In the Club: Strategic Admission, Membership, and Endogenous Splits

Michael Sacks
West Virginia University

January 2017

Abstract

Clubs are an important form of organization in many economic contexts. This is the first study to combine a dynamic analysis of capital formation within clubs with an analysis of competition among clubs, generating several new insights. In particular, individuals with preferences that are far from the objective of the club may not immediately split and form a new club. Instead they may take advantage of the increasing returns from club membership and incubate their new club within an existing one. In equilibrium, clubs may not be able to prevent this type of behavior even if it is undesired. Moreover, there are a range of conditions under which clubs may encourage incubation of future competitors to take advantage of increasing returns themselves and build up their own capital base. The results are applied to the software industry.

Key words: Club formation, endogenous competition, dynamic contributions, capital accumulation, heterogeneous preferences, incubation, open source software.

JEL classifications: C72, C73, D21, D23, H41, L11, L17.
1 Introduction

Communities and clubs play a fundamental role across many branches of economics, including industrial organization, the economics of innovation, political economy, and religion, to name a few. Thus it is not surprising that, over the past sixty years, researchers have labored to understand the complex processes underlying individuals’ participation decisions in clubs, the clubs’ membership rules, and the subsequent production of club goods\textsuperscript{1}. From this long body of literature, three results stand out (besides the obvious free-rider problem): firstly, increased heterogeneity among members decreases the level of contributions (Alesina et al., 1999; Alesina and La Ferrara, 2000), which in turn incentivizes the club to impose membership rules (Ian-naccone, 1992; McBride, 2007\textsuperscript{a}; Aimone et al., 2013). Secondly, the time horizon matters with respect to the level of contributions (Fershtman and Nitzan, 1991; Wirl, 1996; Fujiwara and Matsueda, 2009), and the formation of new, competing clubs. Finally, individuals are strategic in club selection and location (Austin (1995); Alesina and Spolaore (1997); Casella (2001); Haimanko et al. (2004), among others). While models looking at heterogeneous agents, club formation, or time dynamics separately reveal many significant insights, this reductionist approach misses important features stemming from their interaction.

In this paper, I develop a dynamic model analyzing contributions to non-congestible club goods incorporating heterogeneous individuals, capital accumulation within clubs, and competition among clubs. Heterogeneity is modeled horizontally, where each individual is endowed with an ideal location and each club is assigned a location by its members. At any point in time, an individual or coalition of individuals can non-cooperatively orchestrate a schism, forming a child club that splits from the parent club. Since capital accumulates over time, those individuals that split need not leave empty-handed; rather, the amount of capital the coalition leaves with is dependent upon the degree of rivalry and excludability of the contributions, plus the distance between the parent club and the child club. As the degree of heterogeneity between the two clubs increases, the capital becomes less useful, which implies that capital depreciates according to distance.

Individuals strategically choose their club affiliation and contribution schedule,\textsuperscript{1}

\textsuperscript{1}Buchanan (1965) defines a club as an ownership-membership agreement. I follow this definition throughout the remainder of the paper.
where affiliations may change over time as new clubs emerge. I characterize the existence of three types of equilibria: the cohesive equilibrium in which a single group persists in the long run, the splitting equilibrium in which clubs split at the outset, and the incubation equilibrium in which a delayed split occurs. I refer to the underlying strategy played in the incubation equilibrium as strategic membership. There exist conditions under which all three equilibria coexist.

In general, the existence and subsequent payoff ranking of the equilibria are determined by two interdependent tensions. The contribution-scale tension illustrates the underlying tradeoff between economies of scale (club size) and heterogeneity. As individuals are heterogeneous, the introduction of more contributors need not increase the utility of a given member, even though the club good is non-congestible. To illustrate this point, consider the development of open source software (OSS). The development team must choose where to locate the project within the characteristic space (e.g. features and usability), à la Lancaster (1966). Adding another individual to the development team may bias the location decision, which in turn could make a subset of the development team worse off than if the individual was not a member, even with the additional contributions. The mission-scope tension illustrates the underlying tradeoff between depreciation, the chosen location of the child club, and the members’ ideal location. Forming a child club located far away from the parent club is costly, as the accumulated capital becomes less valuable. Reconsider the development of OSS and suppose that an OSS project is forked (the technical term for the splitting of a project). The coalition that forks the project must take into account the relevance of the accumulated source code that the group is leaving with when choosing where to locate the new OSS project on the spectrum of possible projects. Locating too close to the original project will reduce the value of the split because fewer individuals will be developing a similar project. Locating too far from the original project will also reduce the value of a split because the usefulness of the accumulated source code is decreasing in the distance between the two projects, so the development team must essentially start from scratch.

One standard approach to discriminating among equilibria in dynamic settings is to compare steady-state payoffs by assuming that the discount rate is sufficiently small so that individuals are sufficiently patient. An analysis of this kind is rather misleading in my setting. The steady-state payoff in the splitting equilibrium is higher than the steady-state payoff in the incubation equilibrium for all individuals, while
the payoff ranking of the cohesive equilibrium can vary. However, the incubation equilibrium can yield higher payoffs prior to a split occurring, along the path to the steady state, because incubation allows more rapid capital accumulation by both the parent and child clubs. Hence taking into account the full dynamic path, I identify conditions under which the incubation equilibrium payoff dominates both the cohesive and splitting equilibria.

I also identify instruments that the parent club can use to shape decisions at both the extensive and intensive margins. When there is a misalignment of preferences between individuals who split and those who remain in the parent club, or preferences are aligned and incubation is preferred, the parent club can utilize one of two instruments to its advantage. The first involves taxation and subsidization. If the parent club prefers that incubation does (does not) occur, then it can institute a subsidy (tax) to encourage (discourage) such behavior. The club acts strategically to determine who is admitted. Furthermore, if its own members are sufficiently patient and preferences are aligned, then any rents from incubation can be captured fully by the parent club by instituting a delayed benefits scheme, where all members are taxed at the outset, only to have the payments returned post-split. Each individual who remains receives her original payment plus a share of the payments made by those who split. The second instrument involves the use of intellectual property restrictions, which the parent club can use to limit incubation behavior. By controlling the degree of excludability and depreciation through protections such as patents, copyrights, non-disclosure agreements, and non-compete clauses, the parent club can eliminate all benefits from incubation. Taxes and subsidies can then be used to select between the cohesive and splitting equilibria.

The insights of the model are applied to both the internal development of proprietary software (PS) by a for-profit software vendor and the development of OSS by a community. PS vendors developing product lines can benefit by implementing an internal form of incubation, where the development teams work together by first developing shared features as a cohesive unit, then separating to simultaneously and independently develop each product in the line. With respect to OSS, I restrict attention to the history of the derivatives of Unix, namely Linux and the BSD (Berkeley Software Distribution) family of operating systems. Much of the evolution of the Linux kernel, the various Linux distributions, and the BSD operating systems can be better understood in the context of the model developed in this paper. Traditionally,
authors argued that open source software can be best understood by treating the development as the private provision of a public good (Johnson 2002; Polanski 2007). I argue that the development process, illustrated by the development of the Linux kernel, more closely resembles that of a club good.

This paper contributes to three strands of literature in club theory: the effects of heterogeneity, club dynamics, and competition between clubs. These strands of literature have, for the most part, evolved independently over time. This paper analyzes heterogeneity, club dynamics, and competition in a unified framework, providing a link between the strands. The literature on the theory of club goods can be traced back to the works of Buchanan (1965) and Olson (1965). The first models were nonstrategic, in that the analysis was typically conducted considering welfare rather than what are now considered standard game theoretic tools and concepts (e.g. Nash equilibrium). Buchanan formally introduced the idea of a club good - a good located in an intermediate position on the spectrum from purely private goods to purely public goods while Olson’s book provided a treatment of many of the aspects of clubs themselves, rather than the goods they produce. Since those early works, much has been done to advance various branches of the literature. For more details, see Cornes and Sandler (1996); Glazer et al. (1997); Sandler and Tschirhart (1997); Scotchmer (2002), and the references therein.

Fershtman and Nitzan (1991) and Wirl (1996) were among the first to analyze dynamic models of contribution to public goods with capital accumulation and depreciation. They find that the free-rider problem illustrated in static settings can (Fershtman and Nitzan 1991) but need not (Wirl 1996) persist in dynamic settings, depending on the equilibrium strategy chosen. The authors model a single club to which all agents belong, thus abstracting from club issues apart from their public goods component. These findings were later generalized in Fujiwara and Matsueda (2009). The same intuition can be applied to club goods when all membership decisions must be made at the outset, as the setting is isomorphic to analyzing multiple public goods with independent populations of interest. A two-club Tiebout model is considered in Glomm and Lagunoff (1999), where the clubs are differentiated in

---

2See Bergstrom et al. (1986) and Bernheim (1986) for details on the private provision of public goods.

3The free-rider problem persists and becomes more severe when the agents play a Markov perfect equilibrium in linear strategies. Wirl (1996) shows that nonlinear Markov perfect equilibrium strategies exist in which there is no freeriding.
the contribution mechanism: voluntary versus involuntary contributions. The authors find that the involuntary mechanism is more likely to prevail over time when individuals are sufficiently patient. The present paper differs from Glomm and Lagunoff (1999) by making the club formation decision endogenous while holding the contribution mechanism fixed (as voluntary).

Marx and Matthews (2000) take a similar dynamic approach to modeling a club, but focus on a single club with voluntary contributions. The authors model an environment where there is a threshold of contributions that, once met, allows for provision of the good. Their focus is analyzing the time to completion in all possible Nash equilibria, including Bayesian strategies and perfect Bayesian equilibria. Georgiadis (2015) considers a dynamic structure, akin to Marx and Matthews (2000), and finds that when individuals only receive the benefit when the project is complete and the project is sufficiently far from completion, members of a larger club contribute more than members of a smaller club. In my model, whenever incubation is payoff dominant, this finding holds even when individuals receive a benefit at every stage of the project.

Two works, Arnold and Wooders (2005, 2009), are close to this paper. The first paper examines dynamic club formation with myopic homogeneous agents, which places the focus on the equilibrium club size. The second again considers myopic agents, but in a hedonic game setting. In contrast, this paper develops a dynamic model with forward looking agents where the focus is on the members’ decision to split along with the parent club’s reaction to the split.

The literature on competition between clubs has developed along side the literature on club dynamics. Austin (1995) studies coordinated secession from public good jurisdictions in a club goods framework. Alesina and Spolaore (1997), Alesina et al. (1999), and Alesina and La Ferrara (2000) develop models of group selection participation in the context of nation building and local public good provision under (horizontally) heterogeneous agents. The first paper focuses on the optimal size of nations while the latter two develop models which address the membership decisions of individuals. Casella (2001) analyzes the relationship between the number and size of public jurisdictions as a function of the size of the underlying markets when there is heterogeneity among the underlying individuals. Jaramillo et al. (2003) also develop a heterogeneous model, but look at vertical heterogeneity in terms of income inequality. Haimanko et al. (2004) create a generalized model of club formation that
analyzes the differences between competing clubs. Ahn et al. (2008), Polborn (2008),
and Aimone et al. (2013) further this line of work by studying models of endogenous


group formation under varying contexts, and all find multiple clubs as an equilibrium
in the context of religious schism. However, none of the works mentioned above
characterize the timing of competition.

The remainder of the paper is structured as follows. Section 2 outlines a static
version of the model and provides a comprehensive analysis of the underlying tradeoffs. Dynamics are introduced and analyzed in Section 3. In Section 4 I describe two
instruments available to the parent club which can be used to alter the payoff ranking
of the various equilibria: strategic admission and strategic loss. These insights are
then applied in Section 5 to software development, both open source and proprietary.
Section 6 discusses the implications of the paper and concludes.

2 Static Model

There is a community \( \mathcal{N} \) consisting of a finite number of individuals \( N \).4 Each
individual \( i \in \mathcal{N} \) has a “preferred mission” \( A_i \in \mathcal{A} \), where \( \mathcal{A} \) is the set of missions,
or mission space, henceforth referred to as a mission for short.

Assumption 1. There exists a continuous function \( d : \mathcal{A} \times \mathcal{A} \to (0, 1] \), such that for
any \( A, A', A'' \in \mathcal{A} \),

\[
1 - d (A, A') = 0 \iff A = A' \tag{A1}
\]

\[
1 - d (A, A') = 1 - d(A', A) \tag{A2}
\]

\[
1 - d (A, A'') \leq 1 - d (A, A') + 1 - d (A', A'') \tag{A3}
\]

\[
\text{image} (d (\mathcal{A}^2)) = (0, 1]. \tag{A4}
\]

Under assumption\(^4\) specifically \( (A1) - (A3) \), the pair \( (\mathcal{A}, 1 - d) \) forms a metric
space, where \( (A4) \) guarantees that the image of \( d \) spans the entire codomain.\(^5\) The

\(^4\) I use individual as a blanket term to refer entities such as people or organizations.

\(^5\) I also normalize the codomain such that the metric spans the interval \( (0, 1] \), rather than \( \mathbb{R}_+ \cup \{0\} \).

For example, if \( \mathcal{A} \subseteq \mathbb{R}_+^\ell, \ell \geq 1 \), then a suitable metric \( 1 - d \) is \( 1 - \exp (-\|A - A'\|) \), where \( \| \cdot \| \) is the vector norm.
individuals are divided into two types, type 1 and type 2, endowed with missions \( A_m \), \( m = 1, 2 \). Suppose that \( N_1 > N_2 \). I refer to type 1 individuals as majority individuals and type 2 individuals as minority individuals. Let \( \mu \equiv \frac{N_1}{N_1 + N_2} \) denote the distribution of types.

There is a set of clubs \( K \). Each club \( k \in K \) has its own mission \( A^k \) and produces a single, non-congestible club good.\(^6\) The mission of each club is selected by the club’s members through some exogenously given cooperative or noncooperative bargaining process, such as [Nash, Jr.] (1950), [Kalai and Smorodinsky] (1975), [Rubinstein] (1982), [Cho and Duggan] (2009), or [Abreu and Pearce] (2015).\(^7\)

**Definition 1.** \( d(A', A') \) is the match value of an individual/club with mission \( A' \) to a club/individual with mission \( A'' \). A match value of one is referred to as a perfect match.

The cardinality of \( K \) is endogenously determined in a manner to be specified. Each individual \( i \) chooses her affiliation to a single club \( k \) and her contribution level \( x^k_i \geq 0 \). The size of the club is characterized by two measures: the number of members \( N^k \) and new a measure I call the effective number of members.

**Definition 2.** Let \( \hat{N}^k \equiv \sum_{j \in k} d(A^k, A_j) \leq N^k \) denote the effective number of members of club \( k \).

The effective number of members measures club size by aggregating the affiliates and weighting each by her match value.

Each member \( i \) of club \( k \) receives utility from both the aggregate level of contributions to the club \( \sum_{j \in k} x^k_j \) and her own individual contributions to the club. Contributing to a club is costly. Denote the marginal cost of contributing by \( \psi x^k_i \), \( \psi > 0 \). Let \( d(A^k, A_i)b > 0 \) denote the marginal benefit of consumption and let

---

\(^6\)Throughout the paper, superscripts index clubs and subscripts index individuals.

\(^7\)The model developed in [Abreu and Pearce] (2015) is especially useful, as it incorporates an endogenous threat point into a Nash bargaining framework. The endogenous threat point can be treated as the game developed in this paper.
Generically, each individual $i \in k$ receives utility

$$u(x^k_i, x^k_{-i}) = 1\{x^k_i > 0\} d(A^k, A_i) \left( b \sum_{j \in k} x^k_j + \sigma x^k_i \right) - \frac{\psi}{2} (x^k_i)^2,$$

where $1\{\cdot\}$ is a binary indicator function equaling unity if the argument is true and zero otherwise. Contributing $x^k_i = 0$ is akin to nonmembership and generates zero utility. Without loss of generality, I normalize $b = 1$ so that $d(A^k, A_i)\sigma$ is reinterpreted as the relative marginal benefit of contributing. Hence $\sigma = 0$, the typical case studied, corresponds to the marginal benefit of consumption being infinitely more valuable than the marginal benefit of contributing, and $\sigma \rightarrow \infty$ corresponds to the opposite. In addition, when $b = 1$, $\psi x^k_i$ is the relative marginal cost of contributing.

A strategy consists of two components for each individual $i$: a membership decision $i \in k$, and a contribution decision $x^k_i$. In the static analysis, I will be restricting attention to two Nash equilibria. The first is the cohesive equilibrium: a Nash equilibrium in which all individuals belong to a single cohesive club. The second is what I call a splitting (or separating) equilibrium: a Nash equilibrium in which the individuals sort by type into two distinct clubs, denoted by $p$ for parent and $c$ for child. In the splitting equilibrium, I assume that, without loss of generality, all majority individuals join club $p$ while all minority individuals join club $c$.

### 2.1 Static Analysis

Suppose that there is a single cohesive club, endowed with mission $A(\mu)$, which satisfies $d(A(\mu), A_1) > d(A(\mu), A_2)$. That is, the match value is higher for the majority type. This relationship follows from many bargaining models, including those mentioned earlier. For expository convenience, I suppress the affiliation script $k$ as $|K| = 1$. When the community is cohesive, each individual $i$’s objective is given by

$$\max_{x_i} d(A(\mu), A_i) \left( \sum_{j=1}^N x_j + \sigma x_i \right) - \frac{\psi}{2} (x_i)^2,$$

For an interpretation in the context of open source software, see Orman (2008).

8For an interpretation in the context of open source software, see Orman (2008).
It follows that for every $i$, the chosen contribution level is proportional to that individual’s match value:

$$x_i^* = \left(\frac{1 + \sigma}{\psi}\right) d(A(\mu), A_i),$$

(2)

and aggregate contributions are proportional to the effective number of individuals:

$$\sum_{j=1}^{N} x_j^* = \left(\frac{1 + \sigma}{\psi}\right) \hat{N}.$$ 

Thus the indirect utility function for individual $i$ is given by

$$u(x_i^*, x_{-i}^*) = \left(\frac{1 + \sigma}{2\psi}\right) \left[2\hat{N} - (1 - \sigma)d(A(\mu), A_i)\right] d(A(\mu), A_i).$$

(3)

Note that $N$ does not enter the indirect utility function directly, but through the effective number of individuals $\hat{N} = \sum_j d(A(\mu), A_j) = \sum_m d(A(\mu), A_m) N_m$. Furthermore the indirect utility function is increasing and convex in $d(A(\mu), A_i)$.

Taken together, these two findings suggest that increasing the number of members need not increase an individual member’s utility, unless a change in the composition of members does not affect $A(\mu)$. If a change in the composition of members alters $A(\mu)$, then the utility of an individual can actually decrease, even when $N$ increases. For example, suppose that there are 80 majority individuals and 20 minority individuals and that $A(\mu)$ is defined such that $d(A(\mu), A_1) = \mu$ and $d(A(\mu), A_2) = 1 - \mu$. These assumptions imply that $d(A(\mu), A_1) = 0.8$ and $d(A(\mu), A_2) = 0.2$, which yields $\hat{N} = 68$. Now, suppose there are 80 majority individuals and 21 minority individuals, which alters the distribution slightly to $\mu'$, so $d(A(\mu'), A_1) = 0.792$ and $d(A(\mu'), A_2) = 0.208$, which yields $\hat{N} = 67.728$. Therefore $\hat{N}$ decreases and for majority individuals and $d(A(\mu'), A_1) < d(A(\mu), A_1)$, so majority individuals are strictly worse off. Thus when stating that a club benefits from economies of scale, I refer to an increase in utility from a rise in $\hat{N}$, rather than a rise in $N_1$ or $N_2$.

Existence of the cohesive equilibrium can be characterized under many conditions. I focus on two equivalent conditions, one relating to the match value $d(A(\mu), A_i)$ and the other relating to the coefficient $\sigma$ in the relative marginal benefit of contributing. First consider a notion of disagreement. An individual $i$ disagrees with a club $k$ if the
individual and club have different missions. Define a critical level of disagreement by

\[ \bar{d}_i = -d(A(\mu), A_{-i})N_{-i} + \sqrt{[d(A(\mu), A_{-i})N_{-i}]^2 + (1 + \sigma)[2N_i - (1 - \sigma)]} \frac{2N_i - (1 - \sigma)}{2^{N_i - (1 - \sigma)}}. \]  

(4)

I refer to the level of disagreement as having crossed the *club-disagreement threshold* if \( d(A(\mu), A_i) \leq \bar{d}_i \). Next consider the coefficient \( \sigma \). As \( \sigma \) increases, the relative marginal benefit of contributing increases, which implies that individuals are willing to give up economies of scale in exchange for a higher match value, even to the point of being the only member of a club. Denote the cutoff value where the coefficient in the marginal benefit from contributing outweighs economies of scale by

\[ \sigma_i = \frac{2d(A(\mu), A_i)d(A(\mu), A_{-i})N_{-i} + (2N_i - 1)d(A(\mu), A_i)^2 - 1}{1 - d(A(\mu), A_i)^2}. \]  

(5)

Since \( N_1 > N_2 \), it is easier for the minority individuals to cross the thresholds \( \bar{d}_i \) and \( \sigma_i \), as they have a lower match value and thus benefit less from economies of scale than majority individuals. Therefore attention can be restricted to minority individuals.

**Proposition 1.** There exists a cohesive equilibrium if and only if the following two equivalent statements are true: (i) \( d(A(\mu), A_2) \geq \bar{d}_2 \), i.e., the match value is above the club-disagreement threshold; (ii) \( \sigma \leq \sigma_2 \), i.e., the relative marginal benefit from contributing is below the cutoff value.

The proof and all subsequent proofs can be found in Section A of the Supplemental Appendix. Given that the club exhibits economies of scale, if no cohesive equilibrium exists, then a splitting equilibrium must exist. When an individual has the incentive to form a club by herself, she can only be made better off by including more individuals who share her identity (are of the same type). However, this line of reasoning does not complete the analysis. The existence of a cohesive equilibrium need not imply non-existence of a splitting equilibrium. Rather, there may be multiple equilibria, where both the cohesive and splitting equilibria exist under the same set of parameters. For the remainder of the section, I drop the supposition that \( |\mathcal{K}| = 1 \).

If the members of the community select the splitting equilibrium, then each type \( m \) individual belongs to a club consisting of only other type \( m \) individuals. Since all
members of a given club are homogeneous, the bargaining problem is trivial and the club’s mission perfectly matches the mission of its members. The objective of each individual $i \in k$ is

$$\max_{x^k_i} \{ x^k_i > 0 \} \left( \sum_{j \in k} x^k_j + \sigma x^k_i \right) - \frac{\psi}{2} \left( x^k_i \right)^2$$

by choice of $x^k_i$. For every $i$, the chosen contribution level is given by

$$x^k_{i*} = \frac{1 + \sigma}{\psi} \quad (6)$$

if $i \in k$ and zero otherwise, generating indirect utility of

$$u \left( x^k_{i*}, x^k_{-i} \right) = \left( \frac{1 + \sigma}{2\psi} \right) \left[ 2N^k - 1 + \sigma \right].$$

If a type $m$ individual were to unilaterally deviate from the splitting equilibrium, then she does so by joining the competing club. A second notion of disagreement, analogous to club-disagreement, can be introduced: individual-disagreement. Two individuals $i$ and $j$ disagree if they have different missions. When this difference reaches a critical limit, I say the disagreement has crossed the individual-disagreement threshold. Denote this threshold by

$$\hat{d}_{i=m} = \frac{-N_m + \sqrt{N^2_m + (1 + \sigma)[2N_m - (1 - \sigma)]}}{1 + \sigma}. \quad (7)$$

A second cutoff value of $\sigma$ can be defined accordingly. When $\sigma$ is large enough, the losses of economies of scale due to a lack of cohesion are outweighed by the gains in the relative marginal benefit of contributing. Denote the threshold where these effects counterbalance each other by

$$\hat{\sigma}(i,j) = \frac{[2N_j + d(A_i, A_j)]d(A_i, A_j) - 2N_i + 1}{1 - d(A_i, A_j)^2}. \quad (8)$$

If an individual has the incentive to deviate, then she must be leaving the smaller club to join the larger club. This fact implies that attention can again be restricted to a minority individual unilaterally deviating to join the club consisting of majority individuals.

11
Proposition 2. There exists a splitting equilibrium if and only if the following two equivalent statements are true: 

(i) \( d(A_1, A_2) \leq \hat{d}_2 \), i.e., the degree of individual-disagreement crosses the individual-disagreement threshold; 

(ii) \( \sigma \geq \hat{\sigma}(2, 1) \), i.e., the relative marginal benefit of contributing is above the cutoff value.

Careful inspection of Propositions 1 and 2 support the claim of multiple equilibria. Multiple equilibria exist if \( \sigma \in [\hat{\sigma}(2, 1), \sigma(2, 1)] \), or equivalently, if \( d(A(\mu), A_2) \) is not too small while \( d(A_1, A_2) \) is not too large. For example, set \( \sigma = 0 \) and note that \( \hat{\sigma}(2, 1) \leq 0 \) if

\[
d(A_1, A_2)N_1 - N_2 \leq -\frac{1}{2} (1 + d(A_1, A_2)^2)
\]

The right-hand side is bounded between \(-1\) and \(-\frac{1}{2}\), so as \( d(A_1, A_2) \to 0 \), the inequality is necessarily satisfied. \( \sigma(2, 1) \geq 0 \) if

\[
\hat{N} \geq \frac{1}{2} (1 + d(A(\mu), A_2)).
\]

The right-hand side is bounded between \( \frac{1}{2} \) and 1, so for \( d(A(\mu), A_1) \) or \( d(A(\mu), A_2) \) large enough, the left-hand side is weakly larger than the right-hand side as \( N_1 + N_2 > 1 \).

The interesting aspect of multiple equilibria in this context is how the ranking of the payoffs in the two equilibria depend on the match values as well as the distribution of majority and minority individuals. There are conditions under which a minority individual “prefers” the splitting equilibrium over the cohesive equilibrium in terms of utility, and \textit{vice versa}. Similarly, there exist conditions under which a majority individual prefers one equilibrium over the other. The conditions for each type of individual to prefer a specific equilibrium need not coincide, and depend explicitly on the members’ match values. A third cutoff value can be defined representing the point at which a minority/majority individual’s match value is such that the individual is indifferent between the two equilibria. Denote this value for a type \( m \) individual by

\[
\tilde{d}_m(\mu) = \frac{-d(A(\mu), A_{-m})N_{-m} + \sqrt{[d(A(\mu), A_{-m})N_{-m}]^2 + [2N_m - (1 - \sigma)]^2}}{2N_m - (1 - \sigma)}.
\]
Proposition 3. The splitting equilibrium payoff dominates the cohesive equilibrium for a type \( m \) individual if and only if \( d(A(\mu), A_m) \leq \tilde{d}_m(\mu) \), or equivalently, the match value of a type \( m \) individual to the cohesive club is sufficiently small.

If \( d(A(\mu), A_2) \leq \tilde{d}_2(\mu) \) and \( d(A(\mu), A_1) \geq \tilde{d}_1(\mu) \) (or the reverse), then majority preferences and minority preferences are misaligned. To illustrate that preference misalignment is in fact a possibility, note that \( \frac{\partial \tilde{d}_m(\mu)}{\partial \sigma} > 0 \) and \( \tilde{d}_1(\mu) > \tilde{d}_2(\mu) \). Therefore up to three regions of \( \sigma \) exist: one in which preferences are aligned and all individuals prefer the cohesive equilibrium, one in which preferences are aligned and all individuals prefer the splitting equilibrium, and one in which preferences are misaligned. The result in which preferences are misaligned and majority individuals prefer the cohesive equilibrium while minority individuals prefer the splitting equilibrium is illustrated in Figure 1(a). Note that majority individuals preferring the splitting equilibrium and minority individuals preferring the cohesive equilibrium is also a possibility\(^9\). This result is illustrated in Figure 1(b).

Propositions 1-3 illustrate a fundamental tension: the underlying tradeoff between match value and economies of scale. A cohesive club will often possess the greatest economies of scale since all individuals are members and total provision is strictly increasing in the number of members, \textit{ceteris paribus}; however, the value of economies of scale is contingent upon the match value. A cohesive equilibrium implies hetero-

\(^9\)For instance, when \( N_1 \gg N_2 \), so the benefits from an increase in club size are relatively small compared to the losses from a decrease in the match value.
geneity among members.

Imagine that the feasible mission space is $A \equiv [0, 1] \times [0, 1]$, and suppose that $A_1 = (1, 0)$ and $A_2 = (0, 1)$. For each distribution $\mu$, any efficient bargaining process should lead to $A(\mu) = (y, 1 - y)$, for some $y \in [0, 1]$. Therefore the match value is strictly less than one for at least one of the types. By segmenting the community according to type, those individuals with the lower match value necessarily find that the segmentation increases the value of contributions on both the extensive and intensive margins. Therefore total utility may in fact increase by sacrificing economies of scale in exchange for a higher match value. I call this tradeoff the contribution-scale tension.

**Tension 1. (The Contribution-Scale Tension)** Under heterogeneity, there exists a tradeoff between economies of scale and match value. The economies of scale granted by a cohesive community necessarily decreases the match value for a non-degenerate subset of the community.

In summary, there are two interesting equilibria in the static case: the cohesive equilibrium and the splitting equilibrium. There are conditions under which: (i) only a cohesive equilibrium exists; (ii) only a splitting equilibrium exists; (iii) both equilibria coexist. Conditional on both equilibria coexisting, there are conditions under which: (a) all individuals prefer the cohesive equilibrium; (b) all individuals prefer the splitting equilibrium; (c) There is preference misalignment across types, where either majority individuals receive a greater payoff in the cohesive equilibrium than in the splitting equilibrium, while minority individuals receive a greater payoff in the splitting equilibrium than in the cohesive equilibrium, or *vice versa*. In the dynamic extension of the model, to be described shortly, the tradeoffs described in the static setting all carry over. In addition, a second tension is introduced, which gives rise to a new equilibrium outcome of incubation/strategic membership.

### 3 Club Dynamics

In this section, I introduce a dynamic extension that relates directly to the static framework presented in the previous section and illustrates the underlying incentives for incubation and strategic membership. It is not sufficient to focus on the steady
state behavior. Rather, as I shall show, important insights are lost unless one analyzes the entire dynamic path. It turns out that analytical methods are not able to uncover all of the results with respect to incubation. When analytical results are not directly available, numerical methods are employed.

In the Supplemental Appendix, I consider a second dynamic extension that restricts contributions from $\mathbb{R}_+$ to the binary space $\{0, 1\}$ and assumes complete myopia. These two restrictions allow for analytical solutions and more importantly lead to new results. A second form of incubation/strategic membership can be practiced by individuals called passive strategic membership, in which minority individuals do not contribute to any club at the outset. Instead, at some later time, they take the accumulated contributions of the parent club, consisting solely of majority individuals, and uses it to form a new competing child club. It is the discreteness of contributions that drives this new result.

3.1 Dynamic Model

Suppose that time, indexed by $t$, flows continuously with common discount rate $\delta > 0$. Contributions accumulate over time, but depreciate at the rate $\gamma \in (0, 1)$. At any time $t$, the change in accumulated contributions $C^k(t)$ to club $k$, henceforth referred to as installed base, is given by the dynamic

$$\dot{C}^k(t) = \sum_{j(t) \in k} x^k_j(t) - \gamma C^k(t),$$

where $x^k_i(t) \geq 0$ is now interpreted as individual $i$'s contribution to club $k$ at time $t$ and $\dot{C}^k(t) = \frac{dC^k(t)}{dt}$. Individual $i$ contributing to club $k$ at time $t$ is denoted by $i(t) \in k$. The instantaneous utility function is analogous to (1):

$$u \left( x^k_i(t); C^k(t) \right) = 1 \left\{ x^k_i(t) > 0 \right\} d(A^k, A_i) \left( C^k(t) + \sigma x^k_i(t) \right) - \frac{\psi}{2} x^k_i(t)^2.$$

The cohesive and splitting equilibria can still be characterized, as in the static case. When time and capital accumulation are considered, a third interesting equilibrium emerges. Rather than either remaining a cohesive community indefinitely or separating by type at the outset, individuals can incubate a child club within the parent club, leading to what I call an incubation equilibrium. In the incubation equi-
librium, the community begins as a cohesive unit and remains so for a finite period of time, then splitting into multiple clubs. I refer to the period of time in which the community is cohesive as the *incubation period*, and denote the underlying strategy as active strategic membership, where a coalition of individuals temporarily joins a club in order to take advantage of economies of scale, only to then split once the installed base has grown to a sufficiently large level.

In order for active strategic membership to be a viable strategy, individuals must not leave empty-handed. There exists a wide range of interactions under which the installed base is non-rival across clubs. Take software development, where contributions are an amalgamation of source code and expertise. While an individual’s expertise may be rival across clubs, the source code is non-rival\footnote{If individual $i$ is an expert at some process $z$, then if that individual leaves the club, the club no longer has access to the individual with expertise in $z$.}. One’s use of the source code does not preclude another’s use. Let $r \in [0, 1]$ denote the degree of rivalry of the installed base, where $r = 1$ signifies completely rival installed base and $r = 0$ signifies completely non-rival installed base. I assume that $r$ is exogenous and fixed.

There is also a potentially endogenous measure concerning the sharing of installed base: excludability.

Installed base may be excludable across clubs through the use of intellectual property restrictions such as patents, trademarks, licenses, and contracts (e.g., non-compete clauses). For instance, a software development team could choose a typical copyright license, preventing any who leave from retaining any of the source code, or the team could select what is known as a copyleft license, such as the GNU General Purpose License, the Apache License, or the various BSD licenses, which allow individuals who leave to take the source code and alter it to suit their own purposes (Open Source Initiative, 2015\textsuperscript{a,c,b}). Related to excludability is distance-based depreciation.

As the child club moves further away from the parent club, the installed base may not be as relevant, as the match value of the clubs’ missions is poor. Both excludability and distance-based depreciation can be captured by the same measure. Let $1 - \rho(0)$, $\rho(0) \in [0, 1]$, denote the degree of excludability. The proportion of installed base retained by the child club depends on the match value of the two clubs $d(A^p, A^c)$ and $\rho(0)$. The retained proportion is denoted by the function $\rho \left(1 - d(A^p, A^c)\right) \in [0, 1]$. As the distance between the two clubs $1 - d(A^p, A^c)$ increases, the retained proportion of installed base $\rho \left(1 - d(A^p, A^c)\right)$ decreases. For now, suppose that $\rho(0)$ is exogenous.
Thus when a split occurs at time $T^*$, those who leave to form the child club take
\[ \rho \left( 1 - d(A_p, A_c) \right) (1 - \gamma) C(T^*), \]
while the parent club retains $[1 - \rho(0)](1 - \gamma) C(T^*)$, where $C(T^*)$ is the total installed base accumulated during the incubation period\[11\]

3.2 Dynamic Analysis

The model described in section 3.1 allows the interactions to be represented as a linear-quadratic differential game. In the analysis to follow, I will be considering Markov perfect equilibria. In general, there are many, if not infinite Markov perfect equilibria in differential games. Therefore I restrict attention to the solutions found using Pontryagin’s maximum principle and show that these solutions satisfy both the Markov property and subgame perfection. Each individual $i$’s strategy is a pair consisting of a membership decision for every time $t$, and a contribution schedule $x^k_i(t)$ for every time $t$. The equilibrium requires that at almost every $t \in [0, \infty)$ and every possible state $C^k(t)$, each individual’s contribution and affiliation decision satisfies the Nash property and that contributions satisfy the Markov property\[12\]. Derivations of each individual’s equilibrium contribution schedule and the equilibrium state $C^k(t)$ can be found in Supplemental Appendix B.

Suppose that there is a single club with objective $A(\mu)$, which is defined such that
\[ d(A(\mu), A_1) > d(A(\mu), A_2) \]. Each individual’s objective is
\[
\max_{(x_i(t))_t} \int_{t_0}^{\infty} e^{-\delta t} \left( d(A(\mu), A_i) [C(t) + \sigma x_i(t)] - \frac{\psi}{2} x_i(t)^2 \right) dt\[13\]

\[11\] A second cost of splitting may arise along the extensive margin, representing the social cost of splitting. The social cost can be thought of as the price to pay if splitting is considered “taboo,” e.g. a stigma, or alternatively as the price of either forming a new mission or altering an existing mission. For example, religious schisms are socially costly, as the new mission must be justified, while the old mission must be criticized. Furthermore, social ostracization may occur. In OSS development, the act of forking is considered taboo and must be strongly justified (Stewart and Gosain, 2006). This cost is decreasing in the number of splits, à la Thomas Schelling’s tipping point model (Schelling, 1978), or alternatively, Timur Kuran’s preference falsification model (Kuran, 1987, 1995). Once the first split is justified, future splits become easier to justify. Let $S$ represent the number of splits that have already occurred. The social cost of splitting can be written as $f(S)$, where $f(0)$ is the social cost of the first split, and $f(0) \geq f(1) \geq \ldots$. Since only a single split is considered in this paper, $f(0)$ can be normalized to zero without any loss of generality; however, future work involving multiple splits should explicitly consider this cost.

\[12\] “At almost every $t[\ldots]$” means that the contributions must maximize the continuation value at all $t$ except possibly at a finite, measure zero set of times $t$. 

17
for all $s \in [0, \infty)$, subject to the dynamic

$$\dot{C}(t) = \sum_j x_j(t) - \gamma C(t).$$

When the community is cohesive, each individual $i$ chooses contribution level

$$x_i(t) = \frac{1 + \sigma(\delta + \gamma)}{\psi(\delta + \gamma)} d(A(\mu), A_i)$$

at every time $t$, which as in the static case, is proportional to the match value. The installed base evolves along the continuous, strictly concave path

$$C(t) = \frac{1 + \sigma(\delta + \gamma)}{\gamma \psi(\delta + \gamma)} \hat{N} \left(1 - e^{-\gamma t}\right).$$

Since $\gamma > 0$, $C(t)$ converges to a steady-state value of $\frac{1 + \sigma(\delta + \gamma)}{\gamma \psi(\delta + \gamma)} \hat{N}$ in the long run.

Substituting the individual contribution schedule and the installed base path into the utility function at $t = 0$ yields indirect utility of

$$V(d(A(\mu), A_i)) = \frac{(1 + \sigma(\delta + \gamma)) \left(2 \hat{N} + d(A(\mu), A_i)(\sigma(\delta + \gamma) - 1)\right)}{2\delta \psi(\delta + \gamma)^2} d(A(\mu), A_i).$$

Observe the similarity between (12) and (3). $\hat{N}$ enters each indirect utility function in an identical manner, as does the match value $d(A(\mu), A_i)$. Consequently, cutoff values exist akin to those discussed in section 2.1, which characterize existence isomorphically to the static model. A formal dynamic extension of this result is omitted, as no new insights are uncovered.

Next suppose that a split is orchestrated at time $T^* = 0$, where all majority individuals join club $p$ and all minority individuals join club $c$. Each club sets its mission to match that of its members, so $d(A^k, A_i) = 1$ for all $i(t) \in k, k = p, c$. For every $i \in \mathcal{N}$,

$$x_i^k(t) = \frac{1 + \sigma(\delta + \gamma)}{\psi(\delta + \gamma)}.$$ 

13Technically, the objective function is $\int_s^\infty e^{-\delta(t-s)} \left(d(A(\mu), A_i)[C(t) + \sigma x_i(t)] - \frac{\psi}{2} x_i(t)^2 \right) dt$. However, since $e^{-\delta(t-s)} = e^{-\delta t} e^{\delta s} \propto e^{-\delta t}$, $e^{\delta s}$ can be omitted without affecting the optimization. I utilize this expositional simplification for all dynamic optimizations in the text.

14This result is available upon request.
for \( i(t) \in k \) and zero otherwise. Club \( k \)'s installed base evolves along the path

\[
C_k(t) = \frac{1 + \sigma(\delta + \gamma)}{\gamma \psi(\delta + \gamma)} N^k \left(1 - e^{-\gamma t}\right).
\]

As in the cohesive case, there is long run convergence to a steady-state level of

\[
\frac{1 + \sigma(\delta + \gamma)}{\gamma \psi(\delta + \gamma)} N^k.
\]

Conditions isomorphic to those discussed in the static analysis can be described to characterize the existence of a splitting equilibrium, the occurrence of multiple equilibria, and the payoff ranking of the two equilibria. The same tradeoff between economies of scale and match value exists in both the static and dynamic structures. In the cohesive setting, each individual contributes less, but there are more individuals to make up for the smaller contribution level. Total contributions in the cohesive equilibrium are greater than total contributions in the splitting equilibrium at every time \( t \) if \( \hat{N} > N^k \). Recall that \( N_1 > N_2 \), \( N^p = N_1 \), and \( N^c = N_2 \), so there exist conditions on \( d(A(\mu), A_1) \) and \( d(A(\mu), A_2) \) such that either: 1. \( \hat{N} < N_2 < N_1 \); 2. \( N_2 < \hat{N} < N_1 \); or 3. \( N_2 < N_1 < \hat{N} \).

These three conditions illustrate why a minority individual may have the incentive to practice active strategic membership. When \( N_2 < \hat{N} \), the minority individuals benefit from economies of scale. Figure 2 illustrates two possible dynamics under active strategic membership. When assessing the decision to incubate, the distance-based depreciation becomes relevant. As the distance between the parent club and child club \( 1 - d(A(\mu), A^c) \) increases, \( \rho(1 - d(A(\mu), A^c)) \) decreases, which implies that the child club may not set \( A^c \) to perfectly match its members’ missions. The minority individuals may be better off locating in an intermediate position, i.e. \( d(A(\mu), A_2) < d(A^c, A_2) \leq d(A_2, A_2) = 1 \). This relationship illustrates a second tension, which I call the mission-scope tension.

**Tension 2. (The Mission-Scope Tension)** *When splitting, locating the child club’s mission closer to the members’ missions than the parent club’s mission increases the value of the child club to its members, but comes at a cost of increasing the losses due to splitting.*

Line \( b_1 \) illustrates one possible installed base path under the incubation equilibrium, which then falls to \( b_2 \) post-split due to the distance-based depreciation. Lines \( c_1 \) and
Figure 2: Incentives for Active Strategic Membership

c_2 illustrate a second possible incubation equilibrium path. \( \mathcal{C}^k \) represent the steady-state installed base levels for \( k = p, c \), conditional on incubation. Line \( a \) necessarily converges to a level above \( \mathcal{C}^c \).

3.2.1 Active Strategic Membership

Suppose that a split occurs at some finite time \( T^* > 0 \), so the incubation period is given by the non-degenerate interval \([0, T^*]\). \( C(T^*) \) is the installed base of the parent club at time \( T^* \). Consider the subgames initiated post-split at time \( T^* \), where all majority individuals are in club \( p \) and all minority individuals are in club \( c \). There are an infinite number of subgames, one for each pair \((C(T^*), T^*)\). The minority individuals assign mission \( A^c \) to the child club, while the parent club keeps the original mission \( A(\mu) \).

The child club retains \((1 - \gamma)\rho(1 - d(A(\mu), A^c))C(T^*)\) from the parent club, while the parent club retains \([1 - r\rho(0)]C(T^*)\) post-split. For expositional convenience, let

\[
\alpha^k = \begin{cases} 
1 - r\rho(0) & \text{if } k = p \\
\rho(1 - d(A(\mu), A^c)) & \text{if } k = c
\end{cases}
\]

\[\text{If the parent club were to change its mission, then the utility from incubation must be greater than it would be had the parent club retained the original mission } A(\mu). \text{ Therefore assuming that the parent club keeps the original mission represents the most conservative case. Alternatively, one can think of inertia as the reason the mission remains the same. For instance, after a religious schism, it is rare that the original religion changes its practices.} \]
represent the proportion of installed base retained by club $k$ post-split. Each individual $i(t) \in k$ contributes an amount proportional to her match value

$$x_i^k(t > T^*) = \frac{1 + \sigma(\delta + \gamma)}{\psi(\delta + \gamma)} d(A^k, A_i), \quad (14)$$

where the installed base accumulates according to

$$C^k(t > T^*) = \frac{1 + \sigma(\delta + \gamma)}{\gamma \psi(\delta + \gamma)} \hat{N}^k + \left(1 - \gamma\right) \alpha^k C(T^*) - \frac{1 + \sigma(\delta + \gamma)}{\gamma \psi(\delta + \gamma)} \hat{N}^k \right] e^{-\gamma(t-T^*)}.$$

Study of the post-split problem is sufficient in determining the long run implications.

Post-split, each minority individual contributes more than she would in the cohesive equilibrium, while each majority individual’s contributions are unaffected (compare equations (14) and (11)). However, each individual contributes no more than she would in a splitting equilibrium. When incubating, the contributions made by minority individuals are proportional to the contributions in the cohesive equilibrium. That is,

$$x_i^c(t > T^*) = \frac{d(A^c, A_2)}{d(A(\mu), A_2)} x_2(t > T^*),$$

so the difference is determined by the relative match value, which is greater than one. Depending on the relationship between $\hat{N}$ and $\hat{N}^c$, the steady-state level of installed base can be either higher or lower under incubation relative to cohesion. If $\hat{N}^c = d(A^c, A_2)N_2 > \hat{N} = d(A(\mu), A_1)N_1 + d(A(\mu), A_2)N_2$, then the steady-state installed base is greater under incubation. Even if the steady-state level of installed base is lower when incubating, minority individuals can still benefit since $d(A^c, A_2) > d(A(\mu), A_2)$, so both the installed base and individual contributions are more valuable at any given level. The decrease in contributions for minority individuals between the incubation equilibrium and the splitting equilibrium follows from the fact that in the splitting equilibrium, the minority individuals set the child club’s mission such that the match value is one. When incubating, the members of the child club may set the child club’s mission such that the match value is less than one due to distance-based depreciation, as illustrated by the mission-scope tension. Thus in the long run, the splitting equilibrium induces a greater marginal benefit of both consumption and contribution, higher contributions from members, and a greater steady-state level
of installed base. Since future utility is discounted, utility today is more valuable than utility in the future, so individuals are willing to choose actions which increase current utility, but decrease future utility. The benefits of incubation operate under this principle.

**Proposition 4.** Suppose individuals become infinitely patient ($\delta \to 0$) and the cohesive equilibrium, splitting equilibrium, and incubation equilibrium all coexist. The splitting equilibrium always payoff dominates the incubation equilibrium. In addition for each type $m$ individual, $m = 1, 2$: (i) the incubation equilibrium payoff dominates the cohesive equilibrium if and only if $d(A(\mu), A_m)$ is sufficiently small.; (ii) the cohesive equilibrium payoff dominates the splitting equilibrium if and only if $d(A(\mu), A_m)$ is sufficiently large.

In the long run, the incubation equilibrium is always payoff dominated by the splitting equilibrium. This fact does not preclude an incubation equilibrium from existing. While practicing active strategic membership may not be the best option when individuals are sufficiently patient, it is still an option nonetheless.

**Proposition 5. (Active Strategic Membership)** Suppose individuals become infinitely patient. If $d(A^c, A_2)$ and $d(A(\mu), A_1)$ are sufficiently large and $d(A(\mu), A_2)$ is sufficiently small, then there exists an incubation equilibrium in which minority individuals practice active strategic membership.

If the match value of a minority individual to the child club is sufficiently large, then the difference between the indirect utility under incubation equilibrium post-split and the indirect utility under the splitting equilibrium is near zero, so the benefits of economies of scale coupled with the relatively high value of the incubation equilibrium implies, by the contribution-scale tension, that there is no incentive for a minority individual to unilaterally deviate and form a club $k$ with mission $A^k = A_2$. If the match value between a minority individual and the parent club is sufficiently small, then the value of the cohesive equilibrium to the minority individual is low, so no minority individual has the incentive to unilaterally deviate by not splitting at time $T^*$ and instead remaining with the parent club. If the match value between a majority individual and the parent club is sufficiently large, then the benefits from economies
of scale for majority individuals are large enough such that no majority individual wishes to deviate and form a new club \( k' \) with mission \( A^{k'} = A_1 \).

The implications of Proposition 4 only hold in the long run, as the results are valid only if individuals are sufficiently patient. However, the true benefits of incubation are of a short run nature, and are most prevalent when individuals are not too patient. When practicing active strategic membership, individuals are transferring utility from the future to the present. The splitting equilibrium always yields a greater payoff in the steady-state than the incubation equilibrium, but the incubation equilibrium can be more profitable in the short run. The short run incentives can be strong enough that the incubation equilibrium can payoff dominate the splitting equilibrium.

Consider the pre-split contributions by each individual. Each individual \( i \) chooses her pre-split contribution level, knowing that a split is imminent and given a split, she will contribute according to (14) post-split. Thus the pre-split problem for individual \( i \) is given by the family of finite-horizon problems

\[
\max_{(x_i(t))_{t=s}^{T^*}} \int_s^{T^*} \left\{ e^{-\delta t} \left[ d(A(\mu), A_i) (C(t) + \sigma x_i(t)) - \frac{\psi}{2} x_i(t)^2 \right] \right\} dt + B_i(C(T^*)),
\]

for all \( s \in [0, T^*] \), where

\[
B_i(C(T^*)) = \int_{T^*}^{\infty} \left\{ e^{-\delta t} \left[ d(A^k, A_i) \times \left( \frac{1 + \sigma(\delta + \gamma)}{\gamma \psi(\delta + \gamma)} \hat{N}^k + \left(1 - \gamma\right) \alpha^k C(T^*) - \frac{1 + \sigma(\delta + \gamma)}{\gamma \psi(\delta + \gamma)} \hat{N}^k \right] e^{-\gamma(t-T^*)} 
+ \sigma \frac{1 + \sigma(\delta + \gamma)}{\psi(\delta + \gamma)} d(A^k, A_i) - \frac{\psi}{2} \left( \frac{1 + \sigma(\delta + \gamma)}{\psi(\delta + \gamma)} d(A^k, A_i) \right)^2 \right\} dt.
\]

From this system, each individual \( i \) contributes

\[
x_i(t \leq T^*) = \bar{x}_i + \left[ (1 - \gamma) \alpha^k d(A^k, A_i) - d(A(\mu), A_i) \right] e^{-(\delta+\gamma)(T^*-t)},
\]

where

\[
\bar{x}_i = \frac{1 + \sigma(\delta + \gamma)}{\psi(\delta + \gamma)} d(A(\mu), A_i).
\]
The pre-split installed base path is given by

\[ C(t \leq T^*) = \omega_1 \left( 1 - e^{-\gamma t} \right) + \omega_2 e^{-(\delta + \gamma)T^*} \left( e^{(\delta + \gamma)t} - e^{-\gamma t} \right), \]

where

\[ \omega_1 = \frac{1 + \sigma(\delta + \gamma)}{\gamma \psi(\delta + \gamma)} \hat{N}, \quad \omega_2 = \left( \frac{\gamma}{\delta + 2\gamma} \right) \left[ \frac{(1 - \gamma) \sum_k \alpha_k \hat{N}_k - \hat{N}}{\gamma \psi(\delta + \gamma)} \right]. \]

Note that \( \omega_1 \left( 1 - e^{-\gamma t} \right) \) is the installed base path under the cohesive equilibrium. While the pre-split contributions depend explicitly on \( T^* \), the equilibrium choice of \( T^* \) is determined only by the post-split utility and the state at time \( T^*, C(T^*) \).

At \( t = 0 \), a forward looking individual chooses \( T^* \) to maximize the total stream of utility. However, at each instant, each individual reevaluates her decision up until time \( T^* \), the point at which the split occurs. Anticipating this reaction, each individual contributes knowing that the split will occur at time \( T^* \), where \( T^* \) maximizes the discounted post-split stream of utility:

\[ T^* \equiv \arg \max \left\{ B_{i=m=2} \left( \omega_1 \left( 1 - e^{-\gamma T^*} \right) + \omega_2 \left( 1 - e^{(\delta + 2\gamma)T^*} \right) \right) \right\}. \]

This value can only be computed numerically.

Careful analysis of the pre-split contributions yields insight into how the short run benefits of incubation can lead to payoff dominance of the incubation equilibrium. First compare (15) to (11). The pre-split contributions in the incubation equilibrium are equal to the pre-split contributions in the cohesive equilibrium, plus an extra term

\[ \left[ \frac{(1 - \gamma) \alpha_k d(A_k, A_i) - d(A(\mu), A_i)}{\psi(\delta + \gamma)} \right] e^{-(\delta + \gamma)(T^* - t)}. \]

When the above term is positive, each individual contributes more under the incubation equilibrium than under the cohesive equilibrium. The term is positive if

\[ \alpha_k > \frac{d(A(\mu), A_i)}{(1 - \gamma)d(A_k, A_i)}. \]

Recall that \( \alpha_k \) represents the proportion of capital that each club \( k \) retains. If this value is sufficiently large, then anticipating the increased future benefits post-split, each
individual contributes more, pre-split. If \( \alpha^k \) is large relative to the value given in (16), then contributions under the incubation equilibrium are greater than contributions under the splitting equilibrium. This value may or may not be attainable, depending on other values, such as \( N_1 \) and \( N_2 \).

A similar comparison can be made with respect to the installed base path. The path under the incubation equilibrium is equal to the path under the cohesive equilibrium, plus an extra term, and is greater under the incubation equilibrium if \( \alpha^k \) is sufficiently large for \( k = 1, 2 \). The extra term is positive if

\[
\sum_k \alpha^k \hat{N}^k > \frac{\hat{N}}{1 - \gamma}.
\]

(17)

It is clear that (16) and (17) are positively related. Furthermore, this increase is approximately hyperbolic for small \( \delta \), as it includes the term \( e^{\delta t} - e^{-\gamma t} \), so when the term is positive and \( \delta \) is small, the increase can be quite large. Similar conditions exist for the same relationship to hold between the incubation equilibrium and the splitting equilibrium. The above argument illustrates the intuition behind the notion that the incubation equilibrium can payoff dominate both the cohesive equilibrium and the splitting equilibrium.

Unfortunately, the indirect utility function is complicated to the point that an analytical analysis of payoff dominance is infeasible. In order to illustrate the results more formally, I employ numerical methods. In the numerical analysis, I manipulate between one and three parameters at a time. In each round of the analysis, the discount rate \( \delta \) changes. Along with \( \delta \), I manipulate either \( \alpha^p \) and \( \alpha^c \), \( \gamma \), \( d(A(\mu), A^c) \) and \( d(A^c, A_2) \), or \( \sigma \). The rounds with manipulations to \( (\delta, \alpha^p, \alpha^c) \) and \( (\delta, d(A(\mu), A^c), d(A^c, A_2)) \) have three parameters changing simultaneously while the remaining rounds have two parameters changing simultaneously. I normalize \( \psi = 10 \) and keep this parameter fixed throughout. Note that \( \psi \) enters the individual contributions in each of the three equilibria, given by (11), (13), (14), and (15), proportionally, so without loss of generality, \( \psi \) can be kept fixed without affecting the rankings of the equilibrium payoffs.

**Result** There exists an open set of conditions such that the incubation equilibrium payoff dominates both the splitting equilibrium and the cohesive equilibrium.
Parameters | Intuition
---|---
$\delta \gg 0$ | Individuals exhibit a degree of impatience, so they are willing to trade future benefits for current benefits.
$\alpha^p, \alpha^c \gg 0$ | When both $\alpha^p, \alpha^c$ are sufficiently large, both clubs are able to accumulate installed base via incubation.
$\gamma \ll 1$ | Installed base can only accumulate rapidly when the depreciation rate is not too large.
$d(A(\mu), A^c) \ll d(A^c, A_2)$ | If the match values are sufficiently large, then the post-split value of incubation is similar to the value of splitting.
$g(\delta) \leq \sigma \ll \infty$ | If the relative benefit of contributing is too large, then the value of consuming the club good is irrelevant.

**Table 1: Sufficient conditions for the incubation equilibrium to exhibit payoff dominance for minority individuals.**

Table 1 summarizes the conditions under which the incubation equilibrium payoff dominates both the cohesive equilibrium and the splitting equilibrium for the minority individuals. Every condition outlined must be satisfied simultaneously. The exact minimal/maximal value is dependent on the other parameters. Detailing the exact conditions is infeasible since there are an uncountably infinite number of values; for example, there is a unique $\delta$ for each $\sigma$ such that incubation payoff dominates both of the other equilibria, *ceteris paribus*. To illustrate the result, I present some numerical results and comparative statics for a subset of the parameters. Details on the numerical analysis can be found in Supplemental Appendix C.

As shown in Proposition 4, when $\delta$ is sufficiently small, the incubation equilibrium is always payoff dominated, but if individuals are not too patient, then they are willing to give up utility in the future in exchange for utility in the present. It only takes a small degree of impatience to generate payoff dominance for the incubation equilibrium. As there is an approximately hyperbolic component to contributions for a small $\delta$, the utility from incubating increases significantly. The effect is non-monotonic and diminishing as $\delta$ increases (Figure 3(a)). Post-split individual contributions are decreasing in the discount rate, which in turn diminishes the incentive to incubate. Thus as the discount rate increases, the equilibrium incubation period $T^*$ decreases,
Figure 3: Properties of the incubation equilibrium as a function of $\delta$.

Figure 4: Incubation equilibrium utility $-\max\{\text{cohesive equilibrium utility, splitting equilibrium utility}\}$, with respect to $\alpha^p$ and $\alpha^c$ as illustrated in Figure 3(b).

For the remaining numerical results, I restrict $\delta \in \{0.004, 0.05\}$, which corresponds to discount factors of approximately 0.996 and 0.95, respectively.\footnote{The discount factor is given by $e^{-\delta} \approx 1 - \delta$ for small $\delta$.} When $1 - r\rho(0)$ and $\rho(1 - d(A(\mu), A^c))$ are both sufficiently large, both majority and minority individuals are willing to contribute enough to make incubation worthwhile since both the child club and parent club retain a large proportion of the installed base. Figure 4 depicts the region in $(\alpha^p, \alpha^c)$ space under which the incubation equilibrium payoff dominates both the cohesive equilibrium and the splitting equilibrium. The dark-gray region corresponds to the set of parameters under which the incubation equilibrium is payoff dominated while the light-gray region corresponds to the set of parameters in which the incubation equilibrium is payoff dominant. Note that in order for such a set of parameters to exist, $\delta$ must be sufficiently large, as illustrated by the rela-
relationship between Figures 4(a) and 4(b). Conditional on an incubation equilibrium existing, the length of the incubation period is increasing in both $\rho(1-d(A(\mu), A^c))$ and $1-r\rho(0)$ (Figure 5). As $1-r\rho(0)$ increases, the amount a majority individual is willing to contribute pre-split increases. At the same time, when $\rho(1-d(A(\mu), A^c))$ increases, the value of incubation increases as minority individuals have the incentive to contribute more. Once $1-\rho(0)$ and $\rho(1-d(A(\mu), A^c))$ are large enough, the incubation period is a positive measure of time. As these two values increase, pre-split contributions by both type of individuals increase, which makes incubation more profitable, increasing the equilibrium incubation period.

With $\gamma$, there are no surprises. When there is very little depreciation, installed base is able to accumulate at a rapid rate, increasing the value of economies of scale to the members making incubation a valuable strategy. Figures 6(a) and 6(b) show that the more patient individuals are, the smaller $\gamma$ must be to compensate and make incubation payoff dominant. Analysis of Figures 7(a) and 7(b) show that the equi-
librium incubation period decreases as depreciation increases, since larger incubation periods become less valuable when installed base is unable to accumulate at a rapid pace.

The analysis of the match values $d(A(\mu), A_2)$ and $d(A^c, A_2)$ is conducted over a restricted parameter space. There are two conditions which must be satisfied. Firstly, the match value of a minority individual to the parent club must always be weakly lower than the match value of a majority individual to the parent club. Secondly, the match value of a minority individual to the child club must be strictly greater than the match value of a minority individual to the parent club, otherwise by the contribution-scale tension, the benefit from economies of scale dominates the benefit from a higher match value, since there is none.

As $d(A^c, A_2)$ increases, incubating becomes more valuable. However, when holding $d(A^c, A_2)$ fixed, the value of incubation decreases when $d(A(\mu), A_2)$ increases by the contribution-scale tension. For every $d(A(\mu), A_2)$, there is a unique $d(A^c, A_2)$ such that the incubation equilibrium is payoff dominant for the minority individuals, and this value is increasing in $d(A(\mu), A_2)$, as shown in Figures 8(a) and 8(b). The equilibrium incubation period is decreasing in $d(A^c, A_2)$ since the amount minority individuals are willing to contribute both pre-split and post split is increasing in $d(A^c, A_2)$. The opposite is true when $d(A(\mu), A_2)$ increases, as post-split contributions are constant, but pre-split contributions are increasing (see Figures 9(a) and 9(b)).

The relationship between $\sigma$ and payoff dominance is non-monotonic. When $\delta$ and $\sigma$ are both small, incubation is payoff dominated solely due to the patience of minority individuals. If the marginal benefit of contributing increases while $\delta$ remains fixed,
then the relative payoff when incubating actually decreases due to over-reliance on the installed base relative to individual contributions, so the cohesive equilibrium is
the most valuable. If $\sigma$ continues to increase, then the balance shifts and while there is still reliance on the total contributions, the greater future match value under the incubation equilibrium relative to the cohesive equilibrium leads to an increase in the relative payoff under the incubation equilibrium. As $\sigma$ increases, not only does the incubation equilibrium payoff, but the difference between the incubation equilibrium payoff and the payoffs from both the cohesive equilibrium and the splitting equilibrium. However, there exists a cutoff point where eventually, by the contribution-scale tension, the match value effects take over and the relative payoff from incubating is decreasing, to the point where it is no longer payoff dominant (Figure 10(a)). The splitting equilibrium becomes the payoff dominant equilibrium for $\sigma$ sufficiently large.

A similar pattern holds when $\delta$ increases, except the lower range of $\sigma$ under which the incubation equilibrium is payoff dominated collapses to zero. The remaining pattern matches that of the case of $\delta$ small (Figure 11(b)). The equilibrium incubation period is weakly decreasing in $\sigma$, with the equality strict whenever $T^* > 0$ (Figures 11(a) and 11(b)).

4 (Dis)incentivizing Incubation

This section focuses on ways the parent club (as a centralized unit) can react to strategic membership by introducing two instruments the majority individuals can use to exert a degree of control over and profit from strategic membership. The first works through taxation/subsidization and the second through the use of exclusions such as intellectual property restrictions. These instruments can be thought of as
policy parameters. The timing of their implementation is as follows. Prior to the bargaining process leading to the setting of the mission(s), the parent club announces its decisions regarding taxation/subsidization and intellectual property restrictions. The game then proceeds as described in Section 3.1. The results from the previous sections can be reinterpreted as Markov perfect equilibria of the subgame induced by the parent club’s choice of instruments.

4.1 Membership Inclusion and Exclusion: Strategic Admission

Suppose that the parent club can control who may join and who may not. This measure of control leads to the concept of strategic admission, where individuals are either permitted to or prevented from joining the club given that other club members know that this individual will eventually split from the club. Before stating the results, I define some simplifying notation. Let \([0, \hat{T}]\) denote the optimal incubation period from the perspective of the parent club\[^{17}\] Denote \(T' \equiv \max\{\hat{T}, T^*\}\) and denote \(T'' \equiv \min\{\hat{T}, T^*\}\). Let \(V_{i(T')}T''\) and \(V_{i(T')}T'\) represent the indirect utility functions of a member \(i\) of the parent club, given that a split occurs at times \(T''\) and \(T'\), respectively. That is, \(V_{i(T')}T''\) and \(V_{i(T')}T'\) represent the indirect utilities of those individuals who retain their affiliation to the parent club post-split. Let \(W_{i(T')}T''\) and \(W_{i(T')}T'\) represent the indirect utility functions of an individual \(i\) who splits at times \(T''\) and \(T'\), respectively\[^{18}\]. Members of the parent club are indexed by \(j(t) \in p\) and the child club by \(j(t) \in c\).

The timing of a split impacts both minority individuals and majority individuals. Formally, let the impact of delaying a split be the absolute value of the difference between the sum of indirect utilities when a split occurs at \(T'\) and \(T''\). Thus the impact on a club, either parent or child, can be written as

\[
\left| \sum_{j(T') \in k} \left( z_{j(T')T'} - z_{j(T')T''} \right) \right| , \tag{18}
\]

\[^{17}\]If \(\hat{T} = \infty\), then analysis is conducted over the extended nonnegative real line \(\mathbb{R}_+ \cup \{\infty\}\), so \([0, \infty]\) is well defined.

\[^{18}\]For the remainder of this section, the analysis is contingent upon the value of the indirect utility function. Thus the specified assumptions and function forms from Sections 2 and 3 are unnecessary.
for \((z, k) \in \{(V, p), (W, c)\}\). By removing the absolute value restrictions from (18), it can be determined whether preferences are aligned or misaligned. Preferences are misaligned if

\[
\text{sign}\left\{ \sum_{j(T') \in p} (V_j(T') - V_j(T'')) \right\} \neq \text{sign}\left\{ \sum_{j(T') \in c} (W_j(T') - W_j(T'')) \right\}.
\]

Preferences are misaligned if the time at which splitters prefer to split does not coincide with the time at which non-splitters prefer the split occur. Delay is beneficial for a group if

\[
\sum_{j(T')} (z_j(T') - z_j(T'')) > 0,
\]

and harmful if (20) is not satisfied.

**Result (Strategic Admission)** If the impact of delaying the split on non-splitters is larger than the impact of delaying the split on splitters, and the preferences of the two groups are misaligned, then there exists a price \(\phi\), located in the convex hull of the impacts, such that the parent club offers those who split \(\phi\) to:

(i) increase the incubation period from \([0, T']\) to \([0, T'']\) if those who do not split benefit from the delay while those who split are harmed by the delay;

(ii) decrease the incubation period from \([0, T']\) to \([0, T'']\) if those who do not split are harmed by the delay while those who split benefit from the delay.

If there are economies of scale, it may be profitable for the club to keep those members for an extended period of time, i.e. an incubation period. If the benefits from allowing the splitting members to incubate are greater than the losses to the splitters when delaying the split by \(T' - T''\), then the club can offer a fixed payment, divided amongst the splitters, to delay their decision by \(T' - T''\): incubation is subsidized. Similarly, if the splitters would rather wait while the club is worse off with strategic membership, then the club can charge a fee for membership that those who plan on splitting are not willing to pay. Furthermore when the impact of delaying the split on the parent club members is smaller than the impact on the child club members, the parent club may still have an option at its disposal. While the parent club cannot
shift the length of the incubation period to its preferred length, it can pay a price \( \phi \) to adjust the incubation period so that it is closer to the preferred length \([0, \hat{T}]\). For example, if \( \hat{T} < T^* \), but the impacts are such that the parent club is not willing to pay the minimum price \( \phi \) to alter the child club’s behavior, the parent club may be able to offer a price \( \phi \) to adjust the incubation period from \([0, T^*]\) to \([0, \theta T^* + (1 - \theta)\hat{T}]\), \( \theta \in (0, 1) \).

There are two possibilities. Either the club can pay to bring individuals in that otherwise would not join/keep them for an extended period of time, or the club can charge a membership fee to keep undesired members out. Under decentralized clubs, where utility is increasing in the number of effective individuals, it is likely that incubation will occur, where the club temporarily subsidizes the membership of those who would prefer to form their own club. When clubs are centralized and total contributions may decrease when individuals who are not a good fit join, the club can charge a fee large enough to keep those individuals from joining. This result is a generalization of both [Iannaccone (1992, 1994)] and [McBride (2007a)], where the fee can be interpreted as being equivalent to the costly sacrifice in religion [Iannaccone (1992)], or the Mormon church requiring individuals to donate of 10% of their income [McBride (2007a)]. Alternatively, the subsidy is akin to inviting free riders, à la McBride (2007b). This distinction between centralized and decentralized clubs is important, as [Glomm and Lagunoff (1999)] show that centralized clubs are more likely to persist in the long run than decentralized clubs.

There is a corollary to the strategic admission result. Suppose that preferences are aligned, and specifically, both groups prefer strategic membership. Then the parent club can tax those interested in practicing strategic membership and capture the rents from incubating. To illustrate this instrument, suppose that preferences are aligned and minority individuals want to practice strategic membership. Those individuals are willing to pay any tax \( \phi \in [0, W_{j=m=2(T^*)-W_{j=m=2(T^*)}}] \) for the ability to do so. If types are observable, then this result is bad news for those practicing strategic membership. Any rents from strategic membership are captured by the parent club; however, since this tax can be treated as a fixed cost, it has no impact on the marginal contribution of any individual, and thus no efficiency implications.

\[19\] Or equivalently, make joining difficult.

\[20\] Note that the result is stated over the indirect utility functions and is thus independent of the mechanism used to choose the contribution levels.
It is worth noting that if individuals are sufficiently patient, then the parent club can still extract all incubation rents regardless of whether or not types are observable. In particular, if \( \delta \leq -\frac{1}{T^*} \ln \left( \frac{N_1}{N} \right) \), then all individuals will accept a taxation scheme in which: (i) each individual pays a value of \( \phi \) at the outset; (ii) immediately following the split, each individual who remains in the parent club receives their payment \( \phi \); (iii) the payments made by those who split are divided evenly amongst those who remain. At time 0, each agent loses \( \phi \). At time \( T^* \), those who remain receive the payment back, though the value is discounted by \( e^{-\delta T^*} \). This tax scheme is preferable (and thus implementable) if 
\[
-\left(1 - e^{-\delta T^*}\right)\phi + \left(\frac{N_2}{N_1}\right)e^{-\delta T^*}\phi \geq 0,
\]
which simplifies to 
\[
\delta \leq -\frac{1}{T^*} \ln \left( \frac{N_1}{N} \right).
\]

4.2 Excludability of Inputs: Strategic Loss

There are two channels through which strategic loss can be practiced. One is manipulating the degree of excludability through \( \rho(0) \) and the second is through forcing the child club to locate a minimum distance from the parent club, affecting the depreciation through \( d(A(\mu), A^c) \). The first is accomplished through intellectual property restrictions including copyrights, trademarks, and patents and the second is accomplished through the use of non-compete clauses.

With respect to excludability, if incubation is preferred, then \( \rho(0) \) must be set sufficiently high such that there is a large benefit to incubating. Alternatively, if these members are undesirable, then \( \rho(0) \) should be set sufficiently low such that there is no benefit to incubation and those individuals with the desire to split either do so immediately or not at all. \( \rho(0) \) can be affected though the use of intellectual property. For example, assigning a copyright license software prevents other software developers from freely taking, manipulating, and repackaging the source code and branding it as their own. If a coalition of software developers orchestrates a split of copyrighted software, then all the coalition retains when splitting is experience and reputation. The source code in its entirety must remain with the parent organization. If the parent club desires incubation, then the parent club can implement a copyleft license, such as the BSD 2-Clause License, the GNU General Purpose License, and the Apache License. Under these licenses, individuals are free to take, modify, and repurpose the source code, each to a differing extent.\(^{21}\)

\(^{21}\)While each license mentioned requires that contributions made by earlier contributors must
It is impossible to analyze $\rho(0)$ on its own when studying the parent club’s preference for incubation. The degree of rivalry $r$ must be considered simultaneously. Recall that $1 - r\rho(0)$ is the proportion of installed base the parent club retains following a split. If $r = 1$, so contributions are completely rival, then the parent club never wants to set $\rho(0) = 1$, as that implies that post-split, it loses everything. If $r = 0$, then the amount the parent club retains is independent of $\rho(0)$; however, the parent club will not be indifferent between $\rho(0) \in [0, 1]$ because both contributions and the decision on whether or not to the child club has the incentive to incubate are affected. Recall that the child club retains $\rho(1 - d(A(\mu), A^c))$ post-split, which can be rewritten as $\rho(0 + 1 - d(A(\mu), A^c))$. As $\rho(0)$ increases, $\rho(0 + 1 - d(A(\mu), A^c))$ necessarily decreases and as $\rho(0)$ decreases, $\rho(0 + 1 - d(A(\mu), A^c))$ is weakly increasing, strictly so if $\rho(0 + 1 - d(A(\mu), A^c)) > 0$.

By implementing restrictions such as non-compete clauses in membership contracts, the parent club can limit the options for the child club. In non-compete clauses, one party often agrees that it will not form an entity that competes too closely against the other party, where closeness can refer to either the geographic distance or the distance in the product characteristic space. Requiring a noncompete clause is equivalent to imposing a restriction on $A^c$ such that if the child club incubates within the parent club, then the distance between the two clubs must be sufficiently large: $1 - d(A(\mu), A^c) \geq \zeta$, for some $\zeta \in [0, 1]$. Restricting this distance to be sufficiently large places an upper bound on $\rho(1 - d(A(\mu), A^c))$, similar to the impact of copyrights and patents on $\rho(0)$. Requiring this distance to be great enough, plus limiting competition through copyrights can lead to $\rho(1 - d(A(\mu), A^c)) = 0$ for all feasible $A^c$ under the constraint $1 - d(A(\mu), A^c) \geq \zeta$.

The parent club has the incentive to practice strategic loss when either the cohesive equilibrium or splitting equilibrium payoff dominates the incubation equilibrium for majority individuals or alternatively, whenever preferences are misaligned. When the match value of majority individuals to the parent club is sufficiently large, the utility under the cohesive equilibrium will be greater than the equilibrium under the

---

still be attributed to the contributors, the GPL license further requires that any software developed using the source code must also be branded with the GPL license. The Apache and BSD licenses allow the source code to be branded under a new license, even a proprietary (copyright) license. For example, Apple utilized a fork of FreeBSD in the development of its Darwin system, which is a component of the proprietary OS X operating system. Patents and trademarks can be utilized in a similar fashion.
splitting equilibrium. The negative difference in utility caused by a small difference in match values between the cohesive and splitting equilibria is outweighed by the positive difference due to economies of scale given the presence of minority individuals (the contribution-scale tension). If the match value of the minority individuals to the parent club is not too small, then they too will prefer the cohesive equilibrium to the splitting equilibrium, as the benefits from economies of scale outweigh the costs of a poor match value, relative to the splitting equilibrium. If this is the case, then the parent club utilizes strategic loss to ensure that the cohesive equilibrium is payoff dominant for the minority individuals.

The numerical analysis conducted in Section 3.2.1 is useful in illustrating the effectiveness of the strategies under strategic loss. Figure 4 shows that when $\rho(1 - d(A(\mu), A^c))$ is sufficiently small, the incubation equilibrium is payoff dominated for the minority individuals. Therefore utilizing restrictions to place an upper bound on $\rho(1 - d(A(\mu), A^c))$ is sufficient to limit the profitability of strategic membership. Furthermore by Proposition 5, the parent club can utilize its control over depreciation to keep $d(A^c, A_2)$ small enough that an incubation equilibrium no longer exists.

5 Applications

Software development is often characterized by one of two development methodologies. The first is the PS model and the second is the OSS model. I shall now apply the analysis to both methodologies of software development. One feature that both methodologies share is that software development is often a collaborative process, requiring contributions from multiple individuals. PS development is more centralized, often coordinated and enclosed within a firm that takes advantage of the available intellectual property restrictions (often copyrights) to keep its source code proprietary. Nonetheless, the theory developed in this paper is still applicable to help better understand the firm’s options with respect to internal structure through the allocation of the software developers within the firm. The development team can be viewed as a club within the firm, where all clubs (parent and child) fall under the umbrella of the firm. The OSS development methodology corresponds to a decentralized venture, where programmers from all walks of life coordinate in a bazaar-like fashion.

\footnote{A third methodology has also been introduced recently, which is a hybrid of the two methodologies, referred to as the mixed-source model \cite{Casadesus-Masanell and Llanes 2011}.}
Raymond, 1999), and the source code is kept openly available for all to view, edit, and distribute (Open Source Initiative, 2015). An empirical feature of OSS is that projects have been known to fork over time, with development continuing on both the new child project, as well as the original parent project. This section examines each methodology and shows how this theory can be applied to understanding the underlying processes.

With respect to OSS, I use the model to explain the evolution of two open source derivatives of Unix, the BSD family of operating systems and Linux, considering both the Linux kernel and the vast array of Linux distributions. I then apply the model to proprietary software, showing how a firm developing multiple products can choose between three development structures, each corresponding to one of the equilibria described in this paper, to maximize profits.

5.1 Open Source Software Forks

Before considering specific forks, I first provide a brief overview of the process of forking. Forking occurs when an OSS project splits into multiple projects and is typically utilized in two contexts. An individual will often fork a project to provide themselves with a separate code base for developing and testing new features, without affecting the mainline project. When the added developments are ready, the individual makes a request to reintegrate the fork, adding the new features to the original project. Development of the fork then ceases. The second context is the one this paper is concerned with, which are permanent forks. Individuals fork a project with the intent to use the installed base to take the new project in a different direction than the original. The development of both projects continues indefinitely.

While there is only one official Linux kernel, there are many distributions that are built upon the Linux kernel. The Linux kernel represents the operating system, which can be viewed as the central nervous system of a computer. It relays information between the software and hardware. End-users often interact with the distribution, which includes the operating system, plus what are known as “userland” utilities, such as a graphical user interface, a file-system manager, a document typesetter, and an internet browser. In 1993, there were three major Linux distributions: Debian, Red Hat, and Slackware Linux. Presently there are currently over 250 Linux distri-

\footnote{In practice, there are instances where the original project dies out post-fork, though in these cases, the fork is usually a response to an underlying failure that leads to the original project’s death.}
butions listed on Distrowatch, a website dedicated to providing information on OSS distributions. Many of these distributions can be traced back to these three parent distributions and are the product of the splitting of OSS development communities. One such distribution, Ubuntu, is one of the most popular Linux distributions, ranked at #3 by Distrowatch. Ubuntu was forked from Debian in 2004 by Mark Shuttleworth, along with a small team of Debian developers. Shuttleworth and his team felt that Debian was not accessible enough, in terms of ease-of-use, so set out to develop an easy to use distribution based on Debian. Thus the fork can be attributed to differences in missions, on the dimension of usability.

Ubuntu is among the most forked Linux distributions, Debian notwithstanding, as every fork of Ubuntu can also be interpreted as a fork of Debian as well. Ubuntu formally recognizes nine of the forks, which it refers to as “flavors.” However, there are over seventy forks of Ubuntu, and some of those forks, such as Kubuntu, a recognized flavor, have been forked as well. A select few of these forks are given by name, date, and popularity in Table 2. A complete history from Ubuntu’s origin through 2012 is given in Figure 12. It is clear by looking at the names of these forks that they occurred due to heterogeneity among the contributors. I illustrate this

---

24 Slackware itself was forked from SLS, which was the first comprehensive Linux distribution.
25 Rankings are determined by an algorithm based on page traffic.
<table>
<thead>
<tr>
<th>Name of fork</th>
<th>Date</th>
<th>Name of fork</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kubuntu (34)</td>
<td>2005</td>
<td>Easy Peasy (NR)</td>
<td>2008</td>
</tr>
<tr>
<td>Edibuntu (164)</td>
<td>2005</td>
<td>Zorin OS (15)</td>
<td>2008</td>
</tr>
<tr>
<td>Mint (1)</td>
<td>2006</td>
<td>Lubuntu (17)</td>
<td>2009</td>
</tr>
<tr>
<td>Xubuntu (27)</td>
<td>2006</td>
<td>Peppermint OS (38)</td>
<td>2010</td>
</tr>
</tbody>
</table>

* Number in parentheses indicates ranking according to distrowatch.com. NR corresponds to not ranked.

Table 2: Summary of select Ubuntu forks.

point with a brief description of a subset of the forks reported in table 2. Kubuntu and Edibuntu are two of the recognized flavors. Kubuntu took Ubuntu, removed the desktop environment GNOME (distance-based depreciation) and replaced it with the KDE desktop environment. Designed for use in classrooms, Edibuntu adds an array of education-oriented software and is an example where there is minimal distance-based depreciation. Mint Linux removed Ubuntu’s default desktop environment GNOME and replaced it with a fork known as Cinnamon which has been customized for Linux Mint. It also includes some proprietary software such as Adobe Flash. Zorin OS takes the Ubuntu distribution, removes the imbedded userland software (distance-based depreciation) and replaces it with a Windows clone, designed for newcomers to Linux to make the transition easier. The same line of reasoning applies for all of the forks listed in Table 2.

Incubating and forking software for horizontal reasons is not unique to Ubuntu or Linux distributions in general. The BSD family of distributions have also experienced their fair share of forks. The BSD distributions are perceived as direct descendants of AT&T’s Unix operating system. Unlike Linux, no distinction is made between the kernel and the distribution. That is, the BSD kernel and associated software are developed and released together, so each BSD distribution has its own unique kernel. There are many BSD distributions. Most modern BSD distributions are derivatives of the 386BSD distribution. Two well known distributions, FreeBSD and NetBSD, emerged in the early 1990s from 386BSD. FreeBSD represents the continuation of 386BSD under a new name while NetBSD is a distinct fork. The mission (as defined in section 2) of FreeBSD development is features, speed, and stability. The mission of 29BSD was released as open source only after a series of lawsuits between developers and AT&T were settled.
NetBSD development is portability. Portability refers to the ability of software to run many different hardware platforms. Other well known forks of BSD systems include Darwin OS, which was forked from FreeBSD. Darwin OS is developed by Apple and is an input of OS X.\footnote{Darwin was developed as a combination of forks from FreeBSD and BSD 4.4, as well as NextStep 3.3, which was developed by Steve Jobs and his team at NeXT.}

OpenBSD is a distribution forked by Theo de Raadt, a long time developer of NetBSD. OpenBSD was formed with portability in mind, but also emphasizes standardization, correctness, proactive security, and integrated cryptography. It was on these points that de Raadt disagreed with fellow contributors, eventually leading to the fork.\footnote{https://www.freebsd.org/doc/en/articles/explaining-bsd/comparing-bsd-and-linux.html, accessed August 26, 2015.}

### 5.2 Proprietary Software Development: Internal Incubation

Beyond OSS, similar development patterns can be found within individual firms developing PS. It is not uncommon to find proprietary software vendors developing multiple, related products, referred to as product lines. Microsoft’s Office Suite consists of a word processor, a spreadsheet program, a database manager, and an email manager. Apple develops video and audio editing software such as Final Cut Pro and GarageBand. Autodesk develops several products useful in the rendering of three dimensional CGI, including Civil 3D, 3DS Max, and Maya. The important feature of product lines, such as the ones described above, is that the component products consist of shared features (modules). When developing product lines, the vendor must decide how to internally structure development of the software.

In general, there are three internal development structures that can be used to develop software lines, each with its own benefits and costs. The vendor can utilize multiple development teams, where each team specializes in developing a single program. Each development team codes both the common features and the features of their program independently. Alternatively, the vendor could have a single development team manufacture the programs sequentially \cite{Taylor et al, 2010}\footnote{A single team developing the various programs simultaneously is isomorphic to separate teams developing each.} The first approach resembles the underlying strategy in the cohesive equilibrium and the second resembles the underlying strategy in the splitting equilibrium. The third option
is a form of strategic membership, only the split is internal. The developers code modules of each program together as a cohesive unit, then split up to finish each program independently, while still remaining under the umbrella of the vendor.

First suppose that, as in the splitting equilibrium, there are $M$ development teams, each responsible for independently developing one of $M$ programs. Under this structure, development of the software proceeds rapidly since the programs are developed simultaneously. The rapidity comes at a cost, namely through a duplication of effort with the possibility of incompatibilities. Since each product is developed independently, the common features are also coded independently, $M$ times.

A similar relationship exists when there is a single team developing the software, as in the cohesive equilibrium, only development proceeds at a slower pace. Each program is developed separately and sequentially, delaying the income from sales, while still suffering from the inefficiencies of duplication. Figure 13 illustrates how these two structures relate to the vendor’s income and expenditure. During development, expenditure initially exceeds income. Once the software reaches market, income increases and eventually the software becomes profitable. The process is repeated for each of the $M$ programs.

Incubation provides a profitable alternative for the vendor. At the outset, the entire development team works together to code the common features. Once the common features are coded, the team divides up and codes the independent modules for each program. A sample timeline is given in Figure 14. When utilizing the incubation strategy, the vendor eliminates duplicative waste. Not only is the cost of developing each program decreased, but the total cost of development is decreased as well since the development cost is spread among all of the projects. This relationship...
is illustrated for $M = 2$ in Figure 13(b).

In summary, incubation can benefit software vendors in the software industry developing product lines. To end with a concrete example, suppose a software vendor has two programs it wishes to develop: a voice-to-text/text-to-voice program (V2T), and visual voicemail (VV). Each product has its own independent features. The V2T program takes voice and turns it into a text output. The VV program collects voicemails and stores them on the user’s mobile phone, making them available for immediate playback while also creating text versions of each voicemail. However, these two programs also share common features, namely converting voice to text. The vendor must choose between developing these two products separately or incubating, where the development team first works together to develop the text-to-voice libraries, and then splits up. One group then finishes the V2T software while the other finishes the VV software. If the vendor were to develop the projects separately, then there would likely be duplicative waste, as both projects require the development of text-to-voice libraries. By jointly developing these libraries, the vendor eliminates the duplicative waste and is able to save costly resources.

6 Discussion and Concluding Remarks

To summarize, this paper develops a dynamic club goods framework and explores three related equilibria. In the cohesive equilibrium, a single club persists in the long run. In the splitting equilibrium, multiple clubs exist at both the outset and the long run. The third strategy, strategic membership, and its associated equilibrium, the incubation equilibrium, is somewhat of a hybrid between the two. There is a single cohesive club at the outset, but at some point in time, the club splits and individuals partition themselves according to their mission. In the long run, splitting is always preferred to incubation, but due to the impatience of individuals, a tradeoff exists
where it is profitable to exchange future utility for greater utility in the present via incubation.

The parent club can alter the relative payoffs of the various equilibria through the use of various instruments. The two instruments discussed in the paper are taxation/subsidization and intellectual property (exclusionary) restrictions. Through taxation and subsidization, the parent club can incentivize those practicing strategic membership to adjust the length of the incubation period to match the preferred length of the parent club, so long as the impact of the change in the incubation period is greater for the parent club. If the members of the parent club are sufficiently patient, then the parent club can also capture all rents gained by the child club through incubation. Intellectual property restrictions can be used to adjust the excludability of the club, making incubation less desirable as the proportion of installed base retained by the child club can be taken to zero by the use of restrictions including copyrights, patents, and non-compete clauses. All that is necessary for these avenues to be effective is an external enforcement mechanism, such as the court system.

Many of the features and results discussed in this paper are not unique to the software industry. Strategic membership is a strategy prevalent in many environments in which collaborative production occurs and installed base is able to accumulate over time. In some environments, such as R&D, it would be difficult to sustain strategic membership in research joint ventures between firms due to the stochastic nature of payoffs from such ventures. However, when it comes to within-firm R&D, adopting the strategy identified in the proprietary software application in section 5.2 could be worthwhile to a firm. When the R&D program is in its early stages and less focused, the firm concentrates its efforts to a single project, where its researchers work as a single, cohesive team. As the program grows, multiple directions may present themselves, at which point the firm splits the team to pursue each avenue. Alternatively, it could be that the firm splits the team simply to maximize the likelihood of success of the venture. If the probability of success exhibits diminishing marginal returns to group size, two groups of 10 researchers have a better chance of success than one group of 20. As long as the cost of maintaining two groups over one does not exceed the increased expected payoff, incubation can improve the expected outcome in R&D development.

Moving outside of the IT industry and R&D, the model and ideas developed within can also be useful in understanding problems in labor economics (e.g. labor raiding)
and political economy. One avenue worth exploring is the use of incubation to develop political capital in understanding the splitting of political parties in multiparty systems. Furthermore, this paper can provide an explanation for the emergence of extremist groups from mainline groups over time, e.g., the emergence of the Islamic State from Al Qaeda and large scale religious schisms, where religious capital must be built up prior to forming a successful religious denomination.

From a theoretical point of view, there are several avenues for future research. This paper relies on heterogeneity and capital accumulation to generate the underlying incentives. However, there are often multiple groups that appear to be homogeneous. The only way to generate such a situation in equilibrium is if there is congestion. Congestion is often modeled on the demand (consumption) side, but there is an under-explored alternative. The clubs could exhibit supply-side congestion, which opens up the possibility for strategic membership and the splitting of clubs, even with homogeneous individuals.

References


Arnold, Tone and Myrna Wooders (2005), Dynamic Club Formation with Coordination, Vanderbilt University Department of Economics Working Papers 0522, Van-
Arnold, Tone and Myrna Wooders (2009), Dynamic club formation with coalitional deviations.

URL: http://wrap.warwick.ac.uk/46082/


Open Source Initiative (2015a), ‘Apache license 2.0’.

**URL:** http://opensource.org/licenses/Apache-2.0

URL: http://opensource.org/licenses/BSD-2-Clause


URL: http://opensource.org/licenses/GPL-3.0


URL: http://opensource.org/osd


