UC Berkeley
Law and Economics Workshop

Title
Are There Too Many Entrepreneurs? A Model of Client-Based Entrepreneurship

Permalink
https://escholarship.org/uc/item/9cd3v3t7

Author
Watson, Joel

Publication Date
2003-04-01

eScholarship.org Powered by the California Digital Library
University of California

ARE THERE TOO MANY ENTREPRENEURS?
A MODEL OF CLIENT-BASED ENTREPRENEURSHIP

James E. Rauch*
Department of Economics
University of California, San Diego
La Jolla, CA 92093-0508

and

National Bureau of Economic Research
Cambridge, MA 02138

Joel Watson*
Department of Economics
University of California, San Diego
La Jolla, CA 92093-0508

Preliminary and Incomplete — for April 14 seminars only

April 2003

Client relationships create value, which employees may try to wrest from their employers by setting up their own firms. Firms counter by inducing workers to sign contracts that prohibit them from competing or soliciting former clients in the event of termination of employment. Society trades off higher effort by self-employed workers against the cost of establishing redundant businesses, and local governments compete to attract clients. If clients, firms, and workers can renegotiate restrictive employment contracts and make compensating transfers, the socially optimal level of entrepreneurship will be achieved regardless of government policies regarding enforcement of these contracts. If workers cannot finance transfers to firms, however, firms and workers will sign contracts that are too restrictive and produce too little entrepreneurship, and governments can increase welfare by limiting enforcement of these contracts. With or without liquidity constraints, more entrepreneurial locations will attract more clients and have higher employment and output.

*We thank Max Auffhammer for outstanding research assistance. Financial support was provided by NSF grants #SES-0001832 (Rauch) and #SES-0095207 (Watson).

I. Introduction

Entrepreneurs often acquire crucial knowledge by working as employees of businesses similar to the ones they later start. According to the 1992 Economic Census of the United States (1997, p. 86), 45.1 percent of nonminority male business owners "previously worked for a business whose goods/services were similar to those provided by the [current] business." The figure rises to 49.6 percent for Services, the industry group with the largest number of business owners, even though these business owners are the most likely to have a professional or doctoral degree (1997, p. 70) and second most likely (after Finance, Insurance, and Real Estate) to have a bachelor's degree or higher, which might have suggested a relatively diminished need for on-the-job training.

Part of the knowledge that entrepreneurs acquire when employed, especially in services, is knowledge of potential clients for their future businesses. Unfortunately, no aggregate statistics are available on the extent to which, when establishing their own businesses, entrepreneurs serve clients with whom they built relationships during their previous employment.[1] What we know is that employers try to restrict this type of entrepreneurial activity on the part of their former employees by including non-compete or non-solicitation covenants in their employment contracts (Carnevale and Doran 2001). The perception in the business press is that use of these restrictive employment clauses is increasing (Marino 2003), though we are again

---

[1]Rauch and Watson (2002, Table 1) find that, when international trade intermediaries handling differentiated products started their firms, clients outside of the United States with whom they had experience from previous employment accounted for over half of their international business. This information is not directly relevant in the present context, however, since these entrepreneurs are not in direct competition with their previous employers but rather serve to facilitate the sales or purchases of potential competitors.

unaware of any aggregate statistics. In deciding whether non-compete covenants are enforceable, Carnevale and Doran (2001) state that "courts generally consider whether the covenant protects 'trade secrets' to which an employee may have had access or whether the employee's services are 'unique or extraordinary'....With regard to customer relationships courts have found that employers have a legitimate interest in protecting the 'unique' relationship that an employee develops with the employer's clients or an interest in protecting 'customer relations'."

Enforcement of restrictive employment covenants has been controversial. Rulings can be decidedly vague, as in the New York Superior Court ruling "enjoining departing employees from soliciting former clients for a period of three months but refusing to enjoin them from *accepting* business from clients who chose to utilize their services" (Carnevale, Lockhart and Olosunde, 1999, italics in original). The laws on which rulings are based vary across U. S. states. Carnevale and Doran (2001) report, "Although there are common threads of legal analysis throughout the nation, several states have enacted statutes governing unreasonable restraints of trade and non-competition covenants specifically. Some limit or purport to limit the enforceability of such covenants."[2] Gilson (1999) argues that covenants not to compete are much less enforceable in California than in Massachusetts. Examining potential causes of the success of Silicon Valley in California relative to Route 128 in Massachusetts, McMillan (2002, p. 114)

---

[2]Even if the law is clearly on the side of the employer, he or she may be deterred from pursuing enforcement by the time and expense of litigation. We discussed the issue with a provider of audio-visual engineering services (such as configuring corporate boardrooms for videoconferencing) whose firm had been hurt by the departure of an employee who took several important clients with him. The employer asserted that only very large firms with in-house litigation staff would find it worthwhile to pursue this kind of case. His firm was trying to address the problem by building up a discretionary deferred compensation fund, a strategy we will cover in section III below.

claims, citing the work of Gilson, "The post employment covenant lies at the root of the differences between Silicon Valley and Route 128." McMillan is concerned with the ability of restrictive covenants to prevent departing employees from taking with them technological innovations rather than client relationships, but we conjecture that far more entrepreneurs start their businesses on the basis of the latter than the former, and will argue that restrictive covenants and other policies can also affect the economic performance of regions through their impact on client-based entrepreneurship.

The cross-state variation in enforcement of restrictive employment covenants suggests that their effects on entrepreneurship and regional output are an area that is ripe for empirical work, but first we need some theoretical predictions to test. In an influential law review article favoring enforcement of restrictive covenants, Sterk (1993, p. 406) makes the Coasian argument that, "nothing prevents the employee from bargaining with his employer for release from the covenant. If either the employee himself or other prospective employers value the employee's services more than his current employer does, the employee should be willing to pay the employer to release him from the contract." With regard to client-based entrepreneurship, the Coasian argument can be rephrased as follows: the assignment of "property rights" in client relationships to the employer versus the employee will not affect economic outcomes. We evaluate this argument, as well as the consequences for efficiency (if any) of the choice of property rights assignment. We also address the effects on entrepreneurship, regional output, and social welfare of other policies that affect the costs of client-based entrepreneurship. A central theme of our work is be that bargaining between the employer and employee may fail to internalize the interests of the rest of society, in particular the clients. Indeed, Carnevale,

Lockhart and Olosunde (1999) note that "clients of departing employees are now challenging the enforcement of non-compete agreements."

A broader aim of this paper is to integrate the value created by client relationships into economists' thinking regarding determination of income and social welfare. This will only become more important as the service share of GDP continues to grow. Because of conflicting claims to the value created by client relationships that may interact with market failures, (1) the market outcome may not be optimal; (2) the government cannot avoid intervention through the legal system (because of its role in resolving disputes); therefore (3) we need to seek guidelines for that intervention.

In this paper, we analyze a model of the relationships between firms, employees, and clients. The model has three key elements. First, in a given relationship, production relies on the worker exerting effort, but effort is unverifiable and therefore difficult to motivate. The worker's incentive to exert effort is greater when the worker starts his own firm, but to do so he must pay a start-up cost. Second, pairs of workers and firms negotiate their initial employment contract — including any restrictive covenants — prior to matching and contracting with their clients. Third, worker's face liquidity constraints that keep them from borrowing money on the basis of expected future returns. In our model, firms, workers, and clients have the opportunity to renegotiate the terms of their relationship, but the liquidity constraint affects the outcome of negotiation.

We show that, because they negotiate before matching with a client, a firm-worker pair can use a restrictive covenant in the employment contract to expropriate value from the client. This implies a positive relation between entrepreneurial activity and the number of clients in

geographic locations. We find that, when the worker is liquidity constrained, restrictive covenants in the employment contract are inefficiently renegotiated; this leads to an inefficiently low amount of entrepreneurial activity. We show that limits on the enforcement of non-compete and non-solicitation covenants can increase welfare and also increase the number of clients in a geographical region. We comment on other policy implications as well.

In the next section of this paper we survey the most relevant literature. Our benchmark model is presented in section III. In section IV we see how our results are affected by liquidity constraints on workers, and in section V we extend our model to an overlapping generations framework. In our concluding section, we suggest directions for further research in this largely unexplored area.

II. Literature Survey [in process]

III. The Benchmark Model

We model providers of business services and their clients. We focus on business services for two reasons. First, business services are growing even more rapidly than overall services as a share of total output. Business services increased from 7.3% of U.S. GDP in 1987 to 13.7% of U.S. GDP in 2001, raising their share of overall services from 29.1% to 38.6%.[3] Second, the

[3]The overall service share of U.S. GDP increased from 25.0% in 1987 to 35.4% in 2001. Services are calculated as gross aggregate output for SIC 70-89 and business services are defined as SIC 73, Business Services, plus SIC 87, Engineering, Accounting, Research, Management, and Related Services. All figures are from U.S. Department of Commerce, Bureau of Economic Analysis, Industry Economics Division, http://www.bea.gov/bea/dn2/gpo.htm.

clients of business service providers are themselves businesses, making business service quality a key to the ability of locations to attract and retain business in general. For example, corporate headquarters locate where there is easy access to high quality accounting, advertising, consulting, financial, legal, and other business services. Klier and Testa (2002) find that the growth of large company headquarters between 1990 and 2000 across U.S. metropolitan areas is significantly associated only with the growth of metro area population and the 1990 share of metro area nonfarm earnings in business services.[4] Our model can also be applied to manufacturers that need to make relationship-specific investments for their clients if clients tend to locate their headquarters or own manufacturing plants near such subcontractors.[5]

We will assume that a client initially selects a location and then approaches a local business-service provider ("firm") to provide a homogeneous service to support production of a standardized good whose profits have been competed away. For example, a client might ask an advertising firm to update a campaign for an existing product. We further assume that the employee who handles the client's account acquires knowledge of the client's preferences,

---

[4]When explaining their results, Klier and Testa supply a quotation from Lichtenberg (1960) that underscores why relationships with outside business service providers are so important for corporate clients: "Like producers of unstandardized products, the central office executives 'produce' answers to unstandardized problems, problems that change frequently, radically, and unpredictably.... These problems are solved quickly only by consultation with a succession of experts. But ... most central offices would find it inefficient if not impossible to staff themselves internally with all of the specialized personnel and services that they must call on from time to time to solve their problems.... All of these considerations dictate a concentration of central offices ... near their 'suppliers'."

[5]Data collected by Chiu and Ka-chung (2003) show that 14 out of 49 founders of large Hong Kong electronics firms had backgrounds in sales and marketing, nine of whom worked for traders and five of whom worked for other electronics firms, allowing them to develop relationships with the major buying clients of their employers.

capabilities, business strategy, etc. that allows him to provide a more customized service to the client in support of a more novel and ambitious project. The employee of the advertising firm, in our example, can now develop a campaign to launch a new product for the client.[6] The quality of the effort or investment that the employee makes to provide this customized service is not verifiable and therefore not contractible. Moreover, the value of the customized service itself is not verifiable precisely because it cannot be assessed without the deep knowledge of the client's business that is acquired by working closely with it.

We assume that the employer captures part of the value of the customized service provided by the firm to the client. This gives the employee an incentive to establish his own firm, which he must trade off against the time and expense involved in doing so. The employee's incentive to supply his unverifiable effort may also be influenced by whether he continues to work for the employer or becomes self-employed. Other factors that may affect separation versus non-separation are the quality of the match between the employee and the client and the enforceability of non-compete or non-solicitation covenants between the employee and employer.

Although our model is broader than those surveyed in the previous section in that it encompasses both the demand and supply sides of an entire sector (business services), we defer analysis of the interaction of the business services sector with the rest of the economy to future work. Thus we will not attempt to answer the question of whether too much or too little resources are allocated to production of business services rather than other goods and services. We will assume that the opportunity cost of labor is given exogenously, so that employment of

---

[6]In manufacturing, the equivalent to providing a homogeneous service could be making a product with minor modifications to the existing design, and the equivalent to providing a customized service could be making a new product.

labor to produce business services has no impact on availability of labor or wages in the rest of the economy. We will further assume that the total number of clients is given exogenously and we will not model sales to the rest of the economy of goods and services produced by clients. Initially, we will also take the number of firms in each geographical location to be exogenous, but in Section V we will show how this number can be determined.

A. Assumptions, notation, and timing

Our model is populated by three groups of risk-neutral agents: clients, firms, and workers. Within any group, all agents are identical ex ante. Workers are perfectly mobile across locations, hence the opportunity cost of labor to any location is identical. Firms, on the other hand, are attached to a given location. In this section, we analyze one period of interaction.

For the timing of our model, we refer the reader to Figure 1. At the beginning of the period, a fixed measure (quantity) of clients Q is allocated over a fixed number n of locations indexed by i. We let N_i denote the mass of firms in location i. Each client inelastically demands one unit of commercial space. Commercial rent in a given region is increasing in the quantity of clients located there:

$$r_i = r(Q_i),\ r' > 0. \qquad \textbf{(1)}$$

The producer surplus associated with this upward-sloping supply curve is part of local income. Each client chooses the location that maximizes her expected net income $y_{Ci} - r_i$, where y_{Ci} denotes the client's expected value of its relationship with a firm and worker. In equilibrium, therefore, it must be that

$$y_{Cj} - r(Q_j) = y_{Ci} - r(Q_i),\ j \neq i;\ \ \sum_{i=1}^{n} Q_i = Q. \qquad \textbf{(2)}$$

We shall see that y_{Ci} is not a function of Q_i, so our model can be solved recursively: we first determine y_{Ci} and then use equations (2) to solve for Q_i.

After the clients have arrived, in each location i the N_i firms hire workers in anticipation of serving clients, with exactly one worker needed to serve each client. The firm negotiates a contract with each worker that specifies the probability p_i that the worker will be allowed to serve the client if he separates from the firm; low values of p_i represent enforceable non-compete/non-solicitation provisions in the contract. The allowed range of p_i depends on the rules for enforcement of restrictive employment clauses in the location (i.e., the range of values that the law permits and courts enforce). For example, if non-compete clauses are unenforceable in a particular location, then p_i is constrained to be close to one. In addition to selecting p_i, the firm and worker can make an immediate monetary transfer as they reach an agreement.

To simplify our modeling of the bargaining problem between the firm and worker, we assume that each firm correctly anticipates serving $q_i \equiv Q_i/N_i$ clients and interviews q_i workers, without either firms or workers having a second chance to match.[7] Under this assumption, the outside option for a firm in its bargain with each worker is zero and the outside option for each worker is his opportunity cost, which we can set to zero for convenience provided it is small relative to the rents generated by client relationships.

After completing their hiring, each firm accepts q_i clients. Workers then provide clients

[7]We assume away integer problems.

with homogeneous services and acquire the knowledge required to provide them with more valuable, customized services. For simplicity we assume that the homogeneous services do not create any value.

As the worker engages in the provision of homogeneous services to the client, a random draw takes place. The realization this random draw, denoted η, is the idiosyncratic component of the time and expense required for the worker to establish his own business to provide the client with customized services. We assume η is commonly observed by the client, firm, and worker. We suppose η is drawn from a fixed distribution with support $[\underline{\eta},\bar{\eta}]$ and satisfying $\bar{\eta} > \underline{\eta} > 0$. We let μ denote the density function for this distribution and we assume that μ is continuous and positive on $[\underline{\eta},\bar{\eta}]$.

Knowing η, the three parties renegotiate the worker's contract with the firm. In particular, they can jointly alter the probability that the worker will be allowed to serve the client if he separates from the firm, changing this value from p_i to some p_i'. We will see that they have an interest in setting $p_i' = 1$ because a lower value is inefficient when η is small. The parties' agreement at this time may also include immediate monetary transfers. If the negotiation ends in disagreement, then $p_i' = p_i$ and no transfer is made.

After negotiating over p_i', the client, firm, and worker negotiate over whether to stay together and make some immediate monetary transfers between them.[8] Disagreement leads to separation of the client and worker from the firm.

If the client, firm, and worker stay together, the firm employs the worker to provide the

[8]We model renegotiation of the worker's contract and negotiation over whether to stay together as sequential events because it simplifies the analysis of the liquidity-constrained case in Section IV.

customized service to the client. The worker chooses his level of effort e (which is in monetary units), and customized services are produced according to the following production function:

$$v(e) = f(e) + B, \quad \textbf{(3)}$$

where v is the value of the customized service to the client and B is the default value of the customized service if the worker supplies zero effort. We make the standard assumptions that $f(0) = 0, f'' < 0, \lim_{e \to 0} f'(e) = \infty$, and $\lim_{e \to \infty} f'(e) = 0$. Given that the value of the customized service provided to the client is not verifiable to outside parties, the best the client, firm and worker can do is work without a contract and, following the worker's effort decision, rely on their bargaining powers to obtain shares of the value.[9]

If the client and worker separate from the firm, with probability $1 - p_i'$ the firm is able to obtain a preliminary injunction that prevents any provision of the customized service by the worker to the client. As described by Carnevale, Lockhart and Olosunde (1999) and Gilson (1999), this is the most effective and often used way for a firm to use a non-compete or non-solicitation clause in its contract with an employee.[10] We further assume that if a preliminary injunction is obtained the courts will indeed ultimately rule that the worker cannot serve the client. In this case the firm retains the client and provides the customized service without the

---

[9]Here we are assuming that there is a binary, costless "trade decision" that the worker and firm must make to generate the value v; this decision comes after the worker's effort decision and is verifiable. In such a contracting environment, where the worker's investment affects the client's value of trade, the optimal contract specifies "no trade" and the parties anticipate that they will renegotiate it after the worker's effort choice. Che and Hausch (1999) provide the general analysis that yields this conclusion. The contract gives the worker some incentive to exert effort, but the effort is suboptimal. The first-best level of effort cannot be attained.

[10]Carnevale, Lockhart and Olosunde (1999) state, "As attorneys litigating these cases will attest, the process is swift and those prosecuting or defending non-compete actions may be called upon to argue the merits of such clauses within hours of the employee's notice."

input of the worker, yielding the default value B, divided between the client and firm through negotiation.

With probability p_i' the firm is unable to prevent the worker from serving the client. The worker then establishes his own firm (we now call him an entrepreneur), incurring a cost of entry given by:

$$k_i(\eta) = c_i + \eta, \quad \textbf{(4)}$$

where $c_i > 0$ is determined by the location and is common to every worker-client pair. The cost c_i consists of the time and expense required to register the firm, open an office, etc., so well described by Djankov et al. (2002).[11] The idiosyncratic component of entry costs η is the time and expense the worker must incur to be able to replicate the functions of his former employer. These could be administrative functions, or the kind of "finishing" that senior members of a firm sometimes add to the work of junior members in order to complete the service for a client. For example, a provider of architectural services may know what designs will satisfy the client's needs but must learn the procedures for obtaining approval from the appropriate regulatory agencies. A worker will find it easier to learn the functions of his former employer, the better is the match between the needs of the client and his knowledge and talents.

After the worker becomes an entrepreneur, he and the client negotiate over whether the latter will now become a client of the new firm, with disagreement leading to separation that yields zero for both parties. Finally, if the entrepreneur retains the client then he chooses his

---

[11]The costs of entry imposed by government regulations vary widely from country to country, as documented by Djankov et al. (2002). Within the United States, fees such as for business licenses differ across localities, and local chambers of commerce vary in the efficiency with which they guide business start-ups through the legally required procedures. In addition special tax breaks may be offered, such as for locating in an enterprise zone.

effort level and produces a value of customized services given by equation (3), which he and the client divide according to the bargaining weights λ and $1 - \lambda$, respectively.

Figures 2.1 and 2.2 contain the extensive-form diagram of the game described above, from the point at which p_i has been chosen and the number η has been drawn. The circled nodes in the picture are joint-decision nodes; each models a phase in which two or three of the players negotiate. At both individual- and joint-decision nodes, the branches are labeled with the actions that must be taken. At joint-decision nodes, a disagreement point is also indicated, which describes what happens if the players do not reach an agreement.[12] Each party can unilaterally compel disagreement.

Note that Phase (a) in Figure 2.1 is where the firm, worker, and client negotiate over p_i'. Here, an immediate transfer from the client to the firm is denoted $t_{CFi}^{(a)}$, an immediate transfer from the firm to the worker is denoted $t_{FWi}^{(a)}$, and disagreement means $p_i' = p_i$, $t_{CFi}^{(a)} = 0$, and $t_{FWi}^{(a)} = 0$. At Phase (b) the parties negotiate over whether to stay together; here, the possible transfers are denoted with a "(b)" superscript and the disagreement point is separation with no immediate transfers. Figure 2.2 shows the subgame from the point at which the worker and client separate from the firm. This subgame begins with a random event — the move of "nature" that determines whether the worker is allowed to serve the client.

The payoff of each player in the game is simply the player's total monetary gains, which the player is assumed to maximize.

B. Solution

---

[12]For more on games with joint decisions, see Watson (2002).

We solve our model using the standard technique of sequential rationality combined with a bargaining solution.[13] At each joint-decision node, we apply the generalized Nash bargaining solution, where the client has a fixed bargaining weight of $1 - \lambda$ and, wherever the firm and worker are both in the negotiation, the relative bargaining weights of the worker and firm are α and $1 - \alpha$, respectively. Thus, where all three agents negotiate, such as at Phase (c) in Figure 2.1, the bargaining weights for the client, firm, and worker are, in order, $1 - \lambda$, $(1 - \alpha)\lambda$, and $\alpha\lambda$. Where the client negotiates with the entrepreneur, such as at Phase (a′) in Figure 2.2, the client's bargaining weight is $1 - \lambda$ and the entrepreneur's is λ.

At each bargaining phase in the game, the set of feasible payoff vectors has the "transferable utility" property, meaning that, by making monetary transfers, the parties can arbitrarily shift utility between themselves on a one-to-one basis. Because the game has transferable utility, the outcome of negotiation can be viewed in terms of the maximized surplus relative to the disagreement point, with each party obtaining his disagreement value plus his bargaining weight times the surplus. That is, the payoffs to the parties are computed by finding the total surplus relative to the disagreement point, dividing it among the parties according to their bargaining weights, and adding each party's share of the surplus to his individual disagreement (threat) value.[14] Any transfers between the parties implied by these payoffs are

---

[13]That is, we calculate a *negotiation equilibrium* (Watson 2002). This means that the worker maximizes his payoff when making his individual effort decisions and the outcomes of the negotiation phases are consistent with a bargaining solution, assuming that the players accurately anticipate the continuation values from various points in the game tree.

[14]Technically, the generalized Nash solution solves $\max \prod_{i \in I} (u_i - d_i)^{\pi_i}$, where I is the set of players, u_i is player i's utility, d_i is player i's disagreement value, and π_i is player i's bargaining weight. This simplifies to the surplus-division formula in settings of transferable utility. Bargaining weights are nonnegative and sum to one.

computed by subtracting their continuation values (their payoffs in the next phase of the game) from their current payoffs.

Our calculations proceed by backward induction. We begin by solving the subgame from the point at which separation occurs (starting at Phase (a′) in Figure 2.2). With probability p_i this subgame ends with the client receiving a customized service of value $f(e^S) + B$, where e^S is the effort supplied by the entrepreneur (worker) when he and the client separate from the firm. With probability $1 - p_i$ this branch ends with the client receiving a customized service of value B. We will work through the solution to Phase (b′), in which the entrepreneur chooses his effort level and the client and entrepreneur divide the value of the customized service. Solving for the payoffs and transfers in the rest of the separation branch is straightforward.

After the worker has invested $k_i(\eta)$ to become an entrepreneur and has supplied e^S, his disagreement value in the event of a breakdown in bargaining with his client is $-(k_i(\eta) + e^S)$. The client's disagreement value is zero. The gross surplus from proceeding with production of the customized service is therefore $f(e^S) + B - (k_i(\eta) + e^S)$, leaving a net surplus relative to the disagreement point of $f(e^S) + B$. Dividing this according to the bargaining weights of the client and entrepreneur and adding the disagreement points yields the payoffs $y_{Ci}^{S(b')} = (1 - \lambda)(f(e^S) + B)$ and $y_{Ei}^{S(b')} = \lambda(f(e^S) + B) - (k_i(\eta) + e^S)$ shown in Table 1. Comparing these payoffs to the (end-of-game) continuation values $f(e^S) + B$ for the client and $-(k_i(\eta) + e^S)$ for the entrepreneur implies a transfer of $t_{CEi}^{(b')} = \lambda(f(e^S) + B)$ from the client to the entrepreneur, also shown in Table 1. The entrepreneur chooses the effort level that maximizes his anticipated payoff $\lambda(f(e^S) + B) - (k_i(\eta) + e^S)$, so e^S must satisfy

$$\lambda f'(e^S) = 1. \qquad \textbf{(5)}$$

In contrast, the efficient level of effort, which we denote by $\bar{e}$, must satisfy $f'(\bar{e}) = 1$. Even an entrepreneur, therefore, will supply less than the efficient level of effort.

Working backwards to phase (a′), we see that the entrepreneur picks up a transfer $t_{CEi}^{(a')} = (1 - \lambda)e^S$ from the client, yielding client and worker payoffs $(1 - \lambda)(f(e^S) + B - e^S)$ and $\lambda(f(e^S) + B - e^S) - k_i(\eta)$, respectively, that are received with probability p_i'. With probability $1 - p_i'$, on the other hand, the firm is able to prevent the worker from serving the client and the worker gets zero, leaving the client and the firm to divide the surplus B. The expected payoffs from separation for the client, firm, and worker are therefore given by:

$$\begin{aligned} y_{Ci}^S &= p_i'(1 - \lambda)(f(e^S) + B - e^S) + (1 - p_i')(1 - \lambda)B \\ y_{Fi}^S &= (1 - p_i')\lambda B \\ y_{Wi}^S &= p_i'[\lambda(f(e^S) + B - e^S) - k_i(\eta)] \end{aligned} \qquad \textbf{(6)}$$

We now turn to Figure 2.1 and work backward from phase (c) to solve the "stay together" subgame. After the worker has supplied e^T, his disagreement value in the event of a breakdown in bargaining between himself, the client and the firm is $-e^T$. The disagreement values of the client and firm are zero. The gross surplus from proceeding with production of the customized service is $f(e^T) + B - e^T$, leaving a net surplus relative to the disagreement point of $f(e^T) + B$. Dividing this according to the bargaining weights of the client, firm, and worker and adding the disagreement point yields the payoffs $y_{Ci}^{T(c)} = (1 - \lambda)(f(e^T) + B)$, $y_{Fi}^{T(c)} = (1 - \alpha)\lambda(f(e^T) + B)$, and $y_{Wi}^{T(c)} = \alpha\lambda(f(e^T) + B) - e^T$ shown in Table 1. Comparing these payoffs to the (end-of-game) continuation values $f(e^T) + B$ for the client, zero for the firm and $-e^T$ for the worker implies payments of $t_{CFi}^{(c)} = \lambda(f(e^T) + B)$ from the client to the firm and $t_{FWi}^{(c)} = \alpha\lambda(f(e^T) + B)$ from the firm to the worker, also shown in Tabel 1. The entrepreneur chooses the effort level that maximizes his anticipated payoff $\alpha\lambda(f(e^T) + B) - e^T$, so e^T must satisfy

$$\alpha\lambda f'(e^T) = 1. \quad \textbf{(7)}$$

Comparing equation (7) to equation (5) reveals that the worker supplies less effort when working for the firm than he would as an entrepreneur: $e^T < e^S$ since $f'' < 0$, provided $\alpha < 1$.

Given these payoffs in phase (c), in phase (b) the gross surplus to the parties from agreeing to stay together is $f(e^T) + B - e^T$. This must be compared to the total value of disagreement (separation of client and worker from firm), which can be seen from equation (6) to equal $p_i'(f(e^S) + B - e^S - k_i(\eta)) + (1 - p_i')B$. Given efficient bargaining, disagreement will occur when this latter value is at least as large as the former value, in which case all parties receive the payoffs in equation (6). If, on the other hand, the former value is greater, then the net surplus from agreement $f(e^T) - e^T - p_i'(f(e^S) - e^S - k_i(\eta))$ will be divided among the parties. Adding the parties' shares of this surplus to their disagreement points from equation (6) yields the payoffs $y_{Ci}^{T(b)}$, $y_{Fi}^{T(b)}$, and $y_{Wi}^{T(b)}$ shown in Table 1. The transfers from the client to the firm and from the firm to the worker are then given by $y_{Ci}^{T(c)} - y_{Ci}^{T(b)}$ and $y_{Wi}^{T(b)} - y_{Wi}^{T(c)}$, respectively.[15]

Carnevale, Lockhart and Olosunde (1999) note that business service employers are using discretionary deferred pay compensation schemes as a way to supplement non-compete agreements when their enforcement is uncertain. They state, for example, that "firms are beginning to require their employees to defer a portion of their annual bonuses over several years so that employees stand to lose a portion of their bonuses if they leave." Does our model indeed predict that some workers will receive positive transfers from firms when they agree to stay together and p_i' is high? We expect this to be true only for workers for whom entrepreneurship is

[15]For the disagreement point described in the text to always be appropriate, we make the following assumption throughout: $\lambda(f(e^S) - e^S) \geq c_i + \bar{\eta}$. This ensures that the worker wants to become an entrepreneur whenever separation occurs.

attractive ($k_i(\eta)$ is low), but in this case the parties may just separate. Let us consider the case where $p_i' = 1$ and $k_i(\eta)$ is sufficiently low that the surplus from agreement is close to zero. We then have $k_i(\eta) \approx f(e^S) - e^S - (f(e^T) - e^T)$ and $y_{Wi}^{T\,(b)} \approx y_{Wi}^S = \lambda(f(e^S) + B - e^S) - k_i(\eta)$, implying

$$t_{FWi}^{(b)} \approx \lambda(f(e^S) + B - e^S) - k_i(\eta) - [\alpha\lambda(f(e^T) + B) - e^T]$$

$$\approx \lambda(1 - \alpha)B + (1 - \alpha\lambda)f(e^T) - (1 - \lambda)(f(e^S) - e^S).$$

The last expression must be positive when the worker's bargaining weight vis-a-vis the firm α is low but his bargaining weight as an entrepreneur vis-a-vis the client λ is high. Moreover, this positive transfer from the firm to the worker must come out of the firm's pocket as opposed to being entirely funded by the client: with $p_i' = 1$ and the surplus from staying together close to zero, $y_{Fi}^{T\,(b)} \approx 0$ and thus $t_{FWi}^{(b)} - t_{CFi}^{(b)} \approx (1 - \alpha)\lambda(f(e^T) + B)$, the entire continuation payoff of the firm.

In Phase (a) of the game (at the beginning of Figure 2.1), the parties renegotiate p_i. If they foresee that the client and worker will separate from the firm, clearly the surplus from separation will be maximized by choosing $p_i' = 1$, so efficient bargaining will yield this outcome. We must consider three cases:

Case (a1): $p_i(f(e^S) - e^S - k_i(\eta)) \geq f(e^T) - e^T$. Separation is efficient[16] even at the existing p_i, so the disagreement points are the separation payoffs with $p_i' = p_i$ and the surplus from agreement on $p_i' = 1$ is $(1 - p_i)(f(e^S) - e^S - k_i(\eta))$. This gives us the information needed to compute the payoffs shown in Table 1, and the transfers are computed by subtracting from these payoffs the separation payoffs with $p_i' = 1$, which are the continuation values in this case.

[16]"Efficiency" used here is defined with respect to the separation/stay together decision. For example, we say that separation is efficient given p_i' if it yields a higher joint payoff than does staying together.

Case (a2): $p_i(f(e^S) - e^S - k_i(\eta)) < f(e^T) - e^T$, but $f(e^S) - e^S - k_i(\eta) \geq f(e^T) - e^T$. Staying together is efficient at the existing p_i, but separation is efficient with $p_i' = 1$. The disagreement points are $y_{Ci}^{T(b)}$, $y_{Fi}^{T(b)}$, and $y_{Wi}^{T(b)}$ evaluated at $p_i' = p_i$ and the surplus from agreement on $p_i' = 1$ is $f(e^S) - e^S - k_i(\eta) - (f(e^T) - e^T)$. The continuation values are the same as in case (a1).

Case (a3): $f(e^S) - e^S - k_i(\eta) < f(e^T) - e^T$. Nothing is to be gained from renegotiating p_i. The parties' payoffs equal their disagreement points $y_{Ci}^{T(b)}$, $y_{Fi}^{T(b)}$, and $y_{Wi}^{T(b)}$ evaluated at $p_i' = p_i$, and there are no transfers.

The solution for phase (a) reveals that the worker becomes an entrepreneur with $p_i' = 1$ when $k_i(\eta) \leq f(e^S) - e^S - (f(e^T) - e^T)$ and he remains with the firm when $k_i(\eta) > f(e^S) - e^S - (f(e^T) - e^T)$. Given the number of clients Q_i, in each location the constrained socially-efficient level of entrepreneurship is achieved and the constrained socially-efficient level of surplus is generated, regardless of the p_i to which the firm and worker initially agree.

Proposition 1: Given Q_i and the (suboptimal) effort levels e^T and e^S, the parties select p_i' efficiently and they separate if and only if it is efficient to do so.

The key to achieving this Coasian result is the ability of all three parties to renegotiate p_i together.

We conclude this section by solving for the p_i to which the firm and worker agree when they are first matched and for the number of clients Q_i in each location. Let us denote by $\hat{\eta}_i$ the value of the idiosyncratic component of the cost of establishing a new business below which workers will become entrepreneurs. We have:

$$\hat{\eta}_i = \max \{f(e^S) - e^S - (f(e^T) - e^T) - c_i, \underline{\eta}\}. \quad \textbf{(8)}$$

We can now compute the expected client payoff y_{Ci} in equation (2) by summing the client payoffs in phase (a) using the density μ and noting that Case (a3) applies when $\eta > \hat{\eta}_i$ and that Cases (a1) and (a2) yield the same payoffs when $\eta \leq \hat{\eta}_i$. We have:

$$y_{Ci} = \int_{\underline{\eta}}^{\hat{\eta}_i}(1-\lambda)[f(e^S) + B - e^S - (1-p_i)k_i(\eta)]\mu(\eta)d\eta + \int_{\hat{\eta}_i}^{\bar{\eta}}(1-\lambda)[f(e^T) + B - e^T + p_i k_i(\eta)]\mu(\eta)d\eta$$
$$= (1-\lambda)[p_i(c_i + E\eta) + \gamma_i(f(e^S) + B - e^S - c_i) - \tilde{\eta}_i + (1-\gamma_i)(f(e^T) + B - e^T)], \quad \textbf{(9)}$$

where γ_i is the probability that the client and worker separate from the firm in location i, $E\eta$ is the expected value of η, and $\tilde{\eta}_i = \int_{\underline{\eta}}^{\hat{\eta}_i}\eta\mu(\eta)d\eta$. Similar computations yield the expected total payoff for the client, firm, and worker from one client relationship, denoted by y_i:

$$y_i = \gamma_i(f(e^S) + B - e^S - c_i) - \tilde{\eta}_i + (1-\gamma_i)(f(e^T) + B - e^T). \quad \textbf{(10)}$$

Regarding y_{Ci}, we note that its relation to c_i depends on the choice of p_i:

Lemma 1: There is a number $x \in (0,1)$ such that, for a fixed p_i, if $p_i < x$ then y_{Ci} is decreasing in c_i, and if $p_i > x$ then y_{Ci} is increasing in c_i.

The intuition here is that, when p_i is low, the client's payoff is sensitive to c_i only in the event that the parties renegotiate the worker's contract to set $p_i' = 1$ in anticipation of separating, in which case lowering c_i raises the negotiation surplus without changing the client's disagreement value; hence, the client is better off with a lower c_i. On the other hand, when p_i is high, the client's payoff is sensitive to c_i only in the event that staying together is efficient; in this case, raising c_i causes the worker's disagreement value to decrease, which favors the client.

Since y_{ci} is increasing in p_i and y_i is not a function of p_i, the sum of the payoffs to the firm and worker must be decreasing in p_i. It follows that the firm and worker will choose $p_i = 0$, and that the firm will compensate the worker for accepting this iron-clad restrictive employment

clause with an immediate transfer or "signing bonus."

Proposition 2: At the beginning of their relationship, the firm and the worker write a contract that sets p_i at its lowest possible level. If there are no legal constraints, they thus choose $p_i = 0$.

The intuition for this result is that the client benefits from the ability of the worker to serve her as an entrepreneur, so the firm and worker can collectively capture some of the client's surplus by restricting this ability.

Substituting $p_i = 0$ into equation (9) yields $y_{ci} = (1 - \lambda)[\gamma_i(f(e^S) + B - e^S - c_i) - \tilde{\eta}_i + (1 - \gamma_i)(f(e^T) + B - e^T)]$. This result can then be substituted into equations (2) to solve for Q_i. It is immediately clear that $c_i = c_j$ implies $Q_i = Q_j$, and that if c_i is the same across all locations then $Q_i = Q/n$. Moreover, Lemma 1 shows that, with $p_i = 0$, y_{ci} increases as c_i decreases. It follows from equation (2) that $c_i < c_j$ implies $Q_i > Q_j$. Since lower c_i also increases $\hat{\eta}_i$ (and thus increases γ_i), as seen from equation (8), we have the result that more entrepreneurial locations attract more clients.[17]

Proposition 3: In the absence of legal constraints on p_i, the following is true in equilibrium. Locations with more entrepreneurial activity will also have more clients.

It follows that more entrepreneurial locations have greater employment in business services and greater total output of business services.

---

[17]Here we are only considering equilibria in which all locations have positive entrepreneurship, i.e., we assume that c_i is not so large in any location that $\hat{\eta}_i = \underline{\eta}$.

C. Implications for Policy

We have seen that the constrained socially efficient rates of entrepreneurship and expected payoffs from client relationships are achieved without government intervention in our benchmark model. However, it may be possible for policy to improve the allocation of clients across locations. It is easily shown that total expected income from all client relationships plus total producer surplus of all landlords that rent to clients is maximized when

$$y_j - r(Q_j) = y_i - r(Q_i), \ \forall j \neq i. \quad \textbf{(11)}$$

From equations (9) and (10) we see that $y_{Ci} = (1 - \lambda)y_i$ when $p_i = 0$. The equilibrium condition given by equation (2) then yields $(1 - \lambda)y_j - r(Q_j) = (1 - \lambda)y_i - r(Q_i)$, which is not the same as equation (11), except in the special case $c_i = c_j$ which implies $y_i = y_j$. The intuition here is that the clients do not internalize all of the benefits of the decision to locate in a particular region, because this includes benefits to firms and workers who, in our model, do not contract with clients until location choices are made. In general, the allocation of clients will not respond as much as is socially optimal to differences in the cost of entrepreneurship across locations as long as $\lambda > 0$. National welfare can therefore be increased by, for example, subsidizing client rents in locations where entrepreneurship is cheaper.

Before we take this policy recommendation too seriously, we need to ask about the sources of differences in the cost of entrepreneurship across locations. If the sources are differences in government efficiency and regulations, it may be optimal to address these directly rather than taking them as given and setting policy in response.

Another important policy instrument is a government's enforcement of non-compete and non-solicitation clauses. First consider the case in which c_i is the same across all locations. Then

a given local government could refuse to enforce non-compete or non-solicitation clauses in certain circumstances, thereby forcing p_i above zero and attracting clients. This unambiguously reduces national welfare, but it increases the producer surplus of local landlords without changing the expected payoff for any client-firm-worker relationship. The conclusion changes if, for some locations *i* and *j*, intrinsic cost differences (due to, say, geographical variation) imply $c_i < c_j$. Then, because of the externality inherent in the clients' location choice, not enough clients will locate in region *i* than is socially optimal. By forcing p_i up, the government helps to internalize the externality and enhance welfare.

IV. Liquidity-Constrained Workers

Our model predicts that, when hired, workers will agree to give up any rights to serve in their own businesses the clients of whom they acquire deep knowledge while working as employees. A worker who, nevertheless, becomes a client-based entrepreneur will therefore have to *buy from the firm* the right to continue to serve his client. In other words, in the phase of the game where the parties renegotiate the worker's contract, the transfer from the firm to the worker $t_{FWi}^{(a2)}$ will be negative. This is easily verified by substituting $p_i = 0$ into the formula for $t_{FWi}^{(a2)}$ in Table 1. From this formula we can also see that, if the bargaining power α of the worker vis-a-vis the firm is small, then the worker will have to pay the firm almost the entire value of his business, $\lambda(f(e^S) + B - e^S) - k_i(\eta)$, to obtain the rights to serve his client as an entrepreneur.

In reality, there are several barriers that limit the ability of the worker to make a monetary transfer to the firm. First, workers generally do not have the resources to internally (out of pocket) finance a large payment. Even if the firm paid the worker a "signing bonus" at the

beginning of their relationship, little of it may be left to transfer back to the firm. Furthermore, if α is close to zero then, in the benchmark model, the signing bonus is very small relative to the amount that the worker transfers to the firm during renegotiation of his contract. Second, external financing generally is limited due to informational asymmetries between the worker and outside lending institutions. If future returns from the client are unverifiable, the entrepreneur/worker can hide his income and declare that his new firm has failed. If banks cannot easily distinguish between the workers in our model and other, high-risk agents, then the banks will not be willing to lend the worker the money required to buy out his non-compete agreement.[18]

In this section, we propose an extension of our model to investigate the real liquidity constraint that workers face. Because of the technical complexities involved, we do not offer a full model of the constraints inherent in internal and external financing. Instead, to capture the worker's liquidity constraint, we modify our benchmark model by making one additional assumption: that transfers to the worker must be nonnegative. Specifically, we require that $t_{FWi}^{(a)} \geq 0$, $t_{FWi}^{(b)} \geq 0$, and the transfer from the firm to the worker during their initial contract negotiation (which comes before the subgame described in Figures 2.1 and 2.2) is also nonnegative. That is, when the firm and worker first contract and when three parties renegotiate

[18]Similar problems arise in loan arrangements with the client and firm. A loan from the client is subject to hold-up, whereby, after the effort decision, the worker refuses to consummate trade unless the loan is renegotiated. This implies that the entrepreneur's and client's continuation payoffs from the time that they contract are independent of the sunk loan amount, meaning that the client is merely making an immediate transfer to the firm through the worker. A loan from the firm requires the same verification of returns as would a loan from a bank; the worker could hide his returns from the firm and, anticipating that it cannot compel the worker to repay, the firm will not issue the loan.

this contract and then decide whether to stay together, the worker cannot make a positive transfer to the other parties.

Implicitly, we must also assume that the worker's cost of effort e^S or e^T and cost of establishing a new business $c_i + \eta$ consist of time rather than money, so that they do not have to be financed. Regarding the cost c_i imposed by government regulations, Djankov et al. (2002) find for the United States that the typical direct (out-of-pocket monetary) component is less than one-half of one percent of per capita GDP.

A. Assumptions and a note on the bargaining solution

To simplify the analysis and focus our attention on an interesting range of the parameter space, we make the following assumptions on the parameters of the model:

(A1) $\lambda[f(e^S) - e^S] \geq c_i + \bar{\eta}$,

(A2) $e^T \geq c_i + \bar{\eta} + B\alpha\lambda/(1 - \alpha\lambda)$, and

(A3) $\lambda B \geq (1 - \lambda)[f(e^S) - e^S]$.

Assumption A1 was made in the benchmark model. Assumption A2 simplifies the analysis by implying that the worker's liquidity constraint does not bind in the negotiation over whether the three parties stay together or separate (Phase (b) in Figure 2.1). Assumption A3 yields the interesting case in which the worker's liquidity constraint always binds in the negotiation of p_i' (Phase (a) in Figure 2.1). By starting with a specification of f for which $e^T \geq c_i + \bar{\eta} + f(e^S) - e^S$, we can find a number B to satisfy A2 and A3.

The worker's liquidity constraint makes applying the Nash bargaining solution a bit more complicated than is the case in the benchmark model, because the set of payoff vectors over

which the parties negotiate does not necessarily exhibit transferable utility. In other words, the outcome of negotiation cannot always be put in terms of a maximized surplus (relative to the disagreement point) that is divided between the players according to bargaining weights. Thus, we have to employ the general form of the Nash bargaining solution, as described in footnote 14 on page 14. However, in some of the bargaining problems in our model, the outcome of negotiation one obtains by ignoring the liquidity constraint is feasible even with the liquidity constraint. In this case, the liquidity constraint is not binding and so it does not affect the negotiation.

B. Solution

To solve the model with liquidity constrained workers, we use the same backward induction procedure that we used to analyze our benchmark model. In fact, much of the analysis is the same as that in the benchmark model. To start, note that the liquidity constraint does not bind in the negotiation phases shown in Figure 2.2 (between the entrepreneur and the client), nor does it bind in the negotiation at Phase (c) in Figure 2.1. This can be verified by confirming that, in these negotiation phases, the unconstrained Nash bargaining solution specifies a nonnegative transfer to the worker/entrepreneur; note that $t_{FWi}^{(c)}$, $t_{CEi}^{(a')}$, and $t_{CEi}^{(b')}$ in table 1 are all nonnegative. Thus, the values y_{Ci}^{S}, y_{Fi}^{S}, y_{Wi}^{S}, $y_{Ci}^{T(c)}$, $y_{Fi}^{T(c)}$, and $y_{Wi}^{T(c)}$ are the same as they were in the benchmark model. For these values, recall expression (6) and the following paragraph in Section III.B (or refer to Table 1).

Next consider the negotiation at Phase (b) in Figure 2.1, where, given p_i', the parties jointly decide whether to stay together. The disagreement point is separation with no transfers,

which yields the continuation values y_{Ci}^S, y_{Fi}^S, and y_{Wi}^S for the client, firm, and worker, respectively. As we noted in the analysis of the benchmark model, the joint value of separating (the sum of the players' continuation values) is $p_i'[f(e^S) + B - e^S - k_i(\eta)] + (1 - p_i')B$, whereas the joint value of staying together is $f(e^T) + B - e^T$. If the former exceeds the latter then separation is the solution to the bargaining problem, just as in the benchmark model. If the joint value of staying together exceeds the value of separating, then efficiency dictates staying together. However, we must check to see whether the worker's liquidity constraint interferes with this outcome. Ignoring the liquidity constraint for a moment, we recall that, in the Nash solution of the benchmark model, the parties stay together and the worker gets an immediate transfer of $t_{FWi}^{(b)} = y_{Wi}^{T(b)} - y_{Wi}^{T(c)}$, as defined in Table 1. We have the following lemma, which is proved in the Appendix.

Lemma 2: Under the assumptions of this section, the worker's transfer $t_{FWi}^{(b)}$ from the benchmark model is nonnegative when $f(e^T) + B - e^T \geq p_i'[f(e^S) + B - e^S - k_i(\eta)] + (1 - p_i')B$.

This result implies that the worker's liquidity constraint does not bind in the solution to the bargaining problem in Phase (b) of Figure 2.1. In other words, given p_i', the outcome of negotiation over whether to stay together is exactly the same as that computed for the benchmark model. Thus, the values $y_{Ci}^{T(b)}$, $y_{Fi}^{T(b)}$, and $y_{Wi}^{T(b)}$ in the model with the liquidity-constraint are the same as those derived for the benchmark model (shown in Tabel 1).

Our analysis continues with the evaluation of the parties' negotiation over p_i' (Phase (a) in Figure 2.1). Here, the liquidity constraint has an interesting effect and leads to different outcomes than occur in the benchmark model. Let us consider Cases (a1)-(a3) paralleling the analysis in Section III.B.

Case (a1): $p_i(f(e^S) - e^S - k_i(\eta)) \geq f(e^T) - e^T$. Separation would occur even at the existing p_i. It is efficient to set $p_i' = 1$. However, an increase in p_i' causes the firm's continuation payoff $y_{Fi}^{T(b)}$ to drop at the rate of λB and the client's continuation payoff $y_{Ci}^{T(b)}$ to rise at the rate of $(1 - \lambda)[f(e^S) - e^S]$. Assumption A3 implies that the sum of these changes is negative, meaning that, without an immediate transfer from the worker, the firm would not agree to set p_i' above p_i. Because the worker is liquidity constrained, the outcome of negotiation is the disagreement point of $p_i' = p_i$ with no immediate transfers.

Case (a2): $p_i(f(e^S) - e^S - k_i(\eta)) < f(e^T) - e^T$, but $f(e^S) - e^S - k_i(\eta) \geq f(e^T) - e^T$. Staying together is efficient at the existing p_i, but separation is efficient with $p_i' = 1$. It is easy to verify that, as in Case (a1), the sum of the firm's and client's continuation payoffs, $y_{Fi}^{T(b)} + y_{Ci}^{T(b)}$, is *decreasing* in p_i'.[19] (This depends on Assumption A3.) Thus, the firm will not agree to set p_i' above p_i without an immediate transfer from the worker. Because the worker is liquidity constrained, the outcome of negotiation is the disagreement point of $p_i' = p_i$ with no immediate transfers.

Case (a3): $f(e^S) - e^S - k_i(\eta) < f(e^T) - e^T$. Nothing is to be gained from renegotiating p_i. The parties' payoffs equal their disagreement points $y_{Ci}^{T(b)}$, $y_{Fi}^{T(b)}$, and $y_{Wi}^{T(b)}$ evaluated at $p_i' = p_i$, and there are no transfers.

In summary, we have:

---

[19]Taking the derivative of $y_{Fi}^{T(b)} + y_{Ci}^{T(b)}$ with respect to p_i' yields $-\lambda B + (1 - \lambda)[f(e^S) - e^S]$ where p_i' would yield separation. By Assumption A3, this is nonpositive. In the range of p_i' in which the parties will stay together, the derivative is $-\lambda B + \lambda(1 - \alpha)[f(e^S) - e^S] + (1 - \alpha\lambda)k_i(\eta)$, which, given the presumption of Case (a2), does not exceed $-\lambda B + (1 - \lambda)[f(e^S) - e^S] - (1 - \alpha\lambda)[f(e^T) - e^T]$. Again, this is nonpositive.

Proposition 4: When the worker is liquidity constrained, his original contract is not renegotiated (so $p_i' = p_i$) and the parties stay together if and only if it is efficient to do so given p_i — that is, if and only if $p_i[f(e^S) - e^S - c_i - \eta] \leq f(e^T) - e^T$. The equilibrium payoffs from Phase (a) in Figure 2.1 satisfy $y_{Ci}^{T(a)} = y_{Ci}^{T(b)}$, $y_{Fi}^{T(a)} = y_{Fi}^{T(b)}$, and $y_{Wi}^{T(a)} = y_{Wi}^{T(b)}$, for the values of $y_{Ci}^{T(b)}$, $y_{Fi}^{T(b)}$, and $y_{Wi}^{T(b)}$ shown in Table 1 and evaluated at $p_i' = p_i$.

The intuition behind Proposition 4 is simple. Much of the gain of increasing p_i' goes to the worker by way of his anticipated value of directly dealing with the client. Without an immediate transfer from the worker, the firm loses when p_i' is increased. Furthermore, the firm's loss is significant given that B is large (by assumption). Thus, the worker gains in future value, whereas the client and firm jointly lose. Although the *total* value increases, it is value that will be realized only in the future. Because of the liquidity constraint, the worker cannot compensate the firm for its loss.

Proposition 4 has some important observable consequences. First, when clients and workers separate from firms there will be disputes that will be handled by the courts. This clearly inefficient outcome contrasts sharply with the benchmark model result (Proposition 1) in which separation is always accompanied by the complete release of the worker from any non-compete or non-solicitation covenants. Second, the rate of entrepreneurship will be increasing in p_i, again a sharp contrast with the benchmark model in which the rate of entrepreneurship is unaffected by the degree of restrictiveness of the contracts made between firms and workers regarding the ability of the latter to serve clients after leaving their firms. In particular, the cutoff value of η below which workers will become entrepreneurs is now given by

$$\hat{\eta}_i \equiv \max \{f(e^S) - e^S - c_i - [f(e^T) - e^T]/p_i, \underline{\eta}\}. \quad \textbf{(12)}$$

Equation (12) shows that, provided there is any entrepreneurship in equilibrium, $\hat{\eta}_i$ varies

inversely with p_i. Comparing equation (12) to equation (8), we see that the difference comes from the fact that the parties are no longer able to renegotiate p_i to $p_i' = 1$.

Integrating over the random variable η, we write the expected payoffs of the firm and worker as a function of the contract parameter p_i. The client's expected payoff is

$$y_{Ci} = p_i(1-\lambda)[f(e^S) - e^S] + (1-\lambda)B + \int_{\hat{\eta}_i}^{\bar{\eta}}(1-\lambda)[f(e^T) - e^T - p_i(f(e^S) - e^S - k_i(\eta))]\mu(\eta)d\eta.$$

The firm's expected payoff is

$$y_{Fi} = (1-p_i)\lambda B + \int_{\hat{\eta}_i}^{\bar{\eta}}(1-\alpha)\lambda[f(e^T) - e^T - p_i(f(e^S) - e^S - k_i(\eta))]\mu(\eta)d\eta$$

and the worker's expected payoff is

$$y_{Wi} = p_i\lambda[f(e^S) + B - e^S] - p_i(c_i + E\eta) + \int_{\hat{\eta}_i}^{\bar{\eta}}\alpha\lambda[f(e^T) - e^T - p_i(f(e^S) - e^S - k_i(\eta))]\mu(\eta)d\eta.$$

To interpret these payoffs, note that each party obtains the value of separation, plus his share of the surplus of staying together (integrated over those values of η for which the parties stay together).

We conclude our analysis of the liquidity-constrained setting by evaluating the firm and worker's optimal choice of p_i in their initial contract. The following lemma is proved in the Appendix [in process].

Lemma 3: The client's value y_{Ci} is strictly increasing in p_i, whereas the firm's value y_{Fi} is strictly decreasing in p_i. The total value of the parties, $y_{Ci} + y_{Fi} + y_{Wi}$, is increasing in p_i. The sum of the worker's and firm's values, $y_{Fi} + y_{Wi}$, is convex in p_i.

In other words, the client prefers higher values of p_i, the firm prefers lower values of p_i, and.any $p_i < 1$ is inefficient. Given the inability of the parties to renegotiate p_i to $p_i' = 1$, only $p_i = 1$ yields the same, efficient result as the benchmark model.

At the beginning of their relationship, the firm and worker select p_i and can make a positive transfer from the firm to the worker (but not the other way around). Thus, applying the Nash bargaining solution, the firm and worker select p_i and a transfer t to solve

$$\max [y_{Fi}(p_i) - t]^{1-\alpha} [y_{Wi}(p_i) + t]^{\alpha},$$

subject to $p_i \in [0, 1]$ and $t \geq 0$. The disagreement point is zero for both parties. Figure 3 illustrates the bargaining set for two different cases of the parameter values. In the pictures, the dashed line indicates the payoff vectors that are feasible for different values of p_i, holding $t = 0$. The solid line is the frontier of the bargaining set. Note that increasing t from zero moves the payoff vector down and to the right (in the direction of a line with a slope of –1).

The solution to the firm and worker's initial bargaining problem is difficult to characterize in general, but clearly there are three possibilities: a corner solution at $p_i = 0$, a corner solution at $p_i = 1$, and an interior solution with $p_i \in (0, 1)$. Furthermore, the solution will specify $p_i = 0$ if $y_{Fi}(0) + y_{Wi}(0) \geq y_{Fi}(1) + y_{Wi}(1)$, which is the case shown in diagram 2 of Figure 3. Algebra reveals that this inequality simplifies to

$$\lambda(1-\gamma_i)[f(e^S) - e^S - (f(e^T) - e^T)] \leq c_i + E\eta_i - \lambda(1-\gamma_i)E[c_i + \eta_i \mid \eta_i \geq f(e^S) - e^S - c_i - (f(e^T) - e^T)], \quad \textbf{(13)}$$

where γ_i is the probability that separation will be efficient, i.e., that $\eta_i \leq f(e^S) - e^S - c_i - (f(e^T) - e^T)$. The conditional expectation on the right side of this inequality is the expected set up cost for the entrepreneur, conditional on separation *not* being efficient. We have some simple sufficient conditions for $p_i = 0$ to be selected:

Proposition 5: Suppose there are no legal constraints on p_i. If λ is sufficiently close to zero and/or γ_i is sufficiently close to one relative to the other parameters, then the firm and worker will select $p_i = 0$.

In the case in which inequality (13) is not satisfied, the firm and worker face an interesting trade off in their selection of p_i. By lowering p_i from one, they expropriate value from the client, just as in the benchmark model. However, they also reduce the total value of the three parties by creating an environment in which they later stay together when it is efficient for them to separate (with $p_i' = 1$). As Proposition 5 establishes, the former effect dominates when λ is small and/or γ_i is large. In this case, we also know that if local governments force p_i above zero, firms and workers will choose this lower bound when they sign their initial contracts.

Considering this main case, we can conclude for locations with positive entrepreneurial activity that equations (2) and (12) and Lemma 3 imply:

Proposition 6: Comparing locations with equal values of c_i, if the local governments differ in their enforcement of non-compete and non-solicitation agreements, entrepreneurial activity will be higher and the number of clients will be higher where these agreements are less enforced (that is, where p_i is constrained to be higher).

Locations with lower p_i will also have greater employment in business services and greater total output of business services.

C. Implications for Policy

When the worker is liquidity constrained, government intervention plays a heightened role in the pursuit of efficiency. From Lemma 3 we know that increasing p_i implies an increase in the total value generated by each relationship between a client, firm, and worker, and thus enhances efficiency for a fixed Q_i. Considering the case where c_i is identical across all locations,

the optimal policy is therefore for each location not to enforce non-compete and non-solicitation contracts, setting $p_i = 1$ for every location. This policy achieves the constrained socially efficient rates of entrepreneurship and expected payoffs from client relationships and the optimal allocation of clients across locations.

[Analysis of differences in c_i across locations, and of policy regarding c_i taking p_i as given, in process.]

V. Endogenous Determination of the Number of Firms

We can determine the number of firms in each location by embedding our model of the previous two sections in an overlapping generations framework. In order to maintain tractability we will have to make an admittedly artificial simplifying assumption. This will allow us to generate a number of empirically testable predictions.

We now assume that agents in our model live for two periods. Clients are active in the first period of their lives and retire in the second period. Workers either remain employees or become entrepreneurs in the first period of their lives. In the second period of their lives they can retire and receive a benefit R, financed by lump-sum taxation, or they can run business service firms. Workers who became entrepreneurs in the first period of their lives can expand their existing firms by hiring workers from the new generation to serve the new generation of clients, whereas workers who remained employees must invest c_i if they wish to run similar firms. Our key simplifying assumption is that R is sufficiently high that if all previously established firms stay in business, profits per firm will be driven below R. In this case no new firms will be established by workers in the second period of their lives. Some previously established firms

will "mature" into larger firms, and the rest will "fail" in the sense that their entrepreneurs will choose to retire. This assumption insures that a worker's choice whether to become an entrepreneur in the first period of his life has no impact on his income in the second period of his life, which in equilibrium is always R.[20]

Determination of the number of "mature" or large firms in each location is now straightforward. Our model is always in a steady state. The expected firm payoff y_{Fi} gives the profit an entrepreneur can expect from each client he accepts in the second period of his life. Since his cost of entry has already been sunk, his total expected profits from running his firm in the second period of his life are $q_i y_{Fi}$. The equilibrium number of large firms (employers) is then determined by

$$q_i y_{Fi} = R \text{ or } N_i = Q_i y_{Fi}/R. \quad \textbf{(14)}$$

The equilibrium size of mature firms q_i therefore varies inversely with the profitability of client relationships, and the equilibrium number of mature firms N_i varies directly with the profitability of client relationships given the number of clients in the location. Another observable characteristic of locations is the ratio of mature to new firms, given by $N_i/\gamma_i Q_i = y_{Fi}/\gamma_i R$. Note that the number of mature firms has no aggregate welfare consequences because, in the equilibria we consider, all costs of entry have already been sunk and all retirement benefits are funded by lump-sum transfers.

We have seen how, in both our benchmark model and the main case of our model with liquidity-constrained workers, firms and workers will choose to set $p_i = 0$. If locations are to

[20]We assume that workers receive R even if they are unemployed in the first period of their lives, so that expectation of this retirement benefit has no impact on bargaining between firms and workers over their initial contracts.

differ in p_i, then, it will be because local governments have to varying degrees forced p_i above zero by refusing to enforce non-compete or non-solicitation clauses in certain circumstances. We already saw at the end of section III.B that, in the benchmark model, expected payoffs for clients are increasing in p_i and expected payoffs for firms are decreasing in p_i. The same is true in the model with liquidity-constrained workers. We can therefore state:

Proposition 7: The size of mature firms is larger and the average age of firms (ratio of mature to new firms) is lower in locations with higher p_i.

The effect of higher p_i on the average age of firms is greater in the model with liquidity-constrained workers because not only does y_{Fi} fall but γ_i (the rate of entrepreneurship) increases. The effect of higher p_i on the number of mature firms in a location is ambiguous, however, because in both the benchmark model and model with liquidity-constrained workers the number of clients Q_i increases, offsetting the fall of y_{Fi} in equation (14).

Unlike p_i, a lower cost of entrepreneurship c_i does not affect expected firm payoffs in the same direction in the benchmark and liquidity-constrained models. We are therefore unable to use equation (14) to make predictions that are robust across the models. However, the fact that expected payoffs to both clients and firms fall as c_i falls in the main case of our model with liquidity-constrained workers allows us to make a set of unambiguous predictions in this case:

Proposition 8: Consider the liquidity-constrained environment under the assumptions made in Section IV and suppose λ is sufficiently close to zero and/or γ_i is sufficiently close to one relative to the other parameters, so that p_i equals the positive lower bound set by local government. Then there are fewer and larger mature firms and the average age of firms is lower

in locations with lower c_i.

V. Conclusions and Directions for Further Research [in process]

VI. Appendix [in process]

References

Carnevale, Lawrence F. and Doran, Lorraine R. 2001. "Restrictive Employment Clauses." *New York Law Journal*, February 5.

Carnevale, Lawrence F.; Lockhart, Judith A.; and Olosunde, Kikelomo O. 1999. "Common Law and Contractual Restraints on Employee Conduct." http://www.clm.com/pubs/pub-907371_1.html.

Che, Yeon-Koo and Hausch, Donald. 1999. "Cooperative Investments and the Value of Contracting." *American Economic Review*, 89, pp. 125-147.

Chiu, Stephen W. K., and Ka-chung, Wong. 2003. "The Hollowing Out of Hong Kong Electronics: Organizational Inertia and Industrial Restructuring in the 1990s." Working Paper, Department of Sociology, Chinese University of Hong Kong.

Djankov, Simeon; La Porta, Rafael; Lopez-di-Silanes, Florencio; and Shleifer, Andrei. 2002. "The Regulation of Entry." *Quarterly Journal of Economics* 117(1), pp. 1-37.

Gilson, Ronald J. 1999. "The Legal Infrastructure of High Technology Industrial Districts: Silicon Valley, Route 128, and Covenants Not to Compete." *New York University Law Review* 74, pp. 575-628.

Klier, Thomas and Testa, William. 2002. "Location Trends of Large Company Headquarters During the 1990s." *Economic Perspectives* (Federal Reserve Bank of Chicago), 2nd quarter, pp. 12-26.

Lichtenberg, Robert M. 1960. *One-tenth of a Nation*. Cambridge, MA: Harvard University Press.

Lucas, Robert E. 1978. "On the Size Distribution of Business Firms." *Bell Journal of Economics* 9(Autumn), pp. 508-523.

Marino, Vivian. 2003. "So Long. Now Don't Compete." *New York Times*, February 16, Section 3, p. 10.

McMillan, John. 2002. *Reinventing the Bazaar: A Natural History of Markets*. New York: W. W. Norton.

Rauch, James E. and Watson, Joel W. 2002. "Entrepreneurship in International Trade." National Bureau of Economic Research Working Paper 8708, January.

Sterk, Stewart E. 1993. "Restraints on Alienation of Human Capital." *Virginia Law Review* 79, pp. 383-460.

U. S. Department of Commerce. Economics and Statistics Administration. Bureau of the Census. 1997. *1992 Economic Census. Characteristics of Business Owners.* Washington, D. C.: Government Printing Office.

Watson, Joel. 2002. *Strategy: An Introduction to Game Theory*. New York: W. W. Norton.

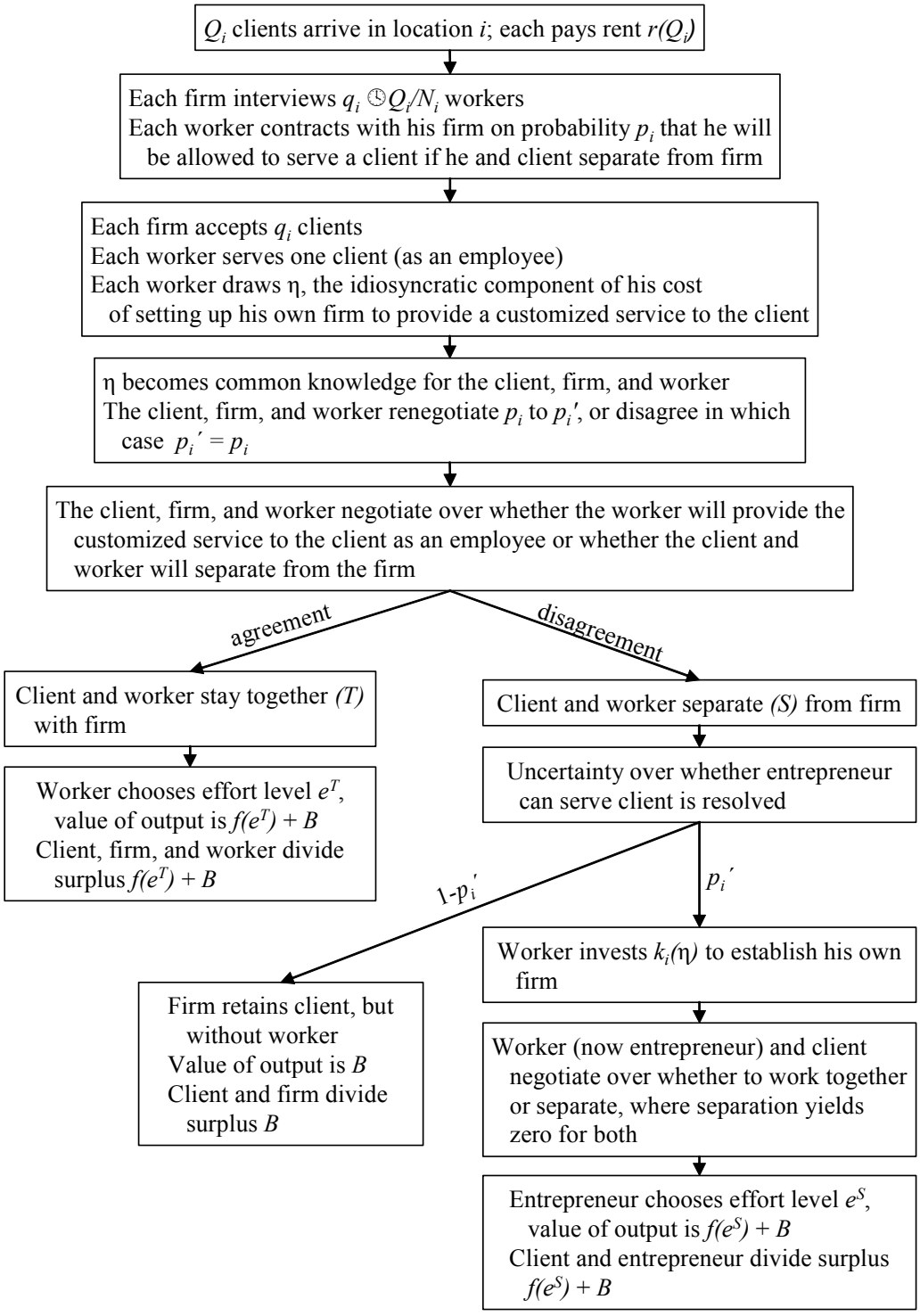

Figure 1: Timing

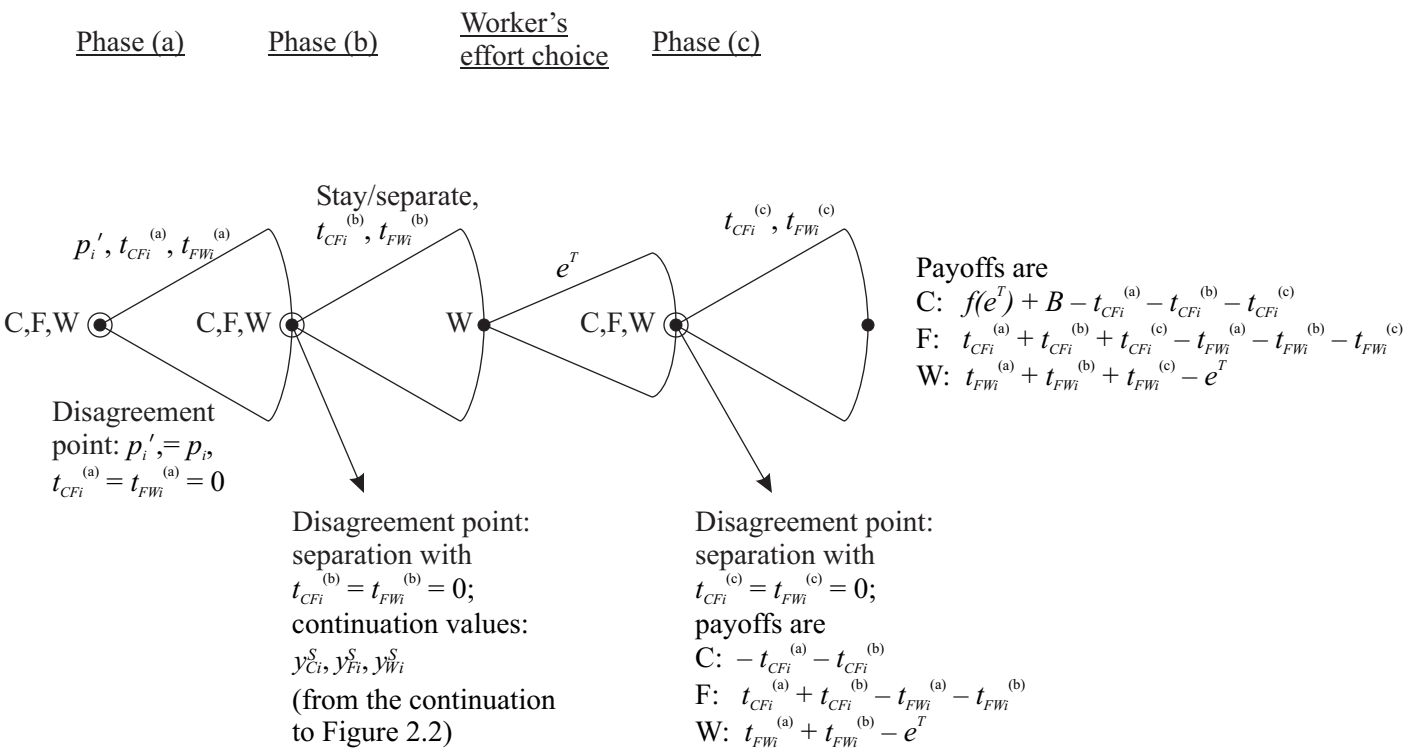

Figure 2.1: The Subgame From the Time η is Observed, Given p_i.

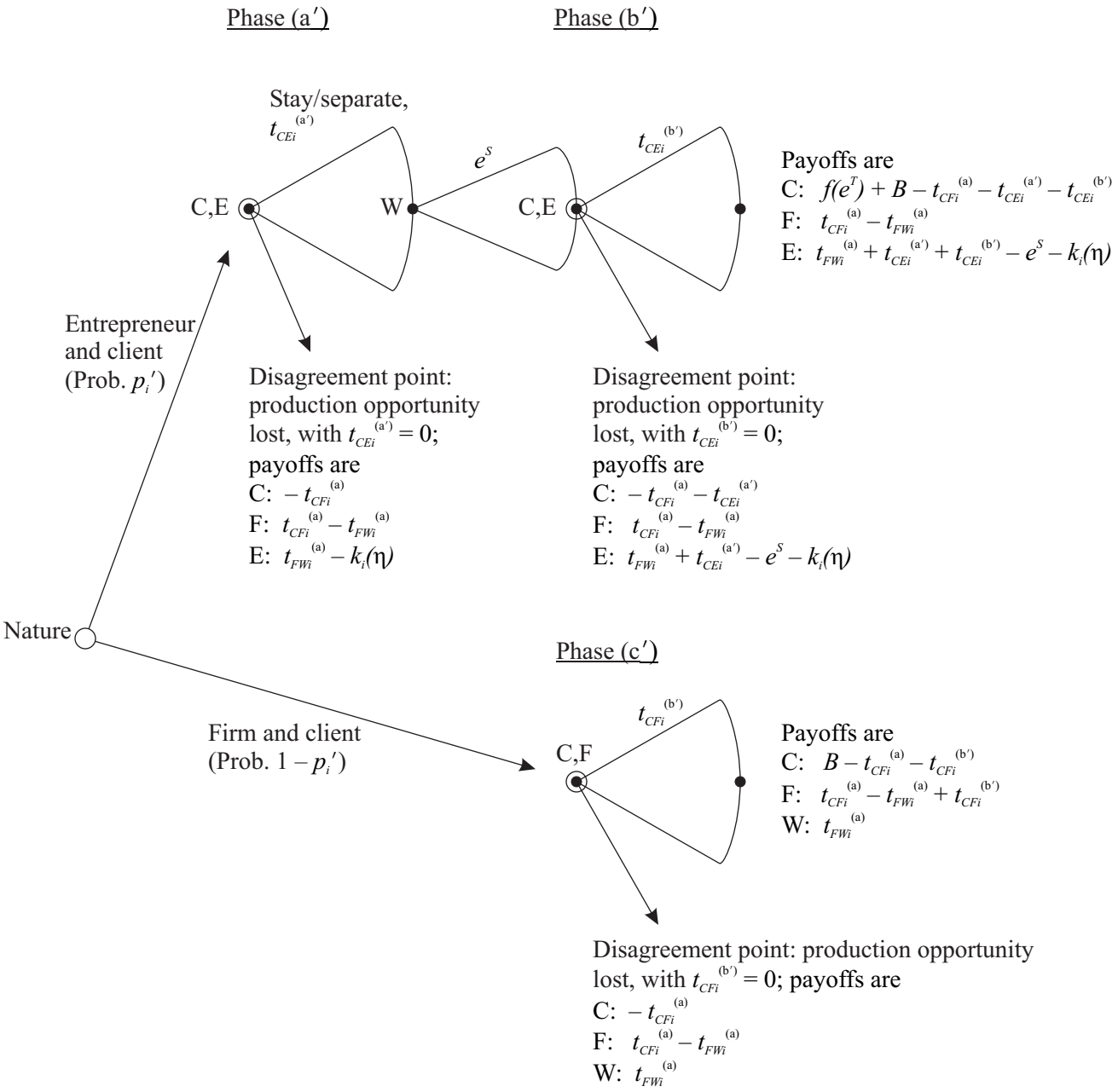

Figure 2.2: The Subgame Following Separation of the Client, Firm, and Worker at Phase (b) in Figure 2.1.

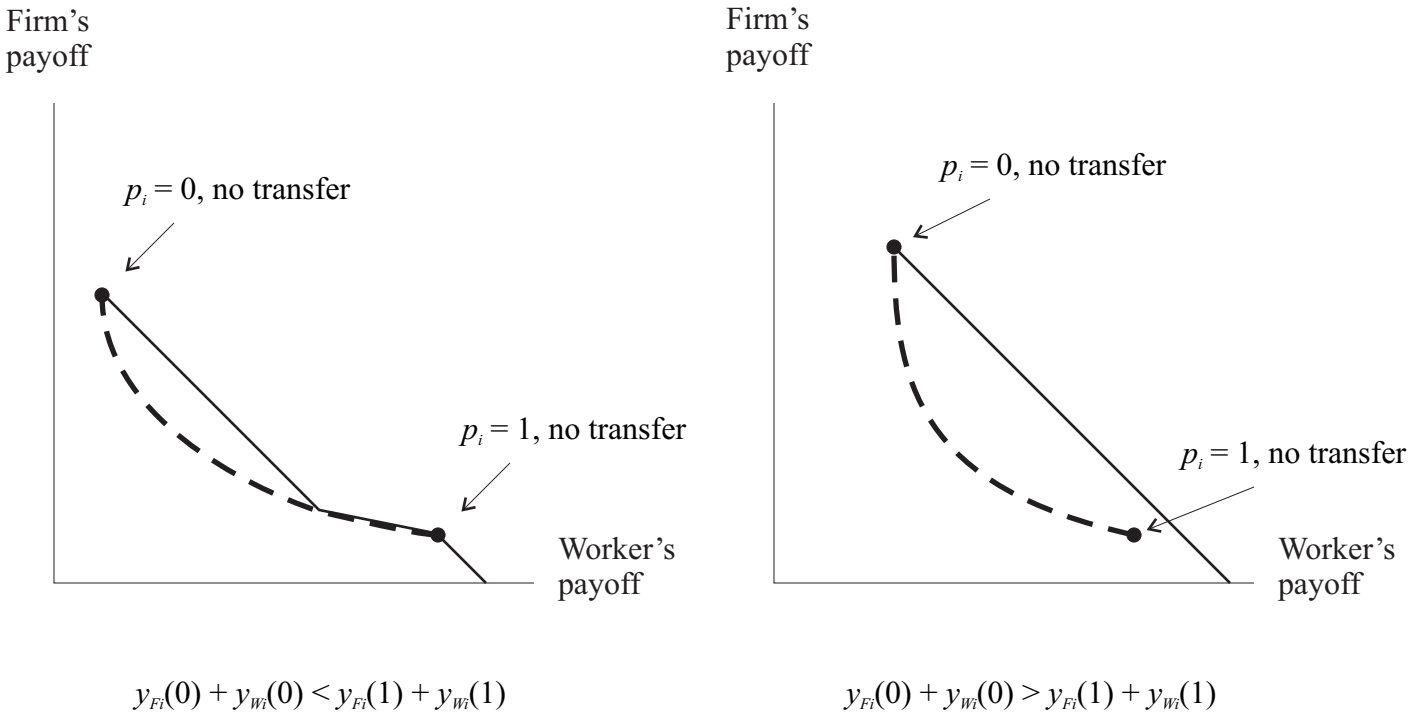

Figure 3: The Bargaining Set for the Firm and Worker's Initial Contract Negotiation.

Table 1: Continuation Payoffs (y) and Transfer Payments (t) for Figures 2.1-2.2

Continuation payoffs and transfers for Figure 2.1

If $f(e^S) - e^S - k_i(\eta) \geq f(e^T) - e^T$:

$$y_{Ci}^{T\,(a)} = (1-\lambda)(f(e^S) + B - e^S - (1-p_i)k_i(\eta))$$
$$y_{Fi}^{T\,(a)} = \lambda(1-p_i)[B + (1-\alpha)(f(e^S) - e^S - k_i(\eta))]$$
$$y_{Wi}^{T\,(a)} = \lambda p_i B + \lambda[p_i + \alpha(1-p_i)](f(e^S) - e^S) - [p_i + \alpha\lambda(1-p_i)]k_i(\eta)$$
$$t_{CFi}^{(a)} = (1-\lambda)(1-p_i)k_i(\eta)$$
$$t_{FWi}^{(a)} = -(1-p_i)\{\lambda[B + (1-\alpha)(f(e^S) - e^S)] - (1-\alpha\lambda)k_i(\eta)\}$$

If $f(e^S) - e^S - k_i(\eta) < f(e^T) - e^T$:

$$y_{Ci}^{T\,(a)} = (1-\lambda)(f(e^T) + B - e^T + p_i k_i(\eta))$$
$$y_{Fi}^{T\,(a)} = \lambda\{(1-p_i)B + (1-\alpha)[f(e^T) - e^T - p_i(f(e^S) - e^S - k_i(\eta))]\}$$
$$y_{Wi}^{T\,(a)} = \lambda p_i B + \alpha\lambda(f(e^T) - e^T) + \lambda p_i(1-\alpha)(f(e^S) - e^S) - p_i(1-\alpha\lambda)k_i(\eta)$$
$$t_{CFi}^{(a)} = 0$$
$$t_{FWi}^{(a)} = 0$$

$$y_{Ci}^{T\,(b)} = (1-\lambda)(f(e^T) + B - e^T + p_i' k_i(\eta))$$
$$y_{Fi}^{T\,(b)} = \lambda\{(1-p_i')B + (1-\alpha)[f(e^T) - e^T - p_i'(f(e^S) - e^S - k_i(\eta))]\}$$
$$y_{Wi}^{T\,(b)} = \lambda p_i' B + \alpha\lambda(f(e^T) - e^T) + \lambda p_i'(1-\alpha)(f(e^S) - e^S) - p_i'(1-\alpha\lambda)k_i(\eta)$$
$$t_{CFi}^{(b)} = (1-\lambda)(e^T - p_i' k_i(\eta))$$
$$t_{FWi}^{(b)} = \lambda(p_i' - \alpha)B + (1-\alpha\lambda)e^T + \lambda p_i'(1-\alpha)(f(e^S) - e^S) - p_i'(1-\alpha\lambda)k_i(\eta)$$

$$y_{Ci}^{T\,(c)} = (1-\lambda)(f(e^T) + B)$$
$$y_{Fi}^{T\,(c)} = (1-\alpha)\lambda(f(e^T) + B)$$
$$y_{Wi}^{T\,(c)} = \alpha\lambda(f(e^T) + B) - e^T$$
$$t_{CFi}^{(c)} = \lambda(f(e^T) + B)$$
$$t_{FWi}^{(c)} = \alpha\lambda(f(e^T) + B)$$

Continuation payoffs and transfers for Figure 2.2

$$y_{Ci}^{S} = p_i'(1-\lambda)(f(e^S) + B - e^S) + (1-p_i')(1-\lambda)B$$
$$y_{Fi}^{S} = (1-p_i')\lambda B$$
$$y_{Wi}^{S} = p_i'[\lambda(f(e^S) + B - e^S) - k_i(\eta)]$$

$$y_{Ci}^{S\,(a')} = (1-\lambda)(f(e^S) + B - e^S)$$
$$y_{Ei}^{S\,(a')} = \lambda(f(e^S) + B - e^S) - k_i(\eta)$$
$$t_{CEi}^{(a')} = (1-\lambda)e^S$$

$y_{Ci}^{S\,(b')} = (1-\lambda)(f(e^S) + B)$
$y_{Ei}^{S\,(b')} = \lambda(f(e^S) + B) - (k_i(\eta) + e^S)$
$t_{CEi}^{(b')} = \lambda(f(e^S) + B)$

$y_{Ci}^{S\,(c')} = (1-\lambda)B$
$y_{Fi}^{S\,(c')} = \lambda B$
$y_{Wi}^{S\,(c')} = 0$
$t_{CFi}^{(c')} = \lambda B$