Essays in Microeconomic Theory

Justin Pappas Johnson

B.A. Mathematics
University of Florida, 1994

Submitted to the Department of Economics in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in Economics at the Massachusetts Institute of Technology.

June 1999

Author: ....
Department of Economics

Certified by: ....
Daron Acemoglu
Professor of Economics

Certified by: ....
Glenn Ellison
Professor of Economics

Accepted by: ....
Peter Temin
Professor of Economics
Chairman, Departmental Committee on Graduate Studies

Essays in Microeconomic Theory

Justin Pappas Johnson

Submitted to the Department of Economics on April 15, 1999 in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Abstract

Three essays are presented which explore how strategic decision-making on a micro level translates into macro effects. Careful attention is paid to how asymmetric information and free-riding exert strong influences on the behavior of individuals.

In the first chapter, a simple model of open-source software development is presented. It is found that either too little development or redundant development effort can occur. While any redundant research effort grows slowly with the size of the community, projects for which user valuations are sufficiently extreme, such as solutions to the Year 2000 Computer Problem (Y2K), will result in significant waste relative to a traditional closed-source environment. Correlations between value and cost are shown to resolve the empirical puzzle as to why some extremely useful and fairly simple software does not get written while more complex software sometimes does. It is shown that a modular design can improve or worsen the performance of an open-source community.

In the second chapter, an industry is considered in which new firms require time to learn whether they have the "right stuff" to grow in size and profitability in the long run. The critical input market (that for skilled labor) is imperfectly competitive. By extending the literature on nonstationary dynamic bargaining, analysis is performed on a set of intertemporal externalities exerted by future parties on today's parties, and vice versa. The results suggest why, even if firms are able to write detailed contingent contracts with their current employees, inefficient levels of firm entry will generally exist. The theory also sheds some light on the continuing debate over the contribution of small firms to economic growth.

In the third and final chapter, players in a war of attrition care about the identity of the winner, even when they lose. In particular, a three player war of attrition is considered. Two "team" players enjoy a fraction of their valuation when their partner wins. The remaining, "solo" player benefits only by winning the war. Imposing team symmetry, the solo player drops out more quickly then either team player. The incentive to avoid fighting costs by free riding on a teammate is outweighed by a strategic commitment effect. Team players thus continue to fight even when they have no chance of winning a subsequent two-player subgame with the solo player. Examining limiting results, when the "caring coefficient" between the team players is small, a selection result obtains: The solo player drops out immediately, allowing the team players to then compete in a standard two-player war of attrition.

Thesis Supervisor: Daron Acemoglu, Professor of Economics

Acknowledgements

I thank the MIT Economics Department as a whole for the unbelievably enriching experience it has provided me. I am honored to have been a member of the MIT community. I gratefully acknowledge the support of the National Science Foundation.

I thank my advisors Daron Acemoglu, Peter Diamond and Glenn Ellison for all of the thoughtful comments and assistance they have given me. Their honesty and encouragement have meant a great deal.

I attribute some of my intellectual growth to my peers. Jonathan Dworak and David Myatt are the finest minds I have known, and I have truly enjoyed the great number of economic discussions we have shared. Also, I acknowledge Jon and David as co-authors of the third chapter of this dissertation.

Travis Broughton has been instrumental in raising my understanding of computers and the computer industry. His comments about open-source software have been especially helpful.

Insofar as graduating from MIT is the end of a critical phase of my life, it is fitting for me to thank my mother, Mary Alice Pappas, for the incalculable contributions she has made to my life thus far. I also thank my sister, E. Barrett Pappas, for her encouragement and support.

Contents

Chapter 1

Economics of Open-Source Software

1.1 INTRODUCTION

Several of the most important software applications available today were developed by people who never expected nor received monetary compensation for their efforts. This software is, has been, and always will be freely available to anyone who wishes to acquire it.

Much of this software was written in a decentralized fashion by a large number of individual programmers around the world, each working in isolation. Yet the sum of these efforts has built an amazing body of useful, well-integrated, reliable, and free software.

For years the most popular web server has been produced by Apache. The server's source code (i.e. the actual typed commands entered by programmers that constitute the logical structure of the program) is freely available, and Apache actively encourages people to acquire and modify it. A second example is the open-source program Sendmail, which moves 80% of email across the Internet every day. Sendmail's email duties are in turn assisted by open-source Domain Name Servers, which (roughly) keep track of all the addresses on the Internet. Third, the source code for Netscape's influential web browser is freely available. A fourth example is the open-source operating system Linux, which since its creation in 1991 has been greatly developed and improved by the labors of a vast collection of programmers worldwide, while a fifth is the open-source programming language Perl, and its libraries. It can fairly be said that the Internet would be vastly different if these programs did not exist.

A competent programmer who has a program's source code can figure out exactly how the program executes. This allows a programmer to alter a program to suit his or her own preferences, to correct bugs in the program, or to use the program as the base for another different program. This ability to use ones own programming skills to change the functionality of a program can be of considerable value to a serious software engineer.

In contrast, most software that one buys has already been compiled to run on a particular operating system. It is exceedingly difficult to invert a compiled program to obtain the underlying source code. Software companies usually have little incentive to release source code, since the knowledge and technical expertise embodied therein is potentially helpful to competing software firms.

The source code of open-source software is freely available. But open-source software is much more than just software for which the source code is available. Open-source programs are distributed under very precise licensing agreements, of which the GNU (Gnu's Not Unix) General Public License is the most popular. The GNU license grants specific legal rights and responsibilities to those who use GNU-licensed products.

The GNU General Public License grants everyone the right to use, copy, distribute, and modify a piece of software (Stallman 1996). However, it also demands that the source code of any changes or enhancements made using the original source code be freely available, and that any such modifications to the software be distributed under the terms of the GNU license itself. Furthermore, modifications or redistributions of open-source software must make the terms of the license apparent to others who might obtain or consider obtaining the software.

All source code that incorporates open source code becomes open itself. Thus open source code cannot be used as an input for a monopolized market, and open source software cannot be made proprietary.

There are many economic issues associated with the workings of open-source licenses in business. In this paper, however, we focus on the positive analysis of an open-source community, and on comparing it to a traditional business model. That is, we ask how well an open-source community can expected to perform with regards to the enhancement and extension of an existing piece of open-source software.

There are many popular arguments for open-source software. Many of these are one-sided or ignore potentially important economic issues. For example, one might say open-source software is good because it is free and therefore more people can use it. This is likely correct, but clearly neglects the important fact that someone needs to invest valuable resources in developing the software.

Another argument is that when source code is freely available, users are more likely to correct problems and add features to the software quickly. The suggestion is that parallel work by many users will lead to resolution of problems very quickly. However, parallel efforts could well involve wasteful duplication of efforts, so that a centralized closed-source model might be preferable from an efficiency standpoint.

We investigate when the open-source model of software development is likely to perform well. We focus mainly on the incentives of individual users, who cannot be compensated for their programming efforts. We further consider the severity of the redundant efforts in the community.

After formulating a simple model that permits the analysis of these issues, we consider how a traditional closed-source or proprietary model performs relative to an open-source model. We show that both models are suboptimal relative to the constrained first-best, but that they fail for different reasons.

We then extend the basic model to explore the importance of incremental improvements (or "modularity") and complexity in open-source software development. Our results are extremely consistent with observed facts.

1.2 Setting up the Model

Open-source software development is modeled as the private provision of a public good. Such models of public good provision have been studied by many people, including Chamberlin (1974), Palfrey and Rosenthal (1984), and Bergstrom, Blume, and Varian (1986). The model presented here is particularly suitable for analysis of the open-source software environment, as will be explained below. Furthermore, asymptotic results are of relatively greater interest in the current context and hence are, perhaps, better developed than those in earlier papers.

Consider the following simultaneous-move game. There are n user-developers in the Internet community. Each knows that an enhancement of a pre-existing software application, the source code of which is open, can potentially be developed. Developing the enhancement of the software takes time, effort, and ingenuity. These costs are summarized for each agent by his or her privately-known cost of development c_i.

Each agent independently decides whether to develop the new application. Any agent i who chooses to develop bears the cost c_i. As long as at least one agent so chooses, the development will occur. We imagine that any developed software can be freely provided over the Internet to the other user-developers, and that it will be so provided if developed (because the terms of the open source contract vastly restrict the developing agent's ability to profit).

Therefore, if the enhancement is developed, all agents receive their respective, privately known valuations v_i. If the software is not developed, all payoffs equal zero.

Suppose that all agents' costs and valuations are independent, identical draws from the joint distribution function $G(c, v)$, with support on the finite rectangle defined by $\{(c, v) : c_L \leq c \leq c_H,\ v_L \leq v_H\}$ where $c_L > 0$ and $v_H \geq 0$. We assume this is a smooth function.

The first object of analysis is the optimal response of any agent i to the strategies of the other agents. Suppose that the agent believes that the probability that the development will take place if he does not innovate himself is π. Then he optimally chooses to develop the program himself if and only if

$$v_i - c_i > \pi v_i$$

which can be rearranged to yield

$$\frac{v_i}{c_i} > \frac{1}{1-\pi}$$

so it is clear that the optimal response is to invest in development if and only if the value-to-cost ratio is sufficiently high.

Assuming a symmetric equilibrium, let $\hat{q}$ denote the critical value-to-cost ratio. Also, let $F(q)$ be the induced distribution of these quotients;[1] that is,

$$F(q) = \Pr\left\{\frac{v}{c} < q\right\}$$

the upper bound of which we will denote by

$$q_H = \frac{v_H}{c_L} < \infty$$

It will also be convenient to define

$$\gamma = \Pr\{\text{No agent develops}\} = F(\hat{q})^n$$

[1] We do not take the value-to-cost distribution as primitive because we will later compare how a closed-source system performs relative to an open-source one. To do so we will need the joint distribution of v and c.

Given these definitions, the probability (from an individual agent's perspective) that none of the remaining agents develops is

$$1 - \pi = F(\hat{q})^{n-1} = \gamma^{\frac{n-1}{n}}$$

and hence we can determine from the agent's optimality condition that he will only be indifferent between developing and not when

$$q_i = \frac{1}{1-\pi} = \gamma^{\frac{1-n}{n}}$$

which of course must equal $\hat{q}$.

Hence, given γ and the imposed symmetry among the remaining agents, we can determine what the critical value of q is for the remaining agent in terms of γ. Plugging this back into the definition of γ, we deduce that γ is an equilibrium value if

$$\gamma = F\left[\gamma^{\frac{1-n}{n}}\right]^n \tag{1.1}$$

which has a unique solution unless the law F places no mass above 1, in which case there is no solution to this equation and the unique equilibrium exhibits no development. To avoid this boring case, we will assume that $F(1) < 1$, or equivalently that $q_H > 1$. Under this assumption, the above condition is both necessary and sufficient for γ to be an equilibrium value.

Having now demonstrated how to solve the model with a fixed population size, we turn to a deeper analysis. In the following sections we will analyze a number of issues of relevance in drawing conclusions about open-source communities.

1.3 The Number of User-Developers

Some open-source projects have a greater number of potential user-developers in the community than other open-source projects do. Here we investigate how the open-source environment is influenced by the population size n. In particular, we examine how the equilibrium probability of development $1-\gamma$ changes, and more generally how social welfare changes. We consider both finite and asymptotic results.

Suppose that the number of user-developers increases. If individuals continued to use their original threshold rule, then clearly γ would decrease. However, when more individuals are present, the incentive to free ride is raised, and any individual will be less likely to develop the application himself in equilibrium. Whether γ falls or rises as a result of including more agents is therefore ambiguous.

On the other hand, social welfare unambiguously rises as the number of user-developers increases. The reason is that, in equilibrium, although the movement of γ_n is ambiguous, π_n must increase (where we now use subscripts to denote the equilibrium values for a given population size n). We know that π_n must increase, because if π_n were to fall with the addition of another agent, each agent would optimally choose to develop more frequently, contradicting the fall in π_n.

Since the probability that any other agent develops the project increases with n it follows that each individual is better off conditional on any realization of his or her own cost and value, and hence is better off unconditionally. As social welfare can be expressed as the sum of individual welfare, society is better off, too. Hence, growth in the population constitutes a Pareto improvement.

It is not true, however, that individuals or society is better off in each state of nature. For example, there are states in which the addition of another user results in the project not being developed when it would have been developed in the absence of the marginal user. This is true precisely because the threshold $\hat{q}_n$ is rising with n. However, individuals are better off in states of the world where the marginal user develops and they do not, and where they themselves would have developed otherwise.

We require the following lemma.

Lemma 1 *The equilibrium probability that one of the first $n-1$ agents develops is increasing in n. That is, π_n is increasing in n. Also, $\hat{q}_n$ is increasing.*

Proof: Recall that an individual agent i is indifferent between developing and not when

$$q_i = \frac{1}{1-\pi_n}$$

In equilibrium, we also have

$$\pi_n = 1 - F(\hat{q})^{n-1} = 1 - F\left[\frac{1}{1-\pi_n}\right]^{n-1}$$

For each value of x the function $F(x)^{n-1}$ is decreasing in n. Therefore, the point π_n at which the above condition holds is strictly increasing in n

(since $\hat{q}_n$ can never equal q_H in equilibrium). This fact plus inspection of the agent's optimization problem reveals that $\hat{q}_n$ is also increasing. ∎

We can now demonstrate our earlier claim that social welfare is increasing in the number of agents present.

Theorem 1 *Expected social welfare is increasing in n. Moreover, the expected welfare of each user is increasing in n.*

Proof: Denote the expected payoff to agent i, conditional on his type and the total number of agents n, by $x_i(v_i, c_i, n)$. Then

$$x_i(v_i, c_i, n) = \max\left[v_i\pi_n,\ v_i - c_i\right] \leq \max\left[v_i\pi_{n+1},\ v_i - c_i\right] = x_i(v_i, c_i, n+1)$$

since $\pi_n \leq \pi_{n+1}$.

Hence agent i's payoff is increasing in n conditional on his type. But since this is true for every type of i, his ex-ante payoffs Ex_i are also increasing in n. Finally, observe that his payoff is never negative.

It follows that expected social welfare with n agents can be expressed as

$$\sum_{i=1}^{n} Ex_i(v_i, c_i, n) \leq \sum_{i=1}^{n} Ex_i(v_i, c_i, n+1) \leq \sum_{i=1}^{n+1} Ex_i(v_i, c_i, n+1)$$

where the last term is expected social welfare with $n+1$ agents. This proves the theorem. ∎

1.4 Limiting Results

One of the major arguments put forth for why the open-source paradigm should be successful is that open source code allows an extremely large labor force (potentially the entire Internet community of programmers) to bring its skills and insight to bear on any problem. The notion prevails among open-source advocates that no single company can hope to match this accumulation of talent.

Thus it is natural to examine the behavior of our model as the pool of user-developers grows very large. We investigate the limiting likelihood of innovation and the asymptotic distribution of redundant development effort.

1.4.1 Development Probability

We begin by considering what happens to the probability of development $1-\gamma_n$, and the probability π_n that any $n-1$ of the n users develops the software, when the population grows large. The following is immediate.

Theorem 2 *Both π_n and γ_n have limiting values. In particular*

$$\gamma^* = \lim_{n\to\infty} \gamma_n = \frac{1}{q_H}$$

$$1-\pi^* = \lim_{n\to\infty} 1-\pi_n = \frac{1}{q_H}$$

Proof: We know that a unique equilibrium exists for each n. Hence, all we need to demonstrate is that for any $\epsilon > 0$ there exists an N such that for $n > N$ the equilibrium value of γ_n lies in $(\gamma^*-\epsilon, \gamma^*+\epsilon)$.

Let $\epsilon > 0$ be given. It is clear that $1/(\gamma^*+\epsilon) < q_H$ and hence for some N_1 we have that $n > N_1$ implies $(\gamma^*+\epsilon)^{(1-n)/n} < q_H$.

Since $F(q_H) = 1$ and F is strictly increasing on its support, it must be that for $n > N_1$

$$F\left[(\gamma^*+\epsilon)^{\frac{1-n}{n}}\right] < 1$$

which implies that

$$F\left[(\gamma^*+\epsilon)^{\frac{1-n}{n}}\right]^n$$

converges to zero. In particular, there is N_3 such that (1.1) can not be satisfied at $\gamma^*+\epsilon$, or for any greater value, since F is an increasing function and since the map $x \mapsto x^{(1-n)/n}$ is decreasing.

Now consider $\gamma^*-\epsilon$. There is some value N_2 such that $n > N_2$ implies that $(\gamma^*-\epsilon)^{(1-n)/n} > q_H$ and hence that $F\left[(\gamma^*-\epsilon)^{(1-n)/n}\right]^n = F(q_H) = 1$ since F is a distribution function. This implies neither $\gamma^*-\epsilon$ nor any point less than it can satisfy (1.1).

We conclude that for $n > \max[N_1, N_2, N_3]$ it must be the case that $\gamma_n \in (\gamma^*-\epsilon, \gamma^*+\epsilon)$. Since ϵ was arbitrary, the result follows. ■

This theorem reveals that it is only the upper bound of the underlying distribution of value-to-cost ratios that has any influence in the limit. The exact shape of the upper tail is of no consequence.

This is intuitive because, in the limit, only the agents with the highest value-to-cost ratios will develop the software. Hence the asymptotic value of γ_n must be such that it keeps an agent of type q_H indifferent. The indifference condition for this agent does not depend upon the shape of the distribution at all, and hence the limiting values depend only on thc upper bound of F.

For example, consider the two following two cumulative distribution functions with support on $[0, 2]$: $x \mapsto \frac{x^2}{4}$ and $x \mapsto x - \frac{x^2}{4}$. The corresponding densities are mirror images around the midpoint of support 1. The first places the greatest density at 2 while the second places zero density at 2. In finite populations, the first exhibits significantly higher likelihood of development. But the probability of development converges in either case to 1/2.

If we were to graph the trajectory of γ for these two distributions, we would observe that the first distribution's γ is converging to 1/2 from below, while the γ associated with the second distribution is converging from above.

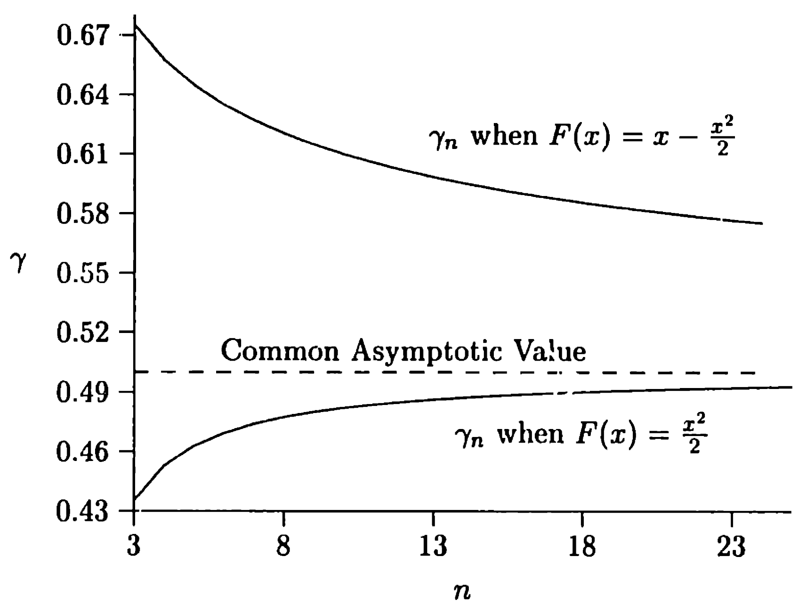

Figure 1.1: Trajectories of γ

The participation threshold must be approaching 2. Hence, heuristically, since the density in the upper tail of the first distribution is very high, small increases in the participation threshold $\hat{q}$ tend to increase γ significantly, whereas such changes have little effect on the second distribution. Thus, since both share the samė limiting γ, they must be converging from opposite directions, with the high-density tail γ rising, and the low-density tail falling.

Theorem 3 *If $f(q_H) > \gamma^*$ then there exists an N such that γ_n is increasing for all $n > N$. If $f(q_H) < \gamma^*$ then there exists an N such that γ_n is decreasing for all $n > N$.*

Proof: The proof relies upon the following fact: If the value-to-cost distribution is uniform, then $\gamma_n = \gamma^*$ for all n. Since the density of a uniform distribution is constant, this allows us to compare any other distribution at some equilibrium point to a suitably-chosen uniform distribution.

To see that this is true, suppose that the distribution is uniform on $[0, q_H]$. All we need to do is show that (1.1) is satisfied at $1/q_H = \gamma^*$ for all n. We see that

$$\frac{1}{q_H} = F\left[\left(\frac{1}{q_H}\right)^{\frac{1-n}{n}}\right]^n = \left[\frac{1}{q_H}\left(\frac{1}{q_H}\right)^{\frac{1-n}{n}}\right]^n = \frac{1}{q_H}$$

which proves our claim about the uniform distribution.

Now, let F be any distribution that satisfies the assumptions of our model, and let γ^* be the limiting value of γ_n for this distribution. We will show that if $f(q_H) > \gamma^*$, then γ_n is eventually increasing. The converse is proven similarly.

Define $g(n,\gamma) = \gamma^{(1-n)/n}$ and then differentiate the right hand side of (1.1) to yield

$$F(g(n,\gamma))^n \left[\log[F(g(n,\gamma))] + \frac{nf(g(n,\gamma))}{F(g(n,\gamma))}\frac{\partial g(n,\gamma)}{\partial n}\right]$$

We know that this value equals zero at γ^* for a uniform distribution with $q_H = 1/\gamma*$. Also, all components of the above equation evaluated at γ^* will equal those of such a uniform, save for the density f. Therefore, the sign of the above equation is determined solely by whether $f(g(n,\gamma^*))$ exceeds γ^*, the density of the uniform at that point. This is in turn equivalent to $f(q_H) > \gamma^*$ since $g(\gamma)$ converges to $1/\gamma$ with n uniformly on any closed subinterval of $(0,1)$.

Finally, since γ_n converges and the above equation is continuous, there must be some N such that for $n > N$ the above equation is always positive when evaluated at γ_n. Therefore, γ_n is increasing eventually. ■

Hence, although the limiting value of γ is determined only by the upper bound of F, the direction of γ_n depends on the density f. For some distributions (for example, ones with monotone density), it will be easy to determine the movements of γ_n even for small n. However, for any distribution we can compute the direction from which the development probability converges asymptotically.

1.4.2 Costs and Redundancy

It is difficult to deny that many open-source applications are more reliable than their proprietary counterparts. For example, Miller, Koski, Lee, Maganty, Murthy, Natarajan, and Steidl (1998) found that failure rates of commercial versions of UNIX utilities ranged from 15-43%, in contrast to failure rates of 9% for Linux utilities and just 6% for GNU utilities.

It is commonly believed that a large number of user-developers working in parallel quickly produce new developments and exterminate bugs. Linus's Law states that "Given enough eyeballs, all bugs are shallow."

Our results so far suggest that any single bug becomes only so shallow, no matter how large the population of user-developers becomes. A proper accounting of the value of an open-source community must, moreover, acknowledge that having a great number of eyeballs looking at source code could result in wasteful duplication of effort.

We have seen that there is a convergent probability that any development occurs. We now ask what can be said about the actual number of development efforts in the limit. Define p_n to be the probability that any individual in a population of size n chooses to develop. Then, for a fixed population, the expected number of developments equals np_n. Of course, p_n is approaching zero as $\hat{q}_n$ approaches q_H, while n becomes infinite. The following theorem characterizes this quantity asymptotically.

Theorem 4 *The expected number of development efforts converges as the population grows. In particular,*

$$\lim_{n \to \infty} np_n = \log(q_H)$$

Proof: Let $\alpha = \log(q_H)$. We show that for any $\epsilon > 0$, there is an N such that for all $n > N$

$$\frac{\alpha - \epsilon}{n} < p_n < \frac{\alpha + \epsilon}{n}$$

Note that this does not simply say that p_n converges to zero, which is trivial. Rather, this is a precise statement regarding an approximation to the rate at which p_n falls to zero.

If this can be demonstrated, then the theorem immediately follows, since for any $\epsilon > 0$ we have eventually

$$\alpha - \epsilon < np_n < \alpha + \epsilon$$

Suppose the upper bound does not hold eventually for some ϵ. Let $\alpha^+ = \alpha + \epsilon$. Then there is a subsequence n_j such that

$$p_{n_j} > \frac{\alpha^+}{n_j}$$

for all j. The definition of γ_n plus this assumption gives

$$\gamma_{n_j} = (1 - p_{n_j})^{n_j} < \left(1 - \frac{\alpha^+}{n_j}\right)^{n_j} \to e^{-\alpha^+}$$

by a commonly known (and easily proven) result. But since γ_n converges, it converges along all of its subsequences. Since $\alpha = \log(1/q_H) = -\log(\gamma^*)$ we deduce that

$$\gamma^* \leq \gamma^* e^{-\epsilon} < \gamma^*$$

a clear contradiction.

Similarly, define $\alpha^- = \alpha - \epsilon$, and suppose that the lower bound fails along some subsequence. Then we have

$$\gamma_{n_j} = (1 - p_{n_j})^{n_j} > \left(1 - \frac{\alpha^-}{n_j}\right)^{n_j} \to e^{-\alpha^-} \Rightarrow \gamma^* \geq \gamma^* e^{\epsilon} > \gamma^*$$

another contradiction.

Therefore, p_n eventually conforms to both of the bounding functions given. The theorem follows. ∎

The incentive to free ride is strong enough to bound the amount of redundant effort in the limit. This fact in turn implies that, regardless of the underlying joint distribution of values and costs, the (random) number of development efforts follows a well-defined distribution asymptotically.

Corollary 1 *The number of development efforts converges to a Poisson random variable with mean* $\log(q_H)$.

Proof: This is a special case of a more general class of theorems regarding the limit of a sum of variables with convergent mean. See, for example, Ash (1972). ■

Our results up to this point allow us to compute the total limiting costs incurred by an open-source community. It turns out that the highest value-to-cost users are also, in the limit, the least-cost developers.

Hence, when development does occur, costs are borne efficiently, although too much work can still take place. This is important, and in contrast to finite sample properties of the model.

Theorem 5 *The total expected costs of development borne by the open-source community converge to* $c_L \log(q_H)$.

Proof: We know that $\hat{q}_n$ converges to $q_H = v_H/c_L$, since the underlying joint distribution $G(c, v)$ has support on the entire rectangle $\{(c, v) : c_L \leq c \leq c_H,\ v_L \leq v_H\}$. This implies that, eventually, the only way an agent can be developing is if both his value and cost are at the extremes. ■

It is important to note that this result relies heavily upon the rectangular support of $G(c, v)$. If the region of support were, for example, circular then it would not be true that the highest values of v/c corresponded to the lowest values of c. But insofar as we accept that being the lowest-cost user does not preclude being the highest-value user, the least-cost users will be the only developers in the limit.

The logarithmic rate of increase in costs is, on the one hand, fairly slow. For example, suppose that the joint value and cost distribution is such that c is constant at c_L and v is uniformly distributed on $[0, q_H]$. Then expected welfare when the population n is large is approximately $n(1 - 1/q_H)q_H/2 - c_L \log(q_H) = n(q_H - 1)/2 - c_L \log(q_H)$. When q_H increases, since the marginal social value is linear while marginal costs are only logarithmic, social welfare improves at a fast rate.

On the other hand, total costs can be quite substantial. In fact, if the support of F were unbounded, the amount of redundant effort would become unbounded as n grew. The reason is that users with extreme valuations will not be able to tolerate even a tiny probability of no development, and hence will be forced to invest their own resources.

We can imagine, therefore, that some types of software developments will result in extreme levels of inefficiency. For example, the Year 2000 Computer Problem, in which a design flaw of some computer software causes the software to become severely dysfunctional, is reasonably seen as something influencing a huge population of users, some of whose values must be gigantic (corporations and governments, for example). If decentralized open-source software development were the only available mechanism for circumventing this problem, we might expect tremendous waste to result relative to a situation in which firms could find and contract with low-cost developers, and in which software firms could credibly assure corporations that the problem would be solved.

1.5 Complexity and Correlation

We have described the workings of a simple model of open-source software development in a fairly complete manner. We have shown how the celebrated problem of private provision of a public good pops up in the economics of open-source software.

Usually, one might have little reason to believe that there is reason to distinguish between the value and cost of individuals in a model of private provision of a public good. In fact, there are two compelling reasons to do so in the current context. First, allowing for such a generalization provides insight into which types of software are most likely to be developed, and thereby provides a simple explanation to an empirical puzzle, as discussed below. Second, allowing for a primitive joint distribution of value and cost allows us to meaningfully compare an open-source model to a closed-source model.

We turn first to the question of what types of software an open-source community is most likely to develop. We might wonder, for example, if there are complexity limits to what a decentralized collection of programmers can achieve.

Empirically, it appears that if there is such a threshold, it is very high indeed. The shining proof is the open-source operating system Linux. The kernel of any operating system is extraordinarily complex, yet that of Linux is widely-regarded as being more stable than that of Windows NT, for example. The Linux kernel has been built collectively since 1991 by a worldwide body of programmers who receive no monetary compensation. In fact, new versions of the operating system tend to appear at a very fast rate, as bugs are identified and corrected, and overall efficiency rises.

A further example is the amazing program Emacs. While Emacs is best known as a text editor, in fact it is capable of doing a great variety of tasks: sending and receiving email, compiling and debugging C and Lisp programs, acting as a front-end for programming in many languages, maintaining a calendar and diary, acting as a command shell, acting as a front-end for telnet and ftp, editing graphics, and much more. The Lisp libraries that support much of Emacs' functionality have received contributions from over 150 programmers, including the eclectic open-source champion and MacArthur Foundation Fellowship recipient Richard M. Stallman.

Linux and Emacs are breathtaking in their complexity and functionality. It is perhaps surprising, therefore, that despite the manifest ability of a loose collection of programmers to surmount extreme technical difficulties, there is a dearth of certain types of software.

A puzzle in the open-source community is why some obviously useful software does not get written. Foı example, there is no open-source word processor available, even though hundreds of other free utilities and applications exist.

An argument put forth by Eric S. Raymond[2] is that open-source programmers wish to establish a reputation for ingenuity in the greater hacker community (Raymond 1998). Thus, projects that are considered more exciting are more likely to be developed.

While it might well be true that hacking the kernel of an operating system is much more glamourous than writing code for a word processor, reputation-building effects do not provide a convincing resolution of this puzzle. The reason is that much difficult work is performed in relative obscurity: most programmers would be hard-pressed to list many contributors to either Emacs or LaTeX, much less the designers of lesser utilities.

Consider instead the following explanation. We can suppose that production processes in the world take both computer software and human capital as inputs. Human capital can be classified as either technical skill or business skill. The production process for software, however, only takes technical skill as an input, as measured by the cost of developing the software.

Imagine there are two potential developments. One is a word processor, and the other is a networking utility. Both are equally valuable as inputs in the ultimate output market, as measured by the marginal distribution of v, and both require the same inputs to develop, as measured by the marginal distribution of c. However, the networking utility is appreciated most by people who will be linking computer networks and establishing mail and web servers, while the word processor is most appreciated by those involved

[2]Eric S. Raymond is a programmer and well known open-source software advocate. He was influential in Netscape's 1998 decision to release its browser source code.

in on the marketing side of the business.

Given that the people who value the networking utility most also are those most able to build it, there is a natural negative correlation between value and cost when considering the network utility, and a positive correlation when considering the word processor. This is not to say that all people who are good programmers care about the networking utility but not about the word processor. Indeed, software engineers who work mostly with graphics applications won't care at all about the network utility, but might like having the word processor, for example. Hence, while the value-to-cost distribution for the networking utility is not, for example, a first-order stochastic shift of the other distribution, we can nonetheless reasonably expect that the networking utility is more likely to be developed.

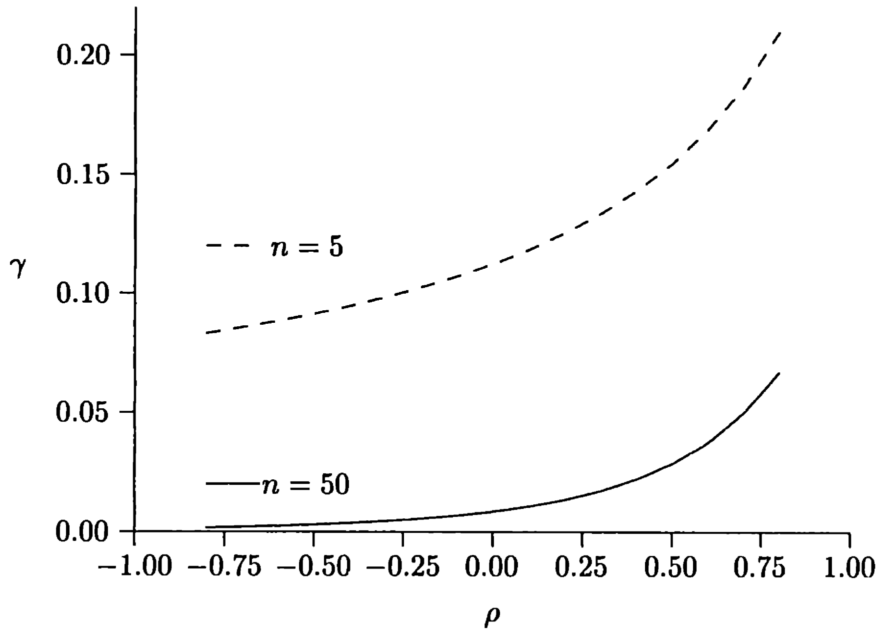

Figure 1.2: Correlation Diagram

In Figure 2 we consider the situation when value and cost have a joint log-normal distribution with correlation coefficient given by ρ, and various population sizes. Changing the correlation coefficient does not alter either marginal distribution, but influences the distribution of the value-to-cost ratio.

We see that as ρ falls, the open-source community performs better, as measured by the decrease in γ. We furthermore see that the marginal changes in the development probability depend significantly on the population size. It appears that when n is quite large, the increase in development probability as the correlation between value costs becomes more negative becomes less pronounced.

The second reason that allowing for a general distribution over both costs and values is useful is because it allows us to address the relative performance of an open-source model of development to that of a closed-source model. We imagine that a software company has already sold a product to n individuals, but has not revealed the source code. Individuals are thus unable to develop independently.

There is a commonly-known potential for product enhancement. The firm has one engineer,[3] whose cost is c. If the firm chooses to write the software enhancement, it will then sell it at the profit-maximizing price. The firm will only produce if its maximum expected revenue exceeds the opportunity cost of having its engineer work on the program.

There are two sources of inefficiency present. The first is standard: monopoly pricing leads to deadweight loss. The second is more subtle: the total amount of resources that can be used in development efforts is too low, because the monopolist has only one engineer.

This is a genuine concern of the open-source community. Not all programmers are equally clever, and moreover not all programmers are equally clever across different development tasks. When source code is unavailable publicly, the substantial human capital and insight present in the community as a whole cannot be harnessed.

The monopolist ignores the distribution of costs in society, being interested solely in the marginal distribution of valuations. On the other hand, a monopolist considers all individual users whose values are sufficiently high, whereas open-source users care only about their own realized cost and value.

Neither system is optimal. In some circumstances one model will tend to outperform the other. For example, we expect that the closed-source system will tend to serve high-value, high-cost individuals much better than the open-source model, so that we are not surprised that many proprietary word processors are available. However, when the development requires a high amount of programming cleverness, and when the overall value of the market is not quite high, then an open-source model will tend to be more successful.

[3]It should become apparent that the important point is that the firm cannot contract with all of the potential developers.

1.6 Modularity

We might wonder whether the decentralized system of software development is more successful when the scope for incremental improvement is large, or when developments are modular. For example, much of the functionality of LaTeX and Emacs comes from the vast variety of add-on packages available. On the other hand, none of the many add-on packages has any value without the core system. Some open-source proponents have argued that this modularity is a key component of the model.

We therefore seek to understand what a proper definition of modular might be. It is certainly true that developments that uniformly cost less for all individuals are, ceteris paribus, more likely to be implemented. This seems a shallow conclusion, especially when we accept that while many add-on developments in the open-source world are not just small in the sense of low cost, but also small in the sense of low value.

What would happen if we altered our model to allow many low-cost, low-value projects to be worked on simultaneously, instead of just one larger project? Suppose we scaled the joint value and cost distribution down by a factor of k, but also allowed up to k potential different developments to coexist. Suppose further that each individual receives different value and cost draws for each project.

Thus, we are considering moving from the situation in which there is a single project whose cost and value is given by a single realization from G, to a situation in which there are many smaller projects whose costs and values are drawn from a scaled-down G.

It is not hard to see that the expected productivity of the community is unchanged. Hence, a useful concept of modularity is not one of many small copies of an original project.

Instead, we now imagine the original project is the sum of many smaller components or "sub-enhancements" that can not be developed independently, and which are of no value if not bundled. We will call this the "non-modular environment".

Costs of communication between developers are likely to make certain developments non-modular. If each subroutine of a main program needs to be crafted so as to integrate well with the other subroutines, then it will be very difficult for many independent, remotely-located developers to make a useful end product.

From the standpoint of an individual user-developer i, the relevant distribution is therefore that of

$$z_i^k = \frac{\sum_{j=1}^k v_i^j}{\sum_{j=1}^k c_i^j}$$

The user will only develop the project if this quantity is high. When the number of different components k is large, this sum will tend to be near the ratio of the expectations of v and c. Heuristically, there will be very little variation and an individual will only choose to develop if he experiences a large number of extreme value and cost realizations, which will be unlikely.

On the other hand, if each individual sub-enhancement can be implemented separately, and if each one has value unto itself, then any individual component will be developed so long as at least one person of n has a draw from the upper tail of F. We will refer to this as the "modular environment".

Of course, a single success in the non-modular case is as good as k successes in the modular case. Nonetheless, when k is large relative to n, and when the size of the community exceeds a certain critical mass, the expected number of distinct developments (i.e. the number of sub-enhancements) in the modular case will exceed the number in the non-modular case. Conceptually, even in an environment with systematic free-riding, it is better to work with a large number of upper tails that correspond to smaller projects than to work with a small number of averages that correspond to larger projects.

We present the following formalization of this intuition. The theorem says that for any fixed population size, the relative performance of the two environments as the degree of modularity grows large is determined by whether or not the population exceeds a critical level.

Theorem 6 *Define N^* as follows:*

$$N^* = 1 + \frac{\log\left(\frac{Ev}{Ec}\right)}{\log\left(F\left(\frac{Ev}{Ec}\right)\right)}$$

For any fixed $n > N^*$ there exists some K such that for all $k > K$ the expected number of sub-enhancements in the modular case with n users and k components exceeds the number in the corresponding non-modular case. For any fixed $n < N^*$, there exists a K such that for $k > K$ the expected number of sub-enhancements in the modular case with n users and k components is less than the number in the corresponding non-modular case.

Proof: Let π^n_{mod} denote the probability that any of $n-1$ users develop any one given project in the modular case (this is independent of k), and let $\pi^{n,k}_{\text{nmod}}$ denote the probability that the development is made in the nonmodular environment.

To prove the theorem, we need only show that when $n > N^*$ there exists a K such that for $k > K$ we have $\pi^{n,k}_{\text{nmod}} < \pi^{n}_{\text{mod}}$, and that when $n < N^*$ there exists a K such that for $k > K$ we have $\pi^{n,k}_{\text{nmod}} > \pi^{n}_{\text{mod}}$. For then elementary facts about expectations of sums of random variables will imply the theorem.

We will show that, as k grows large, the law of z^k_i places arbitrarily large probability on a neighborhood around Ev/Ec. Having shown this, it will follow that for any n the equilibrium of the nonmodular environment converges to the same equilibrium as k grows large. This will in turn allow us to compare the two environments for a given population.

Observe that z^k_i can be expressed as

$$z^k_i = \frac{\frac{1}{k}\sum_{j=1}^{k} v^j_i}{\frac{1}{k}\sum_{j=1}^{k} c^j_i}$$

so that the law of large numbers implies that the corresponding law places arbitrarily high probability on any given neighborhood of Ev/Ec as k grows large.

Now we will show that the equilibrium of the nonmodular environment converges to one in which $\pi = 1 - Ec/Ev$. Since z^k_i is converging in probability to Ev/Ec for each user, any solution to the equation

$$\pi = 1 - F_k\left[\frac{1}{1-\pi}\right]^{n-1}$$

must be arbitrarily close to $1 - Ec/Ev$, since the distribution function $F_k(z^k_i)$ is converging to a function that places an atom at Ev/Ec. Therefore, for fixed n, $\pi^{n,k}_{\text{nmod}}$ converges to $1 - Ec/Ev$ as k grows large.

We now show that $\pi^{n}_{\text{mod}} > \pi^{n,k}_{\text{nmod}}$ if and only if $n > N^*$. The proof is simple. If

$$1 - \pi > F\left[\frac{1}{1-\pi}\right]^{n-1}$$

then it must be that $\pi^{n}_{\text{mod}} > \pi$, whereas the opposite conclusion holds otherwise. Letting $\pi = 1 - Ec/Ev$ we see by directly solving for n that

$$n = 1 + \frac{\log\left(\frac{Ev}{Ec}\right)}{\log\left(F\left(\frac{Ev}{Ec}\right)\right)} = N^*$$

will exactly satisfy the above inequality. This proves the theorem. ■

Emacs is a good example of a modular development. Another good example is the set of applications associated with the Linux operating system (for example the `/usr/bin` and `/usr/sbin` utility directories). On the other hand, a complicated individual program such as a graphical user interface (GUI) is nonmodular, because the inherent difficulty in coordinating the integration of the many elements of the code prohibits a useful division of labor in its production.

1.7 Conclusion

The open-source software movement is not new. However, only in the last few years, with the rising prominence of the Internet and the growing success of major open-source developments such as Linux, has the importance of this movement become apparent.

We present a positive analysis of open-source software development. We focus on issues that are given much credence in the actual developer community. These are the free-riding problem, the costs of redundant parallel effort, and the performance of the system when either the number of developers, the complexity or the modularity of the project changes.

We have found that although there will generally be redundant research efforts, these wasted efforts grow slowly as the upper bound of the distribution grows. Furthermore, only the most efficient programmers deign to develop open-source software. Therefore, the potential cost savings compared to traditional firm with a constrained labor force pool are sometimes quite large.

We have shown that research projects that cannot be broken down into smaller pieces, each of which has individual value regardless of the other pieces, then the open-source system will not perform very well when the population exceeds a critical level. When this level is exceeded, modular projects will be more likely to be implemented, and will be implemented more efficiently. However, when this critical mass is not met, a non-modular environment will tend to outperform a modular one.

Finally, we have also shown how correlations between value and cost can have a large impact on the success of the open-source community. This is true even when the marginal distributions are fixed. This fact allows us to resolve the puzzle of why an open-source community can design a complex operating system, yet consistently fail to develop simpler applications.

Chapter 2

Firm Entry and Growth with an Imperfect Labor Market

2.1 Introduction

Small businesses in the United States enjoy certain privileges that other enterprises do not. Some benefits are tax breaks. For example, the Small Business Administration (1998) estimates that the Taxpayer Relief Act of 1997 will save small businesses $40 billion over the next ten years, over and above those benefits that accrue to all taxpayers. Other benefits take the form of weakened restrictions concerning mandated family and medical leaves, pensions, workplace safety, the filing of affirmative action reports and civil rights (Brock and Evans 1986).

Small firms receive preferential treatment when dealing with the government. In fact, federal regulations demand that each acquisition of supplies or services whose anticipated dollar value lies in the range from $2,500 to $100,000 be set aside exclusively for small businesses, as must larger acquisitions conditional on a reasonable expectation that at least two small firms will bid competitively for the contract (Small Business Administration 1997).

One reason such provisions exist is that there is a widespread belief that small firms are vibrant sources of economic growth which account for a significant portion of job creation. In an empirical study, Birch (1987) purports to confirm this belief. Further evidence is offered by the Small Business Administration, which reports that between 1980 and 1986 firms with fewer than 100 employees accounted for 35 percent of employment, and 53 percent of job growth (Small Business Administration 1987).

The validity of the conclusions reached in these studies, and the sagacity of the underlying conventional wisdom and policy, recently has been challenged forcefully. Davis, Haltiwanger, and Schuh (1996) have attacked the "small business job-creation myth" along a number of empirical and theoretical fronts.

In this paper we respond to some of these criticisms. We provide a theoretical contribution to the study of small firm entry and growth, focusing particularly on industries in which skilled labor is likely to be a critical input.

We assume that young firms enter the industry because they hope, over time, to become more productive as they learn about themselves and the market for their products. Some firms will fail without ever generating significant profits, while others will experience growth in their productivity. A firm whose productivity rises will need to hire additional workers to fully take advantage of its profit potential.

Workers may be unwilling to accept employment with young firms, even when such acceptance would raise the total expected productive value of the firm above the combined outside options of the firm and worker. This phenomenon, which is not driven by credit constraints or sunk costs, leads to suboptimally low entry and inefficiently high unemployment. Standard equity contracts are unable to resolve this problem.

The root of the problem is the inability of firms *today* to contract with unknown prospective workers who might be hired in the *future*. This market failure is related to, but distinctly different from, one discussed by Acemoglu (1997) in the context of job training.

In other words, firms that are learning about their own capabilities don't know if they will need to hire workers in the future, and also may not know which types of skills they would need to search for conditional on needing to hire someone. This means that firms which might become successful in the future cannot negotiate today the share of productivity claimed by future workers.

This limits the willingness of young firms to compensate their current workers. In particular, the zero-profit condition for a young firm involves only a partial transfer of the total expected future productivity of the firm (since the remaining portion has in effect already been claimed by future workers). Ex-ante, however, this reduces the incentive for workers to join new firms, which means that fewer jobs will be created in the future by growing firms. Instead, workers have incentives to free ride on the risks taken by other workers (those who choose to help a young firm grow). This market distortion can not resolved by allowing young firms to write complete contracts with their initial employees, since the externality involves future workers.

We also show that rent-seeking behavior by workers can lead to highly-productive jobs being rejected in equilibrium. Workers instead choose to accept jobs with firms that have inferior future prospects.

Finally, we show that whether too little or too much entry occurs depends intimately, and in a novel manner, on the bargaining power of workers. When workers command a large share of rents, free-riding results as workers with young firms (who help expand the productivity and workforce of the firm) cannot be adequately rewarded for the jobs they provide to future workers. On the other hand, when workers have little bargaining power, excessive entry occurs because entrants see high potential future profits, and also because workers perceive a job with a young firm as a substitute for bargaining power. The reason is that a worker with a young firm can hold up the entire future value of the firm associated with growth and capital gains, whereas a worker with a mature firm can only bargain over productivity directly associated with his job.

This paper therefore constitutes a partial theoretical response to the challenge of Davis, Haltiwanger, and Schuh (1996), who not only have argued that the evidence is mixed regarding small firms' contribution to job creation,[1] but who also have questioned exactly what the market failure is that policies affording special status to small firms are meant to rectify.

Beyond casting doubt on the accuracy of earlier statistical work, Davis, Haltiwanger, and Schuh (1996) argue that even if small firms did create most new jobs in the economy, it would not follow directly that policy intervention is desirable. This criticism is especially poignant given the evidence that jobs created by small firms appear in many ways to be inferior to those created by larger firms.[2] Where is the market failure?

[1]Davis, Haltiwanger, and Schuh (1996) conclude from their study of U.S. manufacturing that while the gross rate of job creation exhibited by small firms exceeds that of large firms, the job-destruction rate of small firms also exceeds that of large firms, so that there is no strong or simple relationship between size and net job creation. However, in a recent study of small and large establishments in Canadian manufacturing, Baldwin and Picot (1995) find that net job creation of small establishments is disproportionately high even after correcting for regression-to-the-mean bias.

[2]There is evidence that small firms tend to offer lower wages and benefits, and less durability than jobs created by larger firms. Davis, Haltiwanger, and Schuh (1996) present evidence that the jobs created by young or small firms tend to survive with lower probability than other jobs. This is consistent with previous empirical work by Audretsch (1991) and Mahmood (1992), for example, which finds that firm propensity to exit is inversely related to firm age. Evidence that wages paid by larger firms tend to be higher, as does the value of benefits provided, is ample (Brown, Hamilton, and Medoff (1990); Acs (1996)).

We present a theory that explains why an intervening social planner might choose to increase the number of small firms in the industry *despite* the fact that they offer lower wages and less job security. While the assumptions underlying the results are surely not applicable to all industries, they are reasonable as approximations for some industries in which skilled labor is a vital input.

The macroeconomics literature has considered both unemployment and wage bargaining. These are likely to be relevant in explaining the role of imperfect labor input markets on firm entry and growth. We feel that this is a natural opportunity to synthesize some elements of macroeconomics and industrial organization, by employing modelling approaches from the first field to address a question from the second.

2.2 Model

Time is continuous, indexed by $t \in [0, \infty)$. There is a mass of workers, whose size is normalized to 1. The flow productivity of a worker employed by a particular firm equals a time-varying parameter idiosyncratic to that firm, $\theta \in \mathbb{R}$.

All firms that enter the market initially have exactly the same flow productivity, given by $\theta_L \geq 0$. Initially, a firm can employ at most one worker.[3] A firm's productivity, and the number of employees it can usefully employ, can only change if the firm is hit by a random shock. Any firm can be hit by at most one shock during its entire existence. A firm that has been hit by a shock experiences two effects. First, the firm immediately enjoys an increase in θ from θ_L to $\theta_H > \theta_L$. Second, with probability $\pi \in [0, 1]$ the firm gains the capacity to employ an additional worker. Whether or not the firm will be able to employ a total of two or merely one worker is therefore determined immediately upon being shocked.[4]

It is unknown when or if a particular firm will ever enjoy a shock. All that is known is that shocks arrive at a fixed, exogenous rate λ, so long as the firm has an employee. We interpret this to mean that new firms face some uncertainty regarding their long-run prospects. Firms could be uncertain of the demand for their products, the technical feasibility of a

[3] It is not necessary to imagine that there are no other inputs in the production process. There may well be capital and unskilled labor inputs. We focus our analysis on the market for skilled labor.

[4] An alternative interpretation that is valid for any $\pi > 0$ is that this parameter gives the average firm scale conditional on the underlying uncertainty being resolved.

research project or the management skills of the owner.[5] Since the arrival rate of shocks is fixed, all firms that have not matured have equal prospects for the future, regardless of how long they have been around. Firms that do not have workers are incapable of receiving a shock.

To summarize, all firms that enter are initially capable of employing at most one worker. That worker's productivity is $\theta_L \geq 0$. A firm that has an employee receives a shock at rate λ. All firms that are shocked immediately experience a boost in productivity to the level θ_H. Also, with probability π a shocked firm immediately becomes capable of opening a second productive post, the productivity of which is equal to θ_H. With probability $1-\pi$, the firm is destined to have at most one productive post for all time.

To ease explication, I will sometimes refer to firms that have not experienced a shock as "young" or "immature" firms, even though they could be quite old. I will also refer to firms that have experienced a shock as "mature" or "old".

All firms die at an exogenous rate δ. At each moment in time, there is an infinitely elastic supply of potential entrants. These firms are inactive until they become matched with a worker. How this matching occurs will be explained later.

There are no contracts of any kind. Instead, firms and workers, all of whom are risk-neutral and discount the future at rate r, divide the surplus of the match according to exogenous bargaining parameters. This bargaining is completely forward-looking, in that the entire present discounted value of the match is bargained over to determine wages today. Workers claim a share β of the surplus generated by a match. This process is described in considerable detail below.

I assume that workers who are separated from an employer for any reason bear a non-contractible cost ϕ. This might represent psychic or financial costs of a job search, training costs necessary to find a job, or moving costs.

2.2.1 Firm Entry and Worker Search

Workers who are not currently paired with a firm look for a new job. Workers have the ability to direct their search. In particular, they are allowed to either apply for a job with a mature firm that has an open post, or to apply for a job with a new firm. If there is an excess demand for jobs

[5]These might be more realistically modelled in a formal Bayesian learning environment, in the spirit of Jovanovich (1982). We do not do so because the present environment allows us to reach rich results in closed form.

of a particular type, however, then a queue develops for those jobs. This queueing process will be described in more detail below.

To simplify the analysis and to focus on the workers, I make the following assumptions about new firms. New firms need pay no sunk costs to enter the market.[6] There is an infinite supply of inactive firms at the fringe of the market at any point in time. They become active if a worker decides to match with them, and if they expect non-negative discounted profits. It turns out that these assumptions guarantee a worker who wishes to match with a new firm can always do so.

2.2.2 Wage Determination

I assume that a particular post held by a worker in a firm is completely specific to that worker. In other words, if a worker leaves a post, the firm can never again productively operate that post. I assume that a firm that has not yet enjoyed a shock disintegrates if its worker leaves. However, if a firm has matured and has two posts, and one of the workers abandons his post, the remaining post is still productive unless the corresponding worker leaves as well.

These assumptions might seem extreme. However, the important issue is that workers are able to threaten to disrupt the everyday flow of production at a firm by leaving, at least to some extent. The qualitative nature of all the results derived below will not be influenced by, for example, assuming instead that a worker who leaves causes a certain positive fraction of the present discounted value of the firm to be lost, or by assuming that it takes time to train a new employee to perform the duties of the previous worker. However, the resulting calculations would become much more cumbersome.

Wages are determined according to forward-looking Nash bargaining with outside options. Labor claims a share β of the surplus value of a match. The technical details of this process are described in the appendix. The results in terms of wages are given here, and described heuristically.

Once a firm has matured wages can be represented as a constant fraction of the worker's outside option and a share of flow profits

$$w^m = (1-\beta)rU + \beta\theta_H \tag{2.1}$$

where U is the (endogenously determined) value to a worker of becoming unemployed, which includes the separation cost ϕ.

[6]This is solely for simplicity. We seek to abstract from issues that have been well-covered in the previous industrial organization literature, such as the role of sunk costs, despite their importance.

The situation is different for workers with young firms. Workers recognize that the firms they are currently working for have the potential to hire more workers in the future. A worker today realizes that the firm will be able to extract surplus from the potential future worker, and he also knows that in the future he himself will not be able to lay claim to any of the surplus generated by the other worker's post (since at that time, the employer and the other worker will bargain separately over the profits generated by that post). However, today the worker can threaten to leave the firm, forever ending any hopes the firm has of operating in the future. Thus, today the worker with a young firm receives a portion of the surplus that is expected to be generated by a future employee of the firm.

To compute the precise wage path of young workers, we need to know how much a firm values receiving a shock that generates a new post. That is, we need to compute the expected capital gain of a firm that is attributable to opening a second post.

A mature firm that is able to open a post of productivity θ_H values the individual post at A, which is given by the expected discounted value of flow profits less wages

$$A = \left(\frac{\gamma(n)}{r+\gamma(n)}\right)\frac{\theta_H - w^m}{r+\delta} = \left(\frac{\gamma(n)}{r+\gamma(n)}\right)\frac{(1-\beta)(\theta_H - rU)}{r+\delta} \qquad (2.2)$$

where $\gamma(n)$ gives the rate at which a mature post is filled. While endogenous to the economy, $\gamma(n)$ is taken as exogenous by individual workers and firms. It will be discussed in more detail later.

Since it is not known when or if a young firm will actually be able to open an extra post, it is not known when or if a firm will be able to accrue the value A. Hence, the bargaining process today only rewards the worker with the young firm a weighted value of A, where the weight depends upon both the rate λ at which shocks arrive and the probability π that a shock leads to an extra post. Results in the appendix show that the wage paid by a young firm is independent of time (in steady state) and given by

$$w^y = (1-\beta)rU + \beta\theta_L + \beta\pi\lambda A \qquad (2.3)$$

The intuition is as follows. In addition to their outside option, workers today claim a share of today's flow profits and also a fraction of the expected capital gain of a firm.

2.2.3 Analysis of the Steady State

In this section we derive equilibrium conditions for a steady state. There are two sets of conditions. The first set relates to the levels of young firms and mature firms, the distribution of workers across firms and unemployment, and also the flow rates that connect these levels. The second set relates to worker and firm optimization.

Stocks and Flows

Denote the number of young (and active) firms at time t by y_t. Keeping in mind that firms only enter when a worker wishes to match, and that a young firm that is abandoned by its worker disintegrates, it follows that there are no young firms without workers.

The evolution of y_t is determined by the number of entrants n_t at time t, and also by the death and growth of firms according to the exogenous parameters δ and λ. Precisely,

$$\dot{y}_t = n_t - (\delta + \lambda)y_t$$

If n_t were constant over time at value n, then y_t would converge to a steady state value

$$y = \frac{n}{\delta + \lambda} \tag{2.4}$$

Sometimes mature firms will experience delay filling extra posts that they develop. To accommodate this, it is necessary to define m_t to be the total stock of mature posts, either filled or vacant. This evolves according to

$$\dot{m}_t = -\delta m_t + \lambda(1 + \pi)y_t$$

where the first term on the right-hand side represents exogenous attrition from the pool of mature posts, and the second term gives the flow rate of new mature posts. The steady-state value is given by

$$m = \frac{\lambda(1 + \pi)y}{\delta} = \frac{\lambda(1 + \pi)n}{\delta(\delta + \lambda)} \tag{2.5}$$

where we have made use of (2.4) above.

Define u_t to be stock of unemployment at time t. In equilibrium, any worker who is unemployed must be queueing for a job with a mature firm, since we have assumed that jobs with young firms can be obtained immediately if desired. Denote by v_t the stock of vacancies associated with mature fi·rms.

It is assumed that whenever there is a positive stock of both vacancies and unemployment, the number of new matches formed is $\min[u_t, v_t]$. This suggests that the matching process must be specified in terms of flows as well as stocks. In a slight abuse of notation, we use the variables u and v to denote a stock unless the stock is zero, in which case they denote a flow. If there is a zero stock of both unemployment and vacancies, then we assume that the number of matches is given by the minimum of the two flows. If, however, there is a positive stock of vacancies, but no stock of unemployment, then all flow unemployment is instantly assigned to jobs with mature firms, save for those n_t workers who are taking jobs with young firms.

Thus, the market for labor is quasi-competitive: stocks of both open jobs and unemployed workers do not coexist. Instead, enough matches are instantaneously formed to reduce at least one of the stocks to zero. However, wages are still subject to ex-post renegotiation. Our model thus differs markedly from more traditional analyses of labor market imperfections.

Hence, when $v_t > 0$ and $u_t = 0$, exactly δ workers become unemployed at t. Of those, n_t form matches with new firms, and the remaining $\delta - n_t$ take posts with mature firms, so that the entire flow into unemployment becomes instantaneously re-employed. Therefore, the evolution of v_t is given by

$$\dot{v}_t = -\delta v_t - (\delta - n_t) + y_t \lambda \pi \tag{2.6}$$

where the first term represents exogenous attrition from the pool of vacancies, the second represents endogenous outflow from the pool as mature firms fill posts, and the third gives the inflow into vacancies as some young firms mature and develop extra posts. The steady-state value is constrained to be non-negative, and so is given by

$$\begin{aligned} v &= \max\left[0, \frac{n - \delta + y\lambda\pi}{\delta}\right] = \max\left[0, \frac{n - \delta + \frac{n}{\delta+\lambda}\lambda\pi}{\delta}\right] \\ &= \max\left[0, \frac{(\delta+\lambda)(n-\delta) + n\lambda\pi}{\delta(\delta+\lambda)}\right] \end{aligned}$$

The unemployment stock u_t is at each time equal to the total mass of workers less the number of those employed by young firms, less those employed by mature firms, subject to the constraint that unemployment cannot be negative. Thus,

$$u_t = \max[0, 1 - y_t - m_t]$$

Whenever $u_t > 0$ and $v_t = 0$ the evolution of unemployment is given by

$$\dot{u}_t = \delta(1 - u_t) - n_t - y_t\lambda\pi \tag{2.7}$$

the first term of which gives flow into unemployment resulting from the exogenous breakup of matches, the second of which gives outflow from unemployment as workers take jobs with young firms, and the third gives outflow resulting from workers matching with mature posts. In steady state the stock of unemployment is given by

$$u = \frac{1}{\delta}\left[\delta - n - \frac{n\lambda\pi}{\delta + \lambda}\right]$$

subject to being non-negative.

Since both y and m are determined completely by n in steady state, and since they are both increasing in n, there is a unique value of n such that $1 - y - m = 0$. Call this critical value n^c. It is given by

$$n^c = \frac{\delta(\delta + \lambda)}{(\delta + \lambda + \lambda\pi)}$$

Because the Leontief matching function does not permit the simultaneous existence of both unemployment and vacancy stocks, n^c is also the smallest value of n such that the stock of vacancies is zero. This observation allows us to recast the derivations given above in a simpler form, at least in steady state. In steady state,

$$u = \begin{cases} \frac{1}{\delta}\left[\delta - n - \frac{n\lambda\pi}{(\delta+\lambda)}\right] & \text{if } n < n^c \\ 0 & \text{if } n \geq n^c \end{cases} \tag{2.8}$$

$$v = \begin{cases} 0 & \text{if } n \leq n^c \\ \frac{(\delta+\lambda)(n-\delta)+n\lambda\pi}{\delta(\delta+\lambda)} & \text{if } n > n^c \end{cases} \tag{2.9}$$

Thus, in steady state it is possible to either have a positive stock of unemployed workers (if $n < n^c$), a stock of vacancies at mature firms (if $n > n^c$), or neither (if $n = n^c$). However, these stocks cannot simultaneously be positive.

We have described the process by which workers are matched to vacancies posted by mature firms. We have also determined the steady-state stocks of unemployment and vacancies in terms of a single endogenous parameter n, and shown that there exists a critical value that determines whether there will be positive stocks of either unemployment or vacancies, or neither. To close the model, we now turn to the decision problem faced by the firm and worker, and identify the appropriate equilibrium conditions.

Firm and Worker Optimization

Firms bargain over the surplus generated by a match, and will not remain in the market if they foresee negative expected profits. However, given the assumption of no sunk costs of entry, and that firms only enter if a worker wishes to match with them, the remainder of their optimal strategy is immediate. Firms that have developed an extra post search for a worker until the post fills. The surplus of any post is split between the firm and the worker occupying that post. The precise manner in which the surplus is split is determined by forward-looking Nash bargaining, as described in the appendix. Those results also imply that, in fact, any firms that choose to enter do expect to make non-negative profits (but recall that they will not choose to enter unless a worker desires to match with that firm).

We now turn to the decision problem of the workers. If $\pi = 0$ then no worker can ever hope to get a job with a mature firm unless that worker was with the firm when young. Hence, the optimal policy for any worker who becomes unemployed is to immediately take a job with a young firm, and to stay with that firm until the match is exogenously separated.

Let $\pi > 0$. Denote by J^m the value expected by an unemployed worker who chooses to direct his search towards a mature post, and denote by J^y the value expected by a worker who chooses to match with a young firm. It is easy to show that, when $\pi > 0$, some workers must choose to take jobs with mature firms.[7] But, likewise, there can not be an equilibrium in which no workers choose to join young firms, for then the steady state level of firms would be zero. Hence, a recently unemployed worker must be indifferent between searching for a job with a mature firm and taking one with a young firm. Thus, the following definition of equilibrium is appropriate.

Definition 1 *A steady-state equilibrium is a quintuple* (n, y, m, u, v) *such that equations* (2.4), (2.5), (2.8) *and* (2.9) *each hold, and such that* $J^y = J^m = U + \phi$.

[7]If workers spurned all jobs offered by mature firms, then we would have $A = 0$, since young firms do not expect to be able to fill exta posts. Hence $w^y < w^m$ since wages would reflect only spontaneous productivity. Clearly, in this case optimizing workers would prefer the mature posts.

We will examine the values J^m and J^y as functions of n, the number of firms that enter the economy each instant. This is a natural approach, since n implicitly determines all of the relevant levels and flows, as demonstrated in the previous section. Our approach to finding an equilibrium is to first exhibit the functional relationships that the value equations must satisfy. Then, we impose the necessary equilibrium conditions and look for a solution.

Assume for the moment that $n < n^c$, so that $u > 0$, $v = 0$ and workers must actually wait to find jobs with mature firms. The total number of open posts that are generated each instant is

$$y\lambda\pi = \frac{n\lambda\pi}{\delta + \lambda} \tag{2.10}$$

There are

$$u = \frac{1}{\delta}\left[\delta - n - \frac{n\lambda\pi}{\delta + \lambda}\right]$$

workers looking for jobs with these posts. We have already assumed that in this case the number of matches is given by (2.10), so it follows that the rate at which an unemployed worker looking for a mature post finds one is given by

$$\rho(n) = \frac{\frac{n\lambda\pi}{\delta+\lambda}}{\frac{1}{\delta}\left[\delta - n - \frac{n\lambda\pi}{\delta+\lambda}\right]} = \frac{n\delta\lambda\pi}{(\delta + \lambda)(\delta - n) - n\lambda\pi}$$

Hence the value J_n^m is

$$J_n^m = \frac{\rho(n)\tilde{J}_n^m}{r + \rho(n)}$$

where $\tilde{J}_n^m$ is the value of actually landing a job. This is given by

$$\tilde{J}_n^m = \frac{w^m + \delta U}{r + \delta}$$

but if the economy is in steady state, then a worker who becomes unemployed must be willing to search for a job with a mature firm, so that $U = -\phi + J_n^m$. Using this, and inserting the definition of the wage from (2.1) we get

$$J_n^m = \frac{\rho(n)\tilde{J}_n^m}{r + \rho(n)} = \left(\frac{\rho(n)}{r + \rho(n)}\right)\frac{(1-\beta)r(J_n^m - \phi) + \beta\theta_H + \delta(J_n^m - \phi)}{r + \delta}$$

which must be solved for J_n^m. Doing so yields

$$J_n^m = \frac{\rho(n)\left[\beta\theta_H - \phi((1-\beta)r+\delta)\right]}{r\left[r+\delta+\beta\rho(n)\right]} \tag{2.11}$$

which in monotonically increasing in n over the interval $[0, n^c)$, with $J_0^m = 0$ and

$$\lim_{n\to n^c} J_n^m = \frac{\beta\theta_H - \phi\left[(1-\beta)r+\delta\right]}{\beta r}$$

since as n approaches n^c, $\rho(n)$ approaches positive infinity and because the payoff to any optimizing worker in a steady-state equilibrium is equal to the payoff of a worker who expects to spend all of his time employed with a mature firm. Because of the separation cost ϕ, the worker is unable to extract the entire flow profit from the match (hence the second term in the numerator).

When $n \geq n^c$, the worker is able to find mature posts at will. Hence, J_n^m is flat beyond n^c.

We now turn to the value equation of a worker with a young firm. Denote this by J_n^y. The flow value is given by the flow wage plus the possible changes in the value of the match, weighted by the appropriate probabilities.

$$\begin{aligned} rJ_n^y &= w^y + \lambda\left(\tilde{J}_n^m - J_n^y\right) + \delta\left(U - J_n^y\right) \qquad (2.12)\\ &= (1-\beta)rU + \beta\left(\theta_L + \pi\lambda A\right) + \lambda\left(\tilde{J}_n^m - J_n^y\right) + \delta\left(U - J_n^y\right) \end{aligned}$$

Recall that the term A is the value to a firm of a filled post of productivity θ_H, weighted by a factor to account for possible delays in filling an extra post that has just become viable. How long a firm expects to wait to fill a mature post depends upon the how many workers are looking for mature posts, and how many other mature posts are available. Both of those quantities, in equilibrium, depend only on n. If $n > n^c$ then there is an oversupply of vacancies. In this case there is no stock unemployment, so the rate at which firms expect to meet workers is given by the ratio of flow unemployment, less those who choose to join young firms, to the stock of vacancies v. Denoting this rate by $\gamma(n)$ gives

$$\gamma(n) = \frac{\delta - n}{v} = \frac{\delta(\delta-n)(\delta+\lambda)}{(\delta+\lambda)(n-\delta)+n\lambda\pi} \tag{2.13}$$

Maintaining the supposition that $n \geq n^c$, the ability of a worker to immediately find employment with either a young firm or a mature firm implies that $\tilde{J}_n^m = J_n^m = J_n^y$. We also know, however, that if the economy were in steady state, then the worker would be willing to work for another young firm if the current match were separated for any reason. Hence, $U = J_n^y - \phi$. Finally, the bargaining equations in the appendix dictate that in this situation the value of a filled mature post to a firm is precisely $(1-\beta)\phi/\beta$. These facts allow us to rewrite (2.12) as

$$rJ_n^y = (1-\beta)r\left(-\phi + J_n^y\right)U + \beta\theta_L + \beta\pi\lambda\left(\frac{\gamma(n)}{r+\gamma(n)}\right)\frac{(1-\beta)}{\beta}\phi - \delta\phi$$

which yields

$$J_n^y = \frac{\beta\theta_L + -\phi(1-\beta)r - \delta\phi + \frac{\phi\pi\lambda(1-\beta)\gamma(n)}{r+\gamma(n)}}{\beta r} \tag{2.14}$$

When $n < n^c$, a slightly different approach must be used to compute J_n^y. The reason is that it is no longer true that $J^y = \tilde{J}^m$. The bargaining equations in the appendix imply that

$$A = \frac{(1-\beta)}{\beta}\left(\tilde{J}_n^m - J_n^y + \phi\right) = \frac{(1-\beta)}{\beta}\left(\phi + \frac{rJ_n^y}{\rho(n)}\right) \tag{2.15}$$

and hence the asset-value equation of a young worker is given by

$$rJ_n^y = (1-\beta)r(J_n^y - \phi) + \beta\theta_L + \pi\lambda(1-\beta)\left[\phi + \frac{rJ_n^y}{\rho(n)}\right] + \frac{\lambda rJ_n^y}{\rho(n)} - \delta\phi$$

the solution to which is

$$J_n^y = \frac{\rho(n)\left(\phi\left[(1-\beta)(\pi\lambda - r) - \delta\right] + \beta\theta_L\right)}{\beta r\rho(n) - \lambda(1+\pi(1-\beta))} \tag{2.16}$$

The properties of this function might seem surprising. For example, it is initially negative, and approaches negative infinity at

$$n^* = \frac{(\delta+\lambda)(1+\pi(1-\beta))}{\delta\pi\beta r + (\delta+\lambda+\lambda\pi)(1+\pi(1-\beta))} < n^c \tag{2.17}$$

at which point it jumps to positive infinity, and falls toward the value $J_{n^c}^y$ derived in (2.14) in the interval (n^*, n^c). It might also seem surprising that neither expression for J^y involves θ_H, despite the fact that in equilibrium workers who join firms have wages that depend on θ_H. The resolution of these oddities lies in the fact that we are imposing equilibrium conditions to facilitate the derivation of J^y, having first written the proper functional forms. Hence, we are actually deducing necessary conditions that the value equations must satisfy. Nonetheless, an intersection of the equations J_n^y and J_n^m derived above does correspond to an equilibrium. It is only the intersection point that has meaning, and the point of intersection is determined in part by θ_H. Hence, the equilibrium value of J^y is determined by θ_H.

This approach may seem confusing, but actually results in the explicit solutions that are simple relative to other feasible approaches. For example, this approach immediately shows that there is a unique steady-state equilibrium.

Lemma 1 *There exists a steady-state equilibrium, and it is unique.*

Proof: By (2.11) $J_0^m = 0$, and the function is finite-valued and strictly increasing up to n^c, at which point it is flat. By (2.16) and (2.14) the function J_n^y is negative beneath $n^* < n^c$, and then falls strictly from positive infinity for $n > n^*$. Thus, there can be no intersection for $n \leq n^*$, and since a strictly decreasing function crosses a weakly increasing function at most once, any equilibrium that exists must be unique.

We know that $J_n^y > J_n^m$ at n^*, and that J^y is continuous above n^*. We can also show that for $n \geq \hat{n} \max[n^c, \delta]$, it is the case that $J_n^y < J_n^m$. We have

$$\begin{aligned}\frac{J_{\hat{n}}^y}{J_{\hat{n}}^m} =& \frac{\beta\theta_L + -\phi(1-\beta)r - \delta\phi + \frac{\phi\pi\lambda(1-\beta)\gamma(\hat{n})}{r+\gamma(\hat{n})}}{\beta\theta_H - \phi\left[(1-\beta)r + \delta\right]} \\ =& \frac{\beta\theta_L - \phi(1-\beta)r - \delta\phi}{\beta\theta_H - \phi(1-\beta)r - \delta\phi} < 1\end{aligned}$$

since $\gamma(\hat{n}) = 0$ and $\theta_L < \theta_H$. Hence, the intermediate value theorem implies the lemma. ∎

Before discussing the qualitative features of different types of equilibria, we consider how a social planner would choose to run the economy. We will then be able to compare the centralized solution to the decentralized equilibrium.

2.3 Welfare Analysis

2.3.1 The Social Optimum

We assume that a social planner is interested in maximizing discounted aggregate productivity in the economy, less separation costs. We prove a result that implies that the social planner of an economy will never allow a stock of vacancies, and that he will never allow stock unemployment. We make the very mild assumption that

$$\theta_L + \lambda \left(\frac{\theta_H - \delta\phi}{r + \delta} \right) - \delta\phi > 0 \tag{2.18}$$

which simply asserts that a worker is more productive always taking a job with a young firm when unemployed than remaining unemployed forever.

Proposition 1 *The constrained first-best of the economy exhibits $(n, u, v) = (n^c, 0, 0)$ asymptotically, for any initial conditions. That is, there is zero stock of unemployed workers and a zero stock of mature vacancies asymptotically.*

Proof: We will prove this by exhibiting a welfare-improving policy from any steady state that involves either u or v greater than zero. This implies the above proposition.

Suppose that $v > 0$. Then $n > n^c$. Consider the policy that sets n_t to n^c for all future times. Since $n > n^c$, the original unemployment level was 0. As a result of this policy, unemployment remains at 0 for all time. This follows because lowering n to n^c causes vacancies to converge to zero from above, by (2.6). But since unemployment is zero whenever $v > 0$, and since unemployment is zero at the steady state induced by n^c, we conclude $u_t = 0$ along the adjustment path, and at the new steady state.

Therefore, this policy has not changed the flow separation costs that are borne. If we can prove that gross productivity is higher along the adjustment path and at the new steady state, then it will follow that welfare has improved. The number of high-productivity posts filled in the original steady state was $1 - y = 1 - n/(\delta + \lambda)$. The flow entry into the pool was $(\delta - n)$. Thus, lowering n raises the flow entry of workers into high-productivity posts. Along the entire path of the economy following the policy change, the number of high-productivity posts is higher than it would have been, and since total employment is unchanged, the economy is strictly better off.

Now suppose that $n < n^c$. Then $u > 0$ and $v = 0$. Consider the policy that raises n_t to n^c for all future time. Then unemployment converges to zero from above, by (2.7). At any point in time, the number of workers with young firms is higher than before, and also the number of workers with mature firms is higher than before. However, by (2.18) a worker employed by a young firm is more productive than an unemployed worker. Hence, a policy that shifts workers from stock unemployment to young firms without lowering the number of high-productivity matches must raise welfare. Since the proposed policy both shifts workers from unemployment to young firms and raises the number of high-productivity posts, it strictly raises the welfare of the economy. ■

This proposition says that the social planner will choose a policy that brings the economy into the steady state in which $n = n^c$, $y = n^c / (\delta + \lambda)$, and with all other workers employed at high-productivity posts. Heuristically, this is true because no production need be lost implementing the new policy.

Given the crispness of this result, it will be easy to describe the properties of decentralized equilibria. We now turn to that issue.

2.3.2 Analysis of Decentralized Equilibrium

The properties of the decentralized equilibrium are determined completely by the steady-state level n of new firms. Thus, the key is to understand whether the decentralized entry level is greater or less than the socially optimal level.

We present the following characterization.

Proposition 2 *Define* $\beta^* = \frac{\phi\pi\lambda}{\theta_H - \theta_L + \phi\pi\lambda}$. *Suppose that* $\beta < \beta^*$. *Then the equilibrium entry level* n *is such that* $n > n^c$. *If* $\beta > \beta^*$, *then the equilibrium entry level is given by* $n < n^c$. *If* $\beta = \beta^*$, *then* $n = n^c$.

Proof: For $n \geq n^c$ consider

$$J_n^m - J_n^y = \frac{\beta\theta_H - \phi\left[(1-\beta)r + \delta\right] - \beta\theta_L + -\phi(1-\beta)r - \delta\phi + \frac{\phi\pi\lambda(1-\beta)\gamma(n)}{r+\gamma(n)}}{\beta r}$$

$$= \frac{\beta(\theta_H - \theta_L) - \frac{\phi\pi\lambda(1-\beta)\gamma(n)}{r+\gamma(n)}}{\beta r}$$

which is increasing in n and increasing in β. Evaluating the expression at n^c gives

$$J_{n^c}^m - J_{n^c}^y = \frac{\beta(\theta_H - \theta_L) - \phi\pi\lambda(1-\beta)}{\beta r}$$

which is equal to zero if and only if

$$\beta = \frac{\phi\pi\lambda}{\theta_H - \theta_L + \phi\pi\lambda} = \beta^*$$

Clearly, if $\beta = \beta^*$ then $n = n^c$ is an equilibrium. We can also deduce that if $\beta > \beta^*$ then no worker would ever choose to join a young firm if mature posts were in excess supply, so that we conclude $n < n^c$. Similar reasoning shows that if $\beta < \beta^*$ workers would always take a job with a young firm if it were available, and if they thought the firm would immediately be able to find a worker upon developing an extra post. Thus, $n > n^c$ so that firms that develop new posts take time to fill them. ■

Hence whether there is over or under entry is determined completely by the magnitude of the bargaining power, given the other parameters, in accordance with a simple threshold rule. In particular, when the bargaining power of workers is relatively low (relative to β^*), there is under entry, while when β is relatively high, there is over entry.

To understand this result, consider again

$$J_n^m = \frac{\rho(n)\left[\beta\theta_H - \phi((1-\beta)r + \delta)\right]}{r\left[r + \delta + \beta\rho(n)\right]}$$

which holds for $n \leq n^c$. The derivative of this with respect to β is

$$\begin{aligned} \frac{\partial J_n^m}{\partial \beta} &= \left(\frac{\rho(n)}{r}\right) \frac{(\theta_H + r\phi)\left[r + \delta + \beta\rho(n)\right] - \rho(n)\left[\beta\theta_H - \phi((1-\beta)r + \delta)\right]}{\left[r + \delta + \beta\rho(n)\right]^2} = \\ &= \left(\frac{\rho(n)}{r}\right)\left[\frac{(\theta_H + r\phi)(r + \delta) + \rho(n)\,\phi(r + \delta)}{\left[r + \delta + \beta\rho(n)\right]^2}\right] > 0 \end{aligned}$$

so that enhanced bargaining position is good for workers who intend to join mature firms. This might seem obvious. In contrast, however, J_n^y can actually be decreasing in β. Since the denominator of (2.16) is increasing, a sufficient condition for J_n^y to be decreasing in β in this region is that the numerator is decreasing. This will occur if

$$\theta_L - \phi\pi\lambda + r\phi < 0$$

which will be the case when, for example, λ is very high. The term $-\phi\pi\lambda$ accounts for a portion of the loss in value of a young post that results from a decreases in A, as given in (2.15). Recall that A is the extra surplus a young firm expects to extract from a future worker, conditional on opening an extra post. This quantity is decreasing in β, as J_n^m is increasing in β.

Hence, a worker with a young firm dislikes more bargaining power insofar as it is also wielded by future employees of the same firm, reducing the profits (associated with the potential expansion of the workforce) that the firm will transfer to the worker today.

The same phenomenon manifests in the fact that J_n^y is decreasing in n in the region of interest $(n > n^*)$. In equilibrium, higher n are associated with enhanced outside options for workers, since mature jobs are relatively plentiful. This means that young firms will be unwilling to pay as much to workers today, since they realize that future workers will exact steep wages. Queueing for jobs with mature firms becomes relatively attractive as a result. If this effect dominates, then there will be under-entry and too much unemployment.

It is important to keep in mind that it is not solely the absence of complete contracts with the current worker that drives under-entry (which is the equilibrium result when $n < n^c$). Rather it is the fact that the gross productive value of an extra post can not be contracted over today if the participants of the contracting relation are unknown today. This means that upon finding a worker to fill an extra post, that worker will demand a share β of the surplus of the match.

If young firms could sign contingent contracts with potential future employees, this externality would be partially mitigated.[8] Future potential employees would agree to a contract today that stipulated lower wages in the future than what they would extract ex-post in the absence of a contract. They would agree to these lower wages because it would decrease the expected time spent in the unemployment pool, in which no wages are earned.

The proposition thus expresses the fact that as β grows, the total compensation that can be granted to workers with young firms falls relative to what they can expect to earn by queuing for a job with a mature firm. The result is under entry, as workers would refuse to work with additional young firms that might enter.

The situation when $\beta < \beta^*$ (and hence $n > n^c$) is markedly different. Since β is relatively low, firms anticipate large profits from each filled high-productivity post.

When firms expect newly-developed posts to be filled quickly, they are

[8] It will generally be difficult to write contracts with future employees. The primary difficulty is that it is not known who they are. If we accept that skilled labor, even within an industry, is heterogeneous, then if firms are unsure of exactly which skills they will need in the future, as is likely if they are learning about their market niche or conducting R&D, for example, the contracting difficulties will become even more severe.

willing to pay high wages to their workers while young (since the workers are the vital input for growth). In fact, as demonstrated in the theorem above, if n were equal to n^c young firms would actually pay wages higher than what mature firms would pay. Equilibration occurs as excessive entry by firms leads to a scarcity of labor for high-productivity posts, and consequent delay in filling mature posts that lowers the expected profit of young firms, and hence leads to falling wages. In equilibrium, of course, workers are indifferent between the two types of jobs.

2.4 Existing Literature

Our work is most closely related to that of Acemoglu (1997), the literature on search and matching, and the macroeconomic literature on quasi-competitive markets. Here we briefly relate these works to ours.

Acemoglu (1997) analyzes training and innovation in an imperfect labor market. The training decisions of workers are more closely related to our work. Young workers have an opportunity to invest in skills that raise their productivity when matched with an employer. Young workers and their employers are able to write complete contracts between themselves, and hence the worker invests in the level of training that maximizes the joint value of the partnership. An externality arises, however, because some workers become separated from their initial employer for exogenous reasons. This externality does not involve the initial employer; the original contract may cover the contingency of separation. When these workers find new employment, the nature of the labor market allows the new employer to extract a positive share of the surplus of the match. Training costs for the worker are at this point sunk (and thus non-contractible), and hence workers know ex-ante that they will not receive the full productive value of their investment in training from later employers. Therefore, young workers agree on a level of training with their initial employer that is too low from a social perspective (although it is jointly privately optimal).

The model of Acemoglu is the first to identify the source of the externalities in our model, which is the inability of parties to contract today with unknown parties of the future. His model always exhibits under-investment in training, however. In contrast, the net effect of the temporal externalities of our model can go in either direction.

Our work heavily draws upon the search framework as exlemplified by Diamond (1982). More specifically, we can compare our Proposition 2 with a result from Hosios (1990). He finds, in a model of firm entry (but not growth) with an imperfect labor market, that there is a threshold level of the Nash bargaining parameter such that there is excessive firm entry above that level, and too little entry beneath that level.

Despite the seeming similarity between the result of Hosios and Proposition 2, there are several important points of contrast. The externalities of Hosios' are driven critically by the matching function, since at any level of vacancies and unemployment, movements in either strictly change the number of matches formed. Our result, on the other hand, is driven by the intertemporal externalities that arise because of the potential of firms to grow in both size and productivity. The matching technology we use ensures that positive stocks of vacancies and unemployment cannot co-exist, and hence that changes in unemployment or vacancies will not affect the matching process when either the stock of unemployment or vacancies, respectively, are positive. Moreover, note that if either growth aspect (size or productivity) were absent from our model, the resulting (decentralized) equilibrium would coincide with the first-best for any value of β, whereas in Hosios' model the first-best is only attained for a single value. As a final point of contrast, in our model it is almost exclusively the decisions made by workers that matter, since young firms face no entry costs, and enter only at the bidding of workers.

The efficacy of the matching technology and the importance of ex-post opportunism in our model place it somewhat close to that of Caballero and Hammour (1998). They investigate a quasi-competitive economy. Certain sectors are unable to write complete contracts when combining their factors of production with other sectors. However, within each sector the market is competitive. They demonstrate a number of results, for example that there will typically be inefficient investment and factor market segmentation.

2.5 Conclusion

Skilled labor is an important input in many industries. In many of these same industries, it is reasonable to suppose that new firms initially face uncertainty as to how successful they will be in the future. This uncertainty could shroud either what the market demand for the firm's product will be, the technical feasibility of actually developing a new product, or the ability of the owner or manager of the firm to make sound strategic decisions. Finally, a firm that has resolved this uncertainty often must hire additional skilled workers to take full advantage of its profit potential.

We identify two distinct phenomena that can develop in such industries.

The first results from an externality between young firms and the workers they will hire in the future (if they prove to be successful). Since future employees will be able to hold up the firm once hired, those workers will extract a portion of the total productivity of the firm. The size of the portion that must be given to future workers can not be changed by writing contracts with today's employee. Since the identity of the worker that might be hired in the future is unknown, the firm cannot write contracts with him today. Hence, a firm that pays today's worker wages that transfer the full expected discounted productivity of the firm will expect to make negative net profits (since the firm is not credit-constrained, it is capable of paying such wages if it so chooses).

The zero-profit condition of young firms thus involves transferring less than the full social value of the firm to its employee (the remainder being claimed by future workers). This can cause a job with a young firm to appear relatively unattractive compared to a job with an established firm. The result of this is an incentive to free ride on the jobs that are being created by those who choose to join young firms. This incentive can only be tempered by a stock of unemployment, which serves to raise the waiting time associated with finding a job with a mature firm. This distortion is therefore associated with low entry, high average productivity among the employed, and a stock of voluntarily unemployed workers.

The second phenomenon that can manifest itself results in over entry. If firms expect to fill vacancies very quickly, and if a firm's shadow valuation of a filled high-productivity post is large, then firms have strong incentives to enter the industry. When young, firms are willing to pay workers high wages (since they anticipate filling any extra positions they develop quickly). An externality arises because firms that open new vacancies lower the rate at which pre-existing vacancies are filled. In fact, the social value of an additional unfilled high-productivity post is zero in this situation.

The distortions of the entry margin described above occur because of the market's failure to correctly price new firms relative to mature jobs. We show that whether too little or too much entry occurs hinges critically on the bargaining power of workers.

When workers are able to demand a large share of the surplus, young firms realize that even if they successfully grow, much of the productive value will be taken by labor. Hence, even if they expect to fill newly-developed posts at once, the value of those posts to the firm is low. As such, the component of wages reflecting the prospect of growth that young firms are willing to pay workers is small. Since the remaining difference in wages between the two types of firms is determined by current productivity, workers always prefer a job with a high-productivity firm over one with a low-productivity firm. As a result, unemployment develops as workers queue for mature jobs, and workers with mature firms earn higher wages.

On the other hand, if firms have a strong bargaining position, then they greatly value filled posts. Since they anticipate high future surplus, they are able to pay high wages when young. In turn, workers realize that a substitute for bargaining power is employment with a young firm. The reason is that workers with young firms are able to threaten to destroy that firm's prospects for growth and capital gains by abandoning the firm, which raises the size of the pie over which workers can bargain. In equilibrium, a stock of vacancies exists, and wages are equalized across the two types of jobs.

Our analysis describes a potential failure in the skilled labor market. We thus provide a theoretical response to the criticism of Davis, Haltiwanger, and Schuh (1996). We show how an imperfect labor market can lead to distortion away from the efficient level of entry. Furthermore, we demonstrate that even if younger and smaller firms pay lower wages than larger and older firms, it can still be the case that social welfare would rise if more such low-wage jobs were accepted by workers. In other words, low observed wages could be symptoms of a market failure, not evidence of the social undesirability of the firms offering such wages. Of course, we do not mean to suggest that all jobs in the real world that offer low wages should be subsidized. Our analysis is predicated on small firms having significant prospects for growth of both productivity and employment.

APPENDIX: NON-STATIONARY BARGAINING

Here we provide a representation for wages paid by a firm to its worker. The representation is valid in a more general context than is required in the paper. Hence we prove the general result, which is of independent interest. Consider a worker and a firm who have chosen to match. The worker at any point in time can choose to dissolve the relationship, earning a reward U. The firm can also choose to dissolve the relationship, earning an outside option of 0. In fact, assume that if either party terminates the relationship, these outside options are realized.

Let $\Omega = \{1, 2, ..., N\}$ index the possible states of the relationship, and denote the time-dependent transition probabilities by $\left\{\lambda_{ij}^t\right\}$ where λ_{ij}^t gives the rate at which the match moves from state i to state j at time t. Denote by π_i^t the flow production of the match in state i at time t.

In addition to the flow production technology, there are lumpy rewards that are accrued by the firm immediately upon transiting from state i to j. Denote these rewards by R^t_{ij}.

We seek a pair of value equations $(V, J) : \Omega \times [0, \infty) \to \mathbb{R} \times \mathbb{R}$, that is, a pair of value equations taking the state of the world and time into the real line. The function V gives firm value, while J gives worker value. We also seek a wage function w^t_i paid by the firm to the worker. We require that the value to the firm at any point equal the expected integrated profits less wages paid, while the worker's value be the expected integrated wages. Agree to define the value to a firm immediately following a transition as net of the reward R^t_{ij}.

We also require that value be split according to the following rule, which must hold for all time and all states where it is optimal to continue the relationship.

$$\beta V^t_i = (1-\beta)(J^t_i - U) \tag{2.3}$$

where $\beta \in (0,1)$. Some states might be such that the parties choose to sever the relationship. The wages we derive here are valid in states where the parties continue the relationship.

Rearrange (2.3) to yield

$$\beta(V^t_i + J^t_i) = J^t_i - U(1-\beta) \tag{2.4}$$

which I will use to derive the wage equation. Note that in general any value equation H can be expressed in its asset-value form

$$rH^t_i = f^t_i + \sum_j \lambda^t_{ij}(H^t_j - H^t_i) + \dot{H}^t_i$$

where f^t_i is the flow payment at time t in state i, H^t_i is the value of being in state i at time t, and so on. Hence if we multiply (2.4) by the discount factor r we get

$$\beta\left(rV^t_i + rJ^t_i\right) = rJ^t_i - rU(1-\beta)$$

which can be written as

$$\begin{aligned}&\beta\left(\pi^t_i + \sum_j \lambda^t_{ij}\left[R^t_{ij} + V^t_j + J^t_j - V^t_i - J^t_i\right] + \dot{V}^t_i + \dot{J}^t_i\right)\\ = \;& w^t_i + \sum \lambda^t_{ij}\left[J^t_j - J^t_i\right] + \dot{J}^t_i - rU(1-\beta)\end{aligned}$$

which can be used to solve for the wage

$$\begin{aligned}
w_i^t &= \beta\left(\pi_i^t + \sum_j \lambda_{ij}^t\left[R_{ij}^t + V_j^t + J_j^t - V_i^t - J_i^t\right] + \dot{V}_i^t + \dot{J}_i^t\right) \\
&\quad - \left(\sum_j \lambda_{ij}^t\left[J_j^t - J_i^t\right] + \dot{J}_i^t - rU(1-\beta)\right) \\
&= \beta\pi_i^t + \sum_j \lambda_{ij}^t\left[\beta R_{ij}^t + \beta V_j^t - (1-\beta)J_j^t - \beta V_i^t + (1-\beta)J_i^t\right] \\
&\quad + \beta\dot{V}_i^t - (1-\beta)\dot{J}_i^t + rU(1-\beta) \\
&= \beta\pi_i^t + \sum_j \lambda_{ij}^t\left[\beta R_{ij}^t + \left(\beta V_j^t - (1-\beta)J_j^t\right) - \left(\beta V_i^t - (1-\beta)J_i^t\right)\right] \\
&\quad + \beta\dot{V}_i^t - (1-\beta)\dot{J}_i^t + rU(1-\beta)
\end{aligned}$$

At this point, impose (2.3) so that the second and third terms in the summation cancel each other, as state-by-state they equal the same quantity, namely $-U(1-\beta)$. This gives

$$w_i^t = \beta\pi_i^t + \beta\sum_j \lambda_{ij}^t R_{ij}^t + \beta\dot{V}_i^t - (1-\beta)\dot{J}_i^t + rU(1-\beta)$$

and now differentiate (2.3), which must hold for all t, with respect to time to give

$$\beta\dot{V}_i^t - (1-\beta)\dot{J}_i^t = 0$$

remembering that U is time-invariant, so that we can further simplify the wage equation to read

$$w_i^t = rU(1-\beta) + \beta\pi_i^t + \beta\sum_j \lambda_{ij}^t R_{ij}^t \qquad (2.5)$$

Thus, the wage at any time can be expressed as a state-and-time-invariant constant $rU(1-\beta)$ plus labor's share of current flow profits, plus labor's share of the expected lumpy payments. The intuition is simple. The pie π can always be bargained over in the future according to the same rule β . Thus, there is no need for forward-looking participants to negotiate payments to account for possible changes in π. However, the rewards R are like cupcakes that the firm can immediately eat upon transit to a new state. Since labor cannot claim any portion of the future cupcakes when they arrive, it claims a portion of the anticipated cupcakes today.

Chapter 3

Team Play in the War of Attrition

3.1 Introduction

A standard modelling assumption in the War of Attrition is that players care only about winning a prize themselves ((Maynard Smith 1974), (Riley 1980)). In particular, if they do not win, then the identity of the eventual victor is of no consequence. In many scenarios this is no doubt reasonable. In certain areas of application, however, this assumption is inappropriate. For example, suppose three political candidates occupy different positions on the political spectrum. Two of these candidates are ideological allies, and differ only on detail rather than broad principle. The third, however, is diametrically opposed to the other two. All candidates wish to determine the political agenda, but incur costs when pursuing their own personal objectives at the expense of consensus; for example, it may be that a lack of agreement produces negative publicity for all politicians. Then, although each politician would most prefer that his own agenda be adopted, the victory of an ally is to be preferred to the victory of the third extreme candidate. A second leading case is that of the adoption of technological standards. We may imagine that negotiations among firms in standard-setting committees approximate wars of attrition (see, for example (Farrell 1993) and (David and Monroe 1994)). It is clear that while agents in such committees have an interest in seeing their own standard adopted, they may be particularly fearful of an industry standard which displays severe incompatibilities with their own technology.

Here we acknowledge the validity of this critique and construct a simple model that incorporates the possibility that losing agents in a war of attrition care about the identity of the winner. Our model is a three-player incomplete-information war of attrition with a single prize. Two of the players will be regarded as partners, each of whom receives a fraction of her own valuation for the prize if her partner wins. The third, solo, player can only receive a payoff by winning the war.

As is well known, the standard three-player war of attrition (without team effects) has no equilibrium in which all players fight — one player must exit immediately. Since the probability that *both* opponents are about to exit is vanishingly small, the marginal value of continuing to fight is zero. With team effects, a partnered player has an added incentive to exit, since her partner may go on to win the war — a free-riding effect. One might thus expect any equilibrium to involve the early exit of team players. In contrast, we show that, in equilibrium, a solo player drops out sooner than a team player with the same valuation.

We also consider the limiting equilibrium as the strength of the bond between partners falls to zero. In this case, the solo player exits the game immediately, under the threat of near-infinite delay by the team players. The team players then go on to fight a standard two-player war of attrition. We interpret this as an equilibrium selection device for the three-player single-prize game with no team effects.

The remainder of the paper is structured as follows. In section 2 we present the model. Section 3 contains the formal analysis. In section 4 we discuss some applications and numerical examples. An appendix contains further technical details.

3.2 The Model

There are three players indexed by $i \in \{1, 2, 3\}$. Each player has a private valuation u_i, distributed independently across players with cumulative distribution function $F(u)$. This distribution function is assumed to have support on some interval of the positive real axis, and corresponding density $f(u)$. Denote the lower bound of this interval by $\underline{u} > 0$. It will prove convenient to denote the underlying hazard rate of F by $\lambda(u) = f(u) / (1 - F(u))$. Assume that this hazard is bounded away from zero. Players $i \in \{1, 2\}$ will be deemed *team players*, and player $i = 3$ a *solo player*. It should be noted that this terminology does not imply collusive play or the transfer of payoffs.

Players pay a (normalized) unit cost per time period while they remain in the game, but may choose to permanently exit at any time, after which they cease to incur this cost. Once all but one player have dropped out, the game ends and payoffs accrue. A strategy for player i is therefore a pair of stopping times $\{t_i, T_i(\tau, j)\}$, where t_i denotes the time at which she drops out if no other player has dropped out yet, and the function $T_i(\tau, j)$ is the time at which she drops out given that player j dropped out at time τ.

Payoffs (gross of fighting costs) are as follows. The last player remaining receives u_i, as in a standard war of attrition. Unlike the standard model, however, a team player may receive a fraction of her valuation, even if she is not the last player remaining. In particular, a "caring coefficient" $\delta \in (0, 2)$ is introduced. If player 1 is the last remaining, then player 2 receives δu_2. Likewise, player 1 receives δu_1 if player 2 wins. The solo player receives a payoff of zero if she is not the winner.

3.3 Analysis

We show that the stopping times in the initial period of play are (weakly) increasing in a player's valuation. It follows that in any continuation game, a two-player war of attrition is played with possibly differing lower bounds on the competitor's valuation, but identical hazard rates. We solve these continuation games, selecting an equilibrium using the perturbation device of Fudenberg and Tirole (1986). Returning to the initial game, we prove the result that the solo player must exit more quickly than the team players. We can then characterize the equilibrium as a solution to a pair of differential equations, and analyze the limiting case where the caring coefficient converges to zero.

To see that strategies in the initial stage must be (weakly) increasing functions of agents' valuations, note that an agent of type u_H can earn at least as high a payoff as an agent of type $u_L < u_H$ by imitating type u_L's strategy for the entire game. In particular, if agent u_H planned to drop out of the game at some time τ strictly less than type u_L's planned exit time τ_L, she could improve her payoff by precisely imitating the lower type's strategy, including continuation game strategies, in the interval $[\tau, \tau_L]$. These considerations underlie the following lemma, whose proof is given in the appendix.

Lemma 1 *Initial exit times are weakly increasing in an agent's valuation.*

Proof: See appendix. ∎

Lemma 1 tells us that, once a player has dropped out, we have a two-player war of attrition with valuations arising from distributions with potentially different lower bounds. There are two possible configurations of such continuation games. In the first, the two team players fight. Since a team player who loses will receive a fraction δ of her valuation, this continuation game becomes a standard war of attrition with valuations scaled by $1-\delta$. In the second configuration, a solo player and a team player compete. There is no chance for the team player to receive a payoff from a win by her partner, and hence both players compete for their full valuations. In either case we may view the players' valuations as having been drawn independently from a common distribution subject to truncation.

As is well known, there are many equilibria of such games.[1] However, it seems natural for any player with a valuation below the lower bound of her opponent to exit immediately. For instance, suppose that players i and j have lower bounds $\underline{u}_i < \underline{u}_j$. One equilibrium is for any player i with valuation $u_i < \underline{u}_j$ to exit immediately, and for both players to employ the standard symmetric exit strategies from $\underline{u}_j$ onwards. Using techniques similar to those of Fudenberg and Tirole (1986), we show that only this equilibrium is robust to a small perturbation in payoffs. Insisting that there is always some probability that a player may find it a dominant strategy to always fight is sufficient to pin down a unique equilibrium.

Lemma 2 *Consider a two-player war of attrition, with valuations u_i and u_j drawn from an initially common distribution truncated below by $\underline{u}_i$ and $\underline{u}_j$ respectively, where $\underline{u}_i < \underline{u}_j$ without loss of generality. Further, suppose that players are restricted to fight forever with some arbitrarily small probability $\xi > 0$. Then the unique equilibrium satisfies:*

$$\tau(u) = \begin{cases} 0 & \underline{u}_i \leq u < \underline{u}_j \\ \int_{\underline{u}_j}^{u} x \frac{(1-\xi)f(x)}{1-(1-\xi)F(x)} dx & \underline{u}_j \leq u \end{cases}$$

In particular, notice that a player with $u_i < \underline{u}_j$ exits immediately, and the equilibrium is symmetric from $\underline{u}_j$ onwards.

Proof: See Appendix. ■

[1] Note however that the possibility of multiple equilibria does not result simply from the failure of the first-order conditions to satisfy Lipschitz continuity at the lower bound of the payoff distribution, since in the continuation game this lower bound will typically be positive for both players. Rather, the multiplicity arises from the inability to pin down the initial conditions when player's lower bounds differ.

Having solved for the continuation strategies above, and thus expected second stage payoffs, we can find the equilibrium stopping rules in the initial stage. Hence we seek a pair of stopping rules $\{s(u), t(u)\}$ carrying private type realizations into stopping times (we allow these times to be infinite). Observe the easy mnemonic that the function s describes the solo player's strategy while t gives that of the team players. It will actually be easier in general to work with the inverses of the stopping rules, that is, the functions carrying times to types. Therefore, agree to define $v(\tau) = s^{-1}(\tau)$ and $w(\tau) = t^{-1}(\tau)$ for $\tau > 0$. By assuming that these inverses exist, we are effectively restricting our analysis to strategies that are strictly increasing in an agent's valuation (save perhaps on a single interval containing $\underline{u}$). In fact, we shall further focus on strategies that are differentiable functions of agent valuations.

The following proposition will be fundamental in the derivations that follow. It is also of interest in its own right.

Proposition 1 *In the first stage of the war of attrition, team players stay in longer than solo players. That is,* $t \geq s$ *everywhere.*

Proof: Suppose not, so that for some u we have $s(u) > t(u)$, and consider a solo player with such a valuation u. She plans to exit at time $s(u)$, and thus at time $s(u) - \varepsilon$ she is willing to pay fighting costs and remain in the game. The only chance for her to win is if one of the team players drops out. If a team player drops out, however, the remaining team player's valuation will be greater than u, for sufficiently small ε, given the continuity of the stopping times. The solo player will thus exit immediately in the continuation game. The solo player has no chance of winning the prize, and hence should exit before time $s(u)$. ∎

In other words, given that team players drop out smoothly, a solo player who realizes that both team players have higher valuations than his own must also know that he can never win the continuation game in equilibrium.

It might seem that similar arguments would imply that $s \geq t$, so that (together with Proposition 1) $s = t$. After all, $s < t$ implies that team players who are certain, with probability one, that their own valuation is inferior to that of the solo player nonetheless continue to play the game. That is, they continue to bear fighting costs despite knowing that they will drop out immediately if their partner does, and that they will drop out immediately if the solo player does. In the case of the solo agent, we argued that similar knowledge must imply that such an agent should already should have dropped out. How can it be that the incentive to free ride on ones teammate is overwhelmed even when one knows that there is no chance to win in the continuation games?

The reason is simple and has a clear economic intuition. At time τ, the other team player (player 2, say) can credibly commit to dropping out as soon as player 1 does, so long as $u_2 < v(\tau)$. It is this possibility that keeps player 1 from dropping out as soon as she realizes that her valuation is the lowest remaining. When player 1 is unsure of her teammate's valuation, she chooses to stay in for fear that her teammate will not be willing to fight in the continuation game that would otherwise ensue. This is the *strategic commitment* effect.

3.3.1 Solution to the First Stage

The Solo Player

A solo player who drops out receives a payoff of zero. By assumption, the cost of choosing to continue to fight in a time interval of length ε is precisely ε. During this interval, one or both of the team players may exit. The probability that they both exit in this interval is of second order. If one team player exits, the solo player proceeds to a continuation game. But we know that the solo player can only win the continuation game if he has the highest valuation of all players involved in the continuation game. Whether or not he does in fact have the highest valuation will be revealed immediately, and hence no extra costs need be borne. The gain from a single player dropping out is hence the probability the remaining team player chooses to drop out immediately as a result. This happens with probability

$$\frac{F(v(\tau)) - F(w(\tau))}{1 - F(w(\tau))}$$

which is just the probability that in the continuation game, a team player has a valuation below the lower bound she perceives on the valuation of the solo player, given that the team player has stayed in the game until time τ.

Thus the indifference condition for the solo player at time τ is thus:

$$\underbrace{1}_{\text{fighting cost}} = \underbrace{2\lambda(w(\tau))w'(\tau)}_{\text{team drop-out rate}} \underbrace{\frac{F(v(\tau)) - F(w(\tau))}{1 - F(w(\tau))}}_{\text{Prob. of instant win}} \underbrace{v(\tau)}_{\text{valuation}}$$

which is a differential equation in the unknown functions $w(\tau)$ and $v(\tau)$.

Further manipulation easily yields

$$w'(\tau) = \frac{[1 - F(w(\tau))]}{2v(\tau)\lambda(w(\tau))[F(v(\tau)) - F(w(\tau))]}$$

which will be one of two key equations we use to determine a complete solution. In the appendix we verify the sufficiency of this first-order condition.

The Team Players

As we have seen, a team player who has reached her equilibrium stopping time plans to drop out immediately in any continuation game. Thus her only reason for fighting is to affect the probability that her partner eventually wins; she is fighting to claim only δu_1. Continuing to fight has no marginal effect on the probability her partner would win a continuation game against the solo player, however. In fact, the only way in which fighting affects the outcome of the game is if a continuation game her partner would have immediately lost is somehow avoided. It follows that a team player gains an extra δu by fighting only when not dropping out induces the solo player to exit, and a continuation game which would have been lost is never entered. The first-order condition is therefore

$$\underbrace{1}_{\text{Fighting cost}} = \underbrace{\lambda(v(\tau))v'(\tau)}_{\text{Rate of solo exit}} \underbrace{\frac{F(v(\tau)) - F(w(\tau))}{1 - F(w(\tau))}}_{\text{Prob. instant loss if 1 drops}} \underbrace{\delta w(\tau)}_{\text{Value if partner wins}}$$

so that the fighting cost is just compensated by the probability that the solo player exits and the other team player would have dropped out of the continuation game immediately if player 1 dropped out before the solo player. This expression can be rearranged to yield

$$v'(\tau) = \frac{[1 - F(w(\tau))]}{\delta[F(v(\tau)) - F(w(\tau))]\lambda(v(\tau))w(\tau)}$$

3.3.2 Solving the System

Taking ratios of these two differential equations gives:

$$\frac{v'(\tau)}{w'(\tau)} = \frac{2v(\tau)\lambda(w(\tau))}{\delta w(\tau)\lambda(v(\tau))}$$

a separable differential equation. We thus obtain:

$$\frac{\delta\lambda(v(\tau))v'(\tau)}{v(\tau)} = \frac{2\lambda(w(\tau))w'(\tau)}{w(\tau)}$$

which can be transformed into an equation involving only v and w by eliminating the time-dependence. This gives

$$\frac{\delta\lambda(x)dx}{x} = \frac{2\lambda(x)dx}{x}$$

which may be integrated, assuming the integrals exist. More precisely, it is necessary that the following integral exists for each $z > \underline{u}$:

$$H(z) = \int_{\underline{u}}^{z} \frac{\lambda(x)}{x} dx \tag{3.1}$$

To ensure existence of the integral, it is sufficient to assume that

$$\lim_{y \downarrow \underline{u}} \frac{\lambda(y)}{y} < \infty,$$

given the assumed continuity of $\lambda(y)/y$ away from $\underline{u}$. Note that the hazard rate is primitive to the game, and hence we need make no assumptions about the strategies *per se* to ensure the existence of the integrals.

We can see that the relationship between v and w can be described by

$$\delta H(v) = 2H(w) + k$$

for some constant of integration k. This equation is fundamental to a comparative statics analysis on the caring parameter δ, as we show below.

Since H is continuous and represents the integral of a positive function that is bounded away from zero, we can be sure that there is a unique solution to the above equation. Let the solution be given implicitly by:

$$v^*(w) = H^{-1}\left[\frac{2H(w) + k}{\delta}\right]. \tag{3.2}$$

So long as $k \geq 0$, this solution has the property that $v^*(w) \geq w$. With any such solution, we can eliminate v from the equation

$$w'(\tau) = \frac{[1 - F(w)]}{2v\lambda(w)[F(v) - F(w)]}$$

giving

$$w'(\tau) = \frac{dw}{d\tau} = \frac{[1 - F(w)]}{2v^*(w)\lambda(w)[F(v^*(w)) - F(w)]}$$

Immediately observe that $w'(\tau)$ is non-negative, so that we have now verified that the equilibrium stopping times are a non-decreasing function of the team players' valuations (identical considerations imply that the solution for the solo player is non-decreasing.)

To explicitly solve for the stopping rule of the team players, $t(u)$, one merely separates the above equation (after having substituted v^* for v) and then integrates both sides to yield,

$$\int_0^t dx = t = \int_{\underline{u}}^{u} \frac{2v^*(z)\lambda(z)[F(v^*(z)) - F(z)]}{[1 - F(z)]} dz \tag{3.3}$$

having imposed the boundary condition $t(\underline{u}) = 0$.

The solo player's strategy can be derived in a similar manner, by first solving for $w^*(v)$ and then integrating the function $v'(\tau)$. The boundary condition is determined by the selection of k and the team boundary condition $t(\underline{u}) = 0$. Observe that larger k correspond to equilibria in which a greater mass of solo players drops out immediately.

3.3.3 Limiting Results

So far we have demonstrated that when the sharing coefficient $\delta < 1$, the incentives of firms to free-ride on the play of their teammates is not so strong as to compensate for their probabilistic advantage ("strength in numbers") We have shown that the solo player always drops out more rapidly than the team players.

Let us now consider what happens as the caring coefficient δ falls to zero. This is a natural limiting case describing the strategic environment when two player's fortunes are only weakly linked, yet still completely severed from those of the third player.

Consider $v^*(u, \delta)$ defined as the solution to the equation, derived above,

$$\delta H(v) = 2H(u) + k$$

for fixed $k \geq 0$. It is clear from this equation that as δ becomes smaller, the required v grows larger (for fixed u), keeping in mind once again that H is an increasing positive function. Hence we conclude that $v^*(u, \delta)$ is increasing without bound as δ falls towards zero. Likewise, observe that $w^*(u, \delta)$ (defined as the solution to $\delta H(u) = 2H(w) + k$) falls monotonically and is bounded below by zero as δ falls to zero. These facts strongly hint at a limiting solution in which the solo player drops out immediately, at which time the team players engage in a traditional war of attrition with a caring coefficient of zero. To prove this, however, we need to consider precisely what happens to the stopping times $s(u, \delta)$ in the limit.

To investigate this issue we consider again the system of differential equations above. Consider the stopping time of a team players of type u, defined by

$$t = \int_{\underline{u}}^{u} \frac{2v^*(z,\delta)\lambda(z)[F(v^*(z,\delta)) - F(z)]}{1 - F(z)} dz$$

and observe that, for fixed u, the integrand grows without bound at each point z, since $v^*(z,\delta)$ does. Hence we have shown that $t(u,\delta)$ grows strictly as δ falls, and $\lim_{\delta \longrightarrow 0} t(u,\delta) = \infty$.

Now consider the stopping time of the solo player with valuation u

$$s = \int_{\underline{u}}^{u} \frac{\delta[F(z) - F(w^*(z,\delta))]\lambda(z)w^*(z,\delta)}{[1 - F(w^*(z,\delta))]} dz$$
$$< \int_{\underline{u}}^{u} \delta\lambda(z)w^*(z,\delta)dz \longrightarrow 0$$

as δ goes to zero, in light of the fact that w^* is decreasing in δ. Hence we have that $\lim_{\delta \longrightarrow 0} s(u,\delta) = 0$ for each u. These results are summarized in the following proposition.

Proposition 2 *As the caring parameter converges to zero, the strategies of the players are such that the solo player drops out immediately, under the threat of infinite delay by the team players. In the continuation game, which is entered without delay, the team players compete as in the standard war of attrition with private information.*

As we have seen, the desire by any team player to free ride is outweighed by the strategic commitment effect, which derives its force from the possibility that a team player's partner drops out immediately in a continuation game. Reducing the caring parameter, by decreasing the incentive to free ride, reinforces the commitment effect. The reason is that the probability of an instant loss in the continuation game increases as the fighting times of the team players increase.

This proposition may be compared to the "twoness" result of Bulow and Klemperer (1999). They modify the standard war of attrition by requiring that players who have dropped out nevertheless continue to pay some reduced costs as long as the game continues. They obtain the result that as these reduced costs fall to zero, all but the two highest valuation competitors exit immediately. Although we also obtain a similar immediate exit property in the limit, which player meets the selection criterion is common knowledge to the participants, while coordinating on an equilibrium in which only the highest valuation players fight would seem to require that players condition on information they do not actually have.

3.4 ILLUSTRATION

Here we present a parametric example that illustrates the force of the results of the previous section. We focus on the Weibull distribution with increasing hazard. In this case, the hazard satisfies

$$\lambda(x) = \lambda x^{\alpha}$$

where $\alpha > 0$. Recall from equation (3.2) that the relationship between $v(\tau)$ and $w(\tau)$ is described by

$$v^{*}(w) = H^{-1}\left[\frac{2H(w)+k}{\delta}\right].$$

where H is as defined in equation (3.1). In this case, we have

$$H(x) = \int_{\underline{u}}^{x} \frac{\lambda(s)}{s} ds = \int_{\underline{u}}^{x} \lambda s^{\alpha-1} ds = \frac{\lambda}{\alpha}(x^{a} - \underline{u}^{\alpha})$$

which may be inverted to find $v^{*}(w)$, which we plug into equation (3.3). Integrating numerically, we may obtain the solution for $t(u)$. In a similar fashion, we can obtain $s(u)$.

3.5 CONCLUSION

Rather than focus on the current model, one could in principle allow arbitrary preferences for each player over the eventual winner; the case of "team effects" seems of particular interest in applications. The conclusion in this case that solo players always exit more quickly hints at the potential for interesting results to be derived from other extensions of the basic model. In addition, we have seen that the introduction of externalities may allow the analyst to avoid the troubling feature of the standard model that at most two players can fight in any equilibrium, which seems to limit its practical application. Our analysis of this special case is a first step toward a general theory of the war of attrition with externalities.

Appendix A: Omitted Proofs

Lemma 3 *Initial exit times are weakly increasing in an agent's valuation.*

Proof: Suppose to the contrary that an agent with valuation u_H prefers to exit in the initial stage at time τ_L while an agent with valuation $u_L < u_H$ prefers to exit at time $\tau_H > \tau_L$. Suppose that the game has reached τ_L without any player dropping out, and consider the decision problem players face. Denote by $\Pi(\tau)$ the probability that an agent wins the prize, given that the agent plans to drop out at time τ and that the agent plans to use the continuation strategy of player u_L in the interval $[\tau_L, \tau]$. Write $P(\tau)$ for the probability that an agent's partner wins the prize, given that the agent plans to drop out at time τ. (This probability may be identically zero if we are considering a solo player). Finally, note that by not dropping out at time τ_L, a player of type u forfeits $u\delta P(\tau_L)$.Then from the fact that the low-valuation agent prefers τ_H to τ_L, it follows that

$$u_L \left[\Pi(\tau_H) - \delta P(\tau_L) + \delta P(\tau_H)\right] \geq C(\tau_H) > 0$$

where $C(\tau_H)$ denotes the additional expected waiting costs associated with staying in until τ_H. Since these costs must be positive, it follows that the term in brackets is positive. We conclude that an agent with valuation $u_H > u_L$ would also prefer to fight until τ_H and follow the continuation strategy of agent u_L if the continuation game begins in the interval $[\tau_L, \tau_H]$. Thus, his original strategy of dropping out at time τ_L is dominated, a contradiction. ■

Lemma 4 *The initial stage first-order conditions characterize an optimal strategy.*

Proof: We consider the solo player's problem. Similar techniques may be applied to the team players' first-order condition.

It is sufficient to check that the first-order conditions imply that the global second-order conditions are satisfied. Refraining from imposing the equilibrium condition $u = v(\tau)$, we have from above

$$0 = -1 + 2u\lambda(w(\tau))w'(\tau)\,\frac{F(v(\tau)) - F(w(\tau))}{1 - F(w(\tau))}$$

which can be differentiated with respect to u to yield

$$2\lambda(w(\tau))w'(\tau)\,\frac{F(v(\tau)) - F(w(\tau))}{1 - F(w(\tau))} > 0$$

if the first-order condition is satisfied. We now exploit this fact to demonstrate that solutions to the above in fact represent maximizers of the player's objective functions. Denote by $U(\tau, u)$ the expected payoff of a type-u solo player choosing to remain in until time τ. Suppose that the first-order condition above is satisfied along the path $s(u)$, and that a player of type-u is considering stopping at time $s(\widehat{u})$. Then

$$\begin{aligned} U(s(u),u) - U(s(\widehat{u}),u) &= \int_{s(\widehat{u})}^{s(u)} U_1(\tau,u)d\tau \\ &= \int_{s(\widehat{u})}^{s(u)} \left[U_1(\tau,\tau) + \int_{\tau}^{u} U_{12}(\tau,y)dy\right] d\tau \\ &= \int_{s(\hat{u})}^{s(u)} \int_{\tau}^{u} U_{12}(\tau,y)dyd\tau \end{aligned}$$

which is never negative. To see that is the case, observe that $s(\hat{u}) < s(u) \Leftrightarrow \tau < u$, and that U_{12} is positive. ∎

Appendix B: Uniqueness of Equilibrium in Continuation Games

In this section, we will consider the uniqueness of our posited equilibrium for continuation games, using a small perturbation argument. The valuations of players will be drawn from a common distribution, subject to truncation below by $\underline{v}$ and $\underline{w}$, where $\underline{v} \geq \underline{w}$. In addition, there is some fixed probability $\xi > 0$ that a player has a dominant strategy to remain in the game forever. For instance, a player might enjoy fighting with some small probability. We will model this by extending the support of the distribution of valuations to include $+\infty$, a value which is taken with probability ξ. Beginning with a valuation described by a density $f(x)$, we define a new density $g(x) = (1-\xi) f(x)$ with associated cumulative distribution function $G(x)$. Since we have an atom at $+\infty$, notice that $\lim_{x\to\infty} G(x) = 1 - \xi$.

We will first show that stopping times in this continuation game are increasing in a player's valuation. We will then show that such stopping time functions are continuous, strictly increasing and differentiable. The key results of relative toughness and uniqueness of equilibrium may then be established.

Strategies are a pair of stopping rules $\{s_2(u), t_2(u)\}$. To be an equilibrium profile, these must satisfy the functional equations:

$$s_2(x) \in \arg\max_{\tau} E_u\left[x\mathbb{I}\left\{\tau > t_2(u)\right\} - \min\left\{\tau, t_2(u)\right\}\right]$$
$$t_2(x) = \arg\max_{\tau} E_u\left[x\mathbb{I}\left\{\tau > s_2(u)\right\} - \min\left\{\tau, s_2(u)\right\}\right]$$

where $\mathbb{I}$ is the indicator function.

Lemma 5 *Stopping times are increasing in a player's valuation.*

Proof: We will show this for $s_2(x)$. Suppose that for $x_L < x_H$ we have $\tau_H = s_2(x_L) > s_2(x_H) = \tau_L$. Then:

$$\begin{aligned} &x_L(E_u[\mathbb{I}\{\tau_H > t_2(u)\} - \mathbb{I}\{\tau_L > t_2(u)\}]) \\ &\geq E_u\left[\min\left\{\tau_H, t_2(u)\right\} - \min\left\{\tau_L, t_2(u)\right\}\right] \end{aligned} \tag{3.4}$$

Observe that the right hand side of this inequality is strictly positive:

$$\begin{aligned} &E_u\left[\min\left\{\tau_H, t_2(u)\right\} - \min\left\{\tau_L, t_2(u)\right\}\right] \\ &\qquad \geq \xi\left(\min\left\{\tau_H, t_2(\infty)\right\} - \min\left\{\tau_L, t_2(\infty)\right\}\right) = \xi\left(\tau_H - \tau_L\right) > 0 \end{aligned}$$

The left hand side of the equation (3.4) must be strictly positive, so that:

$$E_u\left[\mathbb{I}\left\{\tau_H > t_2(u)\right\} - \mathbb{I}\left\{\tau_L > t_2(u)\right\}\right] > 0$$

We can conclude that:

$$\begin{aligned} &x_H E_u\left[\mathbb{I}\left\{\tau_H > t_2(u)\right\} - \mathbb{I}\left\{\tau_L > t_2(u)\right\}\right] \\ &\qquad > x_L E_u\left[\mathbb{I}\left\{\tau_H > t_2(u)\right\} - \mathbb{I}\left\{\tau_L > t_2(u)\right\}\right] \end{aligned}$$

and hence a player with valuation x_H strictly prefers stopping time τ_H to τ_L. It follows that stopping times are increasing in valuations. ∎

Lemma 6 *Stopping times are continuous in players' valuations.*

Proof: Suppose that there is a discontinuity in $s_2(x)$ at the point x^*. Since the function is monotonic, the upper and lower limits $\tau_L = \lim_{x\uparrow x^*} s_2(x)$ and $\tau_H = \lim_{x\downarrow x^*} s_2(x)$ are well defined, where $\tau_H > \tau_L$. The other player will not quit in the interval $(\tau_L, \tau_H]$, since there is no chance of winning in this interval, and hence fighting costs would be saved be exiting at τ_L.

There is a some x such that $s_2(x) = \tau_H + \varepsilon$. This x is bounded above, and this player pays a fighting cost of at least $\tau_H - \tau_L$ to arrive at this point. The probability of winning is:

$$\Pr\left[\tau_H + \varepsilon > t_2(u) \geq \tau_H\right]$$

which is bounded away from 0 since the agent finds it profitable to pay the cost $\tau_H - \tau_L$. Taking limits as ε goes to zero, we find that there must be an atom at τ_H, a contradiction.

■

Lemma 7 *Stopping times are strictly increasing in players' valuations.*

Proof: Suppose not. Then, there exists a $\tau > 0$, such that a positive mass of players drops out at τ. Thus for small ε, dropping out in the interval $(\tau - \varepsilon, \tau]$ is dominated by staying in until just after τ, which contradicts the continuity of stopping times.

■

Since the stopping rules are strictly increasing, they must be differentiable almost everywhere. Since they are strictly increasing, their inverses are well defined. At points of differentiability, inverse stopping rules are characterized by the differential equations:

$$v'(\tau) = \frac{1}{\lambda(v(\tau))\, w(\tau)} \tag{3.5}$$

$$w'(\tau) = \frac{1}{\lambda(w(\tau))\, v(\tau)} \tag{3.6}$$

Lemma 8 *The inverse stopping rules are differentiable for $\tau > 0$.*

Proof: Combine $v(t)$ and $w(t)$ into the vector valued function $y(t)$. This is a Lipschitz continuous function of t, and differentiable almost everywhere. Where differentiable, we can write equations (3.5–3.6) as $dy/dt = h(y(t))$ where h is Lipschitz continuous. It is a standard result that such a function $y(t)$ is differentiable everywhere.

■

Our posited solution is symmetric for valuations above $\underline{v}$. Denoting this solution by $v^*(\tau)$, it is characterized by the differential equation:

$$\frac{dv^*(\tau)}{d\tau} = \frac{1}{\lambda(v^*(\tau))\, v^*(\tau)}$$

Furthermore, the solution to this equation, with boundary condition $v^*(0) = \underline{v}$ is:

$$t^*(u) = \int_{\underline{v}}^{u} \lambda(x)\, x\, dx$$

We now wish to show that this equilibrium is unique. We first do this by showing that any other equilibrium satisfies the Fudenberg-Tirole "relative toughness" criterion.

Lemma 9 *Suppose that $\{\tilde{v}(\tau), \tilde{w}(\tau)\}$ is a solution to the game other than the solution $v = w = v^*(\tau)$. Then, for $\tau > 0$, we have*

$$(\tilde{v}(\tau) - v^*(\tau))(\tilde{w}(\tau) - v^*(\tau)) < 0$$

Proof: Suppose that for some τ the relative toughness condition is true, and without loss for this τ we have $\tilde{v}(\tau) < v^*(\tau)$. We will show that relative toughness holds for all $s > \tau$. Suppose not. Since one of the components of the product in the lemma must change sign, and since the strategies are continuous, we must have either $\tilde{v}(s) = v^*(s)$ or $\tilde{w}(s) = v^*(s)$ for some s. Suppose that s^* is the first such s, and without loss of generality $\tilde{v}(s^*) = v^*(s^*)$. It cannot be the case that $\tilde{w}(s^*) = v^*(s^*)$. To see this, note that the right hand sides of the differential equations (3.5–3.6) are Lipschitz continuous, and hence there is a unique solution through any point. Since we have such a point where $\tilde{v} = \tilde{w} = v^*$, this must be true always and hence $\tilde{v} = \tilde{w} = v^*$ for all s, a contradiction. Hence $\tilde{w}(s^*) > v^*(s^*)$. But inspecting the differential equation for $\tilde{v}(s)$ we have:

$$\tilde{v}'(s^*) = \frac{1}{\lambda(\tilde{v}(s^*))\, \tilde{w}(s^*)} = \frac{1}{\lambda(v^*(s^*))\, \tilde{w}(s^*)} < \frac{1}{\lambda(v^*(s^*))\, v^*(s^*)} = v^{*\prime}(s^*)$$

Thus there exists an $s < s^*$ such that $\tilde{v}(s) > v^*(s)$. This is a contradiction.

It follows that if relative toughness is satisfied for τ, then it is satisfied for all $s > \tau$. We next show that it is satisfied for all $\tau \in (0, \varepsilon)$ for sufficiently small ε. Recall that for any solution either $v(0) = \underline{v}$ or $w(0) = \underline{w}$. Suppose that $\tilde{v}(0) = \underline{v}$, and that $\tilde{w}(0) > \underline{v}$. Then by continuity, for $\tau \in (0, \varepsilon)$ we have $\tilde{w}(\tau) > v^*(\tau)$, since $v^*(0) = \underline{v}$. Next consider the derivative:

$$\tilde{v}'(0) = \frac{1}{\lambda(\tilde{v}(0))\, \tilde{w}(0)} < \frac{1}{\lambda(\underline{v})\, \underline{v}} = v^{*\prime}(0)$$

Hence for $\tau \in (0, \varepsilon)$ we have $\tilde{v}(\tau) < v^*(\tau)$ and hence relative toughness is satisfied. A similar argument applies when $\tilde{w}(0) < \underline{v}$. We cannot have $\tilde{w}(0) = \underline{v}$, since this would give us the solution $v^* = \tilde{v} = \tilde{w}$. ■

Lemma 10 *There is a unique solution to the continuation game.*

Proof: Suppose that there is a solution $\{\tilde{v}(\tau), \tilde{w}(\tau)\}$ other than $v = w = v^*$. The first-order condition gives us:

$$\frac{1}{\tilde{w}(\tau)} = \tilde{v}'(\tau)\frac{f(\tilde{v}(\tau))}{S(\tilde{v}(\tau))} = -\frac{d}{d\tau}\log S(\tilde{v}(\tau))$$

where $S(u) = 1 - F(u)$. Take $\tau > \tau_0 > 0$, and integrate to obtain:

$$\log S(\tilde{v}(\tau_0)) - \log S(\tilde{v}(\tau)) = \int_{\tau_0}^{\tau} \frac{1}{\tilde{w}(s)} ds$$

The relative toughness condition holds, and thus assume without loss that $\tilde{v} < v^*$ and hence $\tilde{w} > v^*$. Hence:

$$\int_{\tau_0}^{\tau} \frac{1}{\tilde{w}(s)} ds < \int_{\tau_0}^{\tau} \frac{1}{v^*(s)} ds = \log S(v^*(\tau_0)) - \log S(v^*(\tau))$$

We next take limits as $\tau \to \infty$. Noting that $S(v(\tau)) \to \xi$, we have that $\log S(v(\tau))$ approaches a finite limit. We may thus conclude that:

$$\log S(\tilde{v}(\tau_0)) < \log S(v^*(\tau_0)) \Rightarrow \tilde{v}(\tau_0) > v^*(\tau_0)$$

This is a contradiction, since we assumed that $\tilde{v} < v^*$ for all $\tau > 0$. ■

References

ACEMOGLU, D. (1997): "Training and Innovation in an Imperfect Labour Market," *Review of Economic Studies*, 64.

ACS, Z. J. (1996): "Small Firms and Economic Growth," in *Small Business in the Modern Economy*, ed. by P. Admiraal. Blackwell, Cambridge, MA.

AKERLOF, G. A. (1982): "Jobs as Dam Sites," *Review of Economic Studies*, 48.

ANDREWS, E. L. (June 23 1994): "FCC Plan for License Diversity: Wireless Auction to Feature Set-Asides," *New York Times*.

ASH, R. B. (1972): *Real Analysis and Probability*. Academic Press, New York.

AUDRETSCH, D. B. (1991): "New Firm Survival and the Technological Regime," *Review of Economics and Statistics*, 73.

BALDWIN, J., AND G. PICOT (1995): "Employment Generation by Small Producers in the Canadian Manufacturing Sector," *Small Business Economics*, 7.

BERGSTROM, T., L. BLUME, AND H. VARIAN (1986): "On the Private Provision of Public Goods," *Journal of Public Economics*, 29.

BIRCH, D. L. (1987): *Job Creation in America: How Our Smallest Companies Put the Most People to Work*. Free Press, New York, NY.

BLANCHARD, O. J., AND P. A. DIAMOND (1992): "The Flow Approach to Labor Markets," *American Economic Review*, 82.

BLISS, C., AND B. NALEBUFF (1984): "Dragon-slaying and Ballroom Dancing: The Private Supply of a Public Good," *Journal of Public Economics*, 25.

BROCK, W. A., AND D. S. EVANS (1986): *The Economics of Small Business: Their Role and Regulation in the U.S. Economy.* Holmes and Meier, New York.

BROWN, C., J. HAMILTON, AND J. MEDOFF (1990): *Employers Large and Small.* Harvard University Press.

BULOW, J., AND P. KLEMPERER (1999): "The Generalized War of Attrition," *American Economic Review*, 89.

CABALLERO, R. J., AND M. L. HAMMOUR (1996): "On the Timing and Efficiency of Creative Destruction," *Quarterly Journal of Economics*, 111.

——— (1998): "The Macroeconomics of Specificity," *Journal of Political Economy*, 106.

CHAMBERLIN, J. (1974): "Provision of Collective Goods as a Function of Group Size," *American Political Science Review*, 68.

DAVID, P. A., AND H. K. MONROE (1994): "Standards Development Strategies under Incomplete Information — Isn't the "Battle of the Sexes" Really a Revelation Game?," All Souls College and International Monetary Fund.

DAVIS, S. J., J. C. HALTIWANGER, AND S. SCHUH (1996): *Job Creation and Destruction.* MIT Press, Cambridge, MA.

DIAMOND, P. A. (1982): "Wage Determination and Efficiency in Search Equilibrium," *Review of Economic Studies*, 49.

FARRELL, J. (1993): "Choosing the Rules for Formal Standardization," University of California, Berkeley, Working Paper.

FUDENBERG, D., AND J. TIROLE (1986): "A Theory of Exit in Duopoly," *Econometrica*, 54.

HECKER, F. (1998): "Setting Up Shop: The Business of Open-Source Software," http://people.netscape.com/hecker/setting-up-shop.html.

HOPENHAYN, H. A. (1992): "Entry, Exit and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 60.

HOSIOS, A. J. (1990): "On the Efficiency of Matching and Related Models of Search and Unemployment," *Review of Economic Studies*, 57.

JOVANOVICH, B. (1982): "Selection and Evolution of Industry," *Econometrica*, 50.

KLEPPER, S. (1996): "Entry, Exit, Growth and Innovation over the Product Life Cycle," *American Economic Review*, 86.

MAHMOOD, T. (1992): "Does the Hazard Rate of New Plants Vary Between High- and Low-Tech Industries?," *Small Business Economics*, 4.

MAYNARD SMITH, J. (1974): "The Theory of Games and the Evolution of Animal Conflicts," *Journal of Theoretical Biology*, 47.

MILLER, B. P., L. FREDRIKSON, AND B. SO (1990): "An Empirical Study of the Reliability of Unix Utilities," *Communications of the ACM*, 33.

MILLER, B. P., D. KOSKI, C. P. LEE, V. MAGANTY, R. MURTHY, A. NATARAJAN, AND J. STEIDL (1998): "Fuzz Revistited: A Re-examination of the Reliability of Unix Utilities and Services," University of Wisconsin Computer Science Working Paper.

PALFREY, T. R., AND H. ROSENTHAL (1984): "Participation and the Provision of Discrete Public Goods: A Strategic Analysis," *Journal of Public Economics*, 24.

RAYMOND, E. S. (1998): "The Cathedral and the Bazaar," http://www.tuxedo.org/~esr/writings/cathedral-bazaar/cathedral-bazaar.html.

RILEY, J. (1980): "Strong Evolutionary Equilibrium and the War of Attrition," *Journal of Theoretical Biology*, 83.

SMALL BUSINESS ADMINISTRATION (1987): "A Report to the President," *The State of Small Business.*

——— (1997): "Federal Acquisition Regulation 19.502-2," SBA Website.

——— (1998): "Small-Business-Friendly Provisions of the Taxpayer Relief Act of 1997," SBA Website.

STALLMAN, R. M. (1996): *GNU Emacs Manual, version 19.33.* Free Software Foundation, Boston, MA.