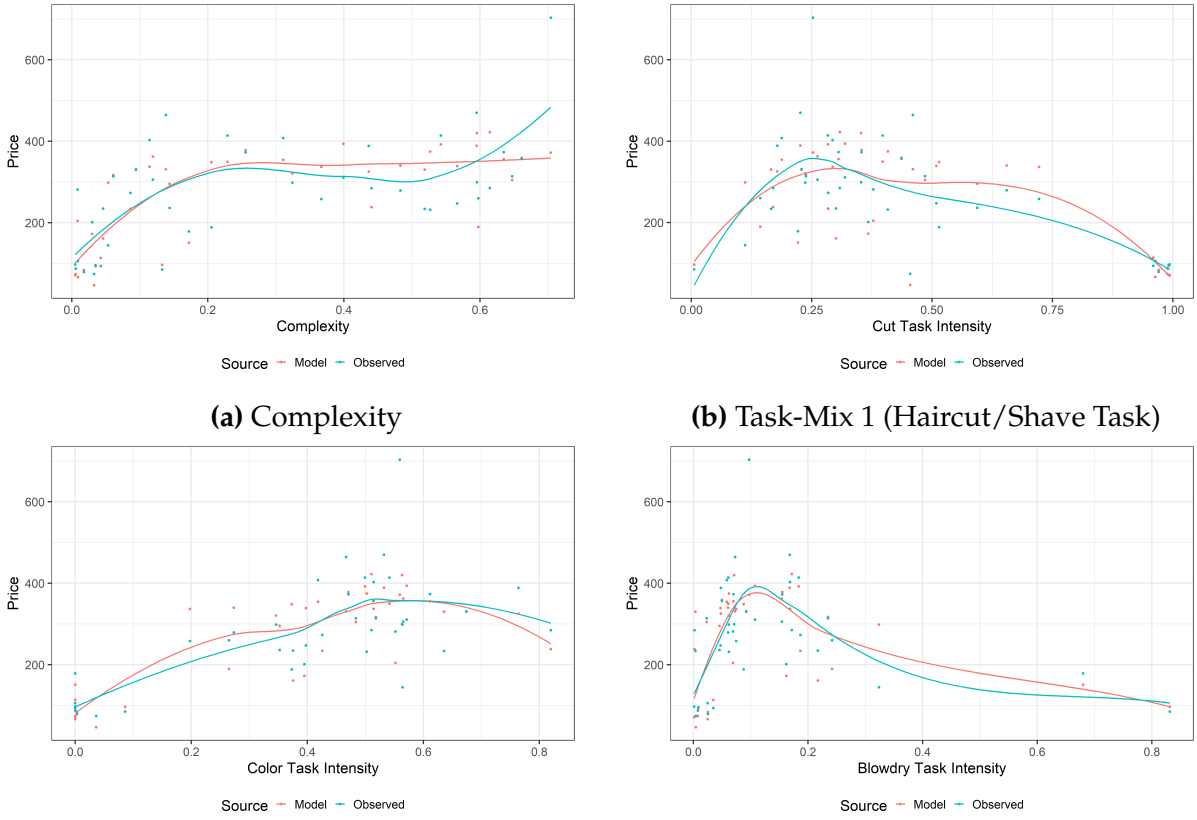
The market parameters of interest are the type-specific wages, specialty skill levels and consumer price sensitivity. These are provided in Table 6. The estimated consumer price sensitivity implies an average own-price elasticity of 11. Excluding the miscellaneous cat-

egory, haircut/shave specialists have the lowest wage while color/highlight/wash specialists have the highest wage. The specialty skill premium is highest for blow dry/style/treatment specialists, meaning that the largest increases in quality come from hiring a specific worker to do these tasks.

8.1 Model Fit and Validation

I assess model fit by comparing the predicted and actual relationship between prices and various organizational variables in Figure 9. I plot smoothed Loess lines to assess the general shape of both the predicted and observed data. The model captures the shape of the relationships.

Figure 9: 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. The model captures the general 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 b_j, and it is a matrix where element i, k denotes the time worker type i spends on task k. Using the model, we can compute b_j for each firm among worker types that it hires. In the data, we 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, we can weight each job by the total amount of labor it represents. Once we combine all J firms, this yields an unobserved and model based distribution of job task content for each of the 5 tasks, where jobs are weighted by their effective labor.

Tables 7, 8 , and 9 compare the model and observed mean, median, variance, 25th percentile and 75th percentile of job task-content. The model exactly matches 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 model decently approximates the total variances of task content and the 75th percentile of the job task-content distribution. It poorly approximates the median and the 25th percentile.

The statistics related to the color/highlight/wash task are the hardest for the model to replicate. The reason for this is that the empirical distribution for this task is triple peaked. The empirical and model generated task-content distributions are presented in full in Appendix Table 22.[20]

8.2 The Determinants of Job Task Content

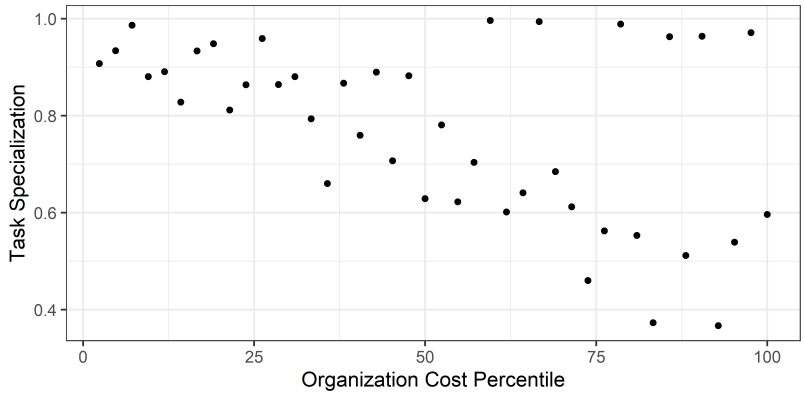
I also analyze how much of the variance in task-specialization is attributable to firm. I measure task-specialization as the amount of time a worker spends on their specialty task. Performing the same decomposition as before, I find that 45% of the variation in task-specialization is attributable to firms.

One problem with the variance decompositions provided so far is that they do not distinguish between job differences due to organization costs or job differences due to task-mix differences across firms. Put another way, it could be that firms intense in coloring hire color specialists, and this is the reason so much of job task content is attributable to firms. To distinguish between these two channels, I compute the average task-specialization at each firm as the fraction of total firm labor spent on any worker's

20. Comparing the entire distribution of content is another, much more rigorous validation test.

specialty. I plot this measure for each firm against its organization cost percentile in Figure 10.

Figure 10: Task-Specialization and Firm Organization Costs

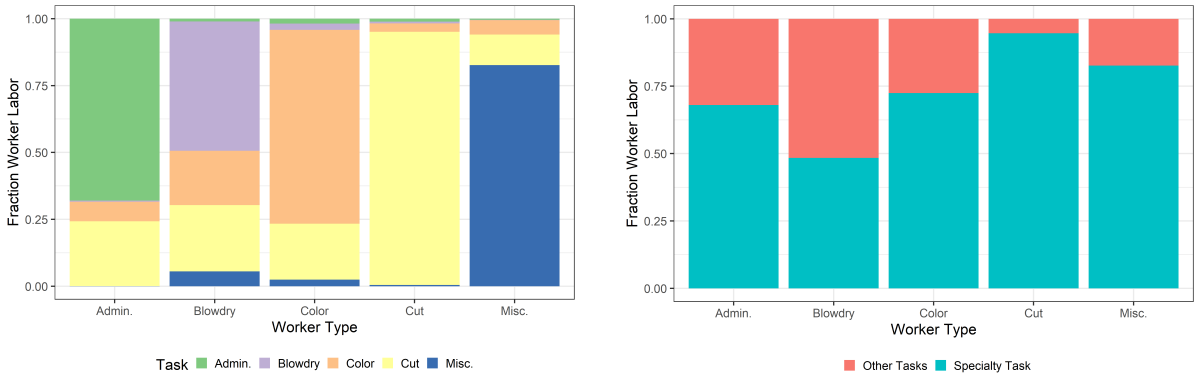


Note: The figure displays the model predicted relationship between task-specialization (fraction of time spent by workers on there focal task) and firm organization cost. Firms with higher organization utilize less task-specialized internal organization structures.

The negative relationship is consistent with the theoretical analysis earlier in the paper. Because complexity generally rises with specialization, firms with higher organization costs exhibit less task-specialization. The magnitude 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) assign on average 67%.

Finally, I plot the equilibrium division of tasks across worker types in Figure 11. Haircut/shave specialists work the most specialized jobs, while blow dry/extension/style specialists work the most generalized jobs. These differences are driven in part by the estimated skill premia of each type as well as the distribution of task-mixtures and organization costs across firms.

Figure 11: Equilibrium Division of Tasks Across Workers



(a) All Tasks **(b)** Specialty v.s. Other Tasks

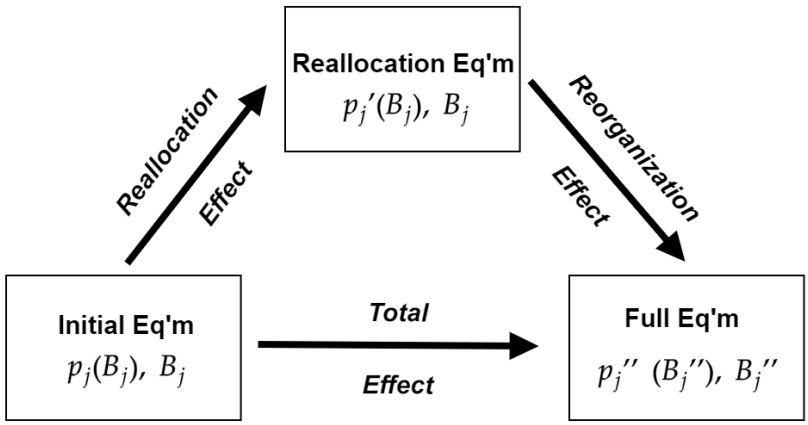
Note: Panel A provides the breakdown of time spent on each task by worker type. Panel B provides the breakdown of time spent on specialty task vs. all other tasks. Haircut/shave specialists are the most task-specialized, while blow dry specialists are the least.

9 Counterfactuals

This section uses the estimated model to study two counterfactual policy changes, one impacting the product market and one impacting the labor market. The counterfactuals illustrate how internal organization mediates shocks in the product and labor market. The procedure used to solve for equilibria and conduct the analyses is described in Appendix Section B.6. I use the structure of the model to distinguish between two effects of any policy: a reallocation effect and a reorganization effect. To do this, I first define the reallocation equilibrium.

The reallocation equilibrium 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 is the change in outcomes between the reallocation equilibrium and the initial equilibrium. It captures changes due to 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 as the change in outcomes from the initial to the full equilibrium. These relationships are summarized in Figure 12.

Figure 12: Reallocation, Reorganization and Total Effect



Note: The diagram shows how the total effect of any counterfactual policy can be decomposed into two parts. The first part fixes the organization structure but allows firms to update prices. This is the reallocation effect, because labor flows across firms but does not change within firms. The second part allows organization structures to adjust. This is the reorganization effect, because firms are now responding to wages and prices by shifting their internal structure.

9.1 Minimum Wage Increase

Two pieces of potential legislation in New York state propose pegging the minimum wage to inflation, and an analysis by the Economic Policy Institute projects this will result in a minimum wage of $21.25 in 2026 for New York City.[21] Because my model is estimated for a single industry, it is not possible to study this policy directly. Instead, I study a counterfactual increase in the minimum wage in 2021 from $15 to $21.25.

To implement the counterfactual, I require that all equilibrium wages be at least $21.25, 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. I believe my model is well-suited for studying large increases in the minimum wage, because I allow 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 discuss these in Appendix Section B.6.2.

The unique equilibrium both with and without adjustment features a binding minimum wage for the haircut/shave specialist only. The wages and employment levels across worker types are given in Table 12. Table 11 summarizes the the impacts of the

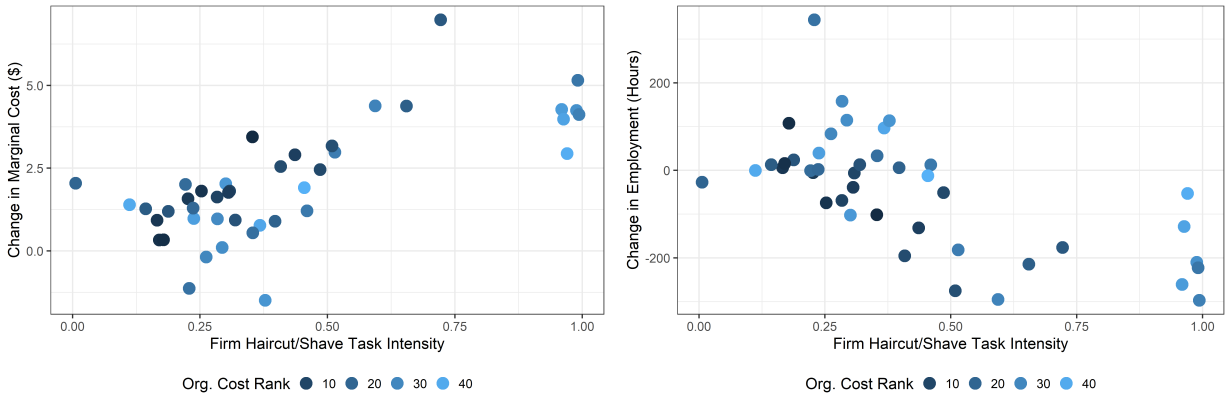
21. Senate Bill S3062C and Assembly Bill A7503B. Analysis by EPI: *Proposed New York state minimum wage increases would lift wages for more than 2 million workers through 2026: Minimum wages would range by region from $16.35 to $21.25 per hour by 2026 — Economic Policy Institute*

policy.

9.1.1 The Reallocation Effect

First, there is the reallocation of labor across firms with different task-mixtures and organization costs. This is the only force operating in the equilibrium without reorganization. It exists because firms are heterogeneous in both their task-mixtures and organization costs. Most of the reductions in consumer welfare, firm profit, wages and employment are due to this effect. I visualize this effect in Figure 13, by looking at changes in employment and marginal cost across firms with different organization costs (γ_j) and task-mixtures in an equilibrium without reorganization. I focus on variation in intensity in the haircut/shave task, because they are the only workers for which the minimum wage binds.

Figure 13: Reallocation Effect Under a Minimum Wage Increase



(a) Marginal Cost Changes **(b)** Employment Changes

Note: In both panels each dot is a firm, with the colors representing the firm's organization cost rank relative to other firms. Darker dots are more efficient firms. Panel A plots the change in marginal cost under the reallocation effect against the firm's intensity in the binding task. Marginal cost increases are highest for firms intense in the binding task that are organizationally inefficient. Panel B plots the change in employment under the reallocation effect against the firm's intensity in the binding task. Employment losses are largest for firms intense in the binding task that are organizationally inefficient.

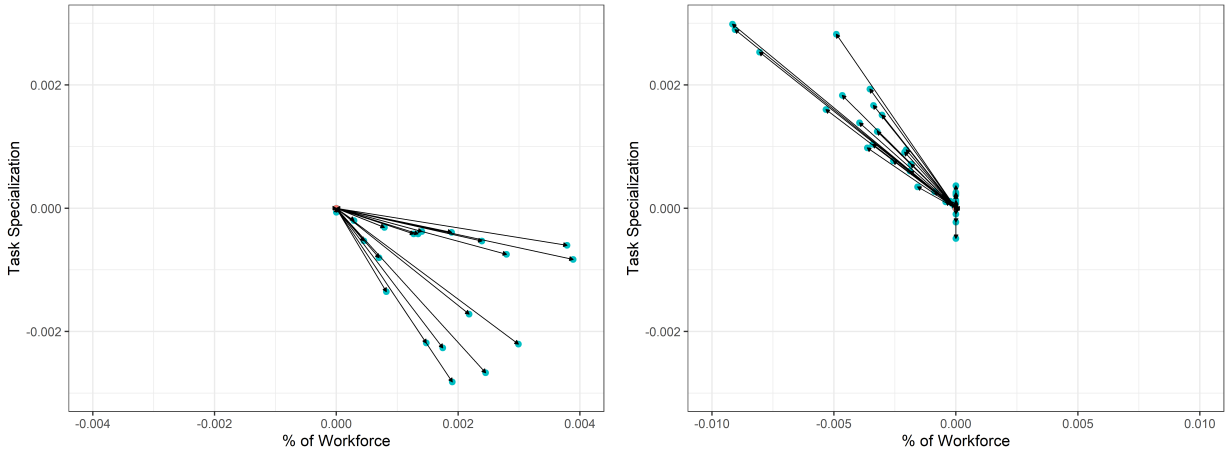
The incidence of the minimum wage is highest for firms that are intense in the haircut/shave task and that have high organization costs. These firms see the largest increases in marginal costs and the largest decreases in employment. Surprisingly, firms which have low organization costs and which are not intense in the haircut/shave task increase employment in response to the minimum wage. Because the minimum wage increases some firm's costs more than others, it changes the competitive position of firms in the product market.

The reallocation effect has different wage consequences for different types of non-binding workers. Table 12 shows that wages decrease for color specialists, but increase for the other types. The magnitudes and signs of the changes depend on the organization costs and task-mixtures of firms in the market.

9.1.2 The Reorganization Effect

Second, there is a reorganization effect, where labor is reallocated within the firm. This force operates only in the full equilibrium and it amplifies the negative consequences of a minimum wage, lowering consumer welfare, employment and profits. It exists because internal organization is allowed to be strategic in the model. To study it, I plot the change in task specialization (fraction of a worker's time spent on their specialty) against the change in the fraction of a firm's workforce a worker type represents. Individual firms are represented by arrows.

Figure 14: 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.

For binding types, salons reduce relative employment and increase task specialization. Intuitively, the minimum wage makes binding workers more costly. Firms layoff binding worker types unless they are performing enough of their specialty task. For non-binding types, salons reduce task specialization and increase relative employment. I call this a "pick-up-the-slack" effect. Intuitively, non-binding workers become relatively less

costly. The firm asks these workers to perform some of the tasks previously performed by the other types. The minimum wage spills-over onto non-binding workers via internal reorganization.

The pick-up-the-slack effect implies task-specialization and wages move in opposite directions. Looking at Table 12, worker types that see a wage decrease (or unemployment increase) from reorganization also see a task-specialization increase. Likewise, worker types that see a wage increase also see a task-specialization decrease.

The reorganization effect tends to moderate the wage changes caused by the reallocation effect. For instance, the reallocation effect increases blow dry specialist wages from \$20.91 to \$22.27. The reallocation effect reduces blow dry specialist wages to \$21.93. This occurs because firms adjust their internal structure so that the marginal costs and benefits of different worker types are equalized.

9.1.3 The Total Effect

The minimum wage increases total welfare by 1.32%. The reason is that total wage increases for employed workers are greater than firm profit losses, consumer welfare losses, and wage losses of unemployed workers.

One interesting aspect is that the reorganization effect makes almost everyone worse off: it lowers firm profit, lowers consumer welfare, lowers employment and lowers wages of employed workers. This is due to a reorganization externality.

Consider why the initial organization structures cannot be sustained after the minimum wage increase. When the minimum wages rises to \$21.25, firms have an incentive to reallocate tasks from the workers for which the wage is binding to workers for which the wage is not binding. We can show formally that such a temptation exists under general conditions, but intuitively, the firm can slightly adjust its structure to use less of the binding workers and more of the non-binding. For small enough adjustments, quality and organization costs remain the same, but the firm's own wage bill strictly falls. Firms do not internalize the fact that this adjustment lowers total labor market demand for binding types and raises total labor market demand for non-binding. To clear the labor market, non-binding wages must rise and binding wages must decrease. The minimum wage prevents binding wages from decreasing further, so unemployment rises.

9.2 Sales Taxes

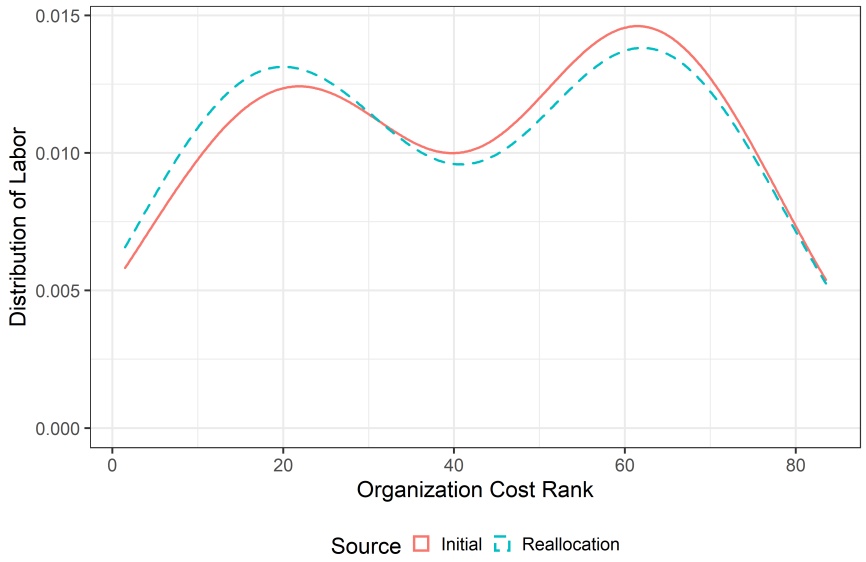
New York City is unique in that it levies a 4.5% sales tax on hair salon services. This section studies the effect of eliminating this sales tax. Formally, I estimate a new equilibrium

with $\tau^{NEW} = 0$. Table 13 summarizes the impacts of the policy.

9.2.1 Reallocation Effect

Eliminating the sales tax confers a competitive advantage to firms producing high quality services in the initial equilibrium. Since salons with low organization costs tend to produce high quality services, this reallocates labor towards organizationally efficient firms, as seen in Figure 15.

Figure 15: Reallocation Effect Under a Sales Tax



Note: The figure plots the smoothed distribution of labor across firms ranked by organization cost. The reallocation effect under a sales tax shifts labor towards lower organization costs (more efficient) firms.

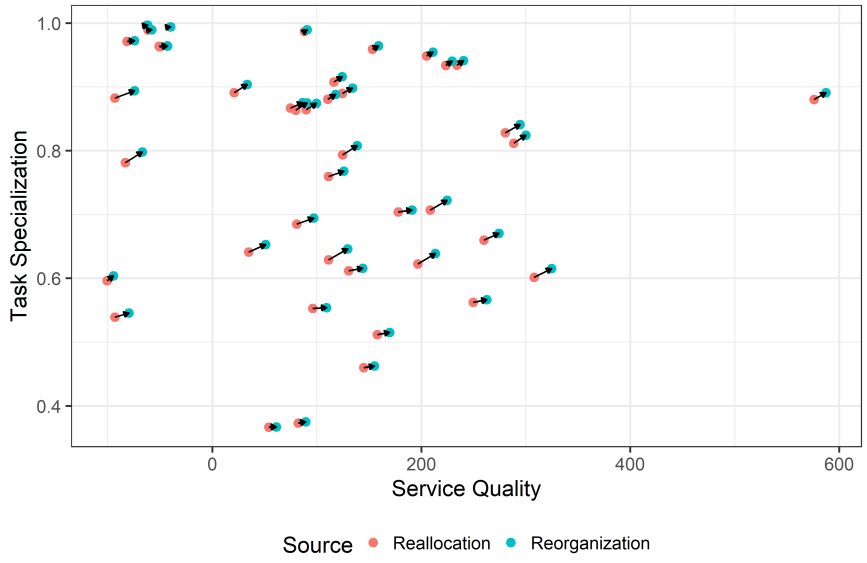
Organizationally efficient salons produce high quality services using a more task-specialized internal structure. Thus the reallocation effect increases market-wide task specialization because more workers are working at task-specialized firms. The reallocation effect accounts for about half of the total increase in task-specialization due to eliminating the sales tax.

9.2.2 Reorganization Effect

Eliminating the sales tax makes producing higher quality product more attractive. In order to produce higher quality products, firms choose internal organizations which are on average 5.5% more complex in order to increase average labor market task specialization by 0.9%. Average firm service quality rises by 10%. This is consistent with the quality-complexity-wage trilemma discussion in the theoretical section.

Figure 16 illustrates that these market-wide patterns also happen at the firm-level. However, the extent to which firms increase quality and increase task specialization depends on the firm's internal organization costs and its particular task-mix. This is why the slopes and lengths of the arrows in Figure 16 differ.

Figure 16: 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 firms 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 particiular organization costs and task-mixture.

The reorganization effect accounts for half of the total increase in labor market task-specialization.

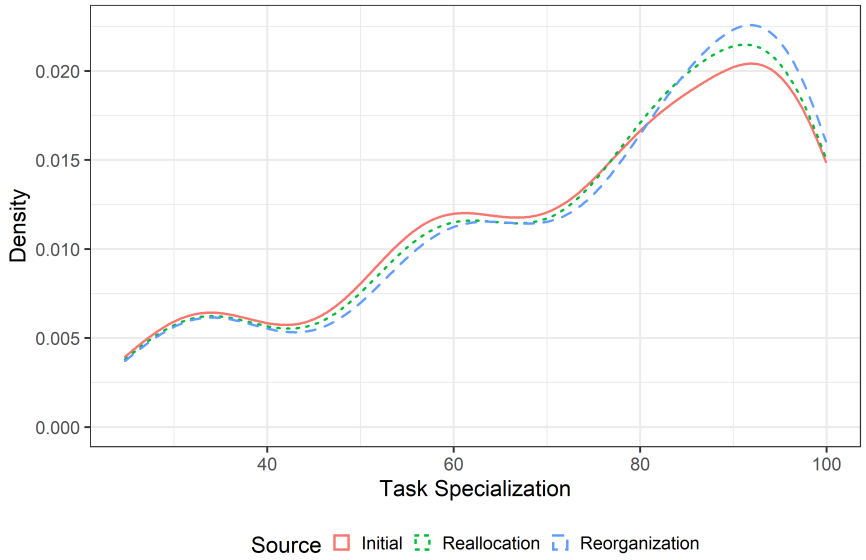
9.2.3 Total Effect

Overall, eliminating the sales tax increases total welfare by 13%. Breaking this down by different actors in the model, it leaves consumer welfare mostly unchanged, while slightly increasing firm profit (+3.7%). The majority of gains accrue to workers through wages which rise by 19.3%. The reason is that eliminating the sales privileges service quality by making consumers less price sensitive.

Figure 17 shows that eliminating the sales tax generates a reallocation and reorganization effect which move task-specialization in the same direction. Labor shifts towards organizationally efficient firms who were already producing high quality products (the reallocation effect). Most firms also increase quality and task specialization (the reorganization effect). Workers are more productive because they are spending more of their time

on their specialty task. Firms capture the surplus from increased quality from consumers via higher prices, and workers capture the surplus from firms via higher wages.

Figure 17: The Task Specialization of Jobs Before and After



Note: The figure plots the distribution of labor across jobs of different task-specialization (amount of time spent on specialty task). Eliminating the sales tax increases the amount of time workers spend on their specialty task, with both the reorganization and reallocation effects contributing equally to this total change.

10 Extensions

Many of the modeling assumptions are made solely for 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.

10.1 Labor Market Power

The current model focuses on situations where firms have product market power but not labor market power. This is realistic when the product market is small relative to the labor market, either geographically or due to worker's ability to work in multiple industries. In many situations this is not realistic and we might expect firms to also hold labor market power.

Introducing labor market power raises an interesting theoretical question, which is why I think this is 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 markdown wages. How does this incentive interact with internal organization, and how does it change competition? Unlike a competitive labor market, firms 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 pigeon-holing for their long term career goals. 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 suggests such talent wars occur. My model provides a way to study the trade-off between over-hiring a certain type of worker while still 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 now impacts marginal costs (through wage markdowns) the characterization in Theorem 1 will no longer hold. We will need new tools to solve and estimate the model. Because the new problem will be a non-linear rational inattention problem, results from Jung et al. (2019) may be helpful in this regard.

10.2 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 $D = -\xi$. Finally, assume all workers have the same wages (skills are not priced by the market).

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:

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

To interpret this, notice 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 their specialty with variance $\sigma^2 - ln(2)\gamma\rho$. Greater organization costs or higher price sensitivity reduce task-specialization. This illustrates two things. First, the model can be easily extended to accommodate the large firm case, where the task space is uncountable (as is the worker type space). Second, the key organizational friction is a deep property of the model, and it does not go away when we make organizations large and less "lumpy."

10.3 A Quantity-Based Model

In some contexts, like manufacturing, we 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: we 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:

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

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:

$$\sum_i W_i \sum_k \alpha_k \sum_i \theta_{i,k} B_{j,i,k}$$

And 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 then come through a reduction in the per-unit wage bill. The logic of 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. Some readers may wonder why I did not use this as the main model. The reason is because the quantity productivity model has the following property.

Proposition 6 *Under the quantity model, prices are decreasing with organizational complexity.*

The proof is provided in Appendix Section A.4. 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 property may be desirable in some settings, like manufacturing, but Stylized Fact 2 shows it is not true for hair salons.

10.4 Non-Additive Quality

The model developed in this paper required that the effect of the quality of each individual task has 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 be accommodate this within the model using multiplicative quality, similar in spirit to Kremer (1993):

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

We can re-write using logarithms as:

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

This is now an f-separable distortion measure, meaning we 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.

10.5 Quality Positioning and Richer Demand Systems

One surprising result from the theoretical section is that the choice of organization structure depends only on other firm choices via wages. The main reason for this is the independence of irrelevant alternatives property of logit demand. In my setting, this property

manifests as consumers substituting uniformly to other products regardless of quality. In some contexts this may be unrealistic, and we 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 give up analytical tractability and greatly increase computational requirements.

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 we specify that $\epsilon_{z,j}$ is type-1 extreme value and q_z is normally distributed it 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, we 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 generalization is tractability. Because of the dependence on the quality choices of other firms, the characterization in Theorem 1 no longer holds. Although the problem is still likely a form of rational inattention problem, it is not a rate-distortion problem. We also dramatically increase the computational complexity of the model. Without equivalence to a rate-distortion problem, estimation requires solving the model for each firm using nonlinear convex optimization, or deriving a new fixed point algorithm. Also, since demand no longer takes a multinomial logit form, there does not exist a closed-form solution relating market shares, prices and unobserved qualities. Inverting market shares to obtain unobserved quality now requires numerical integration and solving a contraction mapping.

11 Conclusion

This paper provides a structural model that can be used to study the effect of internal organization on product and labor market outcomes. The model allows firms to respond to changes in policy by internally reorganization. It also allows firms to differ in their internal organization costs. These features allow the model to capture the strategic and distributional consequences of policies for firms.

An area for future work is the welfare impact of internal organization on workers. This paper shows that both product and labor market policies reallocate labor across firms and change the task-specialization of jobs in an industry. But are such changes beneficial to workers?

This is ultimately an empirical question. On the one hand, greater task specialization could deepen worker experience in certain tasks, improving the production possibilities when workers bring their skills together. Greater task specialization also means workers are less exchangeable with other workers at the same firm. On the other hand, greater task specialization may limit task exploration. It may also make jobs more skill-specific, making it more difficult for a worker to find a new job and limiting a worker's outside option. Workers may also intrinsically value broader jobs more than specialized jobs.

The relative importance of these forces depend crucially on the size of job search frictions, the nature of task-based human capital accumulation, and the amenity value of specialization. Thus, understanding the effect on worker welfare requires extending the model and utilizing better data on worker outcomes across firms.

References

Adenbaum, Jacob. 2021. "Endogenous Firm Structure and Worker Specialization."

Aghion, Phillipe, Nicholas Bloom, and John Van Reenen. 2014. "Incomplete Contracts and the Internal Organization of Firms." *The Journal of Law, Economics, and Organization* 30, no. suppl_1 (May): i37–i63.

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.

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.

Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and Explaining Management Practices Across Firms and Countries." *The Quarterly Journal of Economics* 122, no. 4 (November): 1351–1408.

———. 2011. "Human Resource Management and Productivity." *Handbook of Labor Economics* 4, no. PART B (January): 1697–1767.

Caplin, Andrew, and Mark Dean. 2013. "Behavioral Implications of Rational Inattention with Shannon Entropy" (August).

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.

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

Coase, R. H. 1937. "The Nature of the Firm." *Economica* 4, no. 16 (November): 386–405.

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

Dessein, Wouter, Tano Santos, Gary Becker, Mathias Dewatripont, Jonathan Levin, Kevin Murphy, Canice Prendergast, and John Roberts. 2003. "Adaptive Organizations *."

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.

Garicano, Luis, and Paul Heaton. 2010. "Information technology, organization, and productivity in the public sector: Evidence from police departments." *Journal of Labor Economics* 28, no. 1 (January): 167–201.

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.

Giorcelli, Michela. 2019. "The Long-Term Effects of Management and Technology Transfers." *American Economic Review* 109, no. 1 (January): 121–52.

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

———. 2020b. "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.

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.

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

Lucas, Robert E. 1978. "On the Size Distribution of Business Firms." *The Bell Journal of Economics* 9, no. 2 (23): 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.

New York City Department of Finance. 2022. *New York State Sales and Use Tax*, May.

Ocampo, Sergio. 2018. "A Task-Based Theory of Occupations with Multidimensional Heterogeneity * Click Here for Most Current Version."

Proposed New York state minimum wage increases would lift wages for more than 2 million workers through 2026: Minimum wages would range by region from $16.35 to $21.25 per hour by 2026 — Economic Policy Institute.

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.

Tax Guide for Barbers and Beauty Shops. 2022, May.

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

Yamaguchi, Shintaro. 2015. "Tasks and Heterogeneous Human Capital." *Journal of Labor Economics* 30, no. 1 (July): 1–53.

Tables

Table 1: Variance Decomposition: Without a Model

		Across Firms				Across Quarters	
		Share of Task-Content Variance				Share of Task-Content Variance	
Task	Share of Labor	Firm	Within-Firm	Task	Share of Labor	Quarter	Within-Quarter
1	0.4049	0.3744	0.6256	1	0.4049	0.0057	0.9943
2	0.3902	0.2899	0.7101	2	0.3902	0.0062	0.9938
3	0.0850	0.5056	0.4944	3	0.0850	0.0111	0.9889
4	0.0590	0.4900	0.5100	4	0.0590	0.0193	0.9807
5	0.0610	0.4124	0.5876	5	0.0610	0.0118	0.9882

Table 2: Fixed Effect Regressions of Firm 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

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

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

Table 4: Firm-Quarter Characteristics

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
Organization Complexity	4,558	0.22	0.22	0.00	0.03	0.42	1.02
Share Haircut/Shave	4,558	0.41	0.23	0.00	0.26	0.52	1.00
Share Color/Highlight/Wash/Extensions	4,558	0.38	0.20	0.00	0.29	0.52	1.00
Share Blowdry/Style/Treatment	4,558	0.09	0.12	0.00	0.03	0.11	1.00
Admininstrative	4,558	0.05	0.11	0.00	0.002	0.04	1.00
Nail/Spa/Eye/Misc.	4,558	0.06	0.16	0.00	0.00	0.05	1.00

Table 5: Select Parameter Estimates

Price Sensitivity and Worker Skill Levels

Parameter	Symbol	Estimate
Price Sensitivity	ρ	0.03832
Haircut/Shave Skill	S_1	32.45
Color/Highlight/Wash Skill	S_2	60.03
Blowdry/Style/Treatment/Extension Skill	S_3	141.7
Administrative Skill	S_4	43.29
Nail/Spa/Eye/Misc. Skill	S_5	66.48

Type-Specific Wages

Type	Estimate
Haircut/Shave Specialist	16.96
Color/Highlight/Wash Specialist	37.75
Blowdry/Style/Treatment/Extension Specialist	20.91
Administrative Specialist	26.99
Nail/Spa/Eye/Misc. Specialist	81.16

Table 6: Select Parameter Estimates

	Estimate
Price Sensitivity	0.038
	(0.025)
Haircut/Shave Wage	16.957
	(260.027)
Color/Highlight/Wash Wage	37.751
	(211.503)
Blowdry/Style/Treatment/Extensions Wage	20.912
	(412.630)
Administrative Wage	26.989
	(2,659.288)
Nail/Spa/Eye/Misc. Wage	81.156
	(1,960.048)
Haircut/Shave Skill	32.451
	(31.468)
Color/Highlight/Wash Skill	60.030
	(100.718)
Blowdry/Style/Treatment/Extensions Skill	141.693
	(235.628)
Administrative Skill	43.291
	(123.579)
Nail/Spa/Eye/Misc. Skill	66.480
	(338.955)

Table 7: Model vs. Observed Job Task Mean and Median

	Mean		Median	
Task	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

Table 8: Model vs. Observed Job Task Variance

	Total Variance		Between Firm Variance	
Task	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

Table 9: Model vs. Observed Job Task IQR

	p25		p75	
Task	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

Table 10: Model-Based Decomposition of Job Task-Content Variance

	Share of Task-Content Variance	
Task	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

Table 11: Minimum Wage Increase from $15 to $21.25

Statistic	Reallocation	Reorganization	Total
Avg. Price	2.75%	-0.41%	2.34%
Avg. Complexity	-0.00%	-0.65%	-0.65%
Avg. Quality	0.00%	-0.77%	-0.77%
Avg. Hourly Wage	4.71%	0.28%	4.99%
Std. Dev. Wage	-12.26%	1.47%	-10.79%
Task Specialization	-0.87%	-0.25%	-1.12%
Employment	-2.14%	-0.27%	-2.41%
Market Served	-3.77%	-0.17%	-3.94%
Total Profit	-3.77%	-0.17%	-3.94%
Consumer Welfare	-0.86%	-0.04%	-0.90%
Total Wages	2.47%	-0.01%	2.46%
Total Welfare	1.33%	-0.02%	1.32%

Table 12: Minimum Wage 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	16.957	0.94634	493530	21.25	0.94575	488040	21.25	0.9473
Color/Etc	997050	37.751	0.72452	997050	37.15	0.72272	997050	37.427	0.71939
Blowdry/Etc	444040	20.912	0.48366	444040	22.268	0.48091	444040	21.928	0.48112
Administrative	41860	26.989	0.68013	41860	28.959	0.68097	41860	28.56	0.68126
Nail/Etc	34844	81.156	0.82621	34844	81.804	0.82589	34844	81.923	0.82586

Table 13: Elimination of the 4.5% Sales Tax on Salon Services

Statistic	Reallocation	Reorganization	Total
Avg. Price	0.19%	3.82%	4.00%
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	-0.04%	-0.14%	-0.18%
Total Wages	18.32%	1.02%	19.34%
Total Welfare	12.33%	0.63%	12.96%

A Theoretical Appendix

A.1 Proof of Lemma 1

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). We can re-write the above as:

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

Re-write again:

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

Multiply by ρ and subtract ξ_j:

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

Now denote:

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

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

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

Denote:

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

Then the expression becomes:

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

Everything on the right is exogenous, and the left is the form required by Lambert's W, so we have that the $\tilde{W}$ which solves is given by Lambert's W function of the right hand side by definition. Thus optimal price solves:

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

A property of this function is that $log(W(x)) = log(x) - W(x)$. Using this we have that:

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

Solving for the optimal price:

$$\frac{1}{\rho} + c_j + \rho^{-1} W\bigg(exp\bigg(-1 + \xi_j - \rho c_j \bigg) E_j^{-1}\bigg) = p_j^* \tag{9}$$

This is a closed-form expression for price, because E_j does not depend on p_j. Plugging this into profit gives:

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

Note that profit depends only on the sum of quality less adjusted marginal cost, $\xi_j - \rho c_j$. Denote this v_j, and re-write:

$$\pi_j = \frac{\rho^{-1} exp(-1 + v_j - W(exp(-1 + v_j) E_j^{-1})}{exp(-1 + v_j - W(exp(-1 + v_j) E_j^{-1}) + E_j} (1 + W(exp(-1 + v_j) E_j^{-1})$$

Profit consists of two positive terms. The second term is known to be monotone increasing in v_j because the Lambert's W function is monotone increasing. The first term is essentially $v_j - W(exp(-1 + v_j) E_j^{-1})$ composed within the logistic function, a strictly increasing function. Thus if $v_j - W(exp(-1 + v_j) E_j^{-1})$ is increasing in v_j profit is increasing in v_j. To

do this, we appeal to the well-known result that the derivative of W can be written as:

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

Using this, we have that:

$$\frac{\partial v_j - W(exp(-1+v_j)E_j^{-1})}{\partial v_j} = 1 - W'(exp(-1+v_j)E_j^{-1})exp(-1+v_j)E_j^{-1}$$

$$= 1 - \frac{W(exp(-1+v_j)E_j^{-1})}{exp(-1+v_j)E_j^{-1}(1+W(exp(-1+v_j)E_j^{-1}))}exp(-1+v_j)E_j^{-1}$$

$$= 1 - \frac{W(exp(-1+v_j)E_j^{-1})}{1+W(exp(-1+v_j)E_j^{-1})}$$

W is postive over the positive real numbers, so this derivative is non-negative. It is also strictly positive. To see why note that W can only be 0 if $exp(-1+v_j)E_j^{-1})$ is 0. This cannot be because the space of qualities is bounded for all firms: the maximum achievable quality is θ_H (this occurs when the firm hires only high type workers and assigns them perfectly to their respective specialties) and the minimum price for all firms is 0, so E is never infinite and $exp(-1+v_j)$ is never 0. Thus the second term is strictly less than one, so the overall derivative is strictly positive. Therefore profit is strictly increasing in cost-adjusted quality, v_j. We can relate v_j to quality-adjusted marginal cost in the following way:

$$MC_j - \rho^{-1}\xi_j = -\rho v_j$$

Therefore profit with optimal pricing is strictly decreasing in quality-adjusted marginal cost.

A.2 Proof of Proposition 1 and 2

We have already shown in Theorem 1 that optimal B also solves a rate-distortion problem. In this problem, distortion is adjusted wages less quality. Considering the optimal organization complexity (I) for each quality-adjusted wages (denote it D), we know from information theory that this function is continuous, convex and strictly decreasing in D above some upper bound (Chen, n.d.). At some threshold $\bar{D}$ $I(D)$ becomes 0. Suppose we are below this threshold. Then we have that we can write the firm's problem as solving:

$$\min_D \gamma I(D) + D$$

The first-order condition of this problem is:

$$\gamma I'(D) + 1 = 0 \implies I'(D) = -\gamma^{-1}$$

We notice that because $I(D)$ is a rate-distortion function it is decreasing and convex. Thus the D^* which solves the FOC will be increasing in γ, as depicted in Figure 6. Since D^* is increasing in γ, I will be decreasing in γ, because I is strictly decreasing in D because it is a rate-distortion function.

For profit and output, apply the envelope theorem to $\gamma I(D) + D$:

$$\frac{\partial}{\partial \gamma} \gamma I(D^*) + D^* = I(D^*)$$

This is positive whenever D^* is below some threshold or equivalently whenever complexity is positive. Profit and market share are both decreasing functions of this quantity. Thus profit and output are decreasing in γ.

A.3 Proof of Proposition 3

For the purposes of this proof only, we 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, we have that:

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

Therefore:

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

Therefore:

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

Substituting the value for λ:

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

By the definition fo $h_{i,k}$, we have that:

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

To get to jobs, we 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. We 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). Uniqueness comes from assumption 1, which allows me to apply the result from Caplin and Dean (2013). An increase in the wage corresponds to a decrease in the "pay-off" to the firm of using workers of type i in all tasks (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. We can even say this decrease is strict whenever initially the share of worker i is strictly interior, i.e. $0 < e_i < 1$.

A.4 Proof of Proposition 6

Equation 9 from an earlier proof provides a closed-form expression for price in any Nash Equilibrium under logit demand:

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

Taking the derivative w.r.t. c_j:

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

A property of the LambertW function is that:

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

Thus we can simplify the expression to be:

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

The LambertW function is weakly positive for values which are weakly positive, therefore the derivative is positive, and we have that 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 optimal wage-bill as $D := W(B_j^*)$, we can reformulate the problem as before, as the firm choosing D given some optimal organizaiton cost and wage bill:

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

where we have re-written I and W to be functions of D instead of B_j. Then as before we have that the cross-partial derivative is negative:

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

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 we have that prices should be decreasing as γ decreases, while complexity should be increasing.

A.5 Proof of Proposition 5

As the text states, the proof will consist of three parts. Throughout the firm index j is suppressed. First, we prove that observed organization complexity based on worker identities is equal to unobserved true complexity based on worker skill sets. Consider the augmented model proposed in Section 7.1. In particular, recall that workers with different labor supplies match to firms by some unspecified matching process. Then we prove the below result:

Lemma 3 *All workers with the same skill-set are assigned the same distribution of tasks regardless of their labor supply.*

Proof. In Tian (2019) it is shown that mutual information is strictly distraction free. 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 a assignment always exists: we can just take the total time

spent in 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 organizational costs. Notice that 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 decreased, so profit strictly decreased, contradicting 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 are 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 we have that:

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

Because total amount of each task is fixedc at α_k:

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

Plugging these into organization complexity:

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

Rearranging a bit:

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

Note that the total sum of all $\tilde{B}_{n,k}$ of the same skill set is exactly E_i which is exactly equal to $\sum_{k'} B_{i,k'}$:

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

Therefore organization complexity based on worker identities is equal to organization complexity based on worker skill sets. Since we observe identities, this implies that we can compute organization complexity as the mutual information between worker identi-

ties and tasks.

Second, we appeal to Proposition 2 to say that organization complexity is strictly decreasing in γ, implying a one-to-one mapping whenever Ω is known and $I > 0$. Third, we appeal to the equivalence with a rate-distortion problem to say that we can find the firm's optimal organizational structure, B_j using the Blahut-Arimoto algorithm.

A.6 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 utility of consumer i has the well-known closed form:

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

where C is Euler's Constant. There are a mass M of consumers, therefore total consumer expected utility is $M \cdot V_i$.[22] We then can denominate this in dollar terms by dividing by the coefficient on price, ρ. Our measure of total consumer welfare in dollar terms is:

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

With a sales tax τ, it is:

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

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.7 Organization Complexity As Task-Specialization

This section illustrates that the organizational cost of the firm is proportional to a measure of average task-specialization. To see this, first define a job as a vector, where component

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

k is the fraction of a worker's total labor spent performing task k:

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

We 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 under this assumption 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 we have the following result.

Proposition 7 *Optimal organization cost $I(B_j^*)$ is the weighted-average Kullback-Leibler divergence between 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, we can write out attention cost 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 K-L 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.

In this sense, we are assuming "task-specialization" is costly. On the other hand, the mapping from complexity to task-specialization is not exact. The firm in principle could have a worker spend all of their time on two tasks. Whenever the task-mix is uniform, this is complex, in the sense that is far form the generalist job. However, this is not a

specialized job, since the worker is doing two tasks that may or may not be tasks they are particularly skilled at performing.

A.8 Quality and Price Competition Logit Demand

Elements of the proof of Theorem 1 imply a more general result for oligopoly games with quality and price competition and per-unit quality costs. Suppose there are J firms which compete by choosing prices (p_j) and qualities (ξ_j) simultaneously. Denote the total cost of firm j as: $C_j(q, \xi_j)$. Suppose total cost is multiplicatively separable in the following way:

Assumption 1 *Total costs take the form:*

$$C_j(q, \xi_j) = q \cdot c_j(\xi_j)$$

where $c_j(\xi_j)$ is a strictly increasing, convex function.

There are a mass 1 of consumers interested in purchasing at most one of the J final products. Consumer z's utility for good j is represented by the following utility function:

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

where $\epsilon_{z,j}$ is distributed i.i.d. Type-1 extreme value across consumers and products and $\rho > 0$. a_j represents exogenous quality. Then optimal quality solves:

$$\max_{\xi_j} \xi_j - \rho c_j(\xi_j)$$

Therefore we have the following result:

Proposition 8 *In the unique equilibrium, qualities for all firms are given by:*

$$c_j'(\xi_j^*) = \frac{1}{\rho}$$

Thus, while prices depend on the actions of other firms, equilibrium qualities depend only on consumer price sensitivity. This result also implies that the equilibrium is robust to timing. Suppose the game is sequential: firms first choose quality and then price. Regardless of the pricing strategies of the other firms, every firm will choose ξ_j^*. This result appears to come from the fact that quality and price are perfect substitutes for consumers.

A.9 A Microfoundation for the Organization Complexity Measure

The main text of the paper provides several reasons why the mutual information between tasks and worker types is a reasonable measure of organization costs. Still, a reader may wonder why we can call the mutual information "organization complexity", and if there is a deeper reason why mutual information should be used as the cost function. This section provides such a microfoundation.

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 task categorization system f_T and an employee handbook g_T, which assigns task categories to workers. f_T is a vector-valued function which maps every possible combination of T tasks to an index, where without loss we say the index is of the form $\{1, 2, ..., 2^{T \cdot I}\}$. 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 than complexity times the firm-specific organization cost coefficient γ_j Equivalence to a rate-distortion problem allows us 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, we can think of $I(B_j)$ as measuring the length of the business plan needed to implement a certain organization structure. This is why we call $I(B_j)$ complexity.

B Empirical Appendix

B.1 Task Classification Process: Further Details

A cosmetologist was hired to categorize 20,560 service descriptions. The cosmetologist was hired via UpWork and paid $1,500 for the entire job. As part of the agreement, the person provided a picture of their cosmetology license. This information is available upon request.

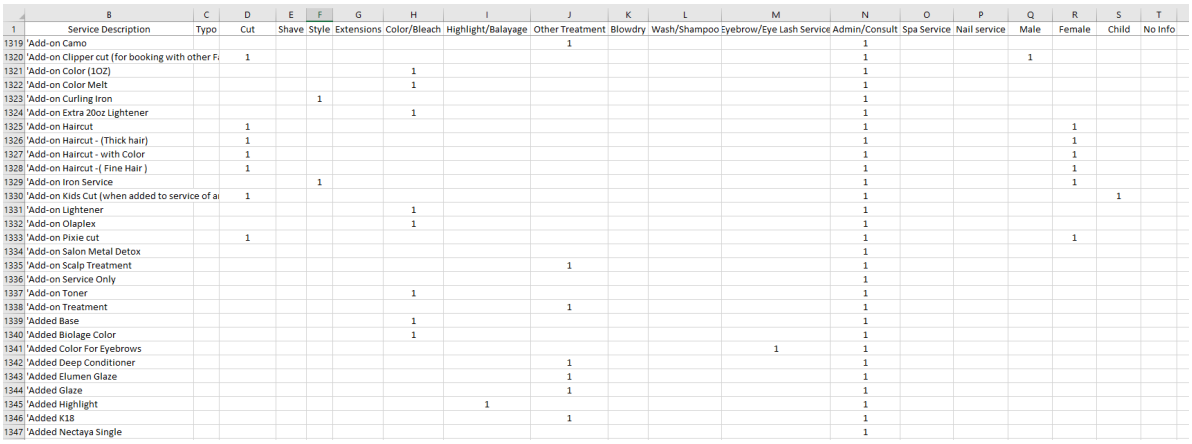
The cosmetologist was provided with a blank spread sheet with predefined subcategories. The cosmetologist was instructed to mark all categories which a description matched with a one. they were instructed that some subcategories may not be mutually exclusive, so they should mark all that applied. The description of what I asked the cosmetologist to do is given verbatim here:

> 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:

> 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, I checked the work, 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 18.



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

Figure 18: Final Task Sub-Categorization Spreadsheet from Cosmetologist

The subcategories are very detailed. To reduce these to six tasks, I hired the same cosmetologist again. I paid the person $100 to classify the subcategories into six task categories. The description of what I asked the cosmetologist to do is given verbatim here:

> 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 19. 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 19: Final Task Sub-Categorization Spreadsheet from Cosmetologist

B.2 An Alternative Task Category Generation Process

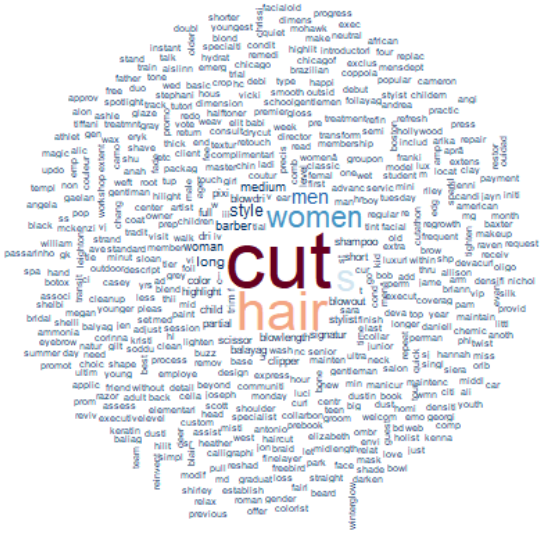
The main analysis in this paper relies on professionally categorized descriptions. This section describes an alternative natural language processing approach used to generate task categories. This process has the benefit of being less labor intensive and not reliant on human judgment. It has the disadvantage of making certain types of classification errors. Much of the analysis in this paper was originally done using this method, and the relationships between revenue and organizational complexity is robust to the classification method used.

In order to minimize issues related to spelling and grammar, I clean 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 are removed from the analysis.

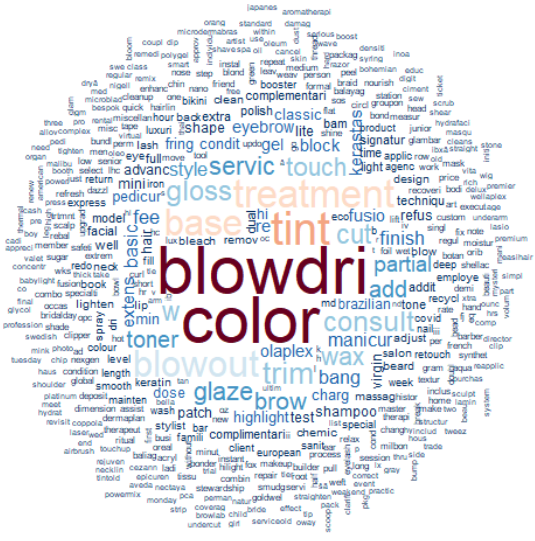
To group descriptions into tasks, I use a hierarchical clustering method where the only choices are the number of categories and the measure of distance. I set the number of categories to be four and the distance measure to be the euclidean distance.[23] Figure 20 presents word clouds which capture the most common words appearing in each category. The size of a word in the cloud is roughly proportional to the number of times the word appeared in a unique description.



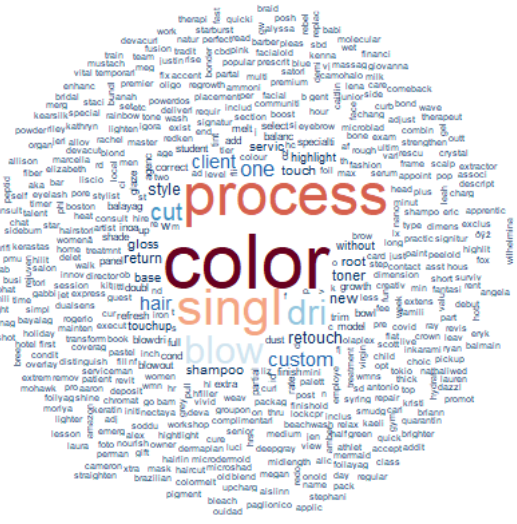
(a) Task 1: "Highlight" N= 1,217,502



(b) Task 2: "Haircut" N=4,308,585



(c) Task 3: "Blowdry" N=3,633,095



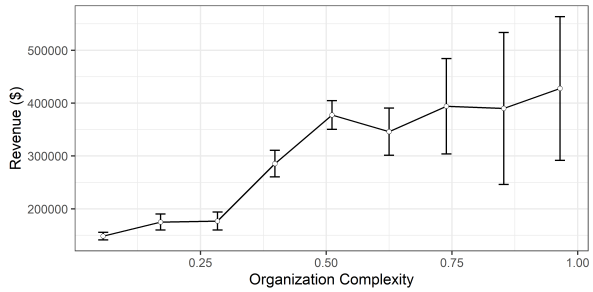
(d) Task 4: "Color" N= 1,378,195

Figure 20: Tasks Generated from Hierarchical Clustering

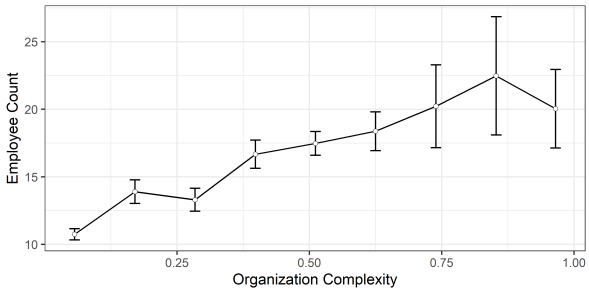
23. The quality-specialization relationship which inspired this paper is robust to both the number of categories and the distance metric.

The four tasks roughly correspond to three services: coloring ('single process'), hair cutting, blow drying and highlighting.[24]

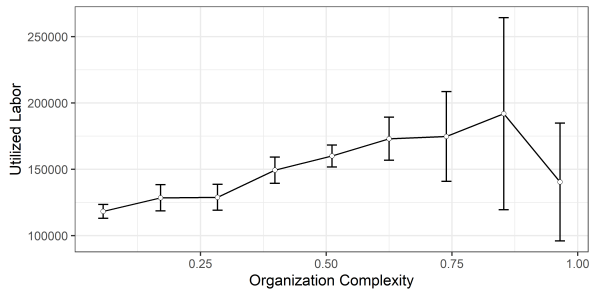
B.3 Firm Size and Complexity Associations: Raw



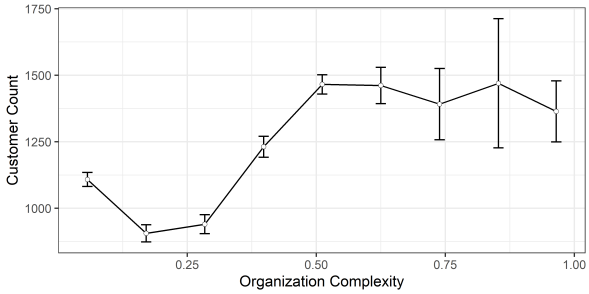
(a) Revenue ($)



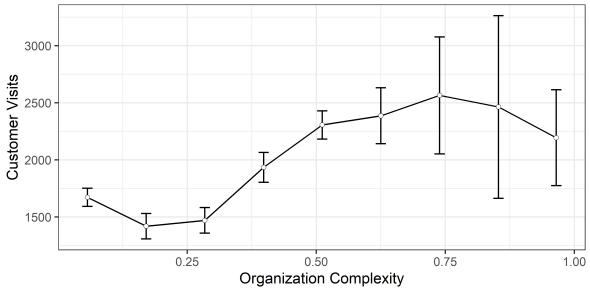
(b) Employees



(c) Utilized Labor (minutes)



(d) Unique Customers



(e) Customer Visits

Figure 21: Organization Complexity and Firm Size

B.4 Task Content Variance Decomposition

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

24. This is born out in the word clouds, but also if one looks at the modal description in each category.

write:

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

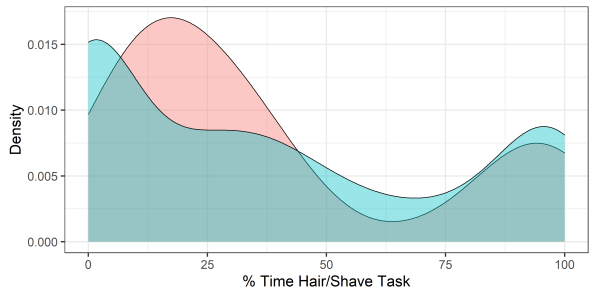
I then adapt the method used by Song et al. (2019) to my setting. Fixing k and taking the distribution to be weighted by effective units of labor, we can then decompose the variance into a worker type component and a within-worker type component, where recall ω_i is the share of total labor represented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

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

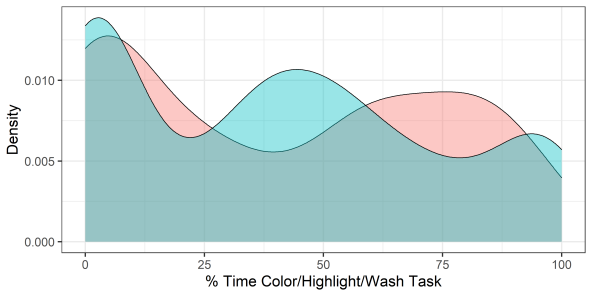
However, because $b_j(i,k)$ is generated by the structural model, and we are considering a single labor and product market, the within-type component comes entirely from variation in firm attributes. Therefore we have decomposed the total variance in task content into a worker and firm component. Dividing through by $Var_{i,j}(b_j(i,k))$, Table 10 shows the share of variance due to each component. For the main tasks (cut, color, blow dry) between 8 and 22 percent of variation in job task content is attributable to firms.

B.5 The Full Distribution of Task Content

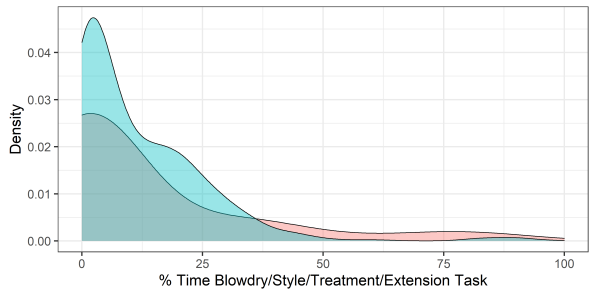
We can also go beyond the variance and compare the entire distribution of model and observed job task content. This is a strong test of the model, because we observe 509 stylists in Manhattan during 2021Q2 (the estimation period), and we are asking the model to match their jobs using only firm-based moments. Figure 22 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, in panel b, we can see 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.



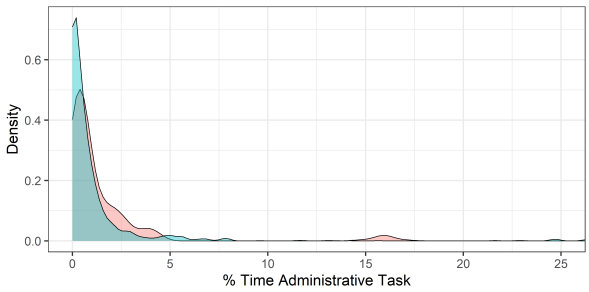
(a) Hair/Shave Task



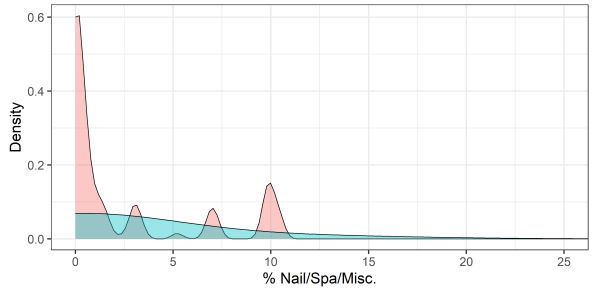
(b) Color/Highlight/Wash Task



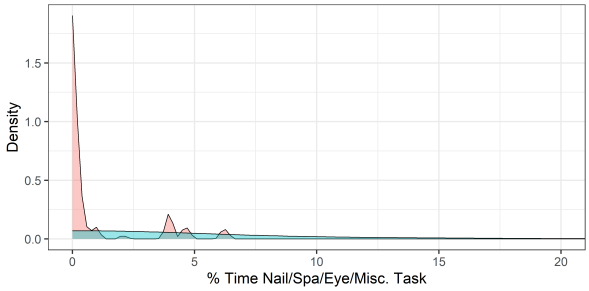
(c) Extension Task



(d) Blowdry/Style/Treatment Task



(e) Administrative Task



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

Figure 22: Model (Red) vs. Observed (Blue) Job Task Content in Manhattan

B.6 Counterfactual Procedures

B.6.1 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.[25]

25. It is necessary to weight firms for counterfactual analysis but not estimation. This is because during estimation we fix an equilibrium, but in counterfactuals we must find a new equilibrium, and firm pricing strategies depend on the number of other firms in the market.

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. If I allow 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. Given organization structures, qualities and costs of all firms are now known. Compute optimal pricing 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. So although each labor demand is monotone decreasing in own-wage, since firms can use other worker types as substitutes, we 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 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.6.2 Minimum Wage Technical Issues

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

The numeraire good in my model is the outside option, which is not getting a haircut. 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 getting a haircut 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 (2020b)).

In general, a minimum wage can result in multiple equilibria. To insure that there are not multiple equilibria, I solve the model under every possible permutation of binding minimum wages. That is, I assume the minimum wage binds for worker types 1 and 2 only, 1,2 and 3 only, etc. With 6 worker types, this amounts to solving the model $2^6 = 64$ times. Each time I solve the model, I fix the wages of the binding types at \$30, and then solve for the wages of the other types which clear the labor market for only those other types. After, I check 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 passes this check is 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 solve for the other 5 wages which clear the market for the other five types. I then check 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 counterfactual, only one of the 64 cases satisfies the checks as a valid equilibrium. One additional case in the full adjustment counterfactual never converged, meaning I could not find wages clear the labor market for the non-binding worker types.

B.7 Job-Level Heterogeneity

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

Table 14: Job Task-Mix

Figure 23: The Job Task-Mix Distribution

