Employee or Contractor? An Incomplete-Information Approach to the Labor Boundaries of the Firm

Michael Raith[1]

January 2023

[1]Michael Raith, Simon Business School, University of Rochester, Rochester, NY 14627, raith@simon.rochester.edu. I would like to thank Patrick Bolton, Matthias Fahn, Guido Friebel, Maija Halonen, Jin Li, Elliot Lipnowski, Jeanine Miklos-Thal, Heikki Rantakari, Mark Spoerer, Giorgio Zanarone, and seminar participants at the 14th Organizational Economics Workshop (ANU/UNSW, 2020), Verein füer Socialpolitik Organizational Economics meeting (2021), NBER Organizational Economics Working Group meeting (2021), SIOE Annual meeting (2022), the Theory and Applications of Contracts Conference (Imperial College, 2022), and the University of Rochester, for helpful comments and suggestions.

Abstract

I develop a theory of employment that builds on, refines, and generalizes Simon's (1951), and compare employment with its closest alternative, sequential spot contracting for labor. In my model, an entrepreneur and a wealth-constrained worker face two-sided incomplete information about the benefits and costs, respectively, associated with different tasks. Employment is a relational contract that requires the worker to provide his service and to follow the entrepreneur's direction, for a wage. I show that consistent with historical and anecdotal evidence, employment helps a firm to secure a predictable supply of labor and to avoid the "transaction costs" of market trade. Employment can be profitable for a firm even if it is inefficient. Adaptation to an uncertain environment may or may not play a role, and when it does, then contracting is no less capable than employment of achieving adaptation, contrary to a common view about the advantage of employment. Nevertheless, employees do tend to have multiple tasks because bundling multiple tasks into jobs may be necessary for employment to be economically viable. Overall, the paper sheds new light on the nature of employment, on job design, and on the origins of capitalist production.

1 Introduction

When should firms employ their workers, and when should they engage in arms-length contracting instead? How does employment differ from contracting anyway? In the 20th century, Coase (1937) and others considered authority relationships to be a hallmark of firms, and employment the institution to establish them.[1] Today, however, firms around the world are increasingly outsourcing labor to independent contractors, and for workers, it is easier than ever to be a freelancer or gig worker rather than a traditional employee.[2] In January 2023, the *Economist* declared that "[t]oday Coasean forces are ushering in a new type of corporate organisation" (Economist 2023). Meanwhile, lawyers are debating how new work arrangements fit, or don't fit, within traditional legal definitions of employee and contractor (Posner 2020).

Understanding today's changes in the organization of work requires understanding why employment exists in the first place. A leading explanation, proposed by Coase (1937) and Simon (1951), is that for a firm (or a boss) to have authority over a worker's actions facilitates a firm's adaptation to an uncertain environment, compared to relying on market transactions with workers.[3] This explanation, however, has always faced two challenges (Hart 1989). One has been to make precise how authority works when, in most modern societies, a worker's right to choose his own actions is inalienable.[4] The second and more fundamental challenge to Coase's argument, and to transaction-cost economics more generally, is that the adaptation advantage of "hierarchy" is not a truism, and that examining it requires a theoretical framework within which market trade for labor is feasible but also costly, thus creating a potential role for authority in firms without already assuming its superiority over the market.[5]

[1] For instance, Simon (1945:177) wrote: "Of all the modes of influence, authority is the one that chiefly distinguishes the behavior of individuals as participants of organizations from their behavior outside such organizations." And Arrow (1974: 63): "the giving and taking of orders ... is an essential part of the mechanism by which organizations function." Throughout, by "authority" I mean "interpersonal authority" (Van den Steen 2010) without using the qualifier "interpersonal," consistent with its use by early organizational scholars and by sociologists (e.g., Coleman 1990).

[2] Weber (2017), Bertrand et al. (2021), Anderson and Cappelli (2021), Oyer (2016), Paychex (2016), Katz and Krueger (2019).

[3] Williamson (1975), Wernerfelt (1997, 2004, 2015), Gibbons (2005), Baron and Kreps (2013), Van den Steen (2010), Rantakari (2021).

[4] Barnard (1938), Alchian and Demsetz (1972), Arrow (1974), Hart (1996), Bolton and Rajan (2003), Marino et al. (2010), Van den Steen (2010), Rantakari (2021).

[5] Hart (1989), Gibbons (2005), Hart and Moore (2008), Tadelis and Williamson (2013). For different approaches to this challenge, see Section 2.

My paper offers new answers to both of these challenges, and in the process leads to a novel explanation for the existence of employment. I propose an economic definition of employment that builds on, refines, and generalizes Simon's (1951), and compare employment with sequential spot contracting within a model of bilateral trade. I derive three sets of results, described in more detail below: First, employment helps firms to secure a *predictable* supply of labor and to avoid costs of the market in the form of failure to trade and rents paid to workers. In fact, employment can be profitable for firms even if it is inefficient. This rationale applies even when adaptation plays no role, such as when a worker has only one task. Second, when adaptation matters, authority has no systematic advantage over market transactions in achieving adaptation. However, bundling different tasks into one job raises the value of a relationship and thus the advantages of employment as well. Third, employment may entail delegating authority to a worker, but doing so is not necessarily profit-maximizing even when it is efficient. My assumptions and results are supported by a range of historical facts about the emergence of wage labor during early Industrialization in the 18th century, as I explain throughout the paper. Overall, my paper sheds new light on the nature of employment, on job design, and on the origins of capitalist production.

In my model (Section 3), a buyer and a seller of labor—an entrepreneur and a worker—interact over time. In each period, the worker can perform one of possibly multiple tasks for the entrepreneur. Actions are contractible; I abstract from effort choice. Long-term contracts are not feasible. What drives the model is that the entrepreneur and the worker are privately informed about the benefits and (opportunity) costs, respectively, of the worker's actions in each period. I focus on the two governance structures at the center of Williamson's (1975, chapter 4) discussion: sequential spot trade and employment. In Appendix A, I argue that within the assumptions of my model, no other governance structures, or contracts, can improve upon these two. At the beginning of the game, the entrepreneur chooses the structure that maximizes her expected profit.[6]

If the worker is a contractor, the entrepreneur and the worker bargain over prices

[6] Like in Hart and Moore (2008), this stage can be thought of as resulting from ex-ante competition among workers to do business with the entrepreneur. Once the entrepreneur chooses a worker and a profit-maximizing governance structure (employment or spot trade), a Williamsonian "fundamental transformation" (1985: 61-63; 1975: 61-64) leads to appropriable quasi-rents and thus in effect a bilateral monopoly between entrepreneur and worker. These quasi-rents could result from deliberate investments or could just be due to "knowledge and skills that are incidentally acquired by the parties while working together" (Tadelis and Williamson 2013).

of different tasks in each period. Because of incomplete information, they may fail to agree on a price even if trade is efficient, as is well known from the bargaining literature (Myerson and Satterthwaite 1983, Ausubel et al. 2002).

As an employee, the worker confers to the entrepreneur the authority to choose the worker's action, in return for a wage. This definition generalizes Simon's (1951) in that, unlike in his model, employment is well-defined and differs from spot trade even if the worker has only one task: The contractor can always decline to perform a service, whereas the employee has promised his services to the entrepreneur. More generally, then, employment amounts to an *obligation to be available to work*. Because the entrepreneur's orders are informed only by her own valuations for each task, employment can be inefficient too if the entrepreneur orders an action whose cost to the worker exceeds the benefit.

Refining Simon's definition and responding to a well-known critique of Coase's and Simon's approach, I assume that authority is not contractible and must therefore be supported by a relational contract, as in Bolton and Rajan (2003). That is, although the worker's transfer of authority to the entrepreneur is voluntary, it is not legally enforceable, and dismissal is the entrepreneur's only sanction for disobedience.[7] The worker obeys in order to keep his job, which is worthwhile for him if the wage is high enough, and worthwhile for the entrepreneur if it is not too high. Envisioning employment as a relational contract is hardly novel (Malcomson 1999, Baron and Kreps 2013, Levin 2010), but the emphasis on its purpose to support obedience defuses Alchian and Demsetz's (1972) famous objection to the notion of authority: Unlike a contractor who refuses to trade, an employee who disobeys is in breach of an agreement he entered into, and stands to lose the benefit of the relationship.

My final key assumption is that the worker is wealth-constrained, reflecting the reality faced by workers throughout history. The worker is therefore unable to pay the entrepreneur upfront for any rents earned either as contractor or as employee. As in standard mechanism design, therefore, the entrepreneur's profit-maximizing structure balances the dual goals of efficiency and rent extraction. Out of a large set of possible bargaining mechanisms that govern spot trade, I focus on the one that maximizes the entrepreneur's

[7] Barnard (1938: 165) argued, "A person can and will accept a communication as authoritative only when ... at the time of his decision, he believes it to be compatible with his personal interest as whole." Consequently (Arrow 1974), "within the firm, the sanctions which authority can use are basically those of hiring and firing" and "The scope of this authority ... is limited by the freedom with which an employee can leave the job."

profit from spot trade, which is for the entrepreneur to make a take-it-or-leave-it price offer to the seller in each period. This assumption maximally biases the entrepreneur's choice in favor of spot trade and thus establishes a lower bound for employment to be more profitable than the market under any (incentive-compatible and interim-rational) bargaining mechanism.[8]

I focus on discount factors that are large enough for employment to be feasible. The entrepreneur's profit-maximizing employment contract then minimizes the worker's wage to satisfy an obedience constraint. The optimal wage equals the worker's expected cost as employee, plus an "obedience wage premium", akin to an efficiency wage markup, that is needed for the worker's continuation utility to exceed any gains from disobedience. The wage premium is decreasing in the discount factor, which leads to the prediction that employment is more profitable relative to spot trade the less the parties discount future payoffs, or equivalently, the more frequently they interact, consistent with evidence.

I show that discrepancies between profit maximization and efficiency can go both ways. Even if employment is more efficient than contracting, the entrepreneur may prefer contracting if the obedience wage premium is too high. This may happen if the discount factor is lower; for instance, if the interaction between the entrepreneur and the worker is less frequent. Conversely, even if employment is less efficient than contracting, the entrepreneur may prefer employment to avoid surrendering too high a rent to the worker with spot trade. That is because with spot trade, the entrepreneur can never capture the full first-best surplus; part of it is either deadweight loss or goes to the worker as rent. With employment, in contrast, ex-ante competition eliminates the worker's rent if the parties are sufficiently patient. This result lends support to claims going back to Karl Marx that hierarchical production and wage labor are "exploitative" and possibly inefficient.[9]

In Section 4, I study a version of the general model with only one task and two types per player. The one-task case is empirically important because many workers' jobs, especially prior to the 21st century, exhibit extreme specialization and rigor instead of being "massively incomplete" (Baron and Kreps 2013). The analysis leads to two main

[8] I show in Appendix A that in spite of the worker's wealth constraint, a stationary contract with a constant wage is optimal within a larger class of employment contracts that allow for non-stationarity.

[9] See Marx (1867), Marglin (1974), Braverman (1974), and Dow (1987, 1993). For further discussion, see Williamson (1985) and Rebitzer (1993).

insights. First, broadly speaking, employment leads to a greater surplus relative to spot trade the larger the quasi-rent at stake. Intuitively, the larger the first-best surplus, the more efficient it is on average for the worker to work on some task than not to work, which is what employment accomplishes. This result provides a simple formal foundation for a central but theoretically elusive prediction of transaction-cost theory, that "larger AQRs make integration more likely" (Gibbons 2005).

Second, by construction, adaptation plays no role in a one-task model. It is also unimportant for the entrepreneur to have any private information; all that matters is for her valuation of labor to be high enough, for employment to be optimal. To illustrate, if the worker's cost is stochastic but the entrepreneur's benefit always exceeds the cost, then spot trade at a price equal to the highest possible cost is efficient but leaves the worker a rent, whereas a lower price may lead to inefficiency. With employment, however, the wage equals the worker's expected cost, plus an obedience premium that shrinks to zero as the discount factor grows to one, enabling the entrepreneur to capture the full first-best surplus.

In Section 5, I extend the base model to two tasks, to investigate a central claim in the literature since Coase (1937) that employment facilitates the adaptation of work to changing business needs. Within my framework but unlike in other theories, whether a job has one or multiple tasks is inessential to the nature of employment. Intuitively, both the employed and the independent IT specialist will manage my database and not my website, if that is what I currently need. *The market is no less capable than authority of achieving efficient adaptation.* The difference between the two is that the contractor can decline my request for service, or demand a surcharge, whereas the employee is relationally required to work for me at an agreed price. This difference, however, is already present with only one task.

Bundling multiple tasks into one job, however, can dramatically raise the profitability of employment if the firm's benefits are negatively correlated across tasks. While the entrepreneur may not have enough demand to hire specialists for database and website management, she may well have demand for one IT person to work on these tasks at different times, especially if the worker cares little about which task is requested (i.e. if the worker has highly correlated costs). In practice, then, employees often do have multiple tasks because their jobs are endogenously designed that way: Firms will pay a regular wage only to people who always have something to do.

Further expanding Simon's (1951) definition, the entrepreneur can also delegate the choice of tasks to the employee. Employment with delegation differs from spot contracting in that the employee is still obligated to provide a service, whereas the contractor is not. I show that even if delegation is efficient (because the costs of tasks vary more than the benefits), the entrepreneur may still prefer to retain authority for herself. This result is driven by the worker's obedience constraint, which limits the entrepreneur's ability to pay a lower wage to a worker with authority.

In Section 6, I examine two situations involving multiple workers. First, I show that complementarity in production between two workers favors employment, because with spot trade, the entrepreneur has to pay more on average to secure both workers' participation. This result helps to explain why wage labor has historically correlated with team production (in the sense of Alchian and Demsetz 1972), but may become less dominant as firms adopt more flexible production methods (Baldwin and Clark 2000, Garud et al. 2009, Matouschek et al. 2022). A very different situation is when multiple workers can perform the same task *ex post*, such as drivers for Uber. Then, with enough workers, spot trade unambiguously dominates employment because the entrepreneur can offer a low price that is likely to be accepted by at least one worker.

Section 7 relates the theory to the (constantly evolving) legal distinction between employees and contractors. Employment laws seek to protect workers who are economically dependent on their employer, which is consistent with both my assumptions (workers without wealth) and results (employment as value capture). Commonly employed legal tests, however, lean heavily toward classifying workers as employees, which may significantly constrain firms' organizational choices. Section 8 concludes.

2 Related literature

My paper relates most closely to a diverse literature that, like mine, models employment as an authority relationship. The biggest challenge in proving any benefits of authority over the market is to devise a framework in which both are feasible options, whose advantages and disadvantages derive from common primitives of the environment. An additional challenge is to explain how authority relationships can exist in the first place, as Alchian and Demsetz (1972) famously questioned. With one exception, my paper is the first to address both challenges.

That exception is Bolton and Rajan (2003, see also Bolton and Dewatripont, 2004: 588, and 2013: 369), who model employment as a relational contract in an environment with incomplete information. The authors assume that the buyer (boss) knows the state of nature, including the seller's (worker's) cost. The purpose of employment is then to utilize the buyer's information, and employment is always efficient if it is feasible. My informational assumptions are both more general and more conventional, and employment is not necessarily more efficient or more profitable than spot trade.

Several papers compare employment with independent contracting but either abstract from obedience as a problem or make reduced-form assumptions about bargaining, or both. Hart and Moore (2008) compare employment with independent contracting in a framework in which the costs of ex-post contracting are derived from behavioral assumptions. Wernerfelt (1997, 2004, 2015) focuses on optimal adaptation in the presence of bargaining costs and other adjustment costs, which are modeled in reduced form. Levin and Tadelis (2010) compare contracting on time (employment) with contracting on quality (outsourcing), under the assumption that time is easy but quality is costly to measure. In contrast to these papers, the costs of contracting in my model result endogenously from incomplete information, and assuring obedience is a binding constraint on employment. In independent work, Schöttner and Upton (2022) distinguish employment and contracting by who has (contractible) authority over the worker's action, which leads to the least costly (respectively, highest benefit) action as the default with contracting (employment). With both structures, a relational contract supports efficient deviations from the default.

Other papers endogenize authority and obedience but do not allow for market trade. In Van den Steen (2010), concentrated asset ownership, authority, and low-powered incentives go hand in hand. Authority is needed for two parties to coordinate on a project when they disagree due to differing priors. In Rantakari (2021), authority can emerge as a relational contract between two symmetric players involved in a coordination problem. The source of the superior's authority is an information advantage gained through endogenous discovery effort. Actions are not contractible in either paper, however; market trade is therefore infeasible. Marino et al. (2010) focus on the constraints that potential disobedience imposes on the allocation of decision rights in firms.

A third strand of literature focuses on how the ex-ante allocation of control rights facilitates adaptation, in contexts other than employment. Like mine, several papers in this literature feature ex-post bargaining under incomplete information. Bajari and Tadelis (2001) study how procurement contracts solve a tradeoff between providing incentives and reducing ex-post transaction costs. Matouschek (2004) and Segal and Whinston (2016) develop property-rights theories of the firm in which the objective is to minimize the inefficiency of ex-post bargaining; see also Segal and Whinston (2013). Chakravarty and MacLeod (2009) show how authority provisions in construction contracts regulate the parties' ex-post bargaining power, in the shadow of legal constraints on enforcement. Baker, Gibbons, and Murphy (2011) study how decision rights support adaptation through the use of relational contracts when the state of the world is observable to both parties. Their Section 4 allows for ex-post contracting with exogenous bargaining costs. Zanarone (2013) shows how franchise contracts facilitate adaptation in the shadow of laws designed to protect franchisees. Powell (2015) compares integrated and non-integrated structures when the goal is to facilitate adaptation and to curtail wasteful influence activities. Like in Van den Steen (2010) and in Rantakari (2021), Powell's setting is a coordination game without market trade.

A very different approach to employment is taken in the property-rights and incentive-system literatures, which define employment by a firm's ownership of the assets that are used by the worker, such as Hart and Moore (1990), Holmström and Milgrom (1991), Baker, Gibbons, and Murphy (2002), and Zanarone (2012). This perspective makes explicit how asset ownership shapes outside options and the resulting equilibria. However, it focuses on ex-ante incentives; adaptation plays no role (see Gibbons 2005), nor does the distinction between hierarchy and market.[10] In addition, empirically, asset ownership

[10] An exception is Zanarone (2012), who points out that contracts involving decisions by one party's fiat can also occur between firms, such as in franchise systems. In his model, asset ownership defines integration or non-integration, whereas fiat (within or between firms) is supported by a relational contract.

is not always a distinguishing characteristic of a worker's contractor status.[11]

Like Alchian and Demsetz (1972), property-rights theory argues that conceptually, firing an employee is no different from firing a grocer. However, the ownership of non-human assets still confers power over people. The fired grocer still owns his assets and therefore has other customers to sell to; the fired employee may have much more limited options, which in turn supports obedience (Hart 1996, Roberts 2004: 104, Van den Steen 2010: 477-478). In my model, by contrast, the worker as employee voluntarily and explicitly vests authority over his actions in the entrepreneur (Coleman 1990, chapter 4), even if the authority is only relational: Irrespective of asset ownership, the employee who refuses to "type this letter" is in breach of an agreement, the grocer who refuses to "sell me this brand of tuna" is not.[12]

Two other arguments relate to the existence of employment but differ from this paper's focus. First, Abraham and Taylor's (1996) predictions for when firms rely on employees or contractors are mostly based on whether firms benefit from long-term relationships with workers or instead can purchase labor in competitive spot markets. My paper, instead, asks why a given long-term relationship would be organized as employment and not a series of spot contracts. As Williamson (1975: 62) argued, we should only ever expect to see employment in "small-numbers exchange conditions;" that is, in relationships with appropriable quasi-rents.

A different argument, sometimes used in legal determinations of employment status (see Section 7), is that contractors are paid for their output or completed tasks, and employees for their inputs, measured by monitoring (Alchian and Demsetz 1972, Fama 1991). Control and monitoring of workers is undeniably an advantage of factory organization. It

[11] For example: (1) Under the inside-contracting system that was common in New England between the Civil War and World War I, capitalists supplied all equipment and raw materials, and sold the final product. "The gap between raw material and finished product, however, was filled not by paid employees ... but by contractors, to whom the production job was delegated" (Buttrick 1952, see also Williamson 1975, 1985). Such "contractor-managers" have been observed in the 21st century as well, see Anderson and Cappelli (2021) and Broughton (2021). (2) As Barrera (2015) details, whether firms obtain legal services from in-house counsel or from outside law firms is driven by factors such as the nature of the required expertise or the ability to handle legal matters quickly, but "does not rely on an asset that is owned by the client when the lawyers are in-house employees and owned by the law firm when the lawyers are outside counsel." (3) What distinguishes today's remote employee from the freelancer is neither the workplace (often, home) nor the productive assets used (say, a computer), but the employee's obligation to work on terms defined by the labor buyer (Van Triest 2021) and, in return, his regular paycheck.

[12] This argument also differs from Williamson's (1991) position that a firm's authority originates from the legal doctrine of forebearance; see Section 7 for further discussion.

seem less relevant in today's knowledge economy, in which firms have to measure their employees' performance when monitoring is ineffective, and in which the performance measures that can be used with contractors barely differ from those used for employees (Bernstein and Peterson 2022).

Finally, within a large and growing recent literature on relational employment contracts, Li and Matouschek (2013) and Li et al. (2020) are closest in that they, too, assume that the worker has limited liability and the principal has private information. However, this literature focuses on the dynamics of employment relationships and is concerned with very different questions.

3 General framework and analysis

3.1 Model

An entrepreneur E and a worker W are engaged in a long-term relationship that extends over potentially infinitely many periods, starting in $\tau = 1$. At the start of the relationship, E makes a take-it-or-leave-it offer that specifies the governance structure for the relationship. Like in Hart and Moore (2008), this situation can be thought of as resulting from ex-ante competition among many workers for E's business, followed by a "fundamental transformation" in the sense of Williamson (1985) that leads to lock-in and thus the creation of appropriable quasi-rents between E and W (see Williamson 1975: 61-64, Williamson 1985: 52-56, Gibbons 2010, Tadelis and Williamson 2013).[13]

Unlike in Hart and Moore, however, W has no wealth; he is therefore unable to pay E upfront for any rents he might earn either as contractor or as employee.[14] Exclusivity is not assumed: E may have multiple workers, and W may work for multiple firms, even as

[13] Thus, for E to have the bargaining power to make a take-it-or-leave-it offer to W, she would more precisely offer her terms and choose W at the same time, while workers are competing.

[14] Williamson (1975: 69-70) makes the same point, adding that "it is seriously to be doubted that [workers] could raise the funds, if their personal assets were deficient, to make the implied full valuation bids." Although many workers were poor throughout history, the trend towards factory production and wage labor that began in the 18th century coincided with "the emergence of an increasingly landless and wage-dependent rural proletariat" (Hudson 1981, see also Mantoux: 74). Town-based workers, lacking land to help make a living, often became dependent on their labor even under the pre-industrial putting-out system (Mantoux 1928, p. 65). In 19th century England, "the most important social issue in England [was] the condition of the working classes, who form the vast majority of the English people... What is to become of these propertyless millions who own nothing and consume today what they earned yesterday?" (Engels 1845, cited in Allen 2009, whose work lends credence to Engels' account).

employee (e.g., in two part-time jobs).

In each period, W can carry out one of possibly multiple tasks in the set $A = \{1, .., M\}$, or alternatively not work for E, which is action 0 ("no work" is a shorthand for pursuing the next best use of time). Each task $k \geq 1$ generates benefit $b^k \in B$ to E and causes cost $c^k \in C$ to W, for finite B and C with strictly positive elements. If W does not work, then $c = b = 0$. Denote the players' type spaces by $\mathcal{B} = B^M$ and $\mathcal{C} = C^M$. Both players have common prior beliefs that $b = (b^1, ..b^M)$ is distributed according to the probability distribution $p : \mathcal{B} \rightarrow [0,1]^M$, and $c = (c^1, ..c^M)$ according to $q : \mathcal{C} \rightarrow [0,1]^M$. I assume that p and q are independent and are i.i.d. across time periods. They can be correlated across tasks but have identical marginal distributions for each task. If W carries out task k, the surplus generated between E and W is $s^k = b^k - c^k$.

If E and W break up, then going forward, both earn 0 in every period. Equating the outside option payoffs with the no-trade payoff within the relationship means that the benefits and costs of trade are to be understood as opportunity benefits and costs, relative to outside options. In particular, changes in W's costs may be primarily due to changing outside options. Both players are risk neutral and discount future payoffs at a common discount factor δ. Each player maximizes, at each point in time, their net present value of payoffs. To abstract from savings dynamics, I assume that W has no access to credit markets. A timeline of events is sketched in Figure 1.

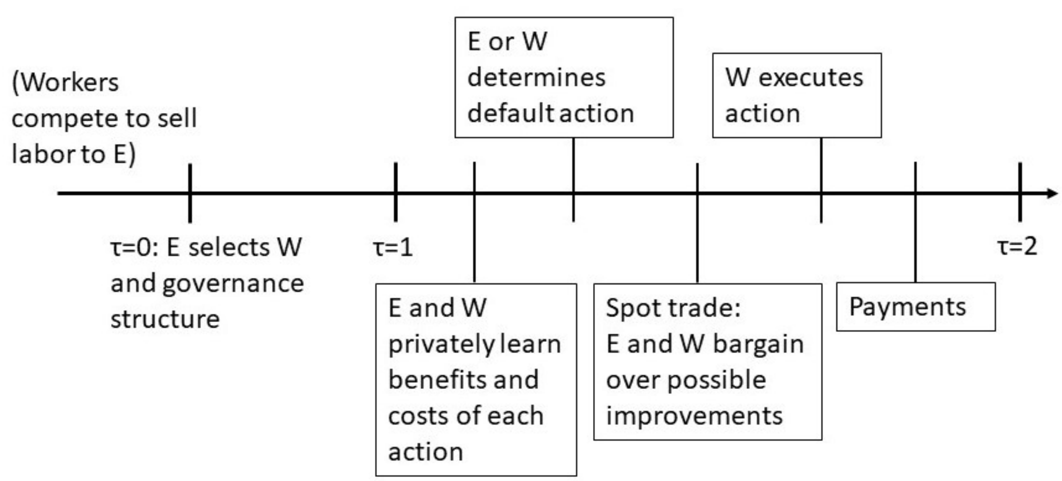

Figure 1: General timeline of events

All tasks are contractible in the short run. However, more for legal than for technological reasons, long-term contracts are infeasible, as are any contracts that deprive W from legal authority over his own actions. I consider two specific governance structures, sequential spot trade and employment. Appendix A discusses to what extent other governance structures can improve upon these two.

Spot trade:

With spot trade, E and W engage in bargaining over all tasks $\{0,..M\}$, with no trade being the default in the absence of an agreement because W has authority over his actions and $\arg\min_{k\in\{0,..M\}} c^k = 0$. Since my focus is on E's profit-maximizing choice of governance structure, I derive a lower bound to the profitability of employment, by focusing on the mechanism that maximizes E's profit under spot trade. That mechanism is to make a take-it-or-leave-it offer of prices $(t^1,\ldots,t^M)$ to W, for any realization of E's type. More formally, a take-it-or-leave-it offer is an ex-ante efficient bargaining mechanism that maximizes the buyer's payoff, see Theorem 5.2 in Satterthwaite and Williams (1989). If employment is profitable even when E gets to make a take-it-or-leave-it offer under spot trade, then for the same parameters, employment is profitable with *any* incentive-compatible and interim-rational bargaining mechanism. W's optimal response to E's price offer $(t^1,\ldots,t^M)$ is to pick action $m^0 \in \arg\max_{k\in\{0,..M\}} t^k - c^k$. This results either in agreeing to task $m^0 \in \{1,\ldots,M\}$, upon which W executes m and receives t^m, or W rejects all prices and no trade takes place. If W faces a tie between optimal choices, I assume that he chooses the action that maximizes the surplus s^k. Although W does not know b, this choice can be thought of as the outcome of an incentive-compatible direct mechanism that selects for Pareto-efficient outcomes subject to W picking a best response to E's price offer.

Employment:

Employment is a relational contract whereby W confers, as a promise, to E the right to control his actions out of the task set A, in return for a constant wage w. In Appendix A, I show that a stationary contract is optimal within a larger class of contracts that give E the authority to choose W's action. Provided that employment is feasible, it is optimal for E to offer the lowest wage that W will accept.[15]

[15] Once in place, there is no scope to renegotiate the wage: although both sides face exit costs, E has no reason to agree to any wage increase, and can credibly fire a worker who demands a raise just as she can fire a worker who shirks. As Malcomson (1999, Section 4.2) argues, even if the firm cannot legally dictate the wage, "employment at will as interpreted by US courts has exactly the same effect."

For any realized benefit vector b, E's optimal action is $\arg\max_{k\in A}\{b^k\}$. Like above for spot trade, in case of a tie, I assume that E chooses the action that maximizes the surplus s^k. Thus, given $b \in \mathcal{B}$ and $c \in \mathcal{C}$, denote E's optimal default action by

$$m^E(b,c) = \arg\max_{l}\{s^l | l \in \arg\max_{k\in A}\{b^k\}\}, \tag{1}$$

which E communicates to W as "order." The set of tasks A corresponds to Simon's (1951) "area of acceptance": While with spot contracting the set of relevant tasks need not be precisely defined, with employment it is essential to define in advance the tasks to which E's authority over W applies.[16]

The relational contract requires W to execute m^E, and if he does, then E pays w, and the relationship continues into the next period. If W disobeys by choosing $l \neq m^E$, then the relationship terminates.[17] Also, between periods, and between all stages within each period, E and W can unilaterally terminate their relationship if they wish. The exception is that E owes W a wage payment if and only if W carries out m^E. Thus, consistent with "employment at will" in practice, E can fire W almost anytime but owes him a wage in the current period if W obeys (see Malcomson 1999 for a discussion of employment at will).

Per Figure 1, the timelines for spot trade and employment are identical except for the bargaining stage, which is central with spot trade but absent with employment. In a previous version of this paper (Raith 2021), I restore symmetry by assuming that bargaining can take place within employment too—for example, the employee may want to negotiate time off for an important personal errand, to be made up another time. Under this assumption, the key difference between spot trade and employment is not the "supersession of the price mechanism" (Coase 1937), but the default outcome that defines

[16] In Raith (2021), I explore the endogenous choice of A within a larger set of tasks. I show, for instance, that if a task such as overtime work is always more costly for W than other tasks, but sometimes efficient, then a contract that still gives E authority but specifies a task-specific wage such as overtime pay may be optimal.

[17] This assumption is consistent with Levin (2003: 840) but is in contrast to Baker et al. (2002: 50), who assume that the parties revert to spot trade if their relational contract terminates. The threat of severance is credible if even after $\tau = 0$ there continue to be other workers available to take W's spot, and if the quasi-rents between E and W (i.e., the costs of turnover) are not so large as to shut out other workers as alternative sellers. It follows, importantly, that W's payoff as employee can be lower than his payoff as contractor, if that is what maximizes E's profit. An employee who were to say, "From now on I wish to be a contractor and charge you higher average rates" would simply be fired. In contrast, Baker et al. (2002) assume reversion to spot trade because there are only two parties.

the disagreement point for bargaining. The difference matters because when bargaining is naturally biased towards the respective disagreement point (Farrell 1987, McKelvey and Page 2002). Bargaining within employment is likely to be relatively more rare, however, if the unilateral adaptation of E's orders to the state of nature leaves less scope for Pareto-improvements than does the no-trade default of the market. Spot adjustments unambiguously benefit E, tipping her optimal choice of structure towards employment. Even if W has the bargaining power to reap all the benefits immediately, E can partially extract part of W's rent through a lower wage, as a compensating differential. However, allowing for spot adjustments within employment complicates the analysis without leading to much different results. Here, therefore, I stick with the simpler version that rules out "haggling" within employment.

With spot trade, the default benefit and cost are $b^0 = c^0 = 0$, and E's and W's expected net gains from bargaining are

$$V^0 = E_{b,c}\left[b^{m^0(b,c)} - t^0(b,c)\right] \text{ and } U^0 = E_{b,c}\left[t^0(b,c) - c^{m^0(b,c)}\right].$$

With employment, given E's optimal order $m^E(b,c)$, the expected benefits and costs are

$$b^E = E_{b,c}\left[b^{m^E(b,c)}\right] = \Sigma_b \Sigma_c p(b) q(c) b^{m^E(b,c)} \text{ and } c^E = E_{b,c}\left[c^{m^E(b,c)}\right],$$

and E's and W's expected employment payoffs in each period are

$$V^E = b^E - w \quad \text{and} \quad U^E = w - c^E. \tag{2}$$

Denote the efficient action for each b and c, which may equal no work, by $m^*(b,c) = \arg\max_{l\in\{0,\ldots M\}}\{s^l\}$. Then the expected first-best surplus is $S^* = E\left[s^{m^*(b,c)}\right]$.

3.2 Enforcement constraints and E's profit

Let $b_L = \min B$ and $c_H = \max C$. By symmetry of p, any action $l \in A$ may be optimal for E, and by symmetry of q and independence of p and q, c_H is a possible realization for any $l \in A$. Therefore, E will order with positive probability an action for which W's cost is c_H. W can quit in this case, and receive a zero payoff now and in the future. The most restrictive condition that ensures S's obedience therefore is

$$w - c_H + \sum_{\tau=1}^{\infty} \delta^\tau U^E \geq 0$$

or simpler,

$$(1-\delta)(w - c_H) + \delta U^E \geq 0. \tag{3}$$

E, in turn, may wish to fire W if her highest benefit is b_L and if $b_L < w$ (but can do so only before W executes the action, otherwise she has to pay w). The most restrictive constraint to prevent that is

$$(1-\delta)(b_L - w) + \delta V^E \geq 0.\text{[18]} \tag{4}$$

Denoting the total surplus from employment by $S^E = U^E + V^E$, and combining (3) and (4) to a joint enforcement constraint, employment as relational contract is feasible if $-(1-\delta)(c_H - b_L) + \delta S^E \geq 0$, or

$$\delta \geq \delta_0 = \frac{c_H - c_L}{S^E + c_H - b_L}. \tag{5}$$

Thus, not surprisingly, employment is feasible only if δ exceeds a lower bound. In what follows, however, I will focus on δ for which (5) is slack, and examine the forces that determine E's choice between spot trade and employment. Given feasibility, the profit-maximizing employment contract pays the lowest possible wage w that satisfies (3).

It will be convenient to express E's profits in relation to the first-best surplus. Define the expected deadweight loss $D^i, i \in \{0, E\}$, as the difference between S^* and the expected surplus from spot trade or employment, respectively.

For spot trade, the identity $S^0 = U^0 + V^0 + D^0$ leads to the profit

$$V^0 = S^* - U^0 - D^0. \tag{6}$$

That is, E's profit is the first-best surplus, minus rents that the worker earns, minus any

[18] See the discussion about commitment to wage payments above. If instead E could fire W after he has worked but before paying a wage, the constraint would be $(1-\delta)(-w) + \delta V^E \geq 0$. This would raise the minimum discount rate needed for employment to be feasible, but would otherwise not affect any of the paper's results.

remaining deadweight loss.

For employment, the optimal wage w is determined by (3) and (2):

$$w = c^E + (1-\delta)(c_H - c^E). \tag{7}$$

According to (7), W needs to be compensated for the expected cost c^E associated with E's orders. In addition, a premium $(1-\delta)(c_H - c^E)$ is required to satisfy W's obedience constraint (3), which resembles an efficiency wage premium (see also Van den Steen 2010).

Plug (7) for w into E's profit (2), and then replace $b^E - c^E$ using the identity $S^E = b^E - c^E + D^E$, to obtain

$$V^E = S^* - D^E - (1-\delta)(c_H - c^E), \tag{8}$$

where the last term is W's payoff, $U^E = (1-\delta)(c_H - c^E)$. Thus, with employment, E can capture the first-best surplus, minus any deadweight loss from inefficient orders, minus the obedience wage premium $(1-\delta)(c_H - c^E)$, which disappears as δ nears 1.

3.3 E's choice between employment and spot trade

Combining (6) and (8), we obtain

Proposition 1. *The difference between E's profits from employment and from spot trade is given by*

$$\Delta V = V^E - V^0 = D^0 - D^E - (1-\delta)(c_H - c^E) + U^0, \tag{9}$$

which is increasing in δ.

Proposition 1 shows that the profitability of employment relative to spot trade is driven by three effects. To begin with, employment is more profitable the more efficient it is relative to spot trade, as reflected by the term $D^0 - D^E$. Without imposing more structure on the model, however, nothing more can be said about this formally, and I postpone a discussion until Sections 4 and 5.

Second, the profit is reduced by the obedience wage premium $(1-\delta)(c_H - c^E)$. As a result, E may prefer spot trade even if employment is more efficient. Thus, employment requires not only a minimal discount factor to be *feasible*, but requires a higher discount factor to be *profitable* for E.

Third, the term U^0 in (9) illustrates a fundamental asymmetry between spot trade and employment when workers are wealth-constrained. A worker's rent as employee converges to zero as δ approaches 1, due to upfront competition that drives down the equilibrium wage. A worker's rent U^0 as contractor, however, is invariably lost to E, and in the limit $\delta \to 1$, ΔV reduces to $D^0 - D^E + U^S$. This result formalizes two connected arguments made by Williamson (1975): first, that spot trade is "impaired by ... problems of opportunism" (p. 72), and second, that workers are unlikely to be able to "submit lump sum bids for jobs at the outset" (p. 69) to compensate E upfront for any rents they might earn later.

It also follows that E may choose employment not just for efficiency reasons but to capture value. This result sheds light on a debate in the 1970s and 1980s about the origins of capitalist production, which to date has lacked any formal analysis.[19] On one side of the debate, Williamson (1985) and others argued that capitalist factory production emerged because it was more efficient, technologically as well as organizationally. On the other side, Marglin (1974) and others contended, in Marxian tradition, that the "social function of hierarchical work organization is not technical efficiency, but accumulation" (Marglin). Proposition 1 shows that these positions are not mutually exclusive. Efficiency obviously matters for E's optimal choice because the surplus difference $\Delta S = D^0 - D^E$ is one part of the profit difference ΔV. However, U^0 is another part of ΔV, which captures E's motive to extract surplus from W by hiring him as employee.

While it may be impossible to quantify efficiency and rent extraction motives empirically, it may well have been on 18th-century capitalists' minds that "the move from the putting-out to the factory system... involved... the disappearance of labourers' opportunities to be opportunists!" (Pitelis 1991, cited in Cowling and Sugden 1998). Especially factory production that did not involve machines could have been less *efficient* than putting-out because of the managerial hurdles of organizing reluctant workers into factories, but more *profitable* if it enabled merchant-manufacturers to hold employed workers to a low wage (see Section 4). Only a model that incorporates wealth constraints, or some other reason for the absence of upfront transfer payments, is capable of capturing Marxian arguments about capitalists' value-capturing motives.

The profit difference ΔV increases with δ ecause the required obedience wage premium is lower the more W values future payoffs. This leads to the prediction that the more

[19] See Marglin (1974), Landes (1986), Williamson (1985), Rebitzer (1993), Clark (1994), Kieser (1994), Cowling and Sugden (1998).

frequently E and W interact, the more likely E prefers employment. This prediction is supported by Abraham and Taylor's (1996) evidence that firms outsource labor to smooth out fluctuations in demand for labor, or when they need specialists but only infrequently.

4 One task

A model with only one task and binary types suffices to derive several main insights of the paper. The case of one task is important because extreme specialization characterized countless jobs from the 18th century until well into the 20th century, and still exists today.[20]

In each period, W can either work for E ($m = 1$) or not work ($m = 0$). E's resulting benefit is mb with $b \in \{b_H, b_L\}$, where $b_H > b_L$ and $\Pr(b = b_H) = \beta$. W's cost is mc with $c \in \{c_L, c_H\}$, where $c_L < c_H$ and $\Pr(c = c_L) = \gamma$. Let $\Delta c = c_H - c_L$, $\Delta b = b_H - b_L$, $\bar{c} = E(c) = \gamma c_L + (1-\gamma)c_H$, and $\bar{b} = E(b) = \beta b_H + (1-\beta)b_L$. I assume $c_L < b_L < c_H < b_H$, which means that trade is efficient with positive probability and inefficient with positive probability. For this case, Bolton and Dewatripont (2005, Section 7.2.1) show that an efficient bargaining mechanism may or may not exist, as it never does with overlapping continuous types, per Myerson and Satterthwaite (1983).

With spot trade, E's optimal take-it-or-leave-it offer to W is straightforward. Type b_L can trade only with worker c_L, and thus optimally offers $t = c_L$. Type b_H can either trade with both worker types at $t = c_H$, or trade only with type c_L at $t_L = c_L$. If low-cost workers are more rare (low γ), then E opts for the higher price. In that case trade is efficient but W earns a rent Δc if $(b, c) = (b_H, c_L)$. For larger γ, E opts for the low price. In that case, W earns no rent but the parties fail to trade if $(b, c) = (b_H, c_H)$. Thus,

Lemma 1. *Let $\widehat{\gamma} = (b_H - c_H)/(b_H - c_L)$.*

(a) If $\gamma \le \widehat{\gamma}$, E's optimal spot-trade prices are $t(b_L) = c_L$ and $t(b_H) = c_H$. Trade is

[20] The best known example of 18th century labor specialization is Adam Smith's description of a pin factory. Much larger and better documented, however, were Josiah Wedgwood's pottery works, in which "his workmen ... were trained to one particular task and they had to stick to it. ... Out of the 278 men, women and children that Wedgwood employed in June 1790, [all but five] were specialists." (McKendrick 1961). The specialization of work reached its peak in the early 20th century with the arrival of mass production, combined with the application of Frederick Taylor's Scientific Management (Lindbeck and Snower 1996, Hu 2013, Kranzberg and Hannan 2017). Efforts to reorganize work from "tayloristic" to "holistic" (Lindbeck and Snower 1996, 2000), including a wider variety of tasks, only gained traction in the 1970s, i.e. during the last 50 of over 250 years of industrial-age wage labor.

efficient ($S^0 = S^*$, $D^0 = 0$), W's expected rent is $U^0 = \beta\gamma\Delta c$ and E's profit is $V^0 = S^* - U^0$.

(b) If $\gamma > \widehat{\gamma}$, E's optimal prices are $t(b_L) = t(b_H) = c_L$. Trade is inefficient ($D^0 = \beta(1-\gamma)(b_H - c_H)$ and $S^0 = S^* - D^0$), W's rent is zero, and E's profit is $V^0 = S^0 = S^* - D^0$.

With employment, E's control over W's action means that W is relationally obligated to work for E in each period, which is efficient except in the state (b_L, c_H). Therefore,

Lemma 2. *Employment leads to an expected deadweight loss of $D^E = (1-\beta)(1-\gamma)(c_H - b_L)$; thus $S^E = S^* - D^E$. E's profit is*

$$V^E = S^* - D^E - (1-\delta)(c_H - c^E) = S^* - (1-\beta)(1-\gamma)(c_H - b_L) - (1-\delta)(c_H - c^E). \quad (10)$$

Combining Lemmas 1 and 2, we obtain

Proposition 2. *Let $\Delta S = S^E - S^0$. Then*

$$\begin{aligned} &\text{If } \gamma \le \widehat{\gamma}: \quad \Delta S = -D^E, \quad \Delta V = U^0 - D^E - (1-\delta)\gamma\Delta c, \\ &\text{If } \gamma > \widehat{\gamma}: \quad \Delta S = D^0 - D^E, \quad \Delta V = D^0 - D^E - (1-\delta)\gamma\Delta c, \end{aligned}$$

with U^0, D^0, and D^E as defined in Lemmas 1 and 2. Both ΔS and ΔV are strictly increasing in $\bar{b}$ and in β, and strictly decreasing in $\bar{c}$, holding Δb and Δc fixed.

All proofs not in the text are in Appendix B. Lemmas 1 and 2 and Proposition 2 (and their counterparts for the two-task case in Section 5), though mathematically simple, have wide-ranging implications:

1. Employment as *predictable* labor supply, not adaptation: The most basic conclusion from Proposition 2 is that employment can be profitable even if W has only one task. In this version of the model, the purpose of employment is not to adapt the worker's labor supply to changing business needs, but to ensure a *predictable* supply of labor, when spot contracting would expose the entrepreneur to either seller opportunism or a shortfall of labor supply.

This result aligns well with historians' accounts of the rise of factory production in 18th century England. Until then, production of many traded goods took place under the putting-out system, inwhich "merchant-manufacturers 'put out' raw materials...to dispersed cottage labor, to be worked up into finished or semi-finished product" (Landes 1966, cited in Williamson 1985: 215). Although the domestic producers sold services not products, they determined their own work hours and were paid by the piece.[21]

But the putting-out system was "plagued by problems of irregularity of production" (Braverman 1974, cited in Williamson 1985: 232). Specifically,

> "As competition increased, production deadlines became both more important and more difficult to ensure in the absence of a regulated workforce. At the same time labour retained some independence in determining the rhythm and pace of work; an independence most frequently exerted and hardest felt in circumstances where the capitalist could have obtained the greatest profits" (Hudson 1981).

Consequently, well before the availability of machines, production in many industries gradually shifted from workers' homes to factories, and workers transitioned from contractors to wage laborers (Mokyr 2009, Chapter 15), consistent with the logic of this section's results.

2. Quasi-rents and integration: Proposition 2 states that both the difference in equilibrium surplus and the difference in E's profit between employment and spot trade are increasing in the first-best surplus (the appropriable quasi-rent) between E and W, as driven by $\bar{b}$ and $\bar{c}$. Proposition 2 thus formalizes a key prediction of transaction-cost economics "that any increase in quasi-rents will increase the likelihood of vertical integration (a finding that is so far consistent with nearly all of the existing empirical literature)" (Whinston 2003). In my model, this prediction holds because the larger quasi-rents from trade, the more likely it is optimal to shift control rights from the seller, where no trade is the default, to the buyer, where trade is the default. In contrast, as Holmström and Roberts (1998), Whinston (2003), and Gibbons (2005) discuss, transaction-cost models tend to *assume* that adaptation is more efficient within firms (or equivalently, that haggling over rents matters less), whereas property-rights theory makes no prediction

[21] See also Landes (1969, chapter 2), Millward (1981), Hudson (1981), Kieser (1994), Mokyr (2007, Chapter 15).

pertaining to levels of quasi-rents.[22]

3. Transaction costs of markets and hierarchies: Economists and organization theorists have for decades debated how the transactions costs of markets relate to the costs of hierarchical organization, see Demsetz (1988), Ghoshal and Moran (1996), and Gibbons (2005), for instance. In my model, the costs of both structures derive from a common information environment and are mirror images of one another. The cost of the market is a failure to trade when trade would be efficient, as a result of bilateral incomplete information. Both sides are "opportunistic", not just the seller; in particular, when E makes a price offer, then inefficiency results if E gambles on too a low price. "Haggling" is simply an attempt to negotiate away from a no-trade default, but one that sometimes fails.

The costs of hierarchy, in this case employment, are by nature the same, except that inefficiency takes the form of excessive trade. Like in the market, bargaining away from the default could improve the outcome but wouldn't always succeed, again due to incomplete information (as discussed in Section 4, allowing for bargaining within employment would be methodologically natural but complicates the analysis without much gain).

That said, there are two sources of asymmetry between market and hierarchy: One is the obedience wage premium that is required to sustain employment; it does not affect efficiency but matters for E's profit. The other asymmetry shows up when W's job has more than one task, see Section 5. There, the market default is still no trade, whereas with employment the default is E's optimal order.

4. The non-role of E's information: The efficiency and profitability of employment does not hinge on any information advantage on part of E, in contrast to a view that runs through most of the literature on employment.[23] To see this, consider the limit case $\beta = 1$, for which the types reduce to $b_H > c_H > c_L$. Here, only W knows the state of the world, and yet employment may be more efficient and more profitable for the same reasons as discussed above: With spot trade, *any* incentive-compatible bargaining mechanism will

[22] On TCE, see Masten (1986), Tadelis (2002), and Tadelis and Williamson (2013, Assumption 1). As the latter note, most assumptions of TCE have various microfoundations, but "a formal integration of these micro-foundations has not been performed" (Tadelis 2002).

[23] See, for example, Bolton and Rajan (2003), (2004), Hart and Moore (2008), Marino et al. (2010), and Rantakari (2021), as well as Baron and Kreps (2013). However, the idea that a boss's information is the rationale for her authority goes back to Knight, Coase, Barnard, and Simon.

lead to either a rent for W or a deadweight loss or generally both, i.e. $U^0 + D^0 > 0$ regardless of δ, whereas with employment, W's rent shrinks to zero as as $\delta \to 1$.

5. The essential role W's opportunity costs: While variations in E's benefit do not matter for the nature of employment, variations in W's cost are essential. For instance, in the limit case $\gamma = 1$, for which the types reduce to $b_H > b_L > c_L$, both employment and spot trade lead to the first-best. It is only when W's cost is stochastic and sometimes exceeds b that both governance structures entail costs from E's perspective: with spot trade, because any bargaining procedure leads to a deadweight loss or a rent for W, and with employment, because work can be inefficient and because it requires an obedience wage markup.

The role of W's costs in my model aligns with historians' accounts of workers' resistance to the introduction of factory production during the 18th and 19th centuries, when workers had to learn to "obey ... that most tyrannical of all masters: the clock" (Mokyr 2009: 340).

> "The feeling of repulsion which [the factory] aroused is easily understood, as, to a man used to working at home, or in a small workshop, factory discipline was intolerable. Even though at home he had to work long hours to make up for the lowness of his wage, yet he could begin and stop at will, and without regular hours. He could divide up the work as he chose, come and go, rest for a moment, and even, if he chose, be idle for days together." (Mantoux 1928, p. 419)

"As a result of this attitude, attendance was irregular, and the complaint of Edward Cave, in the very earliest days of industrialization, was later re-echoed by many others: 'I have not half my people come to work to-day, and I have no great fascination in the prospect I have to put myself in the power of such people." (Pollard 1963). And although workers had little bargaining power, they were not captive to the factories they worked in. Accordingly, workers' resentment to factory discipline translated into "a substantial premium to work under such conditions" (Clark 1981), consistent with the obedience markup in my model.

Today, workers' preferences for control over their work hours are a key supply-side driver of the rise of contract labor. "The IW [independent workers] set their own hours, both in terms of overall quantity and the specific times they work. This allows them to balance work, family, and leisure activities better than traditional employees" (Oyer 2016). Chen et al. (2019) estimate Uber drivers' opportunity costs of working over short time intervals, and find large fluctuations.

6. Drivers of the employee-contractor choice: (a) Firm size. Smaller firms are likely to have less frequent interaction with a supplier, and/or are less likely to have a high valuation for a supplier's service. Firm size is thus likely correlated with δ and $\overline{b}$, and Propositions 1) and 2) thus suggest that small firms are more likely to rely on contractors, consistent with the fact that "80% of buyers using [freelancing platform] Upwork's website are businesses with 10 or fewer employees" (Oyer 2016).

(b) Asset ownership. Although assets are not part of my model, asset ownership still arguably affects E's and W's benefits and costs. Specifically, shifting a productive asset from W to E arguably lowers c (by lowering W's outside options) and raises b (by raising E's demand for labor to complement the asset's productive capacity), which per Proposition 2 increases the relative profitability of employment. This interpretation holds irrespective of whether asset ownership is endogenous to the organizational design problem (as in Holmström and Milgrom (1991) or Van den Steen (2010), for instance), or is instead a symptom of the wealth distribution between E and W and thus a *cause* of employment. On the flip side, just as a lack of productive assets can drive workers into employment, the increasing irrelevance of non-human assets in a knowledge economy may well contribute to the concurrent growth of an independent workforce.

(c) W's spread of costs. Perhaps counterintuitively, the effect of W's spread in costs Δc on ΔS or ΔV is ambiguous: Holding $\overline{c}$ fixed, a larger Δc increases the inefficiency of requiring W (as employee) to work when his costs are high, and requires a larger obedience wage premium as well. However, a larger Δc also raises the contractor's rent when E optimally charges a relatively high price (in the case $\gamma < \hat{\gamma}$) to ensure W's participation. The latter effect dominates, for instance, when β is relatively large.

On the other hand, if E's decision to employ a worker is given, and with workers competing for the job, those with smaller Δc will accept a lower obedience wage premium and will underbid workers whose costs vary more strongly. We should therefore expect to see self-selection among workers between employment and gig-worker status according to the fluctuations in their opportunity costs of working, consistent with Chen et al.'s

(2019) large estimates for Uber drivers, and with Oyer's (2016) finding that "68% of IW respondents to the 2015 FIA survey reported that schedule flexibility is a reason they freelance."

5 Two tasks

Coase (1937) hypothesized that employment makes it easier to adapt production to changing business conditions by enabling a boss to direct a worker between different tasks. This argument runs through the entire literature on employment.[24] To compare authority with prices for this situation, let us expand the model of Section 4 to two tasks. I will show that Coase's hypothesis does not hold if, as in my model, the "transaction costs" of hierarchy and market endogenously result from a given information structure. Instead, there are two other reasons for why most employees do in fact have multiple tasks.

5.1 Setup, equilibria, and the basic irrelevance of task multiplicity

Suppose now that there are two mutually exclusive tasks X and Y. With spot trade, E offers prices (t^X, t^Y). W can accept at most one of them and perform the corresponding task, or reject both prices and not work for E. With employment, E orders W to perform the task that leads to the highest benefit for her.

Distribution of benefits $p(b)$

		Task Y		
		b_L	b_H	
Task X	b_L	p_L	p_M	$\Sigma = 1-\beta$
	b_H	p_M	p_H	$\Sigma = \beta$

Distribution of costs $q(c)$

		Task Y		
		c_L	c_H	
Task X	c_L	q_L	q_M	$\Sigma = \gamma$
	c_H	q_M	q_H	$\Sigma = 1-\gamma$

Table 1: Distributions of E's benefits and W's costs in two-task model

E's and W's valuations for each task are the same as in Section 4. Each player's valuations may be correlated across tasks, and given the symmetry of distributions as

[24] Coase (1937), Simon (1951), Williamson (1975), Wernerfelt (1997, 2004, 2015), Bolton and Rajan (2003), Van den Steen (2010), Rantakari (2021), Baron and Kreps (2013).

assumed in Section 3, both $p(b)$ and $q(c)$ can be described by three parameters each, as shown in Table 1. For comparison with the single-task model, assume that the marginal distributions over b^i and c^i are the same as as in Section 4, as also shown in the table.

The covariances p_M and q_M parameterize the correlation of the benefits and costs across tasks. If $p_M = 0$, b^X and b^Y are perfectly correlated; if $p_M = \beta(1-\beta)$, they are uncorrelated; and if $p_M > \beta(1-\beta)$, then b^X and b^Y are negatively correlated. The largest possible value of p_M (corresponding to the most negative correlation) is given by $\min\{\beta, 1-\beta\}$. Similarly, $q_M \in [0, \min\{\gamma, 1-\gamma\}]$ measures (inversely) the correlation between c^X and c^Y.

As discussed in Section 3, both bargaining under spot trade and E's orders as employer select for Pareto-efficient outcomes. Thus, if E as employer is indifferent between tasks X and Y but W is not, she will order the lowest-cost task. With spot trade, if for instance $b = (b_H, b_L)$ and $t = c = (c_H, c_L)$, then W is indifferent between the tasks but E is not if $\Delta b \neq \Delta c$. The bargaining mechanism then selects the more efficient option.

The following result is the first formalization of what, in the context of an agency relationship with mutually exclusive tasks, a "market" solution looks like, in this case for a take-it-or-leave-it offer made by E. Such a formalization is critical for assessing the purported adaptation benefits of hierarchies.

Lemma 3. *For $\widehat{\gamma}$ as defined in Lemma 1, let $\gamma_0 = (1-q_M)\widehat{\gamma}$ and $\gamma_1 = \widehat{\gamma} - \frac{b_L - c_L}{b_H - c_L} q_M$, such that $\gamma_0 \leq \gamma_1 \leq \widehat{\gamma}$. Then with spot trade, E's optimal prices, W's expected payoff U^0, and the expected deadweight loss D^0 are shown in Table 2. In particular, for any distribution of W's costs (and the corresponding value of γ), E's average price per task is lower than with only one task.*

E's type:	(b_H, b_H)	(b_H, b_L)	(b_L, b_L)	U^0	D^0
If $\Delta b \geq \Delta c$:	E's optimal price (t_H, t_L):				
$\gamma \leq \gamma_0$:	(c_H, c_L)	(c_H, c_L)	(c_L, c_L)	$(1-p_L)\gamma\Delta c$	0
$\gamma \in (\gamma_0, \gamma_1]$:	(c_L, c_L)	(c_H, c_L)	(c_L, c_L)	$2p_M\gamma\Delta c$	$p_H q_H(b_H - c_H)$
$\gamma > \gamma_1$:	(c_L, c_L)	(c_L, c_L)	(c_L, c_L)	0	$(1-p_L)q_H(b_H - c_H) + 2p_M q_M(\Delta b - \Delta c)$
If $\Delta b < \Delta c$:	E's optimal price (t_H, t_L):				
$\gamma \leq \gamma_0$:	(c_H, c_L)	(c_H, c_L)	(c_L, c_L)	$(1-p_L)\gamma\Delta c$	0
$\gamma > \gamma_0$:	(c_L, c_L)	(c_L, c_L)	(c_L, c_L)	0	$(1-p_L)q_H(b_H - c_H)$

Table 2: Prices, W's payoff, and deadweight loss with spot trade over two tasks

Prices, W's payoff, and the deadweight loss vary with γ in the same way as with one task, except that now there are up to three cases to distinguish. If γ is small ($\gamma \leq \gamma_0$), E prices conservatively, and there is no deadweight loss but W earns a rent. If γ is large ($\gamma > \gamma_1$), E prices most aggressively, W earns no rent but there is a deadweight loss. E's prices are unambiguously lower than with one task because E can gamble that W will accept a low price for one *or* the other task. For the same reason, E may optimally offer different prices even if her valuations for the two tasks are both high.

A new source of inefficiency, reflected in the term $2p_M q_M(\Delta b - \Delta c)$ in the D^0-column of Table 2, arises if $\Delta b \geq \Delta c$: If $b = (b_H, b_L)$ and $c = (c_H, c_L)$ or vice versa, and E offers a low price for both tasks, then spot trade will lead to the *wrong* task being chosen because W will optimally pick the task with $(b^i, c^i) = (b_L, c_L)$ even though the other task (with $(b^j, c^j) = (b_H, c_H)$) is more efficient. If $\Delta b < \Delta c$, then W's acceptance decisions are always efficient.

With employment, E orders W to perform task $m^E(b, c)$ as defined in Section 3; that is, the task that maximizes E's benefit, with any ties broken in favor of efficiency.

Lemma 4. *With two tasks, the total surplus with employment is* $S^E = S^* - D^E = b^E - c^E$, *and E's profit is* $V^E = S^* - D^E - (1-\delta)(c_H - c^E)$, *where*

$$
\begin{aligned}
D^E &= \begin{cases} p_L q_H (c_H - b_L) & \text{if } \Delta b \geq \Delta c \\ p_L q_H (c_H - b_L) + 2p_M q_M (\Delta c - \Delta b) & \text{if } \Delta b < \Delta c. \end{cases} \\
b^E &= E_{b,c}\left[b^{m^E(b,c)}\right] = (1 - p_L) b_H + p_L b_L \text{ and} \\
c^E &= E_{b,c}\left[c^{m^E(b,c)}\right] = \bar{c} - (p_L + p_H) q_M \Delta c.
\end{aligned}
$$

By construction, E's expected benefit b^E exceeds $\bar{b}$: E picks a high-benefit task unless *both* tasks yield a low benefit. And in spite of the independence of benefits and costs, because E orders the lower-cost task if she is indifferent, W's expected cost c^E is not $\bar{c}$ but smaller.

Note that neither the equilibrium surplus nor E's profit depend on whether $\Delta b \lesseqgtr \Delta c$; only the first-best surplus S^* does, as expressed through D^E. Out of 16 states of the world, E's order is inefficient in either one or three states. In the state $(b, c) = (b_L, b_L, c_H, c_H)$, W should not work at all but will be ordered by E to perform X or Y. If $b = (b_H, b_L)$ and

$c = (c_H, c_L)$ or vice versa, then E's most profitable task (yielding b_H)is also the efficient task as long as $\Delta b \geq \Delta c \Leftrightarrow b_H - c_H \geq b_L - c_L$. If $\Delta b < \Delta c$ in this state, however, then $b_H - c_H < b_L - c_L$, and E will order the less efficient task, which is represented by the second term of D^E in Lemma 4.

Proposition 2 from Section 4 generalizes to two tasks: Employment is both more efficient and more profitable relative to spot trade the greater the surplus (the appropriable quasi-rent) at stake:

Proposition 3. *For all cases covered by Lemmas 3 and 4,*

(a) ΔS and ΔV are strictly increasing in $\bar{b}$, holding Δb constant.

(b) ΔS and ΔV are strictly decreasing in $\bar{c}$, holding Δc constant.

Comparing Lemmas 3 and 4 with Lemmas 1 and 2, we see that the tradeoffs between employment and spot trade are the same whether W has multiple tasks or just one. The cost (to E) of spot trade is a mix of deadweight loss and rents surrendered to W that depends on the distribution of W's costs; the cost of employment is the deadweight loss of forcing W to work when his cost exceeds the benefit, plus the obedience wage premium. In addition, both structures can lead to the wrong task being picked, depending $\Delta b \lesseqgtr Deltac$.

Employment, then, has no particular adaptation advantage over the market unless one assumes, as Coase seemed to, that determining prices entails higher cognitive costs than giving orders. Both the employed and the independent IT specialist will manage my database and not my website, if that is what I currently need. The real difference between the two is that the contractor can decline my request for service, or demand a surcharge, whereas the employee is relationally required to work for me at an agreed price. This difference, however, is already present with only one task.

More formally, if $b = (b_H, b_L)$, then with employment, E orders W to perform task X. But if W is a contractor, E can direct W just as effectively by offering prices $t = (c_H, c_L)$.[25] E might prefer to offer $t = (c_L, c_L)$ to gamble on W having a low cost, which could cause trade to fail. But that would just be a case of market failure due to "opportunistic haggling" (Williamson 1975), not a failure of the market to adapt to a firm's demands.

[25] As Alchian and Demsetz's (1972) put it, "telling an employee to type this letter rather than to file that document is like my telling a grocer to sell me this brand of tuna rather than that brand of bread."

5.2 The endogenous design of jobs with multiple tasks

If adaptation isn't the main advantage of employment, then why do most employees have jobs with many different tasks? I will show that the likely answer is that often, bundling multiple tasks into one job is the only way for employment to be economical for E.

First, the obedience wage premium creates scale economies in job design. Suppose E could employ two workers, one for each task, who are available whenever their task is requested and who do not work, or get paid, otherwise. The next result show that it is less costly for E to employ one worker instead who can perform tasks X or Y as needed:

Proposition 4. *The total wage cost of employing two workers, one for each task, who do not work when not needed, exceeds the wage of one worker who can perform either task.*

The proof is simple and instructive: For a single worker, the required wage is given by (7). If the job is split symmetrically between two workers, then the expected cost of working is $c^E/2$ and the required wage for each is

$$w = \frac{c^E}{2} + (1-\delta)\left(c_H - \frac{c^E}{2}\right).$$

Clearly, paying twice this wage exceeds the wage (7) for a single worker. □

Intuitively, by splitting tasks among different workers, the total expected cost of work c^E can be split in a cost-neutral way. An obedience premium, however—the extra payment required for a worker to comply with a request to perform his task even if his cost happens to be high—must be paid to each worker, which is more costly than to pay a premium to just one worker.

Proposition 4 suggests that single-task employment relationships are unlikely to exist unless a task is demanded regularly enough to justify a regular wage payment.[26] For small businesses, for instance, the small scale of demand for many individual tasks effectively reduces the choice set of labor arrangements to either hiring employees as "jacks of all trades", or outsourcing tasks to specialists.

Another force towards the bundling of tasks into jobs is the correlation of E's benefits and W's costs across tasks. Coase (1937) hypothesized that authority relationships

[26] The result does not suggest an advantage of full-time over part-time employment. For timed work, the model is best interpreted as applying to one time unit, such as an hour. The resulting wage is then an hourly wage that contains an obedience premium, and the model is silent on the number of hours for which a job is configured.

are most likely if it is "a matter of indifference to the person supplying the service or commodity which of several courses of action is taken, but not to the purchaser of that service or commodity."[27] This turns out to be mostly correct.

Proposition 5. *Spot trade is less profitable relative to the first best ($S^* - V^0 = D^0 + U^0$ is larger) the more W's costs are correlated.*

A high correlation between c^X and c^Y means that W cares relatively little which of the two tasks he performs. The flip side, however, is a high probability that W would prefer to do neither because both costs are high. Consequently, a higher correlation between the costs makes spot trade less efficient (as stated in the proposition) but does not necessarily make employment more efficient or more profitable, leaving the net effect on ΔV ambiguous.

The distinction between the extensive margin (whether to work) and the intensive margin (which task to work on) has not been made in the literature but is critical for understanding employment and alternative market arrangements. For instance, Uber drivers' preferences over destinations presumably vary little over time, but their opportunity cost of working at all vary a lot, as Chen et al. (2019) document. Contrary to Coase's and Simon's intuition, then, indifference across tasks does not necessarily favor employment.

The picture is clearer when we look at the correlation of E's benefits:

Proposition 6. *Spot trade is less profitable relative to the first best ($S^* - V^0 = D^0 + U^0$ is larger) the less E's benefits are correlated. Employment is more profitable relative to the first best ($S^* - V^E = D^E + (1-\delta)(c_H - c^E)$ is smaller) the less E's benefits are correlated across tasks, as long as*

$$q_H(c_H - b_L) + 2q_M(\Delta b - \Delta c) \geq 0. \tag{11}$$

It follows that $\Delta V = v^E - V^0$ is decreasing in the correlation of E's benefits if (11) holds.

Condition (11) is a sufficient condition; it always holds if $\Delta b \geq \Delta c$ or if W's costs are highly correlated (q_M small). Intuitively, when E prefers one task over the other,

[27] Similarly, Simon (1951) hypothesized that "W will be willing to enter an employment contract with B only if it does not matter to him 'very much' which x ... B will choose.".

spot trade is more likely to lead to a rent for W or a deadweight loss. Meanwhile, with employment, E's default orders are more likely to be efficient, and the obedience wage premium is smaller too. In particular, a negative correlation between b^X and b^Y means that E has high demand for X in some periods and for Y in others, or (more loosely interpreted) can shift these actions across time to achieve a negative correlation. Consequently, E has a high valuation for W to perform one or the other task in most periods.

	Example 1		**Example 2a**		**Example 2b**	
β	0.5		0.5		0.5	
(p_H, p_L, p_M)			$(0.25, 0.25, 0.25)$		$(0, 0, 0.5)$	
γ	0.5		0.5		0.5	
(q_H, q_L, q_M)			$(0.25, 0.25, 0.25)$		(0.5,0.5,0)	
b^E	8		9		10	
c^E	6		5.5		6	
FB surplus S^*	2.5		3.625		4	
DWL (D^0, D^E)	0.5	0.5	0.375	0.125	1	0
W's rent (U^0, U^E)	0	$2(1-\delta)$	0	$2.5(1-\delta)$	0	$2(1-\delta)$
Eq. surplus (S^0, S^E)	2	2	3.25	3.5	3	4
E's profit (V^0, V^E)	2	2δ	3.25	$1+2.5\delta$	3	$2(1+\delta)$
δ_0 per (5)	0.5		0.36		0.33	
δ s.t. $\Delta V = 0$	n/a		0.9		0.5	

Table 3: One-task example with different extensions to two tasks. Example 2a: uncorrelated benefits and costs; Example 2b: negatively correlated benefits, positively correlated costs. For all examples, $(b_H, c_H, b_L, c_L) = (10, 8, 6, 4)$ and $\beta = \gamma = 0.5$, and therefore $\gamma > \gamma_1 > \widehat{\gamma} = \frac{1}{3}$.

Propositions 5 and 6 have profound consequences for job design, as they imply that bundling tasks with the right correlation structure into one job may *make* employment more profitable than spot trade. To illustrate, Example 1 in Table 3 is a one-task job as studied in Section 4. Spot trade and employment are equally efficient ($D^E = D^0$) and employment is feasible if $\delta \geq 0.5$, but it is never optimal for E because of the obedience wage premium. Examples 2a and 2b extend Example 1 to two tasks, with the same marginal distributions as in Example 1 but different correlation structures. In Example 2a, the benefits and costs of tasks X and Y are independently distributed. Because of the significantly higher first-best surplus compared to Example 1 (3.625 vs. 2.5), employment can be more profitable than spot trade, but only if the discount factor is quite large ($\delta \geq 0.9$). In eEample 2b, E's benefits are perfectly negatively correlated, and W's costs perfectly positively correlated. Not only is the first-best surplus higher than in Example

2a, but employment now attains the first-best surplus. Consequently, employment is more profitable than spot trade for a much larger range of δ, namely $\delta \geq 0.5$ vs. $\delta \geq 0.9$ in Example 2a. Thus, while a firm may not have enough demand to hire different specialists to (in Alchian and Demsetz's words) "type this letter" or to "file that document," it is more likely to have demand for one worker to do one or the other at different times, especially if the worker cares little which task is requested.

Proposition 6 suggests a causal link between employment and task multiplicity that essentially reverses common intuition. The argument that decision-making by authority is efficient when workers perform different tasks implicitly takes job design as given. Example 2b, by contrast, suggests that hiring people as employees may require bundling tasks that are each demanded only intermittently into one job, to ensure that the employee always has something to do. This argument provides a new economic rationale for how to bundle tasks into jobs that is arguably more basic than incentive-based arguments (which are reviewed in Lazear and Oyer 2013, Section 6.1).

To illustrate the contrast, Wernerfelt (1997) argues that it is efficient for a secretary to be employed because it would be too costly to negotiate spot prices for the many tasks a secretary performs (see also Fama 1991, Simon 1991). The results above, instead, suggest that the job of a secretary only exists because managers, professors, and others have demand for services that individually would be costly to outsource due to low individual demand, and even costlier to employ specialists for (per Proposition 4), but can be performed by the same person who is paid to be available for whenever demand for a secretarial service arises. Supporting this perspective is that due to academics' own use of computers, today's administrative assistants not only perform different tasks than 25 years ago, but also often serve groups of people rather than individuals in response to lower demand. That is, secretaries' and assistants' job descriptions are endogenous and are driven by organizations' goals to ensure that paying a regular salary is economically viable.

5.3 Worker authority

Would E want to employ W but delegate the choice of tasks to him? Simon (1951) briefly mentions this possibility; in Marino et al. (2010), the link between obedience and delegation is the central question. In my model, employment with delegation is distinct from spot trade because the employee with authority is still required to provide a service whereas the contractor is not.[28]

Lemma 4 suggests that the optimality of E- vs. W-authority may hinge on whether $\Delta b \lesseqgtr \Delta c$. Indeed (as stated in the next result), E-authority is most efficient if $\Delta b \geq \Delta c$, and W-authority if $\Delta c > \Delta b$. Both entail inefficient trade in the state $(b, c) = (b_L, b_L, c_H, c_H)$. But when $(b^X, c^X) = (b_H, c_H)$ and $(b^Y, c^Y) = (b_L, c_L)$ or vice versa, E's order is efficient if $\Delta b \geq \Delta c$, and W's own choice is efficient if $\Delta c < \Delta b$. Nevertheless, as shown in the next result, it tends to be more profitable for E to retain authority for herself.

Proposition 7. *(a) With employment, W-authority is most efficient if and only if $\Delta b < \Delta c$. (b) W-authority is most profitable for E if and only if $\Delta b < \Delta c$ and $\delta > \Delta b/\Delta c$.*

Proof: see the Appendix. If $\Delta b < \Delta c$, then delegating authority to W lowers the cost of W's default actions, c^E, by more than it lowers E's benefit b^E, which is efficient. However, the reduction of c^E is offset by a tighter obedience constraint (3) and hence a larger obedience wage premium $(1 - \delta)(c_H - c^E)$. This explains why delegation is optimal only if δ is sufficiently large or the worker sufficiently better informed ($\Delta c > \Delta b/\delta$).[29]

The option to delegate authority to W suggests a further generalization of Simon's (1951) definition, beyond the inclusion of one-task jobs: *Employment is a relational contract that obligates a worker to work for a firm, in return for a wage. Whether the firm or the worker has control over the worker's actions is endogenous.* The significance of Proposition 7 then is that for the "boss" to have authority over the worker's actions—which empirically is the norm—is not definitionally a feature of employment but happens to be optimal for many jobs, even if delegation would be more efficient.

[28] This distinction of employee and contractor differs from Hart and Moore's (2008), who label a contract with buyer authority "employment" and a contract with seller authority "independent contracting." Hart and Moore consider the example of hiring a musician for a gathering and determining who should pick the music, the client or the musician. Intuitively and legally, however, for a one-time engagement the musician would be a contractor irrespective of who chooses.

[29] Proposition 7 contrasts with Marino et al.'s (2010) conclusion that delegation relaxes the worker's obedience constraint, and may thus be optimally chosen by the employer even though centralization would lead to better decisions. Here it is the other way around: Delegation can only ever be optimal if the worker makes better decisions, but it may not be profitable because the obedience constraint is *tighter*. It is tighter because (i) W's highest possible cost is still c_H even if W has authority, and (ii) W's lower costs are reflected in a lower wage, or put differently, W's utility gain from being in control is neutralized by a negative compensating differential. Effect (i) may be specific to this model's symmetry assumptions, but effect (ii) is likely to be present whenever workers compete for a job.

Finally, Proposition 7 complements Proposition 2, which takes E-authority as given. As discussed, employment can be optimal even if E is not better informed than W, contrary to most of the literature since Knight. Proposition 7 additionally shows that E-authority can be optimal even if W is better informed *and* W-authority is an option. Moreover, employment with delegation is still employment, and may dominate spot trade.

These results helps resolve a puzzle posed by Freeland and Zuckerman (2018): Why are hierarchies and employment still so dominant in the 21st century, even though in the knowledge economy, "managers cannot possibly hold sufficient knowledge to direct the firm efficiently because of the importance of adapting to changing local conditions" ? Freeland and Zuckerman propose their own answer, but one that follows from my analysis is that managers' knowledge advantage is not an important condition for employment and hierarchical decision-making to begin with.

6 Multiple sellers

Why did employment became a dominant way to work during the Industrial Revolution, and why may it be losing its dominance today? To explain the former, I show that complementarity in production across workers favors employment. To explain the latter, I show that ex-post competition among workers favors spot trade.

6.1 Complementarity in production

Employment outside of a context of team production has always existed; early examples include domestic servants and farm workers (Thompson 1967). Nevertheless, the proliferation of wage labor during the 18^{th} century appears to correlate with a greater reliance on the coordinated division of labor in factories (McKendrick 1961, Millward 1981), even before the arrival of machines (such as in Adam Smith's pin factory) but certainly thereafter.

To see how this may matter, suppose that there are two workers W1 and W2 whose (one-task) work is complementary in the sense that successful production of a good requires both workers to provide work. Recall that in the model of Section 4, W's work ($a = 1$) or no work ($a = 0$) leads to a benefit ab for E and a cost ac for W. Now, instead, suppose that each of two workers either works or doesn't ($a_i \in \{0, 1\}$), and that E's resulting benefit *per worker* is bY with $Y = a_1 a_2$, with the supports and distributions of b and c being the same as in Section 4. The workers' costs are independent.

With employment, E's profit is the same as with independent production. If E hires two employees, they will always work and will therefore produce the same output, at the same wage, as with independent production. Work is inefficient if $(b, c_1, c_2) = (b_L, c_H, c_H)$. Additionally, work may or may not be efficient if $(b, c_1, c_2) = (b_L, c_H, c_L)$ or (b_L, c_L, c_H), depending on whether $2b_L - c_H - c_L \geq 0$, but this has no effect on E's profit.

With spot trade, E makes simultaneous price offers $t = (t_1, t_2)$ to the two workers, who each accept or reject. To avoid stacking the cards against spot trade, assume that if, say, W1 accepts but W2 rejects and E thus knows that production cannot succeed, E can rescind the offer to W1.

Proposition 8. *With spot trade, for any parameters of the binary model, E's average price per worker is at least as high with complementary as with independent production, and her profit per worker is the same or lower.*

"Average price per worker" refers to the fact that E may wish to offer the two workers different prices, as shown in the proof. The proposition is intuitive: When E knows that both workers must accept their prices for production to be feasible, she offers higher prices on average and ends up paying higher rents to the workers, or she offers the same prices but the workers are less likely to both accept.

It follows immediately that if production requires multiple people whose tasks are complementary rather than independent, then employment is more profitable relative to spot trade, consistent with observed parallel trends during 18th and 19th century Industrialization.[30]

6.2 Ex-post competition among sellers

The analysis so far implicitly assumed that any given task within W's set of tasks can be performed only by W. Gig economy companies such as Uber, TaskRabbit, Upwork and

[30] Still, as Williamson (1975) argued in response to Alchian and Demsetz (1972), team production is only a contributing factor but not an essential feature of employment: "Our assessment of the technological nonseparability argument thus comes down to this: Such conditions are merely symptomatic of a set of underlying transactional factors which, both here and elsewhere, ultimately explain the organization of economic activity as between markets and hierarchies" (p. 61).

others, however, have created platforms where for any service requested by a customer, there are multiple workers available to do the task.

Consider a version of the model of Section 4 where employment works the same way as before, but with a spot market, there are n sellers available to perform a given task, at i.i.d. costs $c \in \{c_L, c_H\}$. A buyer with $b = b_L$ will then offer $t = c_L$, which is accepted by one of the sellers with probability $1 - (1 - \gamma)^n$. If $b = b_H$, B can offer $t = c_H$ and earn payoff $b_H - c_H$. Or she can offer $t = c_L$, which again is accepted by one of the sellers with probability $1 - (1 - \gamma)^n$, leading to payoff $[1 - (1 - \gamma)^n](b_H - c_L)$, which for any γ exceeds $b_H - c_H$ if n is large enough. Clearly, the more sellers compete, the higher the payoff from spot trade, and we obtain

Proposition 9. *For any $\gamma \in [0, 1]$ there exists $\widetilde{n} \geq 1$ such that for any $n \geq \widetilde{n}$, E's profit from spot trade with n competing sellers exceeds the profit with employment.*

Recall that with employment, E's profit falls short of the first-best surplus because of the deadweight loss of inefficient orders and W's obedience wage premium. With spot trade, however, as n grows, not only does the maximal surplus itself increase because $\min\{c_i\}$ converges to c_L, but E also captures all of the surplus in the limit.

The result reinforces the importance that "small numbers" play for employment as an equilibrium governance structure: "[T]ask idiosyncrasies are common, ... these give rise to small-numbers exchange conditions, and ... market contracting is supplanted by an employment relation principally for this reason" (Williamson 1975: 62). Clearly, if firms could each day fill their positions from a large pool of available workers, employment would not exist. Proposition 9 in addition shows that for employment to be optimal requires not only a somewhat durable relationship between a firm and a worker (which certain specific investments would create), but also a degree of job specialization that precludes within-task substitution between workers on the fly.

Information technology, in particular, has made it possible to create ex-post competition for services that used to require hiring employees. Technology has made it possible to (i) find people to perform ad-hoc tasks, such as those offered by TaskRabbit, (ii) for workers to perform many tasks remotely, such as via Amazon Mechanical Turk, and (iii) to help overcome adverse selection and moral hazard with the help of ratings systems. All of of these are demand-side forces towards an expansion of the freelance and gig economy.

7 The legal status of employment

Coase (1937) concluded his essay by observing that his description of authority relationships largely matched the prevailing legal distinction between employer and employee. The same holds for the definition of employment suggested in this paper, compared to, for instance, the U.S. common-law definition: "Under common-law rules, anyone who performs services for you is your employee if you can control what will be done and how it will be done."[31] Like my model, this definition recognizes that how control is exercised depends on the job: "A business may lack the knowledge to instruct some highly specialized professionals; in other cases, the task may require little or no instruction. The key consideration is whether the business has retained the right to control the details of a worker's performance or instead has given up that right."[32]

That said, legal recognition of employment status has consequences that significantly influence firms' decisions on whether to rely on employees or contractors (for previous analyses of the legal environment, see Masten (1988) and Williamson (1991)). Unlike in the early 20th century, employment laws today specify employee protections as well as employers' obligations to pay various taxes and benefits, and as minimum wages. These protections, taxes, and benefits create an incentive for firms to label and pay their workers as contractors even if economically, their work resembles employment. Because of this incentive, a worker's legal status is defined by the economic nature of the relationship between firm and worker, and not by the label chosen by the firm (see, for example, Internal Revenue Service 2017). Misclassification claims by workers can trigger audits and possibly fines by government authorities (in the U.S., the IRS or state or federal Departments of Labor), or play out in court.

Guidelines used by authorities such as the IRS or the DOL err on the side of classifying workers as employees, recognizing as contractors mainly rather obvious cases such as professionals and tradespeople running their own business. The Fair Labor Standards Act's definition of an employee, in particular, is expressly broader than is implied by common law, and includes anyone "who, as a matter of economic reality, is economically dependent on an employer for work—not for income." (DOL 2022). Finally, going even further than the FLSA, the "ABC test" used by many states stipulates as a requirement

[31] https://www.irs.gov/businesses/small-businesses-self-employed/employee-common-law-employee
[32] https://www.irs.gov/businesses/small-businesses-self-employed/behavioral-control

for contractor status that "[t]he worker performs work that is outside the usual course of the hiring entity's business" (Department of Industrial Relations 2019).[33] Had these more comprehensive definitions of employment applied to 18th century Europe, preindustrial home workers under the putting-out system would likely have qualified as employees if they depended on mainly one merchant for their livelihood.

The growth of the gig economy and of alternative work arrangements has led to a proliferation of legal cases against firms, as well as to intense debate in the legal profession over the normative case for distinguishing employees and contractors (see Posner 2020). A prominent example is Uber. The Trump administration's DOL supported Uber's treatment of its drivers as contractors:

> "Drivers' virtually complete control of their cars, work schedules, and log-in locations, together with their freedom to work for competitors of Uber, provided them with significant entrepreneurial opportunity. On any given day, at any free moment, UberX drivers could decide how best to serve their economic objectives: by fulfilling ride requests through the App, working for a competing rideshare service, or pursuing a different venture altogether" (National Labor Relations Board 2019).

A San Francisco court, however, ruled in August 2020 that according to California's more stringent criteria, Uber drivers are employees because their services are central to Uber's business (Siddiqui 2020). Likewise, a London tribunal opined in 2016: "The notion that Uber in London is a mosaic of 30,000 small businesses linked by a common 'platform' is to our minds faintly ridiculous. In each case, the business consists of a man with a car seeking to make a living by driving it." (Wired 2018). In my model (as in Simon's 1951), Uber drivers would unambiguously be contractors due to the flexibility provided by the job. Neither the centrality of drivers to Uber's business nor the degree of drivers' dependence on Uber for work would matter directly, although both contribute to the *optimality* of employment. That said, Uber has at times employed a variety of techniques, known as "algorithmic management," to nudge drivers to be available at certain times, thus

[33] This test led to California's designation in 2020 of Uber drivers as employees. Proposition 22 was a ballot initiative introduced in the November 2020 state election to grant app-based transportation and delivery companies an exemption from California's statute. It passed with 59% of the vote, including overwhelming support by Uber drivers. The U.K.'s "zero-hours contracts" explicitly seek to define a legal middle ground for gig economy workers; for an analysis, see Dolado et al. (2021).

approximating behavioral control more closely associated with employment, see Rosenblat and Stark (2016).[34]

To conclude, the economic definition of employment suggested in this paper matches quite well with the common-law definition of employment. The FLSA's and many states' definitions of employment, however, extend far beyond behavioral control and in effect consider almost any long-term relationship in which a worker relies on a firm for business to constitute employment. These wide-reaching definitions combined with the mandatory costs of employment are likely to significantly influence firms' staffing choices, independent from the determinants studied in this paper. Examples of evidence of this include Bertrand et al (2021) and Dolado et al. (2022).

8 Conclusion

I have proposed a theory of employment that builds on Simon's (1951) seminal work but leads to markedly different conclusions. Simon compared employment with ex-ante contracting, an option that Williamson (1975, p.65) dismissed as "singularly unsuited to permit adaptation." My model, in contrast, compares employment with its main alternative in practice, (sequential) spot contracting for labor, within a framework of bilateral trade with incomplete information about the benefits and costs of different tasks. The key difference between the two structures is the allocation of decision rights: The contractor-worker does not work unless both parties agree on a spot price; the employee has agreed to work at the entrepreneur's direction for a wage set in advance.

It follows, contrary to the prevailing view in the literature, that a boss's ability to tell a worker *what* to do is not an essential feature of employment. Consistent with the ubiquity of jobs exhibiting extreme specialization, employment is well defined and distinct from contract labor even if the worker has only one task. My model thus generalizes Simon's (1951) definition and at the same time suggests a rationale for employment quite different from adaptation: A firm will optimally employ a worker in order to secure the supply of labor at a good price when its provision is valuable to the firm, whether or not adaptation plays a role too. This rationale helps explain why wage labor emerged during early Industrialization concurrent with the move of industrial production from workers'

[34] I thank Erik Madsen for bringing these practices to my attention.

homes into factories. Employment saves both on the deadweight loss of forgone efficient trade, and on the rents extracted by the worker, under spot trade. The latter represents not an inefficiency, but still a cost to the firm if workers' wealth constraints preclude upfront payments to the firm.

Employment can lead to inefficiency too, however, if a firm's order compels an action whose cost to the worker exceeds the benefit. In addition, employment is costly to the firm because authority over a worker's actions can only be relational, which requires paying a wage markup that ensures the worker's obedience. Again, with a wealth-constrained worker who cannot make upfront payments, the relational nature of authority not only constrains the feasibility of employment but adds a cost to the firm.

Both costs of employment help explain why most jobs do indeed have multiple tasks, a fact that most discussions of employment take as given. Bundling the right tasks into one job makes it more likely that there is always a valuable task for the worker to perform, which is necessary for the firm to be willing to pay a constant wage in each period. Thus, while employment is often considered an efficient arrangement to adapt a worker's actions to shifting business needs, the workers' job may well be endogenously designed to make employment viable in the first place.

Combining all results, factors favoring employment are the surplus at stake (magnitude and frequency of demand for a service), the discount factor (frequency of interaction), a worker's lack of funds to make upfront payments, and the ability to bundle tasks into jobs that make paying a regular wage worthwhile for a firm. Also critical for the relevance of employment is a "small numbers" situation with quasi-rents at stake. Three other things don't matter as much in my model, in contrast to hypotheses that run through a 100-year-old literature on the topic. It is not essential for an employee to have many different tasks, or for the boss to have better information than the worker, or for the labor buyer to be the boss in the first place.

The most obvious limitation of my analysis is the omission of effort; i.e., of a distinction between perfunctory performance ("job performance of a minimally acceptable sort") and consummate performance (an "affirmative job attitude"), in Williamson's (1975: 69) terminology. At least when formal incentive contracting is infeasible, this omission does not bias my results in either direction because effort matters both across and within firms, and because both contracts and commands are incomplete (see Cheung 1983, Simon 1991):

"[C]ourts cannot ... determine whether workers put their energy and inventiveness into the job" (Williamson 1975: 69), just like managerial authority "only obligates employees to perform duties ... in accordance with minimum standards" (Blau 1964: 206).

Eliciting consummate performance likely relies on relational contracts both within and between firms: Provided that performance can be observed, the employee will not just show up but also put forth effort in order to keep his job; so will the contractor, in order to earn repeat business, as in Klein and Leffler (1981). Although only a formal analysis can deliver precise answers, it seems likely that the distinction emphasized in this paper remains relevant: No matter how complex the job, the employee has a boss who exercises authority in some form (Simon 1991), the contractor does not.[35] This would also explain why, precisely as Foss (2002) argued, the knowledge economy has not made authority obsolete, contrary to predictions by some organization theorists. That is, although the Coase-Simon view of authority in its simplest statement may evoke "Theory X" (McGregor 1957) work environments, its key ideas, suitably amended, are no less relevant in the 21st century.

Finally, less amenable to formal analysis but no less important, workers' identification with organizational goals is an important driver of motivation, as Barnard and Simon already observed 80 years ago (see also Cassar and Meier 2018, Freeland and Zuckerman 2018). This matters to a firm's choice between employment and contract labor mainly because identification with the firm and its mission requires group membership that possibly only an employee, not a contractor, can have. Firms' ability to harness identification as motivator then depends not, or not only, on how work is organized contractually (as studied in this paper) but on who perceives himself to be a member of the group; see Gartenberg and Zenger (2021).

References

Abraham, Katharine G., & Susan K. Taylor (1996): Firms' use of outside contractors: theory and evidence. *Journal of Labor Economics*, 14(3), 394-424.

Alchian, Armen, and Harold Demsetz (1972): Production, Information Costs, and

[35] In an opinion piece, Asghar (2013) argues that university presidents have a harder job than CEOs because CEOs can and do exercise authority, including by firing executives, to enforce their direction. University presidents have no such powers in dealing with the most powerful group of employees, tenured faculty.

Economic Organization. *American Economic Review* 62(5), 777-795.

Allen, Robert C. (2009): Engels' pause: Technical change, capital accumulation, and inequality in the British industrial revolution. *Explorations in Economic History* 46 (4), 418-435.

Anderson, Tracy, and Peter Cappelli (2020): Management Without Managers: Expanding Our Understanding of Managing Through the Study of Contractors. Available at SSRN 3875634.

Asghar, Rob (2013): The Toughest Leadership Job Of All (It's Not What You Think), Forbes, November 15, 2013

Athey, Susan, and David A. Miller (2007): Efficiency in repeated trade with hidden valuations. *Theoretical Economics* 2 (3), 299-354.

Ausubel, Lawrence, Peter Cramton, and Raymond Deneckere (2002): Bargaining with incomplete information. In: Robert J. Aumann and Sergiu Hart, eds., *Handbook of Game Theory*, Vol. 3, Chapter 50, pp. 1897-1945. Amsterdam: Elsevier

Bajari, Patrick, Steven Tadelis (2001): Incentives versus transaction costs: A theory of procurement contracts. *Rand Journal of Economics*, 387-407.

Baker, George, Robert Gibbons, and Kevin J. Murphy (2002): Relational Contracts and the Theory of the Firm. *Quarterly Journal of Economics* 117 (1), pp. 39–84,

——— (2011): Relational adaptation. Working paper.

Barnard, Chester (1938): *The Functions of the Executive*, Harvard University Press

Baron, James, and David Kreps (2013): Employment as an Economic and a Social Relationship. In: Gibbons and Roberts, pp. 315-341.

Barrero, Jose Maria, Nicholas Bloom, and Steven J. Davis (2021): Why working from home will stick. National Bureau of Economic Research, No. w28731

Barron, Dan, Jin Li, and Michal Zator (2019): Morale and Debt Dynamics. Working paper.

Bernstein, Lisa, and Brad Peterson (2022): Managerial contracting: A preliminary study. SSRN 4036051.

Bertrand, Marianne, Chang-Tai Hsieh, and Nick Tsivanidis (2021): Contract labor and firm growth in India. National Bureau of Economic Research, No. w29151

Blau, Peter M. (1964): *Exchange and Power in Social Life.* Transaction Publishers

Bolton, Patrick, and Mathias Dewatripont (2005): *Contract Theory*, MIT Press

——— (2013): Authority in Organizations. In: Gibbons and Roberts, pp. 342-372.

Bolton, Patrick, and Ashvin Rajan (2003): The Employment Relation and the Theory of the Firm: Arm's Length Contracting vs Authority. Princeton University, working paper.

Braverman, Harry (1974): *Labor and Monopoly Capital: the Degradation of Work in the Twentieth Century*. New York, Monthly Review Press

Broughton, Kristin (2021): Demand for Part-Time CFOs Heats Up as Startups Raise More Money. *Wall Street Journal*, July 26, 2021

Buttrick, John (1952): The Inside Contract System. *Journal of Economic History*, 12(3), 205-221.

Cassar, Lea, and Stephan Meier (2018): Nonmonetary Incentives and the Implications of Work as a Source of Meaning. *Journal of Economic Perspectives*, 32(3), 215-38.

Chakravarty, Surejeet, and W. Bentley MacLeod, W. B. (2009): Contracting in the Shadow of the Law. *The RAND Journal of Economics*, 40(3), 533-557.

Chen, M. Keith, Peter Rossi, Judith Chevalier, and Emily Oehlsen (2019): The Value of Flexible Work: Evidence from Uber Drivers. *Journal of Political Economy* 127(6), pp. 2735-2794.

Chapman, Stanley D. (1974). The textile factory before Arkwright: A typology of factory development. *Business History Review*, 48(4), 451-478.

Clark, Gregory (1994). Factory discipline. *The Journal of Economic History*, 54(1), 128-163.

Coase, Ronald (1937): The Nature of the Firm. *Economica* 4 (16), pp. 386-405

——— (1960): The Problem of Social Cost. *Journal of Law and Economics* 3, pp. 1-44

Coleman, James S. (1990): *Foundations of Social Theory*, Harvard University Press

Cowling, Keith, and Roger Sugden (1998): The essence of the modern corporation: markets, strategic decision-making and the theory of the firm. *The Manchester School* 66 (1), 59-86.

Demsetz, Harold (1988): The theory of the firm revisited. *The Journal of Law, Economics, and Organization* 4 (1), 141-161.

Department of Industrial Relations (2019): Independent contractor versus employee. State of California Department of Industrial Relations, retrieved from https://www.dir.ca.gov/dlse/f on Oct. 14, 2020

Department of Labor (2022): Notice of Proposed Rule: Employee or Independent Contractor Classification Under the Fair Labor Standards Act, RIN 1235-AA43. Retrieved from https://www.dol.gov/agencies/whd/flsa/misclassification/rulemaking on October 27, 2022

Dolado, Juan Jose, Lalé, Etienne and Helene Turon (2021): Zero-Hours Contracts in a Frictional Labor Market. IZA Discussion Paper No. 14979, Available at SSRN: https://ssrn.com/abstract=4114506

Dow, Gregory K. (1987): The function of authority in transaction cost economics. *Journal of Economic Behavior & Organization*, 8(1), 13-38.

——— (1993): Why capital hires labor: A bargaining perspective. *The American Economic Review*, 118-134.

Engels, Friedrich (1845) *The condition of the working class* Otto Wigand, Leipzig

Fama, Eugene. F. (1991): Time, salary, and incentive payoffs in labor contracts. *Journal of Labor Economics*, 9(1), 25-44.

Farrell, Joseph (1987): Information and the Coase theorem. *Journal of Economic Perspectives*, 1(2), 113-129.

Foss, Nicolai J. (2002): 'Coase vs Hayek': Economic Organization and the Knowledge Economy. International Journal of the Economics of Business, 9(1), 9-35.

Freeland, Robert F., & Zuckerman Sivan, Ezra W. (2018): The problems and promise of hierarchy: Voice rights and the firm. *Sociological Science*, 5, 143-181.

Gartenberg, Claudine M., and Todd Zenger (2021). The firm as a subsociety. Available at SSRN.

Gibbons, Robert (2005): Four formal(izable) theories of the firm? *Journal of Economic Behavior & Organization* 58 (2), pp. 200-245.

——— (2010): Transaction-cost economics: past, present, and future? *Scandinavian Journal of Economics* 112, no. 2, pp. 263-288.

——— and John Roberts (eds.) (2013): *Handbook of Organizational Economics*, Princeton University Press

Ghoshal, Sumantra, and Peter Moran (1996): Bad for practice: A critique of the transaction cost theory. *Academy of management Review* 21 (1), 13-47.

Hart, Oliver (1995): *Firms, Contracts, and Financial Structure.* Clarendon Press.

Hart, Oliver (1989): An Economist's Perspective on the Theory of the Firm, *Columbia Law Review* 89, 1757-1774.

——— (1996): An economist's view of authority. *Rationality and Society*, 8(4), 371-386.

——— and John Moore (1990): Property Rights and the Nature of the Firm. *Journal of Political Economy*, 98(6), 1119-1158.

——— (2008): Contracts as Reference Points. *Quarterly Journal of Economics* 123 (1). pp. 1-48.

Sociology, 18(2), 104-132.

Hay, Douglas, and Paul Craven (1993): Master and Servant in England and the Empire: A Comparative Study. *Labour/Le Travail* 31, pp. 175-184.

Holmström, Bengt (1999): The firm as a subeconomy. *Journal of Law, Economics, and Organization*, 15(1), 74-102.

———, and Paul Milgrom (1991): Multitask Principal-Agent Analyses: Incentive contracts, Asset Ownership, and Job Design. *Journal of Law and Economics* 7, pp. 24-52.

——— (1994). The firm as an incentive system. *American Economic Review*, 84 (4), pp. 972-991.

Holmström, Bengt, and John Roberts (1998). The boundaries of the firm revisited. *Journal of Economic Perspectives*, 12(4), 73-94.

Hu, S. Jack (2013): Evolving paradigms of manufacturing: from mass production to mass customization and personalization. *Procedia CIRP* 7, pp. 3-8.

Hudson, Pat (1981): Proto-industrialisation: the case of the West Riding Wool Textile Industry in the 18th and early 19th centuries. In: *History Workshop Journal* (Vol. 12, No. 1, pp. 34-61). Oxford University Press.

Internal Revenue Service (2017): Understanding employee vs. contractor designation. Retrieved from https://www.irs.gov/newsroom/understanding-employee-vs-contractor-designation on Oct. 14, 2020

Jackson, Matthew O., and Hugo F. Sonnenschein (2007): Overcoming incentive constraints by linking decisions. *Econometrica* 75 (1), 241-257.

Katz, Lawrence F., and Alan B. Krueger (2019): The rise and nature of alternative work arrangements in the United States, 1995–2015. *ILR Review* 72.2, pp. 382-416.

Kieser, Alfred (1994): Why organization theory needs historical analyses—and how this should be performed. *Organization Science*, 5(4), 608-620.

Klein, Benjamin, and Keith B. Leffler (1981): The role of market forces in assuring contractual performance. *Journal of Political Economy* 89 (4), 615-641.

Knight, Frank (1921): *Risk, uncertainty and profit*. Houghton Mifflin.

Kranzberg, Melvin and Michael Hannan (2017): "History of the organization of work." *Encyclopedia Britannica*, 2 Jun. 2017, https://www.britannica.com/topic/history-of-work-organization-648000. Accessed 8 July 2021.

Landes, David S., ed. (1966): *The rise of capitalism*. Vol. 36742. Macmillan.

——— (1969): *The unbound Prometheus: Technological change and development in Western Europe from 1750 to the present.* Cambridge University Press.

——— (1986): What Do Bosses Really Do? *Journal of Economic History*, 46, 585-625.

Lazear, Edward P., and Paul Oyer (2013). Personnel Economics. In: Gibbons and Roberts, pp. 479-519.

Levin, Jonathan (2003): Relational Incentive Contracts. *American Economic Review* 93 (3), pp. 835-857.

———, and Steven Tadelis (2010). Contracting for government services: Theory and evidence from US cities. *The Journal of Industrial Economics*, 58(3), 507-541.

Lewchuk, Wayne (1993). Men and monotony: Fraternalism as a managerial strategy at the Ford Motor Company. *The Journal of Economic History* 53(4), 824-856.

Li, Jin, and Niko Matouschek (2013): Managing conflicts in relational contracts. *American Economic Review* 103 (6), 2328-51.

Li, Jin, Arijit Mukherjee, and Luis Vasconcelos (2020): Rules in Relational Contracts. Working paper.

Lindbeck, Assar, and Dennis J. Snower (1996). Reorganization of firms and labor-market inequality. *The American Economic Review*, 86(2), 315-321.

——— (2000): Multitask learning and the reorganization of work: From tayloristic to holistic organization. *Journal of labor economics* 18 (3), pp. 353-376.

Malcomson, James (1999): Individual employment contracts. In: Orley Ashenfelter and David Card, *Handbook of Labor Economics*, Vol. 3B, Elsevier, pp. 2291-2372

Mantoux, Paul (1928): *The Industrial Revolution in the Eighteenth Century.* Harper Torchbooks, revised edition 1961

Marglin, Stephen (1974): What do Bosses do? The Origins and Functions of Hierarchy in Capitalist Production. *Review of Radical Political Economics* 6(2), 60-112.

Marino, Anthony M., John G. Matsusaka, and Ján Zábojník (2010): Disobedience and Authority. *Journal of Law, Economics, and Organization* 26 (3), pp. 427-459.

Marx, Karl (1867): *Capital. A Critique of Political Economy.* Retrieved from https://www.marxists.org/ Volume-I.pdf on Oct. 14, 2020

Masten, Scott E. (1986): Institutional choice and the organization of production: the make-or-buy decision. *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, 493-509.

——— (1988): A legal basis for the firm. *Journal of Law, Economics, & Organization* 4.1 1988, pp. 181-198.

Matouschek, Niko (2004): Ex Post Inefficiencies in a Property Rights Theory of the Firm. *Journal of Law, Economics, and Organization* 20 (1), pp. 125–147.

Matouschek, Niko, Michael Powell, and Bryony Reich (2022): Organizing Modular Production. Working paper.

McGregor, Douglas M. (1957). The human side of enterprise: American management association. *Management Review*, 1, 41-57.

McKelvey, Richard D., and Talbot Page (2002): Status Quo Bias in Bargaining: An Extension of the Myerson–Satterthwaite Theorem with an Application to the Coase Theorem. *Journal of Economic Theory* 107 (2), pp. 336-355.

McKendrick, Neil (1961): Josiah Wedgwood and factory discipline. *The Historical Journal* 4.1, pp. 30-55.

Millward, Robert (1981): The emergence of wage labor in early modern England. *Explorations in Economic History* 18 (1), pp. 21-39.

Mokyr, Joel (2009): *The enlightened economy: an economic history of Britain, 1700-1850.* New Haven, CT: Yale University Press.

Myerson, Roger, and Mark Satterthwaite (1983): Efficient Mechanisms for Bilateral Trade. *Journal of Economic Theory* 29, pp. 265-281.

National Labor Relations Board (2019): Advice Memorandum, Subject: Uber Technologies, Inc., Cases 13-CA-163062 et al. Retrieved from https://src.bna.com/Ibt on Oct. 14, 2020

Oyer, Paul (2016): The Independent Workforce in America. Study commissioned by Upwork. https://s3-us-west-1. amazonaws. com/adquiro-content-prod/documents/paul_oyer_the_inde pdf.

Parsons, Donald O. (1986). The employment relationship: Job attachment, work effort, and the nature of contracts. In: *Handbook of Labor Economics*, 2, 789-848.

Paychex (2016): Goodbye, 9-5! The Growth of the Freelance Economy. Retrieved from https://www.paychex.com/articles/human-resources/goodbye-9-5-growth-of-the-freelance-economy, accessed on 7/21/21

Pitelis, Christos (1991). *Market and Non-market Hierarchies: Theory of Institutional Failure*, Oxford, Blackwell.

Pollard, Sidney (1963): Factory discipline in the industrial revolution. *Economic History Review* 16 (2), pp. 254-271.

Posner, Eric A. (2020): The economic basis of the independent contractor/employee distinction. Working paper, University of Chicago.

Powell, Michael (2015): An influence-cost model of organizational practices and firm boundaries. *The Journal of Law, Economics, and Organization*, 31 (S1), i104-i142.

Rantakari, Heikki (2021): Simon Says? Equilibrium Obedience and the Limits of Authority. Working paper, University of Rochester.

Rebitzer, James B. (1993): Radical political economy and the economics of labor markets. *Journal of Economic Literature* 31 (3), 1394-1434.

Roberts, John (2004). *The modern firm: Organizational Design for Performance and Growth.* Oxford University Press.

Rosenblat, Alex, and Luke Stark (2016): Algorithmic labor and information asymmetries: A case study of Uber's drivers. *International Journal of Communication* 10, 3758–3784.

Satterthwaite, Mark A., and Steven R. Williams (1998): Bilateral trade with the sealed bid k-double auction: Existence and efficiency. *Journal of Economic Theory* 48 (1), 107-133.

Schöttner, Anja, and Harvey Upton (2022): When Should a Firm Employ its Workers? A Relational Contracting Approach. Accounting for Transparency Working Paper Series No. 105, Available at SSRN: https://ssrn.com/abstract=4257746

Segal, Ilya, and Michael D. Whinston. (2013): Property rights. In: Gibbons and Roberts, pp. 100-158

——— (2016): Property Rights and the Efficiency of Bargaining. *Journal of the European Economic Association* 14 (6), pp. 1287–1328

Siddiqui, Faiz (2020): Uber and Lyft must make their drivers in California full employees, judge rules. *Washington Post*, August 10, 2020

Simon, Herbert (1945): *Administrative Behavior*. New York (Free Press)

——— (1951): A Formal Theory of the Employment Relationship. *Econometrica* 19 (3), pp. 293-305.

——— (1991): Organizations and markets. *Journal of Economic Perspectives* 5 (2), 25-44.

Stone, Dan, Edward Deci, and Richard Ryan (2009): Beyond talk: Creating autonomous motivation through self-determination theory. *Journal of General Management*, 34(3), 75-91.

Tadelis, Steven (2002): Complexity, flexibility, and the make-or-buy decision. *American Economic Review*, 92(2), 433-437.

——— and Oliver Williamson (2013): Transaction Cost Economics. In: Gibbons and Roberts, pp. 159-189

Thompson, Edward P. (1967). Time, work-discipline, and industrial capitalism. *Past & Present* (38), 56-97.

Van den Steen, Eric (2010): Interpersonal Authority in a Theory of the Firm. *American Economic Review* 100(1), 466-90.

Van Triest, Sander (2021): Business as Usual: Working from Home During the Pandemic Did Not Lead to Large Changes in Employee-Level Management Control. Available at SSRN 3909809.

Weber, Lauren (2017): The end of employment. *Wall Street Journal*, February 17, 2017

Wernerfelt, Birger (1997): On the nature and scope of the firm: An adjustment-cost theory. *Journal of Business* 70 (4), pp. 489-514.

——— (2004): Governance of adjustments. *Journal of Business* 77 (S2), pp. S3-S24.

——— (2015). The comparative advantages of firms, markets and contracts: a unified theory. *Economica* 82(326), 350-367.

Whinston, Michael D. (2003): On the transaction cost determinants of vertical integration. *Journal of Law, Economics, and Organization* 19 (1), 1-23.

Williamson, Oliver E. (1971). The vertical integration of production: market failure considerations. *The American Economic Review*, 61(2), 112-123.

——— (1975): *Markets and Hierarchies*. New York (The Free Press).

——— (1985): *The Economic Institutions of Capitalism*. New York (The Free Press).

——— (1991): Comparative economic organization: The analysis of discrete structural alternatives. *Administrative Science Quarterly* pp. 269-296.

Wired 2018 https://www.wired.co.uk/article/uber-verdict-court-of-appeal-gig-economy

Zanarone, Giorgio (2012): Coase (1937) revisited: Endogenous fiat in firms and markets. *International Review of Economics*, 59(2), 201-221.

——— (2013): Contract adaptation under legal constraints. *The Journal of Law, Economics, & Organization*, 29(4), 799-834.

Appendix A: Alternative governance structures?

This paper compares two specific and simple governance structures: employment at a constant wage, in which either E or W has the authority to choose W's tasks, and sequential spot contracting. While I am unable to characterize the Pareto frontier of feasible equilibria for the full set of possible contracts, several features of the model suggest that these two structures in fact stand out as relevant options.

To start with, three assumptions already rule out many possible alternatives: (1) formal long-term contracts are infeasible; (2) the legal transfer of authority over W's actions is infeasible; and (3) between E and W as independent parties, no trade is the legal default.

It follows that contracts that require W to be available in each period, or to perform some action every period, or to follow E's direction, can only be relational. The class of relational contracts that require W to work for E, or at least to be available, is large, but within it, two stand out: one in which E chooses W's action and another where W chooses his action. Much of the paper focuses on the first option, and Section 5.3 discussed the second.

If there is only one task, both versions coincide and reduce to W's obligation to work for E in every period. With multiple tasks, either E-authority or W-authority at least achieves unilateral adaptation to the state of nature, whereas any state-independent distribution or sequence of tasks clearly seems suboptimal. Simple authority contracts can be improved by allowing E and W to bargain, with the disagreement point being determined by some initial allocation of authority. I examine such contracts in Raith (2021) but not here.

The other class of governance structures consists of those that do not *require* W to work for E. Here, the infeasibility of long-term contracts rules out mechanisms that formally link periods, such as those suggested by Jackson and Sonnenschein (2007). The analysis of Athey and Miller (2007), in turn, suggests that improving upon sequential spot trade is possible if the periods can be linked monetarily to satisfy a global budget-balance constraint, but not if budget balance is required for each period. Thus, my assumption that W does not have access to credit considerably diminishes E's and W's ability to take advantage of the repeated nature of the relationship. Intuitively, if "no work" is a contractual option for W, then incentive compatibility implies that for any agreed transfer t^m in return for performing task m, W will choose m only if $c^m \leq t^m$. However, given the contractibility of actions, any such agreement is already feasible with spot trade, which leaves nothing for a relational contract to improve upon. This would be different if work had a moral-hazard component, which could lead to an equilibrium as in Klein and Leffler

(1981) in which the seller's (worker's) effort is motivated by repeat business.

A second simplifying assumption in the model is that employment pays a constant wage w in every period. In fact, such a contract is optimal even if the wage can be task- or time-contingent. To see this, suppose that more generally, a relational employment contract defines for each period $\tau \geq 1$ and each history of the game h^τ leading up to τ a wage function $w(h^\tau) : A \to \mathbb{R}$.

Proposition 10. *For any wage function $w(h^\tau)$ there exists a stationary wage $w \in \mathbb{R}$ that satisfies all incentive constraints and leads to the same sequence of orders $\{m_\tau^E\}$. Consequently, there exists an optimal contract that is stationary.*

There are two parts to this result. First, the optimal wage does not depend on the task because under the independence and symmetry assumptions about p and q, a constant wage ensures that E's profit-maximizing action is also efficient, contingent on E's information, and leads to the least restrictive incentive constraints for W. Second, in spite of W's limited liability, an optimal stationary contract exists because in this model, employment is not an incentive contract. In incentive models with limited liability, backloading payments to the agent is often optimal because it strengthens incentives in future periods (see, for instance, Barron et al., 2018), or because it enables the principal to enforce inefficient but profitable actions early on, in order to extract rent from the agent. Here, backloading wages would relax future incentive constraints too, but E cannot gain from that because she already picks the action that is optimal for her in each period. As a result, all incentive-compatible wage profiles lead to the same stochastic sequence of E's orders, and it is weakly optimal for the parties to settle up in each period.

Appendix B: Proofs

Proof of Proposition 10: 1. I first show that an optimal contract will not be task-contingent. First, unlike with relational incentive contracts and private information, there is never a reason to sacrifice surplus in an ongoing relationship. This follows because obedience is the only moral-hazard dimension of the relationship, and instances of disobedience by W are observable and are punished with the strongest possible punishment, termination. Next, E's profit-maximizing orders are already surplus-maximizing conditional on E's information. To see this, note that for surplus $s = b - c$ and tasks k and l,

$$E[s^k|b^k] \geq E[s^l|b^l] \Longleftrightarrow b^k - E[c^k] \geq b^l - E[c^l] \Longleftrightarrow b^k \geq b^l,$$

where the first equivalence follows from independence of p and q and the second follows from the equality of the marginal distributions p^k and p^l. It follows that if E chooses the task that maximizes her profit $b^k - w_\tau(k)$, she will *always* choose $\arg\max_k b^k = \arg\max_k s^k|b$ if and only if $w_t(k) = w_t$ for all k. Any contract with task-contingent $w_t(k)$ can therefore be weakly improved, in terms of E's orders, by switching to $w_\tau = E[w_\tau(k)] = \overline{w}$ for all k, where expectations are taken over b and E's induced orders. Picking this particular wage also preserves incentives (based on continuation payoffs in τ) in prior periods $\tau' < \tau$, which only depend on $E[w_\tau(k)]$ and not on events in t.

To complete this step, we need to check W's obedience constraints. Given that the highest possible cost $c_H = \max C$ is the same for all tasks, the most restrictive obedience constraint for each task k is

$$(1-\delta)(w_\tau(k) - c_H) + \delta U_{\tau+1}(k) \geq 0,$$

with the notation $U_{\tau+1}(k)$ reflecting the possibility that the continuation utility $U_{\tau+1}$ depends on the task chosen in τ. Now if $w_\tau(k) = \overline{w}_\tau$ for all k, then $U_{\tau+1}$ can be task-independent too and can be held to the smallest value that satisfies both current and future incentive and participation constraints. However, if $w(k) > \overline{w} > w(l)$ for some k and l, then while the lower wage $w(l)$ can be made up for by a higher wage later, the reverse is not true for task k: It may not be possible to counterbalance an above-average $w(k)$ in τ by a lower $U_{\tau+1}(k)$, if doing so violates constraints in $\tau + 1$. It follows that in terms of W's obedience constraints, too, there is nothing to gain from differentiating the wage by the task ordered.

2. Next, I show show that a stationary contract is optimal. Key to the result is that E's optimal order, and thus W's eventual action, does not depend on w. Therefore, any feasible wage profile $\{w_t\}$ that satisfies W's obedience constraint in every period,

$$(1-\delta)(w_t - c_H) + \delta \widetilde{U}_{t+1} \geq 0,$$

will lead to the same stochastic sequence of actions m_t^E. In particular, while E could, for any profile $\{w_t\}$, reduce w_1 and pay a higher wage later, doing so would only make later obedience constraints slack without creating any allocative benefit. Thus, any wage profile $\{w_t\}$ that satisfies

$$w_t - c_H + \sum_{\tau=1}^{\infty} \delta^\tau \left[w_{t+\tau} - c^E\right] \geq 0 \quad \text{for all } t \geq 1$$

can also be satisfied for t by the stationary contract

$$w = c_H - \delta\left[c_H - c^E\right],$$

and in that case it satisfies all future incentive constraints $t' > t$ as well. □

Proof of Proposition 2: The expressions for ΔS and ΔV follow immediately from Lemmas 1 and 2. For the comparative statics, note that $\frac{d\Delta S}{db_H} + \frac{d\Delta S}{db_L}$ and $\frac{d\Delta V}{db_H} + \frac{d\Delta V}{db_L}$ represent the effects of equal increases in c_H and c_L that increase $\bar{c}$ while leaving Δc unchanged, and the same holds for changes in $\bar{b}$.

Then, both ΔS and ΔV are increasing in β and $\bar{b}$ because D^0 increasing in β and b_H, D^E is decreasing in β and b_L, and U^0 is increasing in β, and all other derivatives with respect to β, b_H, and b_L are zero. For $\bar{c}$, the result follows because D^0 is decreasing and D^E increasing in c_H. These are also the only terms affecting ΔV because U^0 and the obedience wage premium $(1-\delta)\gamma\Delta c$ depend only on Δc but not the level of costs. □

Proof of Lemma 3: 1. Suppose $\Delta b \geq \Delta c$. (a) Prices: (i) E-type (b_L, b_L) can only trade with a worker with a low cost for at least one task, and therefore offers $t = (c_L, c_L)$. (ii) For type (b_H, b_L), candidates for prices are $t = (c_H, c_L)$ or $t = (c_L, c_L)$. If E offers (c_H, c_L), all types of W accept task A, and E's profit is $b_H - c_H$. If E offers (c_L, c_L), then (resolving ties in favor of A), types (c_L, c_L) and (c_L, c_H) accept A, type (c_H, c_L) accepts B and type (c_H, c_H) rejects both prices. E's resulting expected profit is $\gamma(b_H - c_L) + q_M(b_L - c_L)$. Then $t = (c_L, c_L)$ is optimal if

$$\gamma(b_H - c_L) + q_M(b_L - c_L) > b_H - c_H \quad \text{or} \quad \gamma > \widehat{\gamma} - \frac{b_L - c_L}{b_H - c_L} q_M = \gamma_{1.}$$

(iii) If E-Type (b_H, b_H) offers $t = (c_H, c_H)$, she will trade for sure and receive profit

$b_H - c_H$. If she offers $t = (c_H, c_L)$, then all W-types except (c_H, c_L) except task A, whereas type (c_H, c_L) is indifferent and therefore picks the more efficient task B. E's resulting expected payoff is $(1 - q_M)(b_H - c_H) + q_M(b_H - c_L)$, which always exceeds the payoff from offering $t = (c_H, c_H)$. If E offers $t = (c_L, c_L)$, finally, types (c_L, c_L) and (c_L, c_H) accept task A, (c_H, c_L) accepts B, and type (c_H, c_H) rejects both, leading to a payoff for E of $(1 - q_H)(b_H - c_L)$. This option is preferred to $t = (c_H, c_L)$ if

$$
\begin{aligned}
(1 - q_H)(b_H - c_L) &> (1 - q_M)(b_H - c_H) + q_M(b_H - c_L) \\
\text{or} \quad \gamma(b_H - c_L) &> (1 - q_M)(b_H - c_H),
\end{aligned}
\tag{12}
$$

which is equivalent to $\gamma > (1 - q_M)\widehat{\gamma} = \gamma_0$. It is straightforward to check that $\gamma_0 < \gamma_1$ if and only if $\Delta b > \Delta c$ as assumed, which leads to the three intervals stated in the proposition, and the optimal prices for each type of E.

(b) U^0 and D^0: (i) $\gamma \leq \gamma_0$: Type (b_L, b_L) offers $t = (c_L, c_L)$ and trades with W if and only if it is efficient, leading to $U^0 = D^0 = 0$. Type (b_H, b_L) offers $t = (c_H, c_L)$ and trades with all types of W, which is efficient ($D^0 = 0$) and leaves a rent Δc to a W with low cost for task A (which occurs with probability γ). The same holds analogously for type (b_L, b_H). Type (b_H, b_H) offers $t = (c_H, c_L)$ and all W pick task A, except for type (c_H, c_L) who picks B. Again, the outcome is efficient and types (c_L, c_L) and (c_L, c_H) earn rent Δc. All in all, for $\gamma \leq \gamma_0$, $U^0 = (1 - p_L)\gamma\Delta c$ and $D^0 = 0$.

(ii) $\gamma \in [\gamma_0, \gamma_1)$: For types $(b_L, b_L), (b_H, b_L)$ and (b_L, b_H) the analysis is the same as in case (i). Type (b_H, b_H) now offers $t = (c_L, c_L)$ and all W with at least one low-cost task pick that one, but earn no rent. Type (c_H, c_H) rejects both prices, which is inefficient. Collecting terms, we have $U^0 = 2p_M\gamma\Delta c$ and $D^0 = p_H q_H(b_H - c_H)$.

(iii) $\gamma \in [\gamma_1, 1)$: For types (b_L, b_L) and (b_H, b_H) the analysis is the same as in case (ii). Type (b_H, b_L) now offers $t = (c_L, c_L)$, which forces $U^0 = 0$. W-types (c_L, c_L) and (c_L, c_H) accept, which is efficient. Type (c_H, c_L) accepts task B, but task A is more efficient because $\Delta b > \Delta c$. Type (c_H, c_H) rejects, which is inefficient. The same holds analogously for E-type (b_L, b_H). Overall, we have $U^0 = 0$ and $D^0 = (1 - p_L)q_H(b_H - c_H) + 2p_M q_M(\Delta b - \Delta c)$, where the first term captures inefficient failure to trade with (c_H, c_H) for all E-types except (b_L, b_L), and the second term captures the loss of surplus from ordering an inefficient task.

2. Now suppose $\Delta b < \Delta c$. (a) Prices: For E-types (b_L, b_L) and (b_H, b_H), the analysis

is the same as in 1(a) above. If E-type (b_H, b_L) offers (c_H, c_L), all W-types but (c_H, c_L) accept task X, whereas type (c_H, c_L) is indifferent and chooses the more efficient task Y. E's profit is $(1 - q_M)(b_H - c_H) + q_M(b_L - c_L)$. If E offers (c_L, c_L), then W-types (c_L, c_L) and (c_L, c_H) accept X, type (c_H, c_L) accepts Y and type (c_H, c_H) rejects both prices. E's resulting expected profit is $\gamma(b_H - c_L) + q_M(b_L - c_L)$. Then $t = (c_L, c_L)$ is optimal if

$$\gamma(b_H - c_L) + q_M(b_L - c_L) > (1 - q_M)(b_H - c_H) + q_M(b_L - c_L) \Longleftrightarrow \gamma > \gamma_0.$$

Thus, in the case $\Delta b < \Delta c$ there are only two intervals for γ to distinguish.

(b) For $\gamma \leq \gamma_0$, the analysis of Lemma 3 applies without change: $U^0 = (1 - p_L)\gamma\Delta c$ and $D^0 = 0$. For $\gamma > \gamma_0$, the analysis is the same as in the case $\gamma > \gamma_1$ above, except for E-type (b_H, b_L) who offers $t = (c_L, c_L)$. As before, type (c_H, c_H) rejects, which is inefficient, but all other W-types, including (c_H, c_L), accept the efficient task. Consequently, we have $U^0 = 0$ as before, but the deadweight loss is only $D^0 = (1 - p_L)q_H(b_H - c_H)$. □

Proof of Lemma 4: With probability p_L, both tasks have low benefit, and otherwise at least one task has a high benefit; therefore $b^E = (1 - p_L)b_H + p_L b_L$. As for W's expected cost, if $b = (b_H, b_L)$ or $b = (b_L, b_H)$, all types of W execute X (respectively, Y), at expected cost $\bar{c}$. If $b = (b_L, b_L)$ or (b_H, b_H), E orders a low-cost task except when $c = (c_H, c_H)$. As a result, W's cost in these b-states is lower than $\bar{c}$ by $q_M\Delta c$. Overall, therefore, $c^E = \bar{c} - (p_L + p_H)q_M\Delta c$, which leads to $c_H - c^E = [\gamma + (p_L + p_H)q_M]\Delta c$.

If $b = (b_H, b_H)$, E's order is always efficient. If $b = (b_L, b_L)$, E's order is efficient except when $c = (c_H, c_H)$, which occurs with probability $p_L q_H$. W-Types (b_H, b_L) and (b_L, b_H) order task X and Y, respectively. These orders are efficient for all types of W whenever $\Delta b \geq \Delta c$ because even if W's cost is high, we have $b_H - c_H \geq b_L - c_L > 0$. If $\Delta b < \Delta c$, however, then the low-benefit-low-cost task would be more efficient, and E's choice of the high benefit tasks creates a deadweight loss of $\Delta c - \Delta b$, which occurs with probability $2p_M q_M$. This leads to the expression for D^E stated in the Lemma.□

Proof of Proposition 3: $\frac{d\Delta V}{db_H} + \frac{d\Delta V}{db_L}$ represents the effects of equal increases in b_H and b_L that increase $\bar{b}$ while leaving Δb unchanged, and $\frac{d\Delta V}{dc_H} + \frac{d\Delta V}{dc_L}$ represents the effects of equal increases in c_H and c_L that increase $\bar{c}$ while leaving Δc unchanged. We can examine these changes separately for V^0 and V^E. By inspection of Table 2, $\frac{d(U^0+D^0)}{db_H} + \frac{d(U^0+D^0)}{db_L} \geq 0$ and $\frac{d(U^0+D^0)}{dc_H} + \frac{d(U^0+D^0)}{dc_L} \leq 0$ for all five cases shown, and therefore $\frac{dV^0}{db_H} + \frac{dV^0}{db_L} \leq 0$ and

$\frac{dV^0}{dc_H} + \frac{dV^0}{dc_L} \geq 0$. In fact, all derivatives applied to U^0 alone are zero; the reported derivatives of $U^0 + D^0$ are those of D^0.

With employment, the term $p_L q_H (c_H - b_L)$ in D^E is decreasing in b_L and increasing in c_H, and therefore decreasing in $\bar{b}$ and increasing in $\bar{c}$, holding Δb and Δc fixed. The term $2p_M q_M (\Delta c - \Delta b)$ clearly does not depend on $\bar{b}$ or $\bar{c}$, holding Δb and Δc fixed, nor does the term $c_H - c^E = [\gamma + (p_L + p_H) q_M]\Delta c$ in V^E. Putting the results for V^0 and V^E together establishes the result for ΔV. Since the established derivatives also hold for the deadweight losses D^0 and D^E alone, the same result holds for ΔS. □

Proof of Proposition 5: It is straightforward to compute the correlation coefficient

$$\rho_c = \frac{Cov(c^X, c^Y)}{\sqrt{Var(x)Var(y)}} = \frac{(1-\gamma)^2 q_L + \gamma^2 q_H + 2\gamma(1-\gamma) q_M}{\gamma(1-\gamma)}.$$

For a simultaneous change $\frac{\partial q_H}{\partial \varepsilon_c} = \frac{\partial q_L}{\partial \varepsilon_c} = 1$ and $\frac{\partial q_M}{\partial \varepsilon_c} = -1$, we have $\partial \rho_c / \partial \varepsilon_c = (2\gamma - 1)^2 / [\gamma(1-\gamma)] > 0$, and therefore the described variation corresponds to an increase in the correlation of costs across tasks. It follows that $\frac{\partial \Delta V}{\partial \varepsilon_c} = \frac{\partial \Delta V}{\partial q_H} + \frac{\partial \Delta V}{\partial q_L} - \frac{\partial \Delta V}{\partial q_M}$, and analogous for ΔS.

Now consider the different cases for γ per Lemma 3. For $\gamma \leq \gamma_0$, $\frac{\partial (U^0 + D^0)}{\partial \varepsilon_b} = 0$. For $\gamma \in [\gamma_0, \gamma_1)$, $\frac{\partial (U^0 + D^0)}{\partial \varepsilon_b} = p_H (b_H - c_H) > 0$. For $\gamma \in [\gamma_1, 1)$,

$$\begin{aligned} \frac{\partial (U^0 + D^0)}{\partial \varepsilon_b} &= (1 - p_L)(b_H - c_H) - 2p_M(\Delta b - \Delta c) \\ &> 2p_M(b_H - c_H) - 2p_M(b_H - c_H - (b_L - c_L)) \\ &= 2p_M(b_L - c_L) > 0. \end{aligned}$$

Thus, for any γ, $\frac{\partial (U^0 + D^0)}{\partial \varepsilon_c} > 0$ and therefore $\frac{\partial V^0}{\partial \varepsilon_c} < 0$.

□

Proof of Proposition 6: Consider first spot trade with $S^* - V^0 = U^0 + D^0$, with expressions for the latter stated in Lemma 3. For $\gamma \leq \gamma_0$, $\frac{\partial (U^0 + D^0)}{\partial \varepsilon_b} = -\gamma \Delta c < 0$. For $\gamma \in [\gamma_0, \gamma_1)$, $\frac{\partial (U^0 + D^0)}{\partial \varepsilon_b} = q_H (b_H - c_H) - 2\gamma \Delta c$. Because $\gamma > \gamma_0$, condition (12) holds, which using $b_H - c_L = b_H - c_H + \Delta c$ can be rewritten as $\gamma \Delta c > q_H (b_H - c_H)$. It follows that $\frac{\partial (U^0 + D^0)}{\partial \varepsilon_b} < 0$ for this range of γ as well. For $\gamma \in [\gamma_1, 1)$, finally, $\frac{\partial (U^0 + D^0)}{\partial \varepsilon_b} = -q_H (b_H - c_H) - 2q_M(\Delta b - \Delta c) < 0$. Thus, for any γ, $\frac{\partial (U^0 + D^0)}{\partial \varepsilon_b} < 0$, as stated.

For employment, we have $S^* - V^E = D^E + (1-\delta)(c_H - c^E)$, with expressions for the latter stated in Lemma 4. The derivative of this expression with respect to ε_b is

$$q_H(c_H - b_L) - 2\mathbf{1}_{\Delta c > \Delta b} q_M(\Delta c - \Delta b)] + (1-\delta)2q_M\Delta c, \tag{13}$$

and (11) is a sufficient condition for (13) to be positive. Putting both parts together, $\Delta V = U^0 + D^0 - D^E - (1-\delta)(c_H - c^E)$ is decreasing in ε_b whenever (11) holds.

□

Proof of Proposition 7: Part (a): The analysis of E-authority is covered by Lemma 4. If $\Delta b \geq \Delta c$, then all of E's orders are efficient except when $(b, c) = (b_L, b_L, c_H, c_H)$. If $\Delta b < \Delta c$, E's orders are additionally inefficient when $b = (b_H, b_L)$ and $c = (c_H, c_L)$ or vice versa. If W has authority, then types (c_L, c_H) and (c_H, c_L) pick task X and Y, respectively, whereas types (c_L, c_L) and (c_H, c_H) pick whichever task has the higher benefit (resolving remaining ties in favor of X). If $b = (b_L, b_L)$ and $c = (c_H, c_H)$, neither task is efficient but W is required to provide his service, thus W's choice is inefficient. If $\Delta b \geq \Delta c$, two more inefficient choices occur if $b = (b_H, b_L)$ and $c = (c_H, c_L)$ or vice versa. If $\Delta b < \Delta c$, then W's pick of the lower-cost task is efficient, and only $(b, c) = (b_L, b_L, c_H, c_H)$ leads to inefficiency. It follows that E-authority is more efficient than W-authority if and only if $\Delta b \geq \Delta c$.

Part (b): With E-authority, W's expected cost from E's default orders is $c^E = \bar{c} - (p_L + p_H)q_M\Delta c$, per Lemma 4. With W-authority, W's cost is c_L if at least one cost is low, and c_H if $c = (c_H, c_H)$, which results in the expected cost of $c^E = \bar{c} - q_M\Delta c$, which is smaller than with E-authority.

If $\Delta b \geq \Delta c$, E-authority then leads to both a smaller deadweight loss and a smaller obedience wage premium $(1-\delta)(c_H - c^E)$ than W-authority, and is therefore more efficient and more profitable. If $\Delta b < \Delta c$, then W-authority leads to a smaller deadweight loss, but a larger obedience wage premium. Specifically, with E-authority, E's profit is

$$\begin{aligned} & W^* - D^E - (1-\delta)(c_H - c^E) \\ = \ & W^* - p_L q_H(c_H - b_L) - 2p_M q_M(\Delta c - \Delta b) - (1-\delta)[\gamma + (p_L + p_H)q_M]\Delta c, \end{aligned}$$

whereas with W-authority, her profit is

$$W^* - p_L q_H (c_H - b_L) - (1-\delta)(\gamma + q_M)\Delta c.$$

Delegation is then optimal if

$$2(1-\delta)p_M q_M \Delta c < 2p_M q_M(\Delta c - \Delta b) \Longleftrightarrow (1-\delta)\Delta c < \Delta c - \Delta b \Longleftrightarrow \delta > \Delta b/\Delta c,$$

as stated. □

Proof of Proposition 8: If $b = b_L$, plausible price offers to the two workers are $t = (c_L, c_L)$ and $t = (c_H, c_L)$. (With $t = (c_H, c_H)$, both workers would accept but the entrepreneur would make a loss of $2b_L - 2c_H$.) If $t = (c_L, c_L)$, then both workers—as required for positive output—will accept only if $c_1 = c_2 = c_L$, leading to a profit for E of $2\gamma^2(b_L - c_L)$ as opposed to $2\gamma(b_L - c_L)$ with independent production. Thus, the average price is the same as with independent production, but the profit strictly lower.

If $b = b_L$ and E offers $t = (c_H, c_L)$, then W1 accepts for sure and W2 iff $c_2 = c_L$, leading to profit $\gamma(2b_L - c_H - c_L)$ for E, which is less than $2\gamma(b_L - c_L)$ with independent production. Though this strategy can be optimal for some parameter values, the average price is strictly higher and the profit lower than with independent production.

If $b = b_H$, all three different price pairs are plausible candidates for E to offer. If E offers $t = (c_L, c_L)$, both workers accept iff $c_1 = c_2 = c_L$, and E's profit is $2\gamma^2(b_H - c_L)$. If E offers $t = (c_H, c_L)$, then W1 accepts for sure and W2 iff $c_2 = c_L$, leading to profit $\gamma(2b_H - c_H - c_L)$ for E. This option is more profitable than $t = (c_L, c_L)$ if $\gamma(2b_H - c_H - c_L) \geq 2\gamma^2(b_H - c_L)$ or

$$\gamma \leq \frac{2b_H - c_H - c_L}{2(b_H - c_L)} = \widehat{\gamma}_2.$$

Finally, if $t = (c_H, c_H)$, both workers accept for sure and E's profit is $2(b_H - c_H)$, which is no smaller than $\gamma(2b_H - c_H - c_L)$ iff

$$\gamma \leq \frac{2(b_H - c_H)}{2b_H - c_H - c_L} = \widehat{\gamma}_1.$$

For $\widehat{\gamma}$ as defined in Lemma 1, it is straightforward to show that $\widehat{\gamma} < \widehat{\gamma}_1 < \widehat{\gamma}_2$, which leads to four intervals for γ: (i) If $\gamma \leq \widehat{\gamma}$, then $t = (c_H, c_H)$ with both independent and complementary production, and E's profit in both cases is $2(b_H - c_H)$. (ii) If $\gamma > \widehat{\gamma}$, then with independent production, $t = (c_L, c_L)$ and E's profit is $2\gamma(b_H - c_L)$. Now if $\gamma \in (\widehat{\gamma}, \widehat{\gamma}_1]$, then with complementary production, $t = (c_H, c_H)$ and the resulting profit is $2(b_H - c_H)$, which is less than $2\gamma(b_H - c_L)$ because $\gamma > \widehat{\gamma}$. (iii) If $\gamma \in (\widehat{\gamma}_1, \widehat{\gamma}_2]$, then with complementary production, $t = (c_H, c_L)$ and the resulting profit is $\gamma(2b_H - c_H - c_L) < 2\gamma(b_H - c_L)$. (iv) If $\gamma > \widehat{\gamma}_2$, then with complementary production, $t = (c_L, c_L)$ and the resulting profit is $2\gamma^2(b_H - c_L) < 2\gamma(b_H - c_L)$. Thus, in case (i), prices and profits are the same with independent and complementary production, whereas in all other cases, the average price is higher and the profit lower with complementary production. □

Proof of Proposition 9: With spot trade and E facing n workers competing, type b_L always offers $t = c_L$. Type b_H can offer $t = c_H$ and make profit $b_H - c_H$, or offer $t = c_L$, which is accepted by at least one seller with probability $1 - (1-\gamma)^n$. If follows that $t = c_L$ is optimal if

$$1 - (1 - \gamma)^n > \widehat{\gamma} \text{ or } \gamma \geq 1 - \sqrt[n]{1 - \widehat{\gamma}},$$

where the last expression converges to 0 as $n \to \infty$. It follows that for any γ, there exists n for which E offers $t = c_L$ irrespective of her type. That offer is accepted with probability $1 - (1 - \gamma)^n$, which converges to 1 for large n. Thus, E's profit from spot trade converges to $\bar{b} - c_L$, which exceeds $V^E = \bar{b} - w < \bar{b} - \bar{c} < \bar{b} - c_L$. □