UNIVERSITY OF CALGARY

Coordination in Games with General Network Effects

by

Alexander M. Jakobsen

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF ARTS

DEPARTMENT OF ECONOMICS

CALGARY, ALBERTA

April, 2009

UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "Coordination in Games with General Network Effects" submitted by Alexander M. Jakobsen in partial fulfillment of the requirements for the degree of MASTER OF ARTS.

Supervisor, Dr. B. Curtis Eaton
Department of Economics

Dr. John Boyce
Department of Economics

Dr. Claude Laflamme
Department of Mathematics and Statistics

05/12/2008
Date

Abstract

Network effects exist when there are benefits to aligning one's behaviour with the behaviour of others. There is a large literature on network effects, as issues such as technology adoption, fads, and many others revolve around network effects. Typically, such models are specified in a manner so that multiple purely coordinated and Pareto efficient equilibria exist, which introduces an equilibrium selection problem. Strangely, this selection problem has largely been ignored, and little effort has been made to examine how agents form expectations supporting coordinated outcomes. This thesis attempts to fill this gap by introducing a generalized, dynamic network effect model in which agents make their decisions sequentially and use their observations of previous decisions to form expectations about future decisions. Several results are proven regarding the likelihood and extent of coordination, and numerical examples are provided to complement the formal theory. The results show that even under strong network effects, purely coordinated outcomes are unlikely to occur, and some coordinated outcomes may actually be impossible, even under nontrivial model specifications.

Acknowledgements

I would like to start by thanking my supervisor, Dr. B. Curtis Eaton, for all of his support during my final years as an undergraduate, and especially for his support and encouragement during my time as a graduate student in Economics. I have learned a great many things from him; most importantly, he has set an excellent example for how to be a great professor and supervisor, an example I hope to be able to emulate one day.

I am also indebted to my parents, who have always displayed remarkable faith in me, and who have supported and encouraged me immensely as I have grown up and progressed in my studies. They have also been incredibly patient, as I am now the fourth (out of four) of their children to choose the life of a "professional student"! I am also grateful for the rest of my family, as they have been there with me the whole time, and have always been pleasant and supportive company.

Who could forget friends? There have been so many great people there to support me, it would be difficult to try and name them all: Cameron Greenhall, Paul Huestis, Garrett Mayall, and Chris Watson, to name a few. I also want to thank all of the new friends I made in the graduate economics program, especially Chris McLeod, for the fun (though hectic) year we spent studying together.

Finally, I would also like to thank the many highly influential professors I have had here at the University of Calgary. They have played a major role in shaping who I am, and in my development as an academic. Like Dr. Eaton, they have set an excellent example for me to follow.

Table of Contents

List of Tables

Chapter 1

Introduction

There is a vast array of decision problems in which there are benefits to aligning one's actions with the actions of others. The canonical example involves the choice between incompatible communication networks, such as early telephone networks or modem protocols; if more individuals join a particular network, then the value of the network increases for all of its members, because each member may now use the network to communicate with more people. A modern variation of this includes on-line social networking or dating services, since an individual who joins a service with a greater number of subscribers is more likely to find old (or new) acquaintances there. A software engineer who learns a particular programming language sees a greater return on his investment if additional programmers learn the language as well, because his skills will be more widely applicable and transferable. A DVD player technology increases in value as more consumers purchase it, because more DVDs will be released which support that technology.

All of these are examples of positive *network effects* in that the value of belonging to a particular group increases as the group grows larger. In some cases, the network effect is *direct*, meaning that the value of the network comes about directly from its size, and not through some intermediate channel. The communication network and social networking examples above constitute direct effects. Alternatively, the effect may be *indirect*, in that the value does not derive explicitly from the number of members, but rather through intermediate effects caused by an increase in network size. In the DVD example above, for instance, no individual receives higher utility from knowing that many others have the same technology; instead, his valuation of the network stems from the fact that a more popular player technology will be supported by suppliers of complementary goods

(in this case, manufacturers of movie discs). The root cause of the increase, however, is still network size, so in this thesis the distinction is ignored in favour of a generalized definition wherein network effects exist if the value of the network increases as a function of its size, regardless of the specific mechanism responsible for the effect.

The difficulty with network effect models is that there is a coordination problem. If there are two competing networks, for instance, which network will be more widely adopted? There are many issues to consider in attempting to answer this, but in essence it is an equilibrium selection problem. Typically, network effect models are specified in a static fashion, and have multiple (Pareto efficient) equilibria corresponding to coordination on different networks. Exploring *how* these equilibria may be achieved is the central objective of this thesis. Specifically, a generalized definition of network effect models is given and, using this definition, a number of results regarding dynamic coordination are derived.

There is a large literature on network effects. Rohlfs (1974), in studying how network effects influence demand for a communication network service, is among the very first contributors to this area, although Katz and Shapiro (1985), in a pioneering work, construct a simple but more general model of network effects. Their model is static in that there is only one period in which consumption decisions are made, but they identify multiple coordinated equilibria under the assumption that consumers base their decisions on expected network size. Interestingly, they do not investigate how agents form their expectations, but simply assume that these expectations are fulfilled in equilibrium. Katz and Shapiro (1986) extend this analysis to a two-good, two-period model and find that when coordination does occur, the outcome may not be socially optimal. To that effect, Katz and Shapiro (1994) provide several examples of how network effects may lead to market failure.

Katz and Shapiro (1985) also contrast direct and indirect network effects. The dis-

tinction between the two, and related issues, is analyzed by Liebowitz and Margolis (1994), and also by Clements (2004). While technology adoption problems are the most prevalent examples of real-world network effects (see Farrell and Saloner (1985), Farrell and Saloner (1986), Choi (1994), Cabral (1990), or, in particular, Church and Gandal (1992) for a tangible example of indirect network effects), they are also useful in studying other phenomena. A nice example of this is Church and King (1993), which develops a model of bilingualism and language adoption by noticing that a language becomes more valuable as more people learn it. Economides (1996) explores a variety of sources of network effects, and how they influence both prices and market structure.

There is a related literature on what are usually called "informational cascades" (see Bikhchandani, Hirshleifer, and Welch (1998)). While there are some similarities between network effect models and information cascade models, informational cascades are different in that the behaviour of others is used to make an inference about the information others have concerning the *quality* of a good or technology, and these inferences are what drive individual decisions. For example, Choi (1997) develops a model of technology adoption where the true value of a technology is revealed to all once an agent has adopted it. The models considered in this thesis, however, are "pure" network effect models in that all agents are fully informed about the quality of the networks (and different agent types have different private preferences over the set of possible networks), but they try to infer which network will be more widely *adopted* in order to decide which one to join. That is, the value of a network is determined by its expected *size*, and these values, together with private preferences, are what drive individual behaviour.

This thesis examines how agents in a sequential model, with little information, form expectations about the magnitude of different networks, and how these expectations combine with many other considerations (like private preferences, network values, and population composition) to cause (or inhibit) coordinated behaviour. Analysis is carried

out at a high level of generality, so that results may be applied to a wide variety of network effect models. While some theoretical work has analyzed coordination problems in similar settings (see Crawford (1995), which studies coordination problems in the context of repeated games), little progress has been made. As Farrell and Klemperer (2007) point out, "coordination is hard, especially when different adopters would prefer different coordinated outcomes." Using a model inspired by Eaton and Krause (2005), a variety of new results concerning network effects and coordination are proven, and concrete numerical examples are given to reinforce formal results and provide some insight into just how likely (and how complete) coordination may be.

Chapter 2 begins by introducing relevant mathematical background and notation for use throughout the thesis. Chapter 3 then presents the general network effect model and identifies a set of assumptions needed to represent almost any network effect model. Chapter 4 analyzes the simultaneous-move version of the general network effect model, and formalizes conditions under which equilibria of various degrees of coordination are guaranteed to exist, as well as an analysis of when these outcomes are Pareto efficient. The main focus of the thesis, Chapter 5, introduces a sequential move version of the general model, explores two different decision algorithms agents may use, proves comparative static results regarding the likelihood of coordination in this dynamic setting, and provides several numerical examples to illustrate the comparative static results and other issues of interest. Chapter 6 concludes with a summary and evaluation of the results herein, as well as some possible avenues for future research.

Chapter 2

Mathematical Prerequisites

Most of the concepts in this thesis require rudimentary knowledge of real analysis, set theory, functions, and combinatorics, so a brief digression on these matters, as well as of notational conventions, is presented first. Readers who are familiar with this material may safely skip ahead to Chapter 3.

2.1 Basic Definitions and Notation

2.1.1 Sets and Functions

With minor exceptions, all sets in this thesis are subsets of the real line, $\mathbb{R}$, and so the naïve (rather than the axiomatic) approach to sets is suitable. This means a set may simply be defined to be a collection of objects, called *elements*, which may themselves be numbers, sets, functions, or other constructions deemed necessary. If $\mathcal{S}$ is a set, the statement "$x \in \mathcal{S}$" is read "*x is in $\mathcal{S}$*" and means that x is an element of the set $\mathcal{S}$. Likewise, "$x \notin \mathcal{S}$" means that x is not an element of the set $\mathcal{S}$. The symbol $\emptyset$ denotes the empty set, which is the set containing no elements at all.

Given two sets $\mathcal{A}$ and $\mathcal{B}$, $\mathcal{A}$ is a subset of $\mathcal{B}$ (denoted $\mathcal{A} \subseteq \mathcal{B}$) if every element of $\mathcal{A}$ is also an element of $\mathcal{B}$. Obviously, every set is a subset of itself, and if both $\mathcal{A} \subseteq \mathcal{B}$ and $\mathcal{B} \subseteq \mathcal{A}$, then $\mathcal{A} = \mathcal{B}$. The empty set, for example, is a subset of every set. Also, letting $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$ denote the set of natural numbers and $\mathbb{N}_+ = \{1, 2, 3, \ldots\}$ denote the set of nonzero natural numbers, it is clear that $\mathbb{N}_+ \subseteq \mathbb{N}$. Using similar notation, it is also clear that $\mathbb{R}_+ \subseteq \mathbb{R}$ (here, $\mathbb{R}_+$ is not just the set of nonzero real numbers, but instead the collection of all real numbers greater than or equal to zero).

Next, consider two sets $\mathcal{A}$ and $\mathcal{B}$. The *union* of $\mathcal{A}$ and $\mathcal{B}$, denoted $\mathcal{A} \cup \mathcal{B}$, is the set $\mathcal{S} = \{x \mid x \in \mathcal{A} \text{ or } x \in \mathcal{B}\}$. It consists of all elements in the two sets. Likewise, the *intersection* of $\mathcal{A}$ and $\mathcal{B}$, denoted $\mathcal{A} \cap \mathcal{B}$, is the set $\mathcal{S} = \{x \mid x \in \mathcal{A} \text{ and } x \in \mathcal{B}\}$; it consists only of those elements which are in both $\mathcal{A}$ and $\mathcal{B}$. There is also a "subtraction" operation on sets, which is given by $\mathcal{A} \backslash \mathcal{B} = \{x \mid x \in \mathcal{A} \text{ and } x \notin \mathcal{B}\}$. So, $\mathcal{A} \backslash \mathcal{B}$ contains only those elements which are in $\mathcal{A}$ but not in $\mathcal{B}$. Finally, suppose $\mathcal{A}$ and $\mathcal{B}$ are nonempty sets. Then the *cross product* of $\mathcal{A}$ with $\mathcal{B}$, denoted $\mathcal{A} \times \mathcal{B}$, is given by $\mathcal{A} \times \mathcal{B} = \{(x, y) \mid x \in \mathcal{A} \text{ and } y \in \mathcal{B}\}$; it consists of all *ordered pairs* (x, y) where x belongs to $\mathcal{A}$ and y belongs to $\mathcal{B}$. As a notational convenience, $\mathcal{A} \times \mathcal{A}$ is written $\mathcal{A}^2$. It is easy to extend the notion of cross products to allow ordered triples, quadruples, or, in general, n-tuples. For example, the set $\mathbb{R}^n$ consists of all n-tuples $(x_1, \ldots, x_n)$, where $x_i \in \mathbb{R}$ for every $1 \leq i \leq n$.

Before discussing functions, it is worthwhile to discuss well-ordered sets. Let $\mathcal{X}$ be a nonempty set, and let $\mathcal{R} \subseteq \mathcal{X}^2$. Such a subset is called a *relation* on $\mathcal{X}$. In this thesis, $\mathcal{R}$ is called an *order relation on* $\mathcal{X}$ if $\mathcal{R}$ is irreflexive (for all $x \in \mathcal{X}$, $(x, x) \notin \mathcal{R}$), transitive (for all $x, y, z \in \mathcal{X}$, $(x, y) \in \mathcal{R}$ and $(y, z) \in \mathcal{R}$ implies $(x, z) \in \mathcal{R}$), and complete (for all distinct $x, y \in \mathcal{R}$, either $(x, y) \in \mathcal{R}$ or $(y, x) \in \mathcal{R}$). Often, the notation "$x <_{\mathcal{R}} y$" is taken to mean that $(x, y) \in \mathcal{R}$. If it happens that every nonempty subset $\mathcal{S} \subseteq \mathcal{X}$ has a $<_{\mathcal{R}}$-minimal element (meaning that there is some $m \in \mathcal{S}$ for which $m <_{\mathcal{R}} s$ for every $s \in \mathcal{S} \backslash \{m\}$), then $\mathcal{X}$ is said to be *well-ordered* by $\mathcal{R}$. It is routine to verify that both $\mathbb{N}$ and $\mathbb{N}_+$ are well-ordered by the usual order relation $<$ on the real numbers; note, however, that $\mathbb{R}$ is not well-ordered by $<$. That $\mathbb{N}$ is well-ordered will be used frequently in later sections.

Now, let $\mathcal{X}$ and $\mathcal{Y}$ be nonempty sets. A *function* f *from* $\mathcal{X}$ *to* $\mathcal{Y}$, denoted $f : \mathcal{X} \to \mathcal{Y}$, is a set $f \subseteq \mathcal{X} \times \mathcal{Y}$ such that (1) for every $x \in \mathcal{X}$ there exists some $y \in \mathcal{Y}$ so that $(x, y) \in f$, and (2) for every $x \in \mathcal{X}$, if $(x, y) \in f$ and $(x, z) \in f$, then $y = z$. Here, $\mathcal{X}$ is the *domain* of f, and $\mathcal{Y}$ is the *codomain*. Property (1) asserts that every $x \in \mathcal{X}$ must

have an *image* in $\mathcal{Y}$, and property (2) asserts that each $x \in \mathcal{X}$ has only one image. So, the notation $f(x) = y$ really means that $(x, y) \in f$.

By definition, every $x \in \mathcal{X}$ has an image $f(x) \in \mathcal{Y}$; but, in general, not every $y \in \mathcal{Y}$ has a *preimage* in $\mathcal{X}$ (that is, an element $x \in \mathcal{X}$ such that $f(x) = y$). When every $y \in \mathcal{Y}$ *does* have such a preimage, f is said to be an *onto* function, or a *surjection.* In general, however, preimages are not unique. That is, if $f(x) = y$ for some $x \in \mathcal{X}$, it is possible that there is some $z \neq x$ in $\mathcal{X}$ for which $f(z) = y$ as well. If f satisfies the property that for all $x, z \in \mathcal{X}$, $f(x) = y = f(z) \Rightarrow x = z$, then f is said to be a *one-to-one* function, or an *injection.* If f is both a surjection and an injection, then f is a *bijection.* Bijections are useful because, among other reasons, they send distinct elements of $\mathcal{X}$ to distinct elements of $\mathcal{Y}$ in a manner which "covers" all of $\mathcal{Y}$. Thus, the existence of a bijection between two sets indicates that, in an abstract sense, the two sets have the same size. For a finite set $\mathcal{S}$, $|\mathcal{S}|$ denotes the cardinality of $\mathcal{S}$ (that is, the number of elements in $\mathcal{S}$), and if $\mathcal{T}$ is another set, then $|\mathcal{S}| = |\mathcal{T}|$ if and only if there exists a bijection between $\mathcal{S}$ and $\mathcal{T}$. This is perfectly intuitive for finite sets; and although the same is true of infinite sets, this thesis will not require knowledge of infinite cardinals, so such matters will not be discussed here.

Given a function $f : \mathcal{X} \to \mathcal{Y}$, two related sets may be defined. Let $\mathcal{A} \subseteq \mathcal{X}$. Define the *image of* $\mathcal{A}$ to be the set $f(\mathcal{A}) = \{y \in \mathcal{Y} \mid \text{there exists } x \in \mathcal{A} \text{ for which } f(x) = y\}$; that is, $f(\mathcal{A})$ contains those elements of $\mathcal{Y}$ which are paired with elements of $\mathcal{A}$. $f(\mathcal{X})$ is usually called the *range* of f. Clearly, $f(\mathcal{X}) = \mathcal{Y}$ if and only if f is surjective. Similarly, if $\mathcal{B} \subseteq \mathcal{Y}$, then the *preimage* of $\mathcal{B}$ is defined as the set $f^{-1}(\mathcal{B}) = \{x \in \mathcal{X} \mid \text{there exists } y \in \mathcal{B} \text{ for which } f(x) = y\}$. So, $f^{-1}(\mathcal{B})$ contains all elements $x \in \mathcal{X}$ for which $f(x) \in \mathcal{B}$. If $f^{-1}(\{y\})$ contains exactly one element for every $y \in f(\mathcal{X})$, then f is said to be *invertible* and the function $f^{-1} : f(\mathcal{X}) \to \mathcal{X}$ is given by $f^{-1}(y) = x$, where x is the unique preimage of y.

Sometimes it is useful to think of a function as a "formula" involving real numbers, but quite often the more fundamental (and abstract) idea of a function simply being a pairing between two sets is incredibly helpful. Indeed, most of the fundamental theorems developed in Chapter 5 require this mindset, as well as command of the set-theoretic tools developed above. Still, real-valued functions and sequences play an important role; some of their properties are given next.

2.1.2 Real Functions, Sequences, and Limits

A real-valued function is simply a function $f : D \to \mathbb{R}$, where D is some set (not necessarily of real numbers). Most often, real-valued functions are useful due to the order properties of $\mathbb{R}$, upon which concepts such as limits and derivatives may be built for suitable domains. Unfortunately, the nature of the models in this volume do not permit the use of derivatives. For this reason, the following definition of increasing functions is used throughout:

Definition 2.1.1. *Let $D \subseteq \mathbb{R}^n$ and $f : D \to \mathbb{R}$. Then f is weakly increasing in x_i if for all $(x_1, \ldots, x_i, \ldots, x_n), (x_1, \ldots, x_i', \ldots, x_n) \in D$ with $x_i' > x_i$, it follows that $f(x_1, \ldots, x_i', \ldots, x_n) \geq f(x_1, \ldots, x_i, \ldots, x_n)$. f is strictly increasing if the above statement holds with strict inequality in place of weak inequality, and constant in x_i if the above statement holds with equality in place of weak inequality. Weakly/strictly decreasing functions are defined similarly.*

So, the idea is that f is increasing in a variable x_i if, holding all other variables constant, an increase in x_i causes an increase in f. Note that any strictly increasing function is necessarily weakly increasing, so any function which is defined to be weakly increasing allows for the possibility that the function is actually strictly increasing. Similar statements can be made about decreasing functions.

A *real sequence* is a function $a : \mathbb{N} \to \mathbb{R}$; typically, the notation a_n is used in place of $a(n)$ to denote the n^{th} term of the sequence, and (a_n) refers to the entire sequence. A sequence (a_n) is said to be *constant* if there exists some $c \in \mathbb{R}$ such that $a_n = c$ for every $n \in \mathbb{N}$; it is *non constant* if it is not constant (that is, if there are at least two distinct $n, m \in \mathbb{N}$ for which $a_n \neq a_m$). Since $\mathbb{N} \subseteq \mathbb{R}$, definition 2.1.1 applies to sequences, so increasing/decreasing sequences have already been defined. Finally, a very important concept about real sequences is that of a limit:

Definition 2.1.2. *Let (a_n) be a sequence. Then (a_n) converges to the limit $L \in \mathbb{R}$ if for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $n \geq N \Rightarrow |a_n - L| < \epsilon$. The notation $a_n \to L$ indicates that (a_n) converges to L.*

Obviously, not every sequence has a limit. But if a sequence (a_n) *does* have a limit, then, by definition, for every $\epsilon > 0$, there is some $N \in \mathbb{N}$ for which $n \geq N \Rightarrow L - \epsilon < a_n < L + \epsilon$; that is, (a_n) is eventually bounded by the open interval $(L - \epsilon, L + \epsilon)$.

2.1.3 Basic Combinatorics

In this thesis, *combinatorics* refers to the study of how many ways an object may be constructed. For example, if one wishes to determine how many ways an arrangement of n objects can be made, combinatorics gives the answer: $n!$. This "factorial" function is defined recursively:

$$n! = \begin{cases} 1 & \text{if } n = 0 \\ n \cdot (n-1)! & \text{if } n \geq 1 \end{cases}.$$

Sometimes, it is required to determine how many ways k of n objects can be selected and then arranged. This is easily determined by identifying a "recipe" for constructing such an ordering: there are n choices for the first object in the arrangement, $(n-1)$ choices for the second, and so on, until finally there are $(n-k+1)$ choices for the k^{th}

object. So, letting ${}_nP_k$ denote the number of ways to select and order k of n objects, this demonstrates that

$$ {}_nP_k = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!} \cdot $$

This is the first step toward a formula for how many ways k of n objects may be selected, given that order is irrelevant. For example, such a formula would indicate how many subsets of size k a finite set of n elements has, because sets have no particular order. Let $\binom{n}{k}$ denote the number of ways to select k of n objects without regard to order. This is related to ${}_nP_k$ in that ${}_nP_k = \binom{n}{k} \cdot k!$, because the number of ways of selecting *and* ordering k of n objects is given by first selecting the objects, and then ordering them. Thus

$$ \binom{n}{k} = \frac{n!}{k!(n-k)!} \cdot $$

Of course, all of these equations require that $0 \leq k \leq n$. These formulas, as well as similar counting "recipes", will be used extensively throughout Chapter 5.

Chapter 3

The General Network Effect Model

The principal objective of this chapter is to specify the vital components of a generalized network effect model. Instead of restricting attention to a few select models, the focus is on identifying a minimal list of properties which network effect models must reasonably satisfy. These properties, developed in sections 3.1 and 3.2, serve as basis for all further exposition.

3.1 Agents and their Utility Functions

Every model in this thesis is built upon a set $\mathcal{X}$ of N agents who must choose one of two alternatives: A or B. These could represent goods, technologies, patterns of behaviour, or any other (abstract or otherwise) collection of choices exhibiting positive network effects; they will simply be referred to as *goods*. All questions in this thesis revolve around the composition of $\mathcal{X}$, the preferences of the agents, and the network effects resulting from their decisions. By assumption, agents are one of two types (A or B) and have utility functions $U_i : \{A, B\}^2 \times \mathbb{N}^2 \to \mathbb{R}_+$ of the form

$$U(t_i, c_i, n_A, n_B) = v(t_i, c_i) + e(c_i, n_A, n_B),$$

where $t_i \in \{A, B\}$ is agent i's type, $c_i \in \{A, B\}$ is agent i's decision, and n_j is the total (final) number of agents who adopt good j for $j \in \{A, B\}$. The function $v : \{A, B\}^2 \to \mathbb{R}_+$ represents the direct (private) utility derived from adopting either good, net of any costs from doing so, and the function $e : \{A, B\} \times \mathbb{N}^2 \to \mathbb{R}_+$ represents the network effect from choosing good c_i.

Letting $v_{x,y}$ represent $v(X, Y)$, one of the defining traits of Type A agents is that

$v_{A,A} > v_{A,B}$; similarly, a characteristic of Type B agents is that $v_{B,B} > v_{B,A}$. These relations indicate that a Type i agent has a higher private value of good i than of good $j \neq i$; it does *not* say that a Type i agent has a higher private value of good i than a Type $j \neq i$ agent does. For instance, it could be the case that $v_{A,A} = 2$, $v_{A,B} = 1$, $v_{B,A} = 3$, and $v_{B,B} = 4$. In this case, the required inequalities are satisfied even though Type B agents value *both* goods more than a Type A agent does; it does not matter as long as the original inequalities are satisfied.

One way of interpreting the condition $v_{i,i} > v_{i,j}$ for $j \neq i$ is to say that a Type i agent prefers a small network of Type i over a small network of Type j (specifically, networks of size 1). The other defining characteristic also involves network size; namely, a Type i agent should prefer a "large" Type i network over a "large" Type j network. This, and several other properties of the network function, are made precise in the next section.

3.2 The Network Effect Function

For simplicity, the network function e shall be written as a split function:

$$e(c, n_A, n_B) = \begin{cases} e_A(n_A, n_B) & \text{if } c = A \\ e_B(n_B, n_A) & \text{if } c = B \end{cases}.$$

This is primarily a notational convenience; since n_A and n_B are reversed as inputs in the two functions, only the general function $e_i(n_i, n_j)$, where $j \neq i$, must be characterized. Of course, e_A and e_B could easily be very distinct functions, but to fully capture a wide spectrum of possible network effects, certain conditions must be satisfied. While Swann (2002) examines conditions under which a network effect function may reasonably be assumed to be linear, no such restriction is needed here. Instead, five key properties are identified below. In all cases, assume that $i \neq j$.

(P1) $e_i(n_i, n_j)$ is weakly increasing in n_i and non constant in n_i.

(P2) $e_i(n_i, n_j)$ is weakly decreasing in n_j.

(P3) There exists $M_i \in \mathbb{R}$ such that for all $n_j \in \mathbb{N}$, $\lim_{n_i \to \infty} e_i(n_i, n_j) = M_i$. By (P1), this means that $e_i(n_i, n_j) \leq M_i$ for all $n_i, n_j \in \mathbb{N}$. Furthermore, it is assumed that $v_{i,i} + M_i > v_{i,j} + M_j$.

(P4) For every $n_j \in \mathbb{N}$, there exists $N \in \mathbb{N}$ such that for all $n_i \geq N$, $v_{i,i} + e_i(n_i, n_j) > v_{i,j} + e_j(n_j, n_i)$. (This is actually a consequence of (P3), but it is handy to state it as a separate property; see below).

(P5) There exists a nonempty set $S' \subseteq \mathbb{N}_+$ such that for all $n_i \in S = S' \cup \{0\}$, there exists $N \in \mathbb{N}$ for which $n_j \geq N$ implies $v_{i,j} + e_j(n_j, n_i) > v_{i,i} + e_i(n_i, n_j)$.

(P1) asserts that, given n_j, the value of network i increases as it grows, but that it is not necessarily strictly increasing in n_i; there could be constant regions. But it is also non constant (that is, it is not constant everywhere), which makes network size a nontrivial consideration. Similarly, (P2) states that network i (weakly) diminishes in value as network j increases in size. Since the relation is weak, however, this means that e_i could actually be constant in n_j, or at least have constant regions.

(P3) says that network i approaches some maximum value, M_i, as n_i increases, and that this maximum value is the same regardless of the value of n_j. The intuition is that even if network j is large, network i can eventually grow large enough to approach its maximum value anyway. The second assumption in (P3) ($v_{i,i} + M_i > v_{i,j} + M_j$) is the second defining characteristic of Type i agents: they prefer high-value good i networks over high-value good j networks.

(P4), a direct result of (P3), states that for any given $n_j \in \mathbb{N}$, Type i agents will eventually prefer network i as n_i increases. To see that (P3) $\Rightarrow$ (P4), fix any $n_j \in \mathbb{N}$ and let $\epsilon = (v_{i,i} + M_i) - (v_{i,j} + M_j) > 0$. Since $\lim_{n_i \to \infty} e_i(n_i, n_j) = M_i$, Definition

2.1.2 says there is some $N \in \mathbb{N}$ such that for every $n_i \geq N$, $|e_i(n_i, n_j) - M_i| < \epsilon$. Since $e_i(n_i, n_j) \leq M_i$, this means that $M_i - e_i(n_i, n_j) < \epsilon = (v_{i,i} + M_i) - (v_{i,j} + M_j)$; canceling M_i and rearranging gives $v_{i,j} + M_j < v_{i,i} + e_i(n_i, n_j)$ which, together with $e_j(n_j, n_i) \leq M_j$, yields $v_{i,i} + e_i(n_i, n_j) > v_{i,j} + e_j(n_j, n_i)$ for every $n_i \geq N$, and so (P4) is satisfied.

(P5) states that for some values of n_i and n_j, a Type i agent will find it optimal to adopt good j. But depending on the type of competition between networks, the set S of n_i values where such an n_j exists may be restricted. To rule out completely trivial models, attention is restricted to models where $1 \in S$ by imposing $S' \subseteq \mathbb{N}_+$. By (P1), it is also obvious that 0 must be in S as well. In fact, (P1) also implies that if $k \in S$, then $\ell \in S$ for every $0 \leq \ell \leq k$. Therefore, the union of all sets S satisfying (P5) is again a set satisfying (P5), and is the largest such set (formally, it is the $\subseteq$-maximal set satisfying (P5)). This set, denoted S_i, can be thought of as the "switching set" for Type i agents; it consists of all values of n_i for which it is possible for Type i agents to end up selecting good j. Depending on how the network functions behave, S_i could be all of $\mathbb{N}$, or it could be some proper subset of $\mathbb{N}$, in which case $S_i = \{0, 1, 2, \ldots, n\}$ for some n.

With these properties in place, it is possible to define some useful notation. Let $n_B \in \mathbb{N}$. Then $\alpha_A(n_B)$ denotes the smallest value of n_A for which $v_{A,A} + e_A(n_A, n_B) \geq v_{A,B} + e_B(n_B, n_A)$; that is, $\alpha_A(n_B)$ is the smallest value of n_A for which Type A agents are better off adopting good A. By (P4), there is at least one possible n_A for which this inequality is true; as $\mathbb{N}$ is well-ordered, this means a minimum such n_A exists, and so $\alpha_A(n_B)$ is certain to exist. Next, given $n_A \in S_A$, $\beta_A(n_A)$ denotes the minimum value of n_B for which $v_{A,B} + e_B(n_B, n_A) > v_{A,A} + e_A(n_A, n_B)$; that is, $\beta_A(n_A)$ is the smallest value of n_B for which Type A agents are better off adopting good B. Again, (P5) and the fact that $\mathbb{N}$ is well-ordered guarantees the existence of $\beta_A(n_A)$.

Similarly, for $n_A \in \mathbb{N}$, $\beta_B(n_A)$ is the smallest n_B for which Type B agents are better off adopting good B, and $\alpha_B(n_B)$ denotes the smallest n_A (given $n_B \in S_B$) for which

Type B agents are better off adopting good A.

It is straightforward to prove that $\alpha_A(n_B)$ is weakly increasing in n_B and that $\beta_A(n_A)$ is weakly increasing in n_A (provided $n_A \in S_A$). By symmetry, of course, this means that $\alpha_B(n_B)$ (for $n_B \in S_B$) and $\beta_B(n_A)$ are weakly increasing functions. It is also routine to verify that for all $n_B \in S_B$, $\alpha_B(n_B) \geq \alpha_A(n_B)$; similarly, for all $n_A \in S_A$, $\beta_A(n_A) \geq \beta_B(n_A)$. These facts play an important role several proofs.

3.2.1 Competition Between Networks

It is possible that growth in the size of one network diminishes the quality of the other. This is particularly fitting in cases involving fads: if agents are to choose between two different clothing styles, for example, the value of the network effect may depend only on which style is more widely adopted (in particular, agents may only be concerned about the total fraction of individuals who choose the same style they do). In such cases, the networks are said to *diminish* one another.

Definition 3.2.1. *Suppose $i, j \in \{A, B\}$ and that $i \neq j$. Then network j diminishes network i if there exists some $n_i \in \mathbb{N}$ and $n_j, n_j' \in \mathbb{N}$ with $n_j' > n_j$ such that $e_i(n_i, n_j') < e_i(n_i, n_j)$. Network j strictly diminishes network i if for every $n_i \in \mathbb{N}_+$ and all $n_j, n_j' \in \mathbb{N}$ with $n_j' > n_j$, it follows that $e_i(n_i, n_j') < e_i(n_i, n_j)$.*

So, network j diminishes network i only if (P2) acts non trivially (that is, if e_i is non constant in n_j, which means there is at least one region where e_i decreases strictly in n_j). Network j *strictly* diminishes network i if this is true for all values of $n_i \in \mathbb{N}_+$ and all values of n_j; the restriction to $n_i \in \mathbb{N}_+$ is taken to ensure that $e_i(\cdot) \in \mathbb{R}_+$, since one would often expect $e_i(0, 0) = 0$.

If both networks (strictly) diminish each other, then the networks are said to be *(strictly) competitive*; if they are not competitive (that is, if each function e_i is constant in n_j), then the networks are *independent.* Many types of network effects are independent;

consider, for example, two communications networks. If, other things equal, one more agent joins network A, then network A improves without *diminishing* the quality of network B. Independent networks also have a useful mathematical property, as illustrated in the following theorem:

Theorem 3.2.1. *If networks A and B are independent, then S_A and S_B are bounded.*

Proof. When the networks are independent, each e_i is constant in n_j, $j \neq i$. Since $\lim_{n_i \to \infty} e_i(n_i, n_j) = M_i$, regardless of n_j, and since $v_{i,i} + M_i > v_{i,j} + M_j$, there is some n_i^* for which $v_{i,i} + e_i(n_i^*, n_j) > v_{i,j} + M_j \geq v_{i,j} + e_j(n_j, n_i^*)$ for all $n_j \in \mathbb{N}$. Since $e_i(\cdot)$ is increasing in n_i, this inequality actually holds for all $n_i \geq n_i^*$. But this means a Type i player is always better off with good i than good j whenever $n_i \geq n_i^*$, making n_i^* an upper bound for S_i. □

An immediate consequence of Theorem 3.2.1 is that if each $S_i = \mathbb{N}$, then the networks must be competitive, for otherwise at least one network i would not be diminished by the other, resulting in a bounded S_i for that network. Note that if each $S_i = \mathbb{N}$, then the networks are competitive, but not necessarily strictly competitive; see Appendix A.1.

Intuitively, one might suspect that a bounded S_i puts limits on how likely coordination on one good is, because restricting S_i restricts the number of cases in which it is possible for a Type i agent to switch to good j. On the other hand, coordination might also seem less likely when the networks are competitive, because Type i agents can strengthen their own networks (and diminish the j network) simply by adopting good i. Of course, only one of these results may be valid; this issue will surface again in Chapter 4, but will not be resolved until Chapter 5.

It is worth noting that the converse of Theorem 3.2.1 does not hold; that is, a bounded S_i does not guarantee that the networks are independent. Also, one may conjecture that for strictly competitive networks, the S_i sets are unbounded (that is, $S_i = \mathbb{N}$). But this

is also false; see Appendix A.1 for examples.

Chapter 4

The Simultaneous Move Game

4.1 Nash Equilibria in Simultaneous Moves

The simultaneous move game assumes all of the properties defined in Chapter 3, including the five characteristics the network effect functions must satisfy, and also assumes that the set $\mathcal{X} = \{x_1, \ldots, x_N\}$ consists of a total population of $N = N_A + N_B$ agents, where N_i is the number of Type i agents. Decision are made simultaneously or, equivalently, each agent makes his decision without any knowledge of the choices of other agents. Each agent is assumed to be aware of his own private valuations ($v_{i,i}$ and $v_{i,j}$), as well as the network functions e_A and e_B and the preferences of all agents (hence the values of N_A and N_B).

In this framework, a *strategy profile* is simply an element of the set $\mathbb{S} = \{A, B\}^N$; so, a strategy profile $s = (s_1, \ldots, s_N) \in \mathbb{S}$ represents the choices of all N agents, with $s_i \in \{A, B\}$ denoting the choice of agent x_i. Given an agent $x_i \in \mathcal{X}$ and a profile $s \in \mathbb{S}$, the profile $s_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_N)$ refers to the strategies of all players *other* than agent i; for convenience, the notation (s_i, s_{-i}) refers to the entire profile s. Since $s \in \mathbb{S}$ consists of the choices of all players, the values of n_A and n_B may be inferred from s; let $n_A(s)$ and $n_B(s)$ represent the number of A and B choices in s, respectively. Then agent i's utility function $U(t_i, c_i, n_A, n_B)$ may be written as $U_i(t_i, n_A(s), n_B(s))$, or simply $U_i(t_i, s)$ when the context is clear. This permits the following (standard) definition of Nash Equilibrium:

Definition 4.1.1 (Nash Equilibrium). *Let $s \in \mathbb{S}$. Then s is a Nash equilibrium if, for every $0 \leq i \leq N$ and $s_i' \in \{A, B\}$, it happens that $U_i(t_i, (s_i, s_{-i})) \geq U_i(t_i, (s_i', s_{-i}))$.*

So, a strategy profile s is a Nash equilibrium if no individual agent can improve his outcome by changing his strategy, given the strategies of all other players. The first type of equilibrium in this game is when all players choose the same good. These are called "Pure A" or "Pure B" equilibria, depending on which good is chosen.

Theorem 4.1.1. *If $v_{B,A}+e_A(N_A+N_B,0) > v_{B,B}+e_B(1,N_A+N_B-1)$, then the strategy profile where all players choose good A is a Nash equilibrium in simultaneous moves. A symmetric statement holds for a pure B equilibrium.*

Proof. The given inequality guarantees that no individual agent of Type B will defect and choose good B, given that all other players select good A. Next, since $v_{B,B} > v_{B,A}$, the inequality also shows that $v_{B,B}+e_A(N_A+N_B,0) > v_{B,A}+e_A(N_A+N_B,0) > v_{B,B}+e_B(1,N_A+N_B-1)$, which implies that $e_A(N_A+N_B,0) > e_B(1,N_A+N_B-1)$. Combined with the fact that $v_{A,A} > v_{A,B}$, this gives $v_{A,A}+e_A(N_A+N_B,0) > v_{A,A}+e_B(1,N_A+N_B-1) > v_{A,B}+e_B(1,N_A+N_B-1)$, and so $v_{A,A}+e_A(N_A+N_B,0) > v_{A,B}+e_B(1,N_A+N_B-1)$. This inequality guarantees that no individual agent of Type A will choose good B, given that all other players select good A. Thus, a pure A equilibrium exists. The proof for a pure B equilibrium is similar. □

Theorem 4.1.1 is fairly obvious and, trivially, its converse is also true. Closer examination of the required inequalities, however, reveals that a pure equilibrium of either type exists if N_A and N_B are sufficiently large. This is demonstrated as a corollary to the above theorem.

Corollary 4.1.2. *If $N_A+N_B-1 \geq \alpha_B(1)$, then a pure A equilibrium exists. Similarly, if $N_A+N_B-1 \geq \beta_A(1)$, then a pure B equilibrium exists.*

Proof. First, note that by definition of α_B, the hypothesis gives $v_{B,A}+e_A(\alpha_B(1),1) > v_{B,B}+e_B(1,\alpha_B(1))$. Next, since $e_i(n_i,n_j)$ is weakly increasing in n_i and weakly decreasing

in n_j, this means that $v_{B,A} + e_A(N_A + N_B, 0) \geq v_{B,A} + e_A(N_A + N_B - 1, 0) \geq v_{B,A} + e_A(\alpha_B(1), 0) \geq v_{B,A} + e_A(\alpha_B(1), 1) > v_{B,B} + e_B(1, \alpha_B(1)) \geq v_{B,B} + e_B(1, N_A + N_B - 1)$, and so $v_{B,A} + e_A(N_A + N_B, 0) > v_{B,B} + e_B(1, N_A + N_B - 1)$. But this is exactly the inequality required by Theorem 4.1.1, so a pure A equilibrium exists. Naturally, the proof for a pure B equilibrium is similar. □

Corollary 4.1.2 demonstrates that in all but the most trivial cases, there is both a pure A equilibrium and a pure B equilibrium in the simultaneous move game. Specifically, the corollary reveals an important condition under which the network effect is actually interesting, because without it no Type i agent would ever adopt good j. Therefore, all models from this point forward are required to satisfy the following condition:

(NT) For all $i, j \in \{A, B\}$ with $j \neq i$, $v_{i,i}$, $v_{i,j}$, e_i, and N_i must satisfy $N_A + N_B - 1 \geq \alpha_B(1)$ and $N_A + N_B - 1 \geq \beta_A(1)$.

There is another interesting type of equilibrium which can occur in simultaneous moves: a "split" equilibrium where all agents simply adopt the good corresponding to their own type. Intuition might suggest that a split only occurs under weak network effects (for example, in models where (NT) is not satisfied), but this is not the case. The following theorem provides some general conditions under which a split equilibrium exists:

Theorem 4.1.3. *If $N_A - 1 \geq \alpha_A(N_B + 1)$ and $N_B - 1 \geq \beta_B(N_A + 1)$, then there exists a Nash equilibrium in simultaneous moves where all players adopt the good according to their own type.*

Proof. By definition, $N_A - 1 \geq \alpha_A(N_B + 1) \Rightarrow v_{A,A} + e_A(N_A - 1, N_B + 1) \geq v_{A,B} + e_B(N_B + 1, N_A - 1)$. Since $v_{A,A} + e_A(N_A, N_B) \geq v_{A,A} + e_A(N_A - 1, N_B + 1) \geq v_{A,B} + e_B(N_B + 1, N_A - 1)$, no agent of Type A will choose good B, given that all other players choose the good

according to their own type. Similarly, $N_B - 1 \geq \beta_B(N_A + 1) \Rightarrow v_{B,B} + e_B(N_B, N_A) \geq v_{B,A} + e_A(N_A + 1, N_B - 1)$, so no agent of Type B will choose good A, given that all other agents adopt the good corresponding to their own type. This gives a split equilibrium in simultaneous moves. □

Again, Theorem 4.1.3 is quite obvious; however, the hypothesis of Theorem 4.1.3 may not always be satisfied since $\alpha_A(\cdot)$ and $\beta_B(\cdot)$ are weakly increasing functions. So, for competitive network effects, the hypothesis can fail; see Appendix A.2 for examples. Independent networks, however, will always have a split equilibrium provided there are enough agents of each type. This is presented as a corollary to Theorem 4.1.3.

Corollary 4.1.4. *If the networks are independent and N_A and N_B are sufficiently large, then there is a split equilibrium in simultaneous moves.*

Proof. Since the networks are independent, Theorem 3.2.1 asserts that S_A and S_B are bounded sets; in particular, there are constants $s_A, s_B \in \mathbb{N}$ such that $S_A = \{0, \ldots, s_A\}$ and $S_B = \{0, \ldots, s_B\}$. So, if $N_A \geq \min(\mathbb{N} \backslash S_A)$ and $N_B \geq \min(\mathbb{N} \backslash S_B)$ then, by definition of S_i, all agents will choose the good corresponding to their own type, and so a split equilibrium exists. □

Having established the existence of multiple equilibria, it is worthwhile to ask which, if any, are Pareto efficient. The easiest way to test for Pareto efficiency in this model is to use the *Pareto dominance* criterion; this is defined next.

Definition 4.1.2. *Let $s, s' \in \mathbb{S}$. If $U_i(t_i, s') \geq U_i(t_i, s)$ for every $1 \leq i \leq N$ and $U_i(t_i, s') > U_i(t_i, s)$ for at least one i, then s' Pareto dominates s; if $p \in \mathbb{S}$ is not dominated by any other profile, then p is Pareto efficient.*

A surprising fact is that under (P1)-(P5) and (NT), none of the equilibria identified thus far are guaranteed to be Pareto efficient; see Appendix A.2 for examples. If, however,

each e_i is strictly increasing in n_i, then at least one pure equilibrium is efficient. Before stating and proving this fact, a useful lemma is required.

Lemma 4.1.5. *If e_A is strictly increasing in n_A and e_B is strictly increasing in n_B, then any profile $s \in \mathbb{S}$ with $n_A(s) > 0$ and $n_B(s) > 0$ does not Pareto dominate any pure profile $(i) = (i, i, \ldots, i)$ for $i \in \{A, B\}$.*

Proof. In the profile (i), the payoff to agents of Type i is $v_{i,i} + e_i(N, 0)$, and the payoff for Type $j \neq i$ agents is $v_{j,i} + e_i(N, 0)$. Now, suppose $s \in \mathbb{S}$ with $n_A(s) > 0$ and $n_B(s) > 0$. There are two cases. First, if at least one Type i agent has chosen i in the profile s, then this agent receives utility equal to $v_{i,i} + e_i(n_i(s), n_j(s)) < v_{i,i} + e_i(N, n_j(s)) \leq v_{i,i} + e_i(N, 0)$, so this agent is worse off under s than under (i), and so s does not Pareto dominate (i) in this case. On the other hand, if all Type i agents choose j in s, then at least one Type j agent chooses i in s, because $n_i(s) > 0$. Then this Type j agent receives utility equal to $v_{j,i} + e_i(n_i(s), n_j(s)) < v_{j,i} + e_i(N, n_j(s)) \leq v_{j,i} + e_i(N, 0)$, so he is also worse off under s. Hence s does not Pareto dominate (i). □

Lemma 4.1.5 demonstrates that a pure equilibrium cannot be dominated by an "interior" profile where some agents choose A and some choose B. This is essential for demonstrating that at least one pure equilibrium is Pareto efficient; this is made precise in the next theorem.

Theorem 4.1.6. *Suppose e_A is strictly increasing in n_A and e_B is strictly increasing in n_B. Then at least one pure equilibrium is Pareto efficient.*

Proof. Suppose (A) is not efficient. It suffices to show that (B) is efficient. Since (A) is not efficient, this means (A) is dominated by some other profile $s \in \mathbb{S}$; in particular, $s \neq (A)$ implies that either $s = (B)$ or $n_A(s) > 0$ and $n_B(s) > 0$. If $n_A(s) > 0$ and $n_B(s) > 0$, then (by Lemma 4.1.5) s does not dominate (A), and therefore it must be the

case that $s = (B)$ dominates (A). But this means (B) is efficient, for any $s' \in \mathbb{S}$ is either (A) (which (B) dominates) or $n_A(s') > 0$ and $n_B(s') > 0$, which (by Lemma 4.1.5) means s does not dominate (B). Thus no profile $s' \in \mathbb{S}$ dominates (B), so (B) is efficient. □

Theorem 4.1.6 guarantees that *at least* one of the pure equilibria is efficient in any given model, but are they *both* efficient? Indeed, it is possible that they are, but it turns out that even with strictly increasing network functions, one of the equilibria *may* be dominated by the other; see Appendix A.2 for examples. Similar issues arise for split equilibria: sometimes the split outcome is efficient, but one or both of the pure equilibria may dominate it. Again, see Appendix A.2 for an example.

The above analysis suggests that there is a difficult equilibrium selection (coordination) problem to be solved. There are many possible equilibria, and even with the non triviality condition, criteria for the existence of the qualitatively different types of outcomes are similar (namely, large N_i). Furthermore, the equilibria are not guaranteed to be Pareto efficient, so this does not help solve the selection problem; indeed, which equilibria are optimal is highly circumstantial and may vary under changes to the population composition or the relative value of the networks. Finally, the simultaneous move game may not be a realistic choice problem, because quite often agents are able to observe the actions of some of the other agents prior to making their own decision. For this reason, two different types of sequential move games are considered in Chapter 5; these allow much more to be said about the likelihood of the various outcomes.

Chapter 5

Sequential Choice Games

The elementary components of the sequential move game are the same as those of the simultaneous game; there is a set $\mathcal{X}$ of $N = N_A + N_B$ agents who must choose between good A or good B, agents of type i have private valuations satisfying $v_{i,i} > v_{i,j}$ for $i \neq j$, and there are network functions e_A and e_B satisfying (P1)-(P5) and (NT). Of course, choices are now made sequentially, so the order in which agents appear is relevant. A bijection $\omega : \{1, \ldots, N\} \rightarrow \mathcal{X}$ is called a *permutation*, or *ordering*, of the agents, and may be represented by an N-tuple $\omega = (\omega_1, \ldots, \omega_N)$, where ω_i is shorthand for $\omega(i)$. The collection Ω consists of all possible permutations ω; obviously there are $N!$ permutations in Ω.

Exactly which permutation is realized is assumed to be exogenous; letting agents choose when to make their decisions opens up some interesting possibilities, but this thesis does not address them. Instead, the emphasis is on deriving comparative static results on the likelihood of eventual coordination among the agents, which requires an analysis of which, and how many, permutations result in coordination. Since the object of interest is a likelihood or probability, a *random* determination of which permutation is realized is required. However, no assumption about the probability distribution $\Gamma : \Omega \rightarrow [0, 1]$ imposed on Ω is needed to achieve the theoretical results in section 5.3, but for simplicity a uniform distribution is used in numerical examples and algorithms.

Rather than appealing to standard equilibrium concepts such as subgame perfect Nash equilibrium (requiring perfect information) or variations on sequential equilibrium notions (requiring arbitrary beliefs to support equilibrium strategies), the problem is simply viewed as one of dynamic choice. Agents are assumed to know the total number

of agents, N, but not the values of N_A and N_B. So, given an ordering $\omega \in \Omega$, each agent observes the choices of all previous agents (formally, the i^{th} agent observes the choices of all agents in the set $\{x \in \mathcal{X} \mid \omega^{-1}(x) < i\}$) and uses this information, along with his own type, to form an expectation about the population composition and future decisions. From this information, agents compute their expected utility from selecting A or B, and simply act to maximize their expected utility (as a convention, a Type i agent will adopt good i if the expected utilities to choosing A and B are equal). From this, one may deduce which permutations result in coordination on good A, good B, or neither, allowing the probability of such outcomes to be unambiguously determined.

The tricky bit is in deciding *how* agents form their expectations. This is a matter of taste, and subject to a variety of considerations contingent on what is being modeled. This thesis contrasts two opposing cases: the "sophisticated" case (requiring $e_{A,}$, e_B, $v_{A,A}$, $v_{A,B}$, $v_{B,A}$, and $v_{B,B}$ to be known by *all* agents), and the "naïve" case (requiring Type i agents only to be aware of e_A, e_B, $v_{i,i}$, and $v_{i,j}$). These models, as well as their motivation, are discussed in sections 5.1 and 5.2, respectively, and generalized comparative static results for both cases (including "mixed" models which have some sophisticated and some naive agents) are presented in section 5.3. Finally, numerical examples are discussed in section 5.4.

5.1 Sequential Decision Making for Sophisticated Agents

As noted above, sophisticated agents are aware of most relevant parameters in the model (N, $e_{A,}$, e_B, $v_{A,A}$, $v_{A,B}$, $v_{B,A}$, and $v_{B,B}$), but they do not know the values of N_A and N_B. They are also unaware of which order $\omega \in \Omega$ the set $\mathcal{X}$ is given, so that they cannot always anticipate future decisions perfectly. Instead, each agent k observes the decisions already made, and from this determines the values A_k and B_k of A and B adoption

decisions made prior their own. Since sophisticated agents know the private valuations of both player types, they are able to determine how different *types* will respond to their own decision. That is, agent k, given A_k and B_k, decides the values of A_{k+1} and B_{k+1}, and uses this to anticipate how agent $k+1$ (and subsequent agents) will behave. This requires two constructions: first, agent k must form a probability distribution over what sequence of *types* (not decisions) will occur after him; second, he must be able to evaluate each of these sequences to compute the final values of n_A and n_B, given his own decision and the fact that subsequent agents behave optimally. In this way, agent k may compute the expected utility to choosing A or B (denoted $EU_k(A)$ and $EU_k(B)$), and acts to maximize his expected utility.

For a given agent k, a sequence of types for the remaining $N-k$ agents is called a *form*, and is simply a member of the set $\mathcal{F}_k = \{A, B\}^{N-k}$. To help distinguish forms from other mathematical objects, they will be represented using angled brackets; for example, $\langle A, B, B\rangle \in \mathcal{F}_{N-3}$. To set a probability distribution over the set $\mathcal{F}_k$, agent k uses the values A_k and B_k, together with his own type, to estimate the probability of an agent being Type A or Type B; in particular, if agent k is Type A, then $P_k(A) = \frac{A_k+1}{k}$ and $P_k(B) = \frac{B_k}{k}$; if he is Type B, then the probabilities are $P_k(A) = \frac{A_k}{k}$ and $P_k(B) = \frac{B_k+1}{k}$. Finally, let $t_A(F)$ and $t_B(F)$ denote the number of Type A and Type B agents, respectively, in the form $F \in \mathcal{F}_k$. Then agent k will assign probability $P_k(F) = P_k(A)^{t_A(F)} \cdot P_k(B)^{t_B(F)}$ to the form F.

Before proceeding further, some explanation of these probabilities is needed. The probabilities $P_k(A)$ and $P_k(B)$, by construction, only give consistent estimates for the likelihood of player types if each player chooses the good corresponding to his own type. But this need not be the case; once sufficiently many Type A agents have selected good A, for example, a Type B agent may find it optimal to choose good A also. But then any subsequent agents will have inconsistent probabilities $P_k(A)$ and $P_k(B)$. The difficulty

here is that subsequent agents have no way of knowing if everyone else who chose A is Type A or not; it could be the case that many consecutive Type A agents appeared, and only they selected A's, or it could be the case that some Type B agents along the way selected A as well. The given probabilities, however, will result in consistent *behaviour*, in that an agent of Type i will only adopt good $j \neq i$ if all subsequent agents also adopt j.

Next, for every $F \in \mathcal{F}_k$, let $A(F|c)$ and $B(F|c)$ denote the number of A and B *choices* in F, given that agent k selects $c \in \{A, B\}$. Only by computing $A(F|c)$ and $B(F|c)$ for each $c \in \{A, B\}$ and $F \in \mathcal{F}_k$ can agent k determine his expected utility from choosing A or B. But determining these values requires a dynamic approach. Let $\underline{A}_N$ be the smallest value of A_N for which agent N will select good A, regardless of type. This is equivalent to the smallest value of A_N for which a Type B agent in position N will select A, since a Type A agent will also select A if a Type B agent does. So, $\underline{A}_N$ is the smallest A_N satisfying

$$v_{B,A} + e_A(A_N + 1, B_N) > v_{B,B} + e_B(B_N + 1, A_N) \ .$$

Of course, it is also possible to define a parameter $\underline{B}_N$, which is the smallest value of B_N for which agent N of Type A will select good B. It follows that agent N will choose A if $A_N \geq \underline{A}_N$, B if $B_N \geq \underline{B}_N$, and t_N (his own type) otherwise. But agent $N-1$ is aware of the values $\underline{A}_N$ and $\underline{B}_N$, and since agent $N-1$ (given A_{N-1} and B_{N-1}) determines the values of A_N and B_N through his own actions, he is able to evaluate $A(F|c)$ and $B(F|c)$ for $c \in \{A, B\}$ and $F \in \mathcal{F}_{N-1} = \{\langle A \rangle, \langle B \rangle\}$. This means he can compute $EU_{N-1}(A)$ and $EU_{N-1}(B)$, and thereby the values $\underline{A}_{N-1}$ and $\underline{B}_{N-1}$ may be defined similarly to $\underline{A}_N$ and $\underline{B}_N$. Then agent $N-2$ is aware of the values $\underline{A}_{N-1}$ and $\underline{B}_{N-1}$ and, by a similar process, this allows $\underline{A}_{N-2}$ and $\underline{B}_{N-2}$ to be defined. In general, agent $k < N$ uses the values $\underline{A}_\ell$ and $\underline{B}_\ell$ for $\ell > k$ to evaluate $A(F|c)$ and $B(F|c)$ for each $F \in \mathcal{F}_k$, and in this way computes $EU_k(A)$ and $EU_k(B)$ in order to make an adoption decision. Note that

since $A_k + B_k + 1 = k$, $\underline{A}_k$ and $\underline{B}_k$ may not be defined for small values of k; in that case, they will simply be set to ∞, with the understanding that no history of choices are sufficient for agent number k to switch his choice.

Given $1 \leq k \leq N$, let $\mathbb{N}_k^2 = \{(p,q) \in \mathbb{N}^2 \mid p+q+1 = k\}$. Then, agent k's *choice function* $\mathfrak{c}_k : \{A,B\} \times \mathbb{N}_k^2 \to \{A,B\}$ is given by

$$\mathfrak{c}_k(t_k, A_k, B_k) = \begin{cases} A & \text{if } A_k \geq \underline{A}_k \\ B & \text{if } B_k \geq \underline{B}_k \\ t_k & \text{otherwise} \end{cases} .$$

Note that, given A_k, B_k, and $\underline{A}_{k+1}, \ldots, \underline{A}_N$, $EU_k(A)$ is

$$EU_k(A) = v_{t_k,A} + \sum_{F \in \mathcal{F}_k} P_k(F) e_A(A_k + 1 + A(F|A), B_k) ,$$

and that $EU_k(B)$ may be defined similarly. $\underline{A}_k$ is therefore the smallest value of A_k for which $EU_k(A) > EU_k(B)$, and $\underline{B}_k$ is the smallest value of B_k for which $EU_k(B) > EU_k(A)$.

With this notation in place, algorithms may be devised to compute how many permutations of the N agents result in an A cascade, a B cascade, or a split outcome. An A cascade is simply an ordering $\omega \in \Omega$ in which at least one Type B agent chooses A, because then all subsequent agents choose A. B cascades are defined similarly. The notion of a cascade, however, relies on the property that if a Type i agent chooses $j \neq i$, then *all* subsequent agents will choose j as well. This is inarguably the most important property for a network effect model to have. Let $(N_i) = (N_A, N_B)$, $(v_{i,j}) = (v_{A,A}, v_{A,B}, v_{B,A}, v_{B,B})$, $(e_i) = (e_A, e_B)$, and $(\mathfrak{c}_i) = (\mathfrak{c}_1, \ldots, \mathfrak{c}_N)$ be a vector of choice functions (not necessarily the sophisticated choice functions described thus far). Then the collection $\mathcal{M} = \langle (N_i), (v_{i,j}), (e_i), (\mathfrak{c}_i) \rangle$ is a *model*, for it contains all relevant information for constructing a sequential choice model. The Cascade Property is therefore defined as follows:

Definition 5.1.1 (The Cascade Property). *A model* $\mathcal{M} = \langle (N_i), (v_{i,j}), (e_i), (\mathfrak{c}_i) \rangle$ *satisfies*

the Cascade Property if for every permutation $\omega \in \Omega$, *every* $i, j \in \{A, B\}$ *with* $i \neq j$, *and every* $1 \leq k \leq N$ *with* $t_k = i$, *it happens that* $\mathfrak{c}_k(i, A_k, B_k) = j \Longrightarrow \mathfrak{c}_\ell(t_\ell, A_\ell, B_\ell) = j$ *for every* $k \leq \ell \leq N$.

Luckily, any model with sophisticated agents (that is, any model where the functions $(\mathfrak{c}_i)$ are those given above) will satisfy the Cascade Property.

Theorem 5.1.1. *Let* $\mathcal{M}$ *be a model with sophisticated agents. Then* $\mathcal{M}$ *satisfies the cascade property.*

At first glance, the Cascade Property seems like a trivial matter to verify, for intuition strongly suggests that it must be true. If, for example, one Type B agent optimally selects A, then his expected payoff to choosing B must be fairly bleak compared to choosing A. A subsequent Type B agent will then have an even worse expectation from choosing B, because now one more A choice is locked in (diminishing the quality of a B cascade if one were to occur), and his subjective probability of a Type B agent occurring is also lower, so that the probability of a Type B cascade is also diminished. These statements are much more difficult to verify rigorously, however, so the (lengthy) proof is given in Appendix A.3 instead.

As mentioned previously, subjective probabilities are inconsistent once, for example, a Type B agent selects A. But if all agents are using these probabilities, then their behaviour will be consistent due to the Cascade Property; the first Type B agent to select good A will only do so if he finds it optimal to do so, and he assumes that all others are using the same decision procedure, so that by selecting A he is guaranteeing for himself that all subsequent agents select A as well, regardless of their types, and regardless of how probable those types are. Formally, a consequence of the Cascade Property is that $\underline{A}_k + 1 \geq \underline{A}_{k+1}$. So, by using this decision algorithm, sophisticated agents are able to properly deduce outcomes, even though their estimates of the population composition

will be inconsistent.

With the Cascade Property in place, it is now possible to formulate algorithms for finding out how many permutations result in A cascades, B cascades, or split outcomes. This is detailed in the next section.

5.1.1 Pure Cascades, General Cascades, and Split Outcomes for Sophisticated Agents

A *pure cascade* is one in which every agent chooses the same good. This possibility, of course, is motivated by the equilibrium results in the simultaneous move game. Given that sophisticated agents satisfy the Cascade Property, counting the number of permutations which result in a pure A cascade, for instance, is fairly simple. All that is required is to find out how many leading Type A agents are needed to ensure that a subsequent Type B agent will select A. Let A^* be the minimum number of leading Type A agents required to cause an A cascade (formally, A^* is the smallest integer k so that $k \geq \underline{A}_{k+1}$). Then, by the Cascade Property, *any* combination of agent types may follow the initial A^* Type A agents, so that the number of *permutations* which result in a pure A cascade is

$$ {}_{N_A}P_{A^*}(N - A^*)! = \binom{N_A}{A^*} A^*!(N - A^*)! = \frac{N_A!(N - A^*)!}{(N_A - A^*)!} \, . \tag{5.1} $$

The term ${}_{N_A}P_{A^*}$ gives the number of ways to select and arrange the initial A^* Type A agents, given that there are N_A Type A agents available to choose from, and $(N - A^*)!$ counts the number of ways to arrange all remaining agents.

Some caution is needed, however, because this construction assumes that 1) the value A^* exists (that is, a value k exists so that if the first k agents choose A, then all subsequent agents will choose A), and 2) $A^* \leq N_A$. Whether these considerations hold or not depends on *all* parameters of the model. It is therefore necessary to restrict attention to models where cascades are actually possible. More precisely, attention must be restricted to the

set $\mathbb{M}(N)$ of models $\langle (N_i), (v_{i,j}), (e_i), (\mathfrak{c}_i) \rangle$ where $N_A + N_B = N$ and both types of pure cascades are possible. Of course, every model $\mathcal{M} \in \mathbb{M}(N)$ also satisfies (P1)-(P5) and (NT).

Given that expression 5.1 yields the number of permutations resulting in a pure A cascade, and assuming a uniform distribution over the $N!$ possible permutations, the probability of a pure A cascade is given by

$$\frac{N_A!(N_A + N_B - A^*)!}{(N_A + N_B)!(N_A - A^*)!} \,. \tag{5.2}$$

Using this expression, simple comparative static results may be derived. For example, routine algebra verifies that the probability is strictly decreasing in A^*, strictly increasing in N_A, and strictly decreasing in N_B. This is intuitive, because if more leading A's are required for a pure A cascade, then one would expect such a cascade to be less likely. Similarly, if the proportion of Type A to Type B agents increases (decreases), then one would expect the probability of a pure A cascade to increase (decrease). Comparative statics are treated separately in section 5.3, so at present it is best to think of expressions 5.1 and 5.2 simply as algorithms for computing the probability of a pure A cascade in a given model. Naturally, similar formulas may be derived for pure B cascades.

A *general* A (or B) cascade is simply an A (or B) cascade; the term *general* is used to distinguish this case from pure cascades, although every pure cascade obviously qualifies as a general cascade. For this reason, no modification to the definition of $\mathbb{M}(N)$ is needed, because guaranteeing the existence of pure cascades trivially guarantees the existence of general cascades as well. The procedure for counting the number of general cascades, however, is considerably more involved than that for pure cascades. Once again, the algorithm will be outlined for the case of A cascades; the procedure for B cascades is analogous.

Given a model $\mathcal{M} \in \mathbb{M}(N)$, let $\mathcal{F}_A \subseteq \{A, B\}^N$ denote the set of forms which result

in an A cascade. Then F may be represented by the sequence

$$F = \langle (a_1), (b_1), (a_2), (b_2), \ldots, (a_{n-1}), (b_{n-1}), (a_n), * \rangle \ ,$$

where (a_i) represents a sequence of a_i consecutive A's, (b_i) represents a sequence of b_i consecutive B's, and $*$ represents any combination of A's and B's so that F has length N. Some restrictions are needed for this representation to be useful. In particular, each b_i is strictly greater than zero, and for every $1 < i \leq n$, $a_i > 0$ as well. But $a_1 \geq 0$ since $a_1 = 0$ is needed to describe forms which begin with a B. In addition, it is assumed that the final segment of A's, (a_n), is the segment which causes an A cascade, and that given (a_i) and (b_i) for $i < n$, a_n is the exact *minimum* number of additional Type A agents needed to cause the cascade. To ensure that the outcome is not a split outcome, it must also be the case that $\sum_{i=1}^{n-1} b_i < N_B$; in this way, at least one Type B agent will choose A. There are also only N_A Type A agents available, so $\sum_{i=1}^{n} a_i \leq N_A$ is needed as well.

Since (a_n) is the segment which causes an A cascade, all prior segments may not cause cascades of either type. In particular, a_1 is restricted so that $a_1 < \underline{A}_{a_1+1}$, b_1 is restricted so that $b_1 < \underline{B}_{a_1+b_1+1}$, a_2 is restricted to $a_1 + a_2 < \underline{A}_{a_1+b_1+a_2+1}$, and so on. In general, a_i must satisfy $\sum_{j=1}^{i} a_j < \underline{A}_{\sum_{j=1}^{i}(a_j+b_j)-b_i+1}$ and b_i must satisfy $\sum_{j=1}^{i} b_j < \underline{B}_{\sum_{j=1}^{i}(a_j+b_j)+1}$. Finally, (a_n), being the segment which causes an A cascade, must satisfy $\sum_{j=1}^{n} a_j = \underline{A}_{\sum_{j=1}^{n-1}(a_j+b_j)+a_n+1}$. A segment (a_i) for $i < n$ is said to be *maximal* if $\sum_{j=1}^{i} a_j = \underline{A}_{\sum_{j=1}^{i}(a_j+b_j)+1} - 1$; that is, (a_i) is maximal if adding one more Type A agent to it causes an A cascade. Maximal (b_i) segments are defined similarly. Note that $\mathcal{F}_A$ contains all possible forms which result in an A cascade, and therefore will contain forms which have maximal A segments, maximal B segments, or both, provided that non-pure cascades exist.

Given a form $F \in \mathcal{F}_A$, how many *permutations* in Ω will satisfy F? Letting $\mathrm{Perm}_A(F) \subseteq \Omega$ be the set of permutations satisfying form F and $|\mathrm{Perm}_A(F)|$ the number of permu-

tations in this set, observe that there are ${}_{N_A}P_{a_1}$ ways of selecting and arranging the first segment (a_1), there are ${}_{N_B}P_{b_1}$ ways of selecting and arranging the segment (b_1), there are ${}_{N_A-a_1}P_{a_2}$ ways of selecting and arranging the segment (a_2), and so on, until finally there are $\left(N-\sum_{i=1}^{n-1}(a_i+b_i)-a_n\right)!$ ways of arranging all remaining agents after the segment (a_n). Therefore, $|\mathrm{Perm}_A(F)|$ is equal to

$$\begin{aligned}&\left({}_{N_A}P_{a_1}\right)\left({}_{N_B}P_{b_1}\right)\left({}_{N_A-a_1}P_{a_2}\right)\left({}_{N_B-b_1}P_{b_2}\right)\left({}_{N_A-(a_1+a_2)}P_{a_3}\right)\left({}_{N_B-(b_1+b_2)}P_{b_3}\right)\cdots \\ &\quad\cdots\left({}_{N_B-\sum_{i=1}^{n-2}b_i}P_{b_{n-1}}\right)\left({}_{N_A-\sum_{i=1}^{n-1}a_i}P_{a_n}\right)\left(N-\sum_{i=1}^{n-1}(a_i+b_i)-a_n\right)!\,.\end{aligned} \tag{5.3}$$

Noting that ${}_nP_k=(n)(n-1)\ldots(n-k+1)$ and rearranging the terms, expression 5.3 becomes

$$(N_A)\ldots\left(N_A-\sum_{i=1}^{n}a_i+1\right)\cdot(N_B)\ldots\left(N_B-\sum_{i=1}^{n-1}b_i+1\right)\left(N-\sum_{i=1}^{n-1}(a_i+b_i)-a_n\right)!\,.$$

Simplifying this expression yields

$$|\mathrm{Perm}_A(F)|=\frac{N_A!N_B!\left(N-\sum_{i=1}^{n-1}(a_i+b_i)-a_n\right)!}{\left(N_A-\sum_{i=1}^{n}a_i\right)!\left(N_B-\sum_{i=1}^{n-1}b_i\right)!}\,. \tag{5.4}$$

With the function $|\mathrm{Perm}_A(F)|$ in place, it is possible to compute the total number of permutations which result in an A cascade. Letting

$$\mathcal{A}=\bigcup_{F\in\mathcal{F}_A}\mathrm{Perm}_A(F)$$

denote the set of all permutations for which an A cascade occurs, and noting that $\mathrm{Perm}_A(F)\cap\mathrm{Perm}_A(F')=\emptyset$ whenever $F\neq F'$, it follows that

$$|\mathcal{A}|=\sum_{F\in\mathcal{F}_A}|\mathrm{Perm}_A(F)|\,,$$

and so $\frac{|\mathcal{A}|}{N!}$ is the probability of a general A cascade occurring. Naturally, the sets $\mathcal{F}_B\subseteq\{A,B\}^N$, $\mathrm{Perm}_B(F)\subseteq\Omega$, and $\mathcal{B}\subseteq\Omega$ may be defined similarly, so that $|\mathrm{Perm}_B(F)|$ and $|\mathcal{B}|$ may also be determined.

For split outcomes, it is possible to characterize all forms in a manner similar to that for general cascades. But since the only possible outcomes are general A cascades, general B cascades, or split outcomes, it is easier to simply compute the number of splits as $|\mathcal{S}| = N! - |\mathcal{A}| - |\mathcal{B}|$, and so the probability of a split is just $\frac{|\mathcal{S}|}{N!}$.

While it was possible to deduce some comparative static results for the probability of a pure cascade in a combinatorial manner, there is little hope of doing so for general cascades; the dynamics are more involved, and the distinction between the number of *forms* versus the number of *permutations* adds another layer of complexity in computing how the probability reacts to various parameter modifications. A higher-level approach is needed; before delving into this, however, a similar framework is developed for naïve agents.

5.2 Sequential Decision Making for Naïve Agents

In many ways, agents in the sophisticated model are very "ideal"; they have perfect knowledge of the private valuations of *both* types of players, and they use this information to anticipate how *all* remaining agents will respond to their decision. But the informational assumption is questionable in many applications, and the notion that each agent performs a complex series of dynamic computations is troublesome as well, especially when N is large. Sophisticated agents use these calculations to recognize whether or not they are in a cascade (or about to start one through their own decisions); naïve agents, on the other hand, do not recognize if they are in a cascade (or about to start one). Instead, they take their subjective probabilities of player types as given properties of the world, and use these probabilities to examine the likelihood of different outcomes (that is, the different final values of n_A and n_B) using the binomial distribution. Once again, these subjective probabilities will be inconsistent once an agent chooses to switch

networks. The difference now is that naïve agents will not even correctly identify all outcomes resulting from their actions, because they do not recognize cascades. The term naïve, then, is quite fitting.

5.2.1 Utility Maximization for Naïve Agents

Suppose agent i ($1 \leq i \leq N$) has observed A_i agents select A, and B_i agents select B. What is agent i to do? Letting $P_i(A)$ denote the subjective probability of an agent being Type A (from agent i's perspective), it is clear that if agent i is Type A, then $P_i(A) = \frac{A_i+1}{i}$ and that $P_i(B) = 1 - P_i(A) = \frac{B_i}{i}$; if i is Type B, then $P_i(A) = \frac{A_i}{i}$ and $P_i(B) = \frac{B_i+1}{i}$. These probabilities can be used to form a binomial distribution on the set of possible outcomes, for after i makes his adoption decision, there are only $N - A_i - B_i - 1 = N - i$ agents remaining. Then, if $t_i \in \{A, B\}$ denotes agent i's type, his expected payoff from choosing A (denoted $EU_i(A \mid t_i, N, A_i, B_i)$) is

$$v_{t_i,A} + \sum_{j=0}^{N-i} \underbrace{\binom{N-i}{j} P_i(A)^j P_i(B)^{N-i-j}}_{\text{Prob. of } j \text{ additional } A\text{'s}} \underbrace{e_A(A_i + 1 + j, N - A_i - 1 - j)}_{\text{Utility derived when } n_A = A_i + 1 + j} . \tag{5.5}$$

Similarly, the expected payoff from choosing B, denoted $EU(B \mid t_i, N, A_i, B_i)$, is

$$v_{t_i,B} + \sum_{j=0}^{N-i} \binom{N-i}{j} P_i(B)^j P_i(A)^{N-i-j} e_B(B_i + 1 + j, N - B_i - 1 - j) . \tag{5.6}$$

In this way, all agents can make their decision by comparing $EU(A \mid t_i, N, A_i, B_i)$ and $EU(B \mid t_i, N, A_i, B_i)$ (by convention, a Type i agent will adopt good i if the two quantities happen to be the same). Like the sophisticated case, variables $\underline{A}_i$ and $\underline{B}_i$ may be defined as the smallest values of A_i and B_i which will cause agent i to choose A or B, respectively, regardless of type. So, for instance, $\underline{A}_i$ is the smallest value of A_i for which $EU(A \mid B, N, A_i, B_i) > EU(B \mid B, N, A_i, B_i)$; noting that $B_i = i - 1 - A_i$, this means

$\underline{A}_i$ is the smallest value of A_i for which

$$v_{B,A} + \sum_{j=0}^{N-i} \binom{N-i}{j} P_i(A)^j P_i(B)^{N-i-j} e_A(A_i + 1 + j, N - A_i - 1 - j) > \tag{5.7}$$
$$v_{B,B} + \sum_{j=0}^{N-i} \binom{N-i}{j} P_i(B)^j P_i(A)^{N-i-j} e_B(i - A_i + j, N - i + A_i - j).$$

Of course, $\underline{B}_i$ may be defined analogously. This allows the *choice function* $\mathfrak{c}_k : \{A, B\} \times \mathbb{N}_k^2 \to \{A, B\}$ for naïve agents to be given by

$$\mathfrak{c}_k(t_k, A_k, B_k) = \begin{cases} A & \text{if } A_k \geq \underline{A}_k \\ B & \text{if } B_k \geq \underline{B}_k \\ t_k & \text{otherwise} \end{cases},$$

and so the same combinatorial approach to pure and general cascades will work, provided the choice function $\mathfrak{c}$ for naïve agents satisfies the Cascade Property. Fortunately, it does:

Theorem 5.2.1 (Cascade Theorem for Naïve Agents). *Let $\mathcal{M} \in \mathbb{M}(N)$ be a model with naïve agents. Then $\mathcal{M}$ satisfies the cascade property.*

Like the sophisticated case, intuition strongly suggests that this theorem must be true, but it is actually a nontrivial matter to prove it rigorously. Due to length and technicality considerations, the formal proof is given in Appendix A.4.

5.2.2 Pure Cascades, General Cascades, and Split Outcomes for Naïve Agents

Like the sophisticated case, the number of permutations which result in a pure A cascade is given by

$${}_{N_A}P_{A^*}(N - A^*)! = \binom{N_A}{A^*} A^*!(N - A^*)! = \frac{N_A!(N - A^*)!}{(N_A - A^*)!},$$

where A^* is the smallest integer k so that $k \geq \underline{A}_{k+1}$. The only difference now is the specific condition this implies. Note that for the first Type B agent after k initial

Type A's, $P_k(A) = \frac{k}{k+1}$ and $P_k(B) = \frac{1}{k+1}$. Thus $P_k(A)^j P_k(B)^{N-k-1-j} = \frac{k^j}{(k+1)^{N-k-1}}$ and $P_k(B)^j P_k(A)^{N-k-1-j} = \frac{k^{N-k-1-j}}{(k+1)^{N-k-1}}$. For naïve agents, therefore, A^* is the smallest integer k so that

$$v_{B,A} + \sum_{j=0}^{N-k-1} \binom{N-k-1}{j} \frac{k^j}{(k+1)^{N-k-1}} e_A(k+1+j, N-k-1-j) > \quad (5.8)$$
$$v_{B,B} + \sum_{j=0}^{N-k-1} \binom{N-k-1}{j} \frac{k^{N-k-1-j}}{(k+1)^{N-k-1}} e_B(1+j, N-1-j).$$

For general cascades, the counting procedure for sophisticated agents will also work for naïve agents; the only difference is that the values $\underline{A}_i$ and $\underline{B}_i$ satisfy different equations here than they do in the sophisticated case. For naïve agents, then, the set $\mathcal{A}$ of permutations resulting in A cascades also originates from the sets $\text{Perm}_A(F)$ for every $F \in \mathcal{F}_A$, with the understanding that agents now use the binomial distribution.

These combinatorial properties suggest that the naïve and sophisticated models, despite differences in informational and behavioural assumptions, behave similarly to one another. In the next section, the Cascade Property is shown to be the vital link between the two, and that this property ensures both models will have the same comparative static properties.

5.3 General Comparative Statics

The sophisticated and naïve models are, in a sense, complete opposites of each other. Sophisticated agents have more information available to them, and they use this information to recognize cascades. In contrast, naïve agents have limited information and do not recognize cascades. Despite this, both types satisfy the Cascade Property. This section shows how any network effect model satisfying the Cascade Property will behave under various parameter changes.

The main object of interest is the *probability* of coordination in a generalized model

$\mathbb{M}(N)$. Whereas the models in the sophisticated and naïve cases were restricted to contain *only* sophisticated or *only* naïve agents, respectively, the generalized model $\mathbb{M}(N)$ may contain *any* collection of agents satisfying the Cascade Property, including mixtures of sophisticated, naïve, or even other kinds of agents, provided they all satisfy the Cascade Property.

Demonstrating these results relies on the fact that the sets $\mathcal{A}$ and $\mathcal{B}$ of permutations resulting in A and B cascades, respectively, are both finite, so that if these sets change in size, the probability of coordination changes accordingly, regardless of the probability distribution imposed on Ω. To avoid notational difficulties, it is assumed that all models in $\mathbb{M}(N)$ originate from the same set $\mathcal{X} = \{x_1, \ldots, x_N\}$ of N agents, so that each agent may be uniquely identified in each model, but that agent *types* may vary across models. For example, if $\mathcal{M}, \mathcal{M}' \in \mathbb{M}(N)$ are models where $N'_A > N_A$ (forcing $N'_B < N_B$ since there must be a total of N agents), then both models still have the same set of agents, but the types associated with some of the agents have changed. This is useful for directly converting a permutation $\omega \in \Omega$ of model $\mathcal{M}$ into a permutation $\omega' \in \Omega'$ of $\mathcal{M}'$. For a subset $\mathcal{S}$ of Ω, let $\phi : \mathcal{S} \to \Omega'$ be the identity map. To understand exactly what this map does, recall that a permutation ω is really a bijection $\omega : \{1, \ldots, N\} \to \mathcal{X}$, with $\omega_i = \omega(i)$ representing the particular *agent* which appears in position i. Thus ϕ sends an *ordering* of $\mathcal{X}$ to the same ordering. Obviously $\Omega = \Omega'$, but it is sometimes useful to distinguish between them so that it is clear which model is being analyzed. Note that ϕ is an injection, and that every permutation ω may be identified with a sequence $\Big((\omega_1, t_{\omega_1}), \ldots, (\omega_N, t_{\omega_N})\Big)$, where ω_i is the i^{th} agent and t_{ω_i} is that agent's type.

The basic principle behind the following demonstrations is that if $\phi(\mathcal{A}) \subseteq \mathcal{A}'$, then $\mathcal{A} \subseteq \mathcal{A}'$, which means $|\mathcal{A}| \leq |\mathcal{A}'|$. This means the probability of an A cascade is at least as large in model $\mathcal{M}'$ as it is in $\mathcal{M}$. Being the identity map, the condition $\phi(\mathcal{A}) \subseteq \mathcal{A}'$ may seem trivial, but the usefulness of this method is apparent when an agent changes

type, so that his position may be kept constant for comparison purposes.

The first comparative static result involves an increase in how valuable one network is relative to the other. For example, the value of network A could increase relative to network B if private valuations $v_{t,A}$ have increased or if the network function e_A has transformed to attain higher values over the entire domain $\mathbb{N}^2$. Or, it could be that private valuations for network B have decreased, or that e_B has decreased, or any combination of these events. This is made precise in the following definition:

Definition 5.3.1. *Let $\mathcal{M}, \mathcal{M}' \in \mathbb{M}(N)$. Then switching from model $\mathcal{M}$ to $\mathcal{M}'$ is said to increase the value of network i relative to network $j \neq i$ if each of the following conditions is satisfied:*

1. *for every $k \in \{A, B\}$, $v'_{k,i} \geq v_{k,i}$;*
2. *for every $k \in \{A, B\}$, $v'_{k,j} \leq v_{k,j}$;*
3. *for all $n_A, n_B \in \mathbb{N}$, $e_i(n_i, n_j)' \geq e_i(n_i, n_j)$; and*
4. *for all $n_A, n_B \in \mathbb{N}$, $e_j(n_j, n_i)' \leq e_j(n_j, n_i)$.*

The increase is said to be nontrivial when at least one of the above statements is satisfied with strict inequality in place of weak inequality.

Intuitively, one would expect an increase in the value of one network to increase the probability of a cascade on that network. This is indeed correct. Let $P(i_{\text{Pure}})$ and $P(i_{\text{Gen}})$ denote the probability of a pure i cascade and a general i cascade, respectively, for some $i \in \{A, B\}$. Consider the following:

Theorem 5.3.1. *Let $\mathcal{M}, \mathcal{M}' \in \mathbb{M}(N)$ be models where each choice function satisfies the Cascade Property. If switching from $\mathcal{M}$ to $\mathcal{M}'$ increases the value of network i relative to network $j \neq i$, then $P(i_{Pure} \mid \mathcal{M}') \geq P(i_{Pure} \mid \mathcal{M})$ and $P(i_{Gen} \mid \mathcal{M}') \geq P(i_{Gen} \mid \mathcal{M})$.*

Proof. For simplicity, the proof is given for the case of A cascades. Given a permutation $\omega \in \mathcal{A}$ which results in an A cascade (pure or otherwise), let ω_i be the first Type B agent in ω to choose A. This means $EU_i(A) > EU_i(B)$ in model $\mathcal{M}$. The same permutation $\phi(\omega)$ in model $\mathcal{M}'$, however, must also have $EU_i(A) > EU_i(B)$ for the same agent ω_i, because the relative increase in value for network A implies that every possible outcome from selecting A is at least as good in $\mathcal{M}'$ as it is in $\mathcal{M}$, and that every possible outcome resulting from the choice of B in $\mathcal{M}$ is at least as good as it is in $\mathcal{M}$'. Hence, agent ω_i will select A in model $\mathcal{M}'$. By the cascade property, this means ω will also result in an A cascade in $\mathcal{M}'$, so both $\mathcal{A}_{\text{Pure}} \subseteq \mathcal{A}_{\text{Pure}}{}'$ and $\mathcal{A}_{\text{Gen}} \subseteq \mathcal{A}_{\text{Gen}}{}'$. If the relative change is sufficiently large and if $\mathcal{B}_{\text{Gen}} \backslash \mathcal{B}_{\text{Pure}} \neq \emptyset$, then there is also a permutation $\omega \in \mathcal{B}$ where a Type B agent who, in $\mathcal{M}$, selected B, but will now select A (for example, the first Type B agent after a maximal segment of Type A agents will switch to A if the relative increase is sufficiently large). So, in some cases, the previous subset relations are strict. □

In Chapter 3, the concept of *competitive networks* was introduced. Central to this concept was the notion of one network *diminishing* the other (see Definition 3.2.1). It is possible to compute some basic comparative statics in relation to how the networks diminish each other, provided one can define what it means for networks to become more competitive. This is difficult and subjective, but the following definition is suitable in a variety of cases:

Definition 5.3.2. *A competitive diminishing of network i by network $j \neq i$ is a transformation of e_i to e_i' so that for all $n_i, n_j \in \mathbb{N}$, $e_i(n_i, n_j)' \leq e_i(n_i, n_j)$. The diminishment is said to be nontrivial if $e_i'(n_i, n_j) < e_i(n_i, n_j)$ for at least one pair (n_i, n_j).*

The idea is that, for example, network B will diminish network A *more* if it weakly decreases the value of e_A on its entire domain. Of course, it could be that network A has simply decreased in value without the networks being more "competitive", so some

caution is needed in applying this definition. One example where network B diminishes A in a competitive sense is if $e'_A(n_A, n_B) = e_A(\max\{0, n_A - n_B\}, n_B)$. This is simply a rightward "shift" of e_A, resulting in lower values, and the size of the shift is proportional to the increase in n_B. Thus, the greater in magnitude network B is, the more network A suffers (although, holding n_B constant, network A can still approach the same maximum value M_A that it could before the transformation).

To see how such a diminishment affects the probability of coordination on A or B, notice that a competitive diminishing satisfies the definition of an increase in network B relative to network A (or, equivalently, a decrease in the value of network A relative to network B). So Theorem 5.3.1 applies, and the probability of A cascades will decrease while the probability of B cascades will increase. The size of these changes depends on the size of the diminishment. Consequently, the effect is ambiguous if *both* networks transform so that they competitively diminish each other. Without specifying all functions and parameters of the model, including the transformed functions, it is impossible to determine which effect will dominate. Thus no general conclusion may be drawn regarding how competition between networks affects the probabilities of the various outcomes.

Finally, comparative static results may be deduced for changes to the population composition, N_A and N_B. If, for instance, an agent changes from Type A to Type B, one would expect the probability of a B cascade to increase. This intuition is correct, as the following theorem demonstrates.

Theorem 5.3.2. *Let $\mathcal{M}, \mathcal{M}' \in \mathbb{M}(N)$ be models where each choice function satisfies the Cascade Property. If $N'_i > N_i$, then $P(i_{Pure} \mid \mathcal{M}') \geq P(i_{Pure} \mid \mathcal{M})$ and $P(i_{Gen} \mid \mathcal{M}') \geq P(i_{Gen} \mid \mathcal{M})$.*

Proof. For simplicity, the proof is once again given for A cascades. Let ω be a permutation of the agents. Note that if $N_A < N'_A$ (under the interpretation that the set of Type A

individuals has expanded in model $\mathcal{M}'$), then any of the new Type A agents who originally selected good A in ω will still select A, so that if ω was an A cascade in $\mathcal{M}$, it will still be an A cascade in $\mathcal{M}'$. □

All of these results indicate that the Cascade Property is, indeed, an essential property for a network effect model to have, and that it is likely the relevant characteristic to look at when exploring new conjectures at this level of generality. Ignoring this property, for example, makes proving the above results for heterogeneous models (that is, those which include both sophisticated and naïve agents) extremely difficult, if not impossible. Indeed, the dynamic aspect makes proving comparative static results for general cascades even in *homogeneous* models extremely difficult as well. It follows that identifying primitives like (P1)-(P5), which help ensure the Cascade Property, should be made a priority when constructing new network effect models.

5.4 Numerical Examples

The comparative statics derived in section 5.3 show how the probabilities of various outcomes change as different model parameters are modified, but due to the level of generality, nothing has been said about the *magnitudes* of these changes (or even initial probability sizes), and nothing has been said about the degree of coordination in non-pure cascades. In this section, a number of concrete examples are given using a Java implementation of the algorithms derived in section 5.3.

Four different network effect functions are used. Consider first the function

$$e_i^1(n_i, n_j) = \begin{cases} 0 & \text{if } n_i = n_j = 0 \\ \frac{n_i}{n_i+n_j} \cdot 5 & \text{otherwise} \end{cases}.$$

This is a simple, strictly-competitive network function where agents are only concerned about the *ratio* of n_i relative to the total number of decisions made. This is useful for

constructing strictly competitive models. Note that $e^1(n_i, n_j)$ never attains the maximum value of 5 if $n_j > 0$. Next, consider the function

$$e_i^2(n_i, n_j) = \frac{n_i}{n_i + 1} \cdot 12 \ .$$

This is similar to e^1, except that the value of network i does not depend on the the size of the other network. This will be used for constructing examples of independent networks. Note that e^2 *never* attains a maximum value, although it has a least upper bound of 12. Another function for creating independent networks is

$$e_i^3(n_i, n_j) = \begin{cases} n_i - 1 & \text{if } n_i < 5 \\ 5 & \text{if } n_i \geq 5 \end{cases} \ .$$

The main difference between e^3 and e^2 is that e^3 actually attains, and remains constant at, a maximum value of 5 once 5 agents have joined the network. Clearly, then, this function is only suitable for models with a relatively small population size (such as the $N = 5$ and $N = 10$ cases illustrated first). For larger populations (like $N = 20$ and $N = 25$, examined afterwards), a similar function is given by e^4:

$$e_i^4(n_i, n_j) = \begin{cases} \frac{n_i}{3} & \text{if } n_i < 15 \\ 5 & \text{if } n_i \geq 15 \end{cases} \ .$$

The only other ingredient needed is a set of private values $v_{i,i}$, $v_{i,j}$ for each $i, j \in \{A, B\}$. All models will consider two cases: a "symmetric" case where $v_{A,A} = 4 = v_{B,B}$ and $v_{A,B} = 1 = v_{B,A}$, and an "asymmetric" case where $v_{A,A} = 4$, $v_{A,B} = 1$, $v_{B,B} = 3$, and $v_{B,A} = 2$. Various population compositions are then imposed, and outcome probabilities computed. As the reader may verify, each model satisfies the properties (P1)-(P5) and (NT).

Table 5.1 gives outcome probabilities for naïve agents in the symmetric model with $N = 5$ and $N = 10$. The labels are self-explanatory, except for $N_B(A)$ and $N_A(B)$; $N_B(A)$

represents the average (expected) percentage of Type B agents in the given population who choose good A, given that an A cascade occurs. $N_A(B)$ is defined similarly. Similar results are given in table 5.2, which gives outcome probabilities for sophisticated agents.

The first thing to notice is that, in a number of cases, the outcomes are identical for sophisticated and naïve agents, and that sophisticated agents generally are less likely to have split outcomes. Indeed, for a given model, the probability of a split for sophisticated agents *never* exceeds that for naïve agents. This is a common theme throughout all of the simulation results, suggesting that a general theorem in this regard may be provable.

Table 5.1: Outcomes for Naïve Agents and Small N.

$(v_{A,A}, v_{A,B}) = (4,1)$		Population Composition (N_A, N_B)							
$(v_{B,B}, v_{B,A}) = (4,1)$		$(4,1)$	$(3,2)$	$(2,3)$	$(1,4)$	$(9,1)$	$(7,3)$	$(5,5)$	$(2,8)$
$e_A = e_A^1$ $e_B = e_B^1$	$P(A_{\text{Pure}})$	0.4	0.1	0	0	0.6	0.1667	0.02381	0
	$P(A_{\text{Gen}})$	0.4	0.1	0	0	0.6	0.2	0.02381	0
	$E[N_B(A)]$	100%	100%	-	-	100%	94.4%	100%	-
	$P(B_{\text{Pure}})$	0	0	0.1	0.4	0	0	0.02381	0.3333
	$P(B_{\text{Gen}})$	0	0	0.1	0.4	0	0	0.02381	0.5111
	$E[N_A(B)]$	-	-	100%	100%	-	-	100%	82.6%
	$P(\text{Split})$	0.6	0.9	0.9	0.6	0.4	0.8	0.9524	0.4889
$e_A = e_A^2$ $e_B = e_B^2$	$P(A_{\text{Pure}})$	0.4	0.1	0	0	0.6	0.1667	0.02381	0
	$P(A_{\text{Gen}})$	0.4	0.1	0	0	0.6	0.1667	0.02381	0
	$E[N_B(A)]$	100%	100%	-	-	100%	100%	100%	-
	$P(B_{\text{Pure}})$	0	0	0.1	0.4	0	0	0.02381	0.3333
	$P(B_{\text{Gen}})$	0	0	0.1	0.4	0	0	0.02381	0.4222
	$E[N_A(B)]$	-	-	100%	100%	-	-	100%	89.5%
	$P(\text{Split})$	0.6	0.9	0.9	0.6	0.4	0.8333	0.9524	0.5778
$e_A = e_A^3$ $e_B = e_B^3$	$P(A_{\text{Pure}})$	0.6	0.3	0.1	0	0.7	0.2917	0.08333	0
	$P(A_{\text{Gen}})$	0.6	0.3	0.1	0	0.7	0.4417	0.09524	0
	$E[N_B(A)]$	100%	100%	100%	-	100%	88.7%	97.5%	-
	$P(B_{\text{Pure}})$	0	0.1	0.3	0.6	0	0.008333	0.08333	0.4667
	$P(B_{\text{Gen}})$	0	0.1	0.3	0.6	0	0.008333	0.09524	0.7333
	$E[N_A(B)]$	-	100%	100%	100%	-	100%	97.5%	81.8%
	$P(\text{Split})$	0.4	0.6	0.6	0.4	0.3	0.55	0.8095	0.2667

Tables 5.1 and 5.2 also provide concrete examples of how changes to the population composition affect the likelihood of different outcomes. Models where Type A agents greatly outnumber Type B agents have greater probabilities of A cascades, and this probability gradually diminishes (and the probability of B cascades gradually increases) as the population composition shifts to favour Type B agents. As one would expect,

Table 5.2: Outcomes for Sophisticated Agents and Small N.

$(v_{A,A}, v_{A,B}) = (4,1)$ $(v_{B,B}, v_{B,A}) = (4,1)$		Population Composition (N_A, N_B)							
		(4,1)	(3,2)	(2,3)	(1,4)	(9,1)	(7,3)	(5,5)	(2,8)
$e_A = e_A^1$ $e_B = e_B^1$	$P(A_{\text{Pure}})$	0.6	0.3	0.1	0	0.8	0.4667	0.2222	0.02222
	$P(A_{\text{Gen}})$	0.6	0.3	0.1	0	0.8	0.5667	0.2302	0.02222
	$E[N_B(A)]$	100%	100%	100%	-	100%	94.1%	99.3%	100%
	$P(B_{\text{Pure}})$	0	0.1	0.3	0.6	0	0.06667	0.2222	0.6222
	$P(B_{\text{Gen}})$	0	0.1	0.3	0.6	0	0.06667	0.2302	0.8
	$E[N_A(B)]$	-	100%	100%	100%	-	100%	99.3%	88.9%
	$P(\text{Split})$	0.4	0.6	0.6	0.4	0.2	0.3667	0.5397	0.1778
$e_A = e_A^2$ $e_B = e_B^2$	$P(A_{\text{Pure}})$	0.4	0.1	0	0	0.6	0.1667	0.02381	0
	$P(A_{\text{Gen}})$	0.4	0.1	0	0	0.6	0.1667	0.02381	0
	$E[N_B(A)]$	100%	100%	-	-	100%	100%	100%	-
	$P(B_{\text{Pure}})$	0	0	0.1	0.4	0	0	0.02381	0.3333
	$P(B_{\text{Gen}})$	0	0	0.1	0.4	0	0	0.02381	0.4222
	$E[N_A(B)]$	-	-	100%	100%	-	-	100%	89.5%
	$P(\text{Split})$	0.6	0.9	0.9	0.6	0.4	0.8333	0.9524	0.5778
$e_A = e_A^3$ $e_B = e_B^3$	$P(A_{\text{Pure}})$	0.8	0.6	0.4	0.2	0.7	0.2917	0.08333	0
	$P(A_{\text{Gen}})$	0.8	0.6	0.4	0.2	0.7	0.4417	0.09524	0
	$E[N_B(A)]$	100%	100%	100%	100%	100%	88.7%	97.5%	-
	$P(B_{\text{Pure}})$	0.2	0.4	0.6	0.8	0	0.008333	0.08333	0.4667
	$P(B_{\text{Gen}})$	0.2	0.4	0.6	0.8	0	0.008333	0.09524	0.7333
	$E[N_A(B)]$	100%	100%	100%	100%	-	100%	97.5%	81.8%
	$P(\text{Split})$	0	0	0	0	0.3	0.55	0.8095	0.2667

the probability of a split outcome is greatest when there are roughly the same number of Type A and Type B agents, and lowest when one type significantly outnumbers the other.

It is worth mentioning that even though each of these models satisfies (P1)-(P5) and (NT) (in particular, conditions which guarantee purely coordinated equilibria of both types in the simultaneous move game), some of these models in fact have a 0 probability of either an A cascade or a B cascade. In all cases, the reason for this is that there are simply not enough agents of that type to start even a pure cascade, which therefore rules out the possibility of a general cascade of that type. This happens for at least one model in the $N = 20$ and $N = 25$ cases as well (and for less trivial population compositions), suggesting that even "reasonable" looking models may fail to exhibit certain outcomes once dynamics come into play.

It is also interesting to note that the degree of coordination does not necessarily

improve as the probability of a cascade improves. Consider, for example, the B cascades in tables 5.1 and 5.2 where both network functions are given by e^3. As the population composition changes from $(7,3)$ to $(5,5)$ to $(2,8)$, the probability of a B cascade increases significantly, but the degree of coordination diminishes from 100% to 81.8%. This is due to the fact that the probability of a *general* B cascade increases at a much greater rate than a *pure* cascade for those models, resulting in relatively more outcomes which are not purely coordinated.

These results are not restricted to models with symmetric preferences. Tables 5.3 and 5.4 give outcomes for the case where $v_{A,A} = 4$, $v_{A,B} = 1$, $v_{B,B} = 3$, and $v_{B,A} = 2$. The same population compositions and network functions are used to allow comparisons with the symmetric case.

Table 5.3: Outcomes for Naïve Agents and Small N (Asymmetric Preferences).

$(v_{A,A}, v_{A,B}) = (4,1)$ $(v_{B,B}, v_{B,A}) = (3,2)$		Population Composition (N_A, N_B)							
		$(4,1)$	$(3,2)$	$(2,3)$	$(1,4)$	$(9,1)$	$(7,3)$	$(5,5)$	$(2,8)$
$e_A = e_A^1$ $e_B = e_B^1$	$P(A_{\text{Pure}})$	0.6	0.3	0.1	0	0.8	0.4667	0.2222	0.02222
	$P(A_{\text{Gen}})$	0.6	0.5	0.1	0	0.8	0.8833	0.4762	0.02222
	$E[N_B(A)]$	100%	80.0%	100%	-	100%	78.0%	81.3%	100%
	$P(B_{\text{Pure}})$	0	0	0.1	0.4	0	0	0.02381	0.3333
	$P(B_{\text{Gen}})$	0	0	0.1	0.4	0	0	0.02381	0.5111
	$E[N_A(B)]$	-	-	100%	100%	-	-	100%	82.6%
	$P(\text{Split})$	0.4	0.5	0.8	0.6	0.2	0.1167	0.5	0.4667
$e_A = e_A^2$ $e_B = e_B^2$	$P(A_{\text{Pure}})$	0.6	0.3	0.1	0	0.8	0.4667	0.2222	0.02222
	$P(A_{\text{Gen}})$	0.6	0.5	0.1	0	0.8	0.8083	0.2897	0.02222
	$E[N_B(A)]$	100%	80.0%	100%	-	100%	78.7%	93.4%	100%
	$P(B_{\text{Pure}})$	0	0	0.1	0.4	0	0	0.02381	0.3333
	$P(B_{\text{Gen}})$	0	0	0.1	0.4	0	0	0.02381	0.4222
	$E[N_A(B)]$	-	-	100%	100%	-	-	100%	89.5%
	$P(\text{Split})$	0.4	0.5	0.8	0.6	0.2	0.1917	0.6865	0.5556
$e_A = e_A^3$ $e_B = e_B^3$	$P(A_{\text{Pure}})$	0.8	0.6	0.4	0.2	0.8	0.4667	0.2222	0.02222
	$P(A_{\text{Gen}})$	0.8	0.7	0.4	0.2	0.8	0.8833	0.4722	0.02222
	$E[N_B(A)]$	100%	92.9%	100%	100%	100%	78.0%	81.7%	100%
	$P(B_{\text{Pure}})$	0	0.1	0.3	0.6	0	0.008333	0.08333	0.4667
	$P(B_{\text{Gen}})$	0	0.1	0.3	0.6	0	0.008333	0.09524	0.7333
	$E[N_A(B)]$	-	100%	100%	100%	-	100%	97.5%	81.8%
	$P(\text{Split})$	0.2	0.2	0.3	0.2	0.2	0.1083	0.4325	0.2444

All of the patterns and regularities for the symmetric case also show up in the asymmetric case, although the probability of a split is lower and the likelihood of A cascades

Table 5.4: Outcomes for Sophisticated Agents and Small N (Asymmetric Preferences).

$(v_{A,A}, v_{A,B}) = (4,1)$ $(v_{B,B}, v_{B,A}) = (3,2)$		Population Composition (N_A, N_B)							
		(4,1)	(3,2)	(2,3)	(1,4)	(9,1)	(7,3)	(5,5)	(2,8)
$e_A = e_A^1$ $e_B = e_B^1$	$P(A_{\text{Pure}})$	0.8	0.6	0.4	0.2	0.9	0.7	0.5	0.2
	$P(A_{\text{Gen}})$	0.8	0.8	0.5	0.2	0.9	0.9833	0.8611	0.4
	$E[N_B(A)]$	100%	87.5%	93.3%	100%	100%	88.7%	89.6%	93.1%
	$P(B_{\text{Pure}})$	0	0.1	0.3	0.6	0	0.008333	0.08333	0.4667
	$P(B_{\text{Gen}})$	0	0.1	0.3	0.6	0	0.008333	0.08730	0.5556
	$E[N_A(B)]$	-	100%	100%	100%	-	100%	99.1%	92.0%
	$P(\text{Split})$	0.2	0.1	0.2	0.2	0.1	0.008333	0.05159	0.04444
$e_A = e_A^2$ $e_B = e_B^2$	$P(A_{\text{Pure}})$	0.8	0.6	0.4	0.2	0.9	0.7	0.5	0.2
	$P(A_{\text{Gen}})$	0.8	0.8	0.5	0.2	0.9	0.925	0.6468	0.2222
	$E[N_B(A)]$	100%	87.5%	93.3%	100%	100%	90.1%	95.2%	98.8%
	$P(B_{\text{Pure}})$	0	0	0.1	0.4	0	0	0.02381	0.3333
	$P(B_{\text{Gen}})$	0	0	0.1	0.4	0	0	0.02381	0.4
	$E[N_A(B)]$	-	-	100%	100%	-	-	100%	91.7%
	$P(\text{Split})$	0.2	0.2	0.4	0.4	0.1	0.075	0.3294	0.3778
$e_A = e_A^3$ $e_B = e_B^3$	$P(A_{\text{Pure}})$	0.8	0.6	0.4	0.2	0.9	0.7	0.5	0.2
	$P(A_{\text{Gen}})$	0.8	0.8	0.5	0.2	0.9	0.9417	0.6944	0.2222
	$E[N_B(A)]$	100%	87.5%	93.3%	100%	100%	89.1%	92.1%	98.8%
	$P(B_{\text{Pure}})$	0	0.1	0.3	0.6	0	0.008333	0.08333	0.4667
	$P(B_{\text{Gen}})$	0	0.1	0.3	0.6	0	0.008333	0.09127	0.6444
	$E[N_A(B)]$	-	100%	100%	100%	-	100%	98.3%	86.2%
	$P(\text{Split})$	0.2	0.1	0.2	0.2	0.1	0.05	0.2143	0.1333

is higher in the asymmetric models. This comes as no surprise, because the the change in private valuations for Type B agents is a relative increase in the value of network A, and the comparative static theorems of section 5.3 guarantee that A cascades will be more likely and B cascades will be less likely after such a change. What is surprising is that, once again, this increase in the likelihood of an A cascade does not always improve the average degree of coordination in A cascades. In both the naïve and sophisticated cases, switching from the symmetric to asymmetric models with population composition $(7,3)$ *reduces* the average percentage of Type B agents who end up choosing A in A cascades. Once again, this is because the probability of a *general* cascade has increased more significantly than the probability of a *pure* cascade, resulting in a lower overall degree of coordination.

Finally, observe that the three different network effect combinations (namely, the cases where both network functions are e^1, both are e^2, or both are e^3) can be used

to examine how "competitiveness" affects the likelihood of coordination. When both functions are e^1, the networks are strictly competitive; for the other two, the networks are independent. But a moments' observation reveals that there is no regular pattern between the likelihood of split outcomes between competitive and independent networks, and that there also is not a regular pattern for how well coordinated the cascades are. This is consistent with the ambiguity result hinted at in section 5.3.

For larger populations (namely, $N = 20$ and $N = 25$), outcomes for symmetric preferences are given in tables 5.5 and 5.6, and asymmetric outcomes are given in tables 5.7 and 5.8. The same comparative static patterns (and ambiguities) for outcome probabilities are present in these outcomes, as are the ambiguities relating to the level of coordination in general cascades.

Table 5.5: Outcomes for Naïve Agents and Large N.

$(v_{A,A}, v_{A,B}) = (4,1)$		Population Composition (N_A, N_B)							
$(v_{B,B}, v_{B,A}) = (4,1)$		$(18,2)$	$(14,6)$	$(10,10)$	$(4,16)$	$(20,5)$	$(15,10)$	$(12,13)$	$(8,17)$
$e_A = e_A^1$ $e_B = e_B^1$	$P(A_{\text{Pure}})$	0.6316	0.2066	0.0433	2.1E^{-4}	0.3830	0.1079	0.0391	5.5E^{-3}
	$P(A_{\text{Gen}})$	0.8842	0.3124	0.0481	2.1E^{-4}	0.7806	0.1259	0.0406	5.5E^{-3}
	$E[N_B(A)]$	85.7%	92.9%	99.0%	100%	77.1%	98.2%	99.7%	100%
	$P(B_{\text{Pure}})$	0	3.1E^{-3}	0.0433	0.3756	4.0E^{-4}	0.0166	0.0565	0.1881
	$P(B_{\text{Gen}})$	0	3.1E^{-3}	0.0481	0.7529	4.0E^{-4}	0.0168	0.0602	0.2561
	$E[N_A(B)]$	-	100%	99.0%	77.6%	100%	99.9%	99.4%	95.3%
	$P(\text{Split})$	0.1158	0.6845	0.9038	0.2469	0.2190	0.8574	0.8992	0.7383
$e_A = e_A^2$ $e_B = e_B^2$	$P(A_{\text{Pure}})$	0.4105	0.0443	1.6E^{-3}	0	0.1165	6.0E^{-3}	4.6E^{-4}	9.3E^{-7}
	$P(A_{\text{Gen}})$	0.5579	0.0443	1.6E^{-3}	0	0.1187	6.0E^{-3}	4.6E^{-4}	9.3E^{-7}
	$E[N_B(A)]$	86.8%	100%	100%	-	99.6%	100%	100%	100%
	$P(B_{\text{Pure}})$	0	0	1.6E^{-3}	0.1476	0	4.2E^{-5}	1.2E^{-3}	0.0225
	$P(B_{\text{Gen}})$	0	0	1.6E^{-3}	0.1534	0	4.2E^{-3}	1.2E^{-3}	0.0225
	$E[N_A(B)]$	-	-	100%	99.1%	-	100%	100%	100%
	$P(\text{Split})$	0.4421	0.9557	0.9969	0.8466	0.8813	0.9940	0.9984	0.9775
$e_A = e_A^4$ $e_B = e_B^4$	$P(A_{\text{Pure}})$	0.7158	0.3193	0.1053	3.5E^{-3}	0.3830	0.1079	0.0391	5.5E^{-3}
	$P(A_{\text{Gen}})$	0.9211	0.6449	0.1205	3.5E^{-3}	0.8666	0.1479	0.0442	5.6E^{-3}
	$E[N_B(A)]$	88.9%	77.5%	98.4%	100%	75.4%	96.5%	99.1%	99.9%
	$P(B_{\text{Pure}})$	0	0.0175	0.1053	0.4912	4.0E^{-4}	0.0166	0.0565	0.1881
	$P(B_{\text{Gen}})$	0	0.0176	0.1205	0.9005	4.0E^{-4}	0.0174	0.0673	0.3158
	$E[N_A(B)]$	-	100%	98.4%	79.6%	100%	99.7%	98.5%	91.7%
	$P(\text{Split})$	0.079	0.3375	0.7589	0.096	0.1330	0.8347	0.8885	0.6786

One noteworthy result is that despite the larger number of agents, the level of coordination in cascades is typically quite high. The overall probability of coordination is also

Table 5.6: Outcomes for Sophisticated Agents and Large N.

$(v_{A,A}, v_{A,B}) = (4,1)$		Population Composition (N_A, N_B)							
$(v_{B,B}, v_{B,A}) = (4,1)$		$(18,2)$	$(14,6)$	$(10,10)$	$(4,16)$	$(20,5)$	$(15,10)$	$(12,13)$	$(8,17)$
$e_A = e_A^1$ $e_B = e_B^1$	$P(A_{\text{Pure}})$	0.8053	0.4789	0.2368	0.0316	0.6333	0.35	0.22	0.0933
	$P(A_{\text{Gen}})$	0.9632	0.6881	0.2964	0.032	0.9395	0.5635	0.3428	0.1315
	$E[N_B(A)]$	91.8%	93.5%	97.8%	99.9%	90.9%	95.5%	97.0%	98.3%
	$P(B_{\text{Pure}})$	5.3E^{-3}	0.079	0.2368	0.6316	0.0333	0.15	0.26	0.4533
	$P(B_{\text{Gen}})$	5.3E^{-3}	0.0844	0.2964	0.9123	0.0412	0.2241	0.4113	0.7263
	$E[N_A(B)]$	100%	99.5%	97.8%	89.1%	99.0%	97.7%	96.6%	94.1%
	$P(\text{Split})$	0.0316	0.2276	0.4072	0.0557	0.0193	0.2125	0.2459	0.1422
$e_A = e_A^2$ $e_B = e_B^2$	$P(A_{\text{Pure}})$	0.4789	0.0775	5.4E^{-3}	0	0.1613	0.0134	1.6E^{-3}	1.7E^{-5}
	$P(A_{\text{Gen}})$	0.6053	0.0775	5.4E^{-3}	0	0.1632	0.0134	1.6E^{-3}	1.7E^{-5}
	$E[N_B(A)]$	89.6%	100%	100%	-	99.8%	100%	100%	100%
	$P(B_{\text{Pure}})$	0	2.6E^{-5}	5.4E^{-3}	0.2066	0	2.5E^{-4}	3.6E^{-3}	0.0405
	$P(B_{\text{Gen}})$	0	2.6E^{-5}	5.4E^{-3}	0.2190	0	2.5E^{-4}	3.6E^{-3}	0.0405
	$E[N_A(B)]$	-	100%	100%	98.6%	-	100%	100%	100%
	$P(\text{Split})$	0.3947	0.9225	0.9892	0.7810	0.8368	0.9864	0.9948	0.9595
$e_A = e_A^4$ $e_B = e_B^4$	$P(A_{\text{Pure}})$	0.8053	0.4789	0.2368	0.0316	0.6333	0.35	0.22	0.0933
	$P(A_{\text{Gen}})$	0.9632	0.7873	0.3081	0.032	0.9155	0.4239	0.2415	0.0951
	$E[N_B(A)]$	91.8%	87.2%	97.1%	99.9%	88.7%	97.8%	99.2%	99.9%
	$P(B_{\text{Pure}})$	5.3E^{-3}	0.079	0.2368	0.6316	0.0333	0.15	0.26	0.4533
	$P(B_{\text{Gen}})$	5.3E^{-3}	0.0845	0.3081	0.9408	0.0334	0.1574	0.2938	0.5983
	$E[N_A(B)]$	100%	99.5%	97.1%	87.5%	100%	99.7%	98.9%	95.5%
	$P(\text{Split})$	0.0316	0.1282	0.3837	0.0272	0.0511	0.4187	0.4647	0.3066

quite high, except when both network functions are e^2, in which case the likelihood of a split is very high under a number of population compositions. This is due to the fact that the networks are independent, but also (and more importantly) the fact that the network approaches a higher maximum value (12) than it does in other models. As such, agents are only likely to switch networks if the other network already has a large size and if their own (privately preferred) network has a small size. This is why most of the cascades in those models are purely coordinated and are fairly unlikely to come about.

Although these examples are only a few among (infinitely) many possible network effect models, they help to reinforce the point that purely coordinated outcomes are not particularly likely to occur, and that split outcomes are a real possibility even under different population compositions and models which favour coordination on one good (such as the asymmetric models, which favour coordination on good A). They also highlight the importance of population composition as a real factor in determining different out-

Table 5.7: Outcomes for Naïve Agents and Large N (Asymmetric Preferences).

$(v_{A,A}, v_{A,B}) = (4,1)$ $(v_{B,B}, v_{B,A}) = (3,2)$		Population Composition (N_A, N_B)							
		(18,2)	(14,6)	(10,10)	(4,16)	(20,5)	(15,10)	(12,13)	(8,17)
$e_A = e_A^1$ $e_B = e_B^1$	$P(A_{\text{Pure}})$	0.8053	0.4789	0.2368	0.0316	0.6333	0.35	0.22	0.0933
	$P(A_{\text{Gen}})$	0.9737	0.9719	0.4714	0.0382	0.9955	0.8931	0.4323	0.1377
	$E[N_B(A)]$	91.4%	81.7%	89.8%	98.9%	88.4%	78.0%	92.1%	97.7%
	$P(B_{\text{Pure}})$	0	3.1E^{-3}	0.0433	0.3756	4.0E^{-4}	0.0166	0.0565	0.1881
	$P(B_{\text{Gen}})$	0	3.1E^{-3}	0.0481	0.7406	4.0E^{-4}	0.0168	0.0602	0.2550
	$E[N_A(B)]$	-	100%	99.0%	78.4%	100%	99.9%	99.4%	95.4%
	$P(\text{Split})$	0.0263	0.025	0.4805	0.2213	4.1E^{-3}	0.0902	0.5075	0.6073
$e_A = e_A^2$ $e_B = e_B^2$	$P(A_{\text{Pure}})$	0.7158	0.3193	0.1053	3.5E^{-3}	0.4957	0.1978	0.0957	0.0244
	$P(A_{\text{Gen}})$	0.9368	0.6769	0.1488	3.5E^{-3}	0.9383	0.2807	0.1105	0.02479
	$E[N_B(A)]$	88.2%	82.5%	96.3%	100%	80.5%	95.3%	98.7%	99.9%
	$P(B_{\text{Pure}})$	0	0	1.5E^{-3}	0.1476	0	4.2E^{-5}	1.2E^{-3}	0.0225
	$P(B_{\text{Gen}})$	0	0	1.5E^{-3}	0.1534	0	4.2E^{-5}	1.2E^{-3}	0.0225
	$E[N_A(B)]$	-	-	100%	99.1%	-	100%	100%	100%
	$P(\text{Split})$	0.0632	0.3231	0.8497	0.8431	0.0617	0.7193	0.8884	0.9527
$e_A = e_A^4$ $e_B = e_B^4$	$P(A_{\text{Pure}})$	0.8053	0.4789	0.2368	0.0316	0.6333	0.35	0.22	0.0933
	$P(A_{\text{Gen}})$	0.9737	0.9685	0.5562	0.03922	0.9965	0.9524	0.5297	0.1584
	$E[N_B(A)]$	91.4%	83.2%	86.3%	98.6%	88.8%	78.0%	87.8%	96.2%
	$P(B_{\text{Pure}})$	0	0.0175	0.1053	0.4912	4.0E^{-4}	0.0166	0.0565	0.1881
	$P(B_{\text{Gen}})$	0	0.0176	0.1203	0.8801	4.0E^{-4}	0.0174	0.0672	0.3122
	$E[N_A(B)]$	-	100%	98.5%	80.6%	100%	99.7%	98.5%	92.0%
	$P(\text{Split})$	0.0263	0.0139	0.3235	0.0807	3.1E^{-3}	0.0301	0.4030	0.5294

comes. Models where a particular type of cascade is literally impossible, for example, illustrate this point well. Consider the naïve case for $N = 20$ and asymmetric preferences. This model has likelihood 0 of coordination on B when e^2 is used and the composition is $(14, 6)$, even though Type B agents constitute 30% of the population! For sophisticated agents using e^1, switching from composition $(18, 2)$ to $(4, 16)$ reduces the probability of an A cascade from 0.96 to about 0.03 (and yet, the average level of coordination in A cascades *increases*). These results, together with the comparative static theorems of section 5.3 suggest that outcome probabilities behave as expected, but the level of coordination (hence related welfare considerations) do not.

Table 5.8: Outcomes for Sophisticated Agents and Large N (Asymmetric Preferences).

$(v_{A,A}, v_{A,B}) = (4,1)$ $(v_{B,B}, v_{B,A}) = (3,2)$		Population Composition (N_A, N_B)							
		$(18,2)$	$(14,6)$	$(10,10)$	$(4,16)$	$(20,5)$	$(15,10)$	$(12,13)$	$(8,17)$
$e_A = e_A^1$ $e_B = e_B^1$	$P(A_{\text{Pure}})$	0.9	0.7	0.5	0.2	0.8	0.6	0.48	0.32
	$P(A_{\text{Gen}})$	0.9947	0.9825	0.8947	0.5088	0.9957	0.9478	0.8757	0.7043
	$E[N_B(A)]$	95.2%	94.2%	94.1%	94.5%	95.5%	95.3%	95.3%	95.5%
	$P(B_{\text{Pure}})$	0	0.0175	0.1053	0.4912	4.3E^{-3}	0.0522	0.1243	0.2957
	$P(B_{\text{Gen}})$	0	0.0175	0.1053	0.4912	4.3E^{-3}	0.0522	0.1243	0.2957
	$E[N_A(B)]$	-	100%	100%	100%	100%	100%	100%	100%
	$P(\text{Split})$	5.3E^{-3}	0	0	0	0	0	0	0
$e_A = e_A^2$ $e_B = e_B^2$	$P(A_{\text{Pure}})$	0.9	0.7	0.5	0.2	0.8	0.6	0.48	0.32
	$P(A_{\text{Gen}})$	0.9842	0.9044	0.5852	0.2033	0.9868	0.7209	0.5445	0.3392
	$E[N_B(A)]$	95.7%	93.7%	98.1%	99.9%	93.7%	97.8%	99.0%	99.7%
	$P(B_{\text{Pure}})$	0	2.6E^{-5}	5.4E^{-3}	0.2066	0	2.5E^{-4}	3.6E^{-3}	0.0405
	$P(B_{\text{Gen}})$	0	2.6E^{-5}	5.4E^{-3}	0.2169	0	2.5E^{-4}	3.6E^{-3}	0.0405
	$E[N_A(B)]$	-	100%	100%	98.8%	-	100%	100%	100%
	$P(\text{Split})$	0.0158	0.0956	0.4094	0.5798	0.0132	0.2789	0.4519	0.6203
$e_A = e_A^4$ $e_B = e_B^4$	$P(A_{\text{Pure}})$	0.9	0.7	0.5	0.2	0.8	0.6	0.48	0.32
	$P(A_{\text{Gen}})$	0.9947	0.9825	0.8947	0.5088	0.9956	0.9437	0.8462	0.6144
	$E[N_B(A)]$	95.2%	94.2%	94.1%	94.5%	95.4%	94.7%	95.0%	96.3%
	$P(B_{\text{Pure}})$	0	0.0175	0.1053	0.4912	4.3E^{-3}	0.0522	0.1243	0.2957
	$P(B_{\text{Gen}})$	0	0.0175	0.1053	0.4912	4.4E^{-3}	0.0558	0.1401	0.3541
	$E[N_A(B)]$	-	100%	100%	100%	100%	99.6%	99.0%	97.4%
	$P(\text{Split})$	5.3E^{-3}	0	0	0	7.5E^{-5}	5.4E^{-4}	0.0137	0.0315

Chapter 6

Conclusion

This thesis constructed a generalized network effect model for the two good case and explored a number of issues related to coordination. In particular, properties have been identified which guarantee multiple coordinated equilibria in simultaneous move games, and these conditions have been shown to be insufficient for guaranteeing the existence of highly coordinated outcomes in a dynamic setting. To arrive at these results, two very distinct decision algorithms (the sophisticated and naïve algorithms) were developed and shown to satisfy the Cascade Property, which asserts that a good i cascade will form if an agent of Type $j \neq i$ optimally selects good i. Comparative static results regarding the likelihood of various outcomes were then derived for arbitrary models which satisfy the Cascade Property, demonstrating that even "mixed" models consisting of some sophisticated and some naïve agents will have the same comparative static properties. Using the Java programming language, numerical examples were computed for a variety of cases, showcasing both the comparative static results already derived, as well as some initial results about how well-coordinated cascades will be on average.

There are a number of possible directions for future research stemming from this work. An obvious drawback to the models in this thesis is that they are restricted to two goods only. While it is likely that similar results would hold in an arbitrary n-good case, it would be interesting to identify a suitable generalization of the Cascade Property to n goods; indeed, this seems like the central contribution that an n-good extension could make.

The numerical examples of section 5.4 identified an interesting regularity: models consisting only of sophisticated agents never have a greater probability of a split outcome

than models consisting only of naïve agents do. A rigorous proof of this statement, if possible, would be quite interesting, especially if one could prove comparative static results in this regard as a population shifts from having more naïve agents to having more sophisticated agents.

There are several interesting issues related to the timing of decisions. A nice generalization of this work would be to allow groups of agents to simultaneously make adoption decisions, and to consider all possible sequences of groups. This thesis has laid important groundwork in this respect, for the choice functions derived do not depend on how previous decisions were timed. Indeed, generalizing to allow groups would really only change the combinatorial aspects of this thesis. This could be difficult, but it would surely be a worthwhile endeavor.

Another timing issue would be to allow agents the opportunity to defer their decisions to a later point in time, or to attempt to make their decisions sooner. Such a model would be fundamentally different from this one, although, once again, this work may prove useful as a starting point.

At present, very little is understood about how coordination problems may be resolved. Hopefully, this thesis has made a useful new contribution, and provided an interesting new perspective on, the general theory of network effects.

Bibliography

Bikhchandani, S., D. Hirshleifer, and I. Welch (1998). Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades. *The Journal of Economic Perspectives 12*(3), 151–170.

Cabral, L. (1990). On the Adoption of Innovations withNetworkExternalities. *Mathematical Social Sciences 19*(3), 299–308.

Choi, J. (1994). Irreversible Choice of Uncertain Technologies with Network Externalities. *Rand Journal of Economics 25*(3), 382–401.

Choi, J. (1997). Herd Behavior, the" Penguin Effect," and the Suppression of Informational Diffusion: an Analysis of Informational Externalities and Payoff Interdependency. *Rand Journal of Economics 28*(3), 407–425.

Church, J. and N. Gandal (1992). Network Effects, Software Provision, and Standardization. *Journal of Industrial Economics 40*(1), 85–103.

Church, J. and I. King (1993). Bilingualism and Network Externalities. *Canadian Journal of Economics 26*(2), 337–45.

Clements, M. (2004). Direct and indirect network effects: are they equivalent? *International Journal of Industrial Organization 22*(5), 633–645.

Crawford, V. (1995). Adaptive Dynamics in Coordination Games. *Econometrica 63*(1), 103–43.

Eaton, B. and D. Krause (2005). Coordination Cascades: Sequential Choice in the Presence of a Network Externality. *Mimeo*.

Economides, N. (1996). The economics of networks. *International Journal of Industrial Organization 14*(6), 673–699.

Farrell, J. and P. Klemperer (2007). Coordination and Lock-in: Competition with Switching Costs and Network Effects. *Handbook of Industrial Organization.*

Farrell, J. and G. Saloner (1985). Standardization, Compatibility, and Innovation. *Rand Journal of Economics 16*(1), 70–83.

Farrell, J. and G. Saloner (1986). Installed Base and Compatibility: Innovation, Product Preannouncements, and Predation. *American Economic Review 76*(5), 940–955.

Katz, M. and C. Shapiro (1985). Network Externalities, Competition, and Compatibility. *The American Economic Review 75*(3), 424–440.

Katz, M. and C. Shapiro (1986). Technology Adoption in the Presence of Network Externalities. *The Journal of Political Economy 94*(4), 822.

Katz, M. and C. Shapiro (1994). Systems Competition and Network Effects. *Journal of Economic Perspectives 8*, 93–93.

Liebowitz, S. and S. Margolis (1994). Network Externality: An Uncommon Tragedy. *The Journal of Economic Perspectives 8*(2), 133–150.

Rohlfs, J. (1974). A Theory of Interdependent Demand for a Communications Service. *Bell Journal of Economics 5*(1), 16–37.

Swann, G. (2002). The functional form of network effects. *Information Economics and Policy 14*(3), 417–429.

Appendix

A.1 The Network Effect Function

$S_A = S_B = \mathbb{N}$ does not imply strictly competitive networks.

For example, suppose $i \neq j$ and take $v_{i,i} = 2$ and $v_{i,j} = 1$. Consider the network function

$$e_i(n_i, n_j) = \begin{cases} \left(\max\left\{0, \frac{2n_i - n_j}{2n_i+1}\right\}\right) \cdot 10 & \text{if } n_j \text{ is even} \\ \left(\max\left\{0, \frac{2n_i - n_j - 1}{2n_i+1}\right\}\right) \cdot 10 & \text{if } n_j \text{ is odd} \end{cases}.$$

It is routine to verify that this model satisfies (P1)-(P5); to see that $S_A = S_B = \mathbb{N}_+$, fix any $n_i \in \mathbb{N}$ and observe that $\lim_{n_j \to \infty}(v_{i,j} + e_j(n_j, n_i)) = 11 > 2 = \lim_{n_j \to \infty}(v_{i,i} + e_i(n_i, n_j))$, so there is some n_j^* for which $n_j \geq n_j^* \Rightarrow v_{i,j} + e_j(n_j, n_i) > v_{i,i} + e_i(n_i, n_j)$. But, given $n_i \in \mathbb{N}_+$, e_i is not strictly decreasing in n_j, because if n_j is even, then $e_i(n_i, n_j) = e_i(n_i, n_j + 1)$. (Trivially, sufficiently large n_j such as $n_j > 2n_i$ will also cause constant regions due to the $\max\{\cdot\}$ component of e_i, but the other case is more interesting).

Bounded switching sets do not imply independent networks.

To see this, suppose $i \neq j$ and take $v_{i,i} = 2$ and $v_{i,j} = 1$. Consider the network function

$$e_i(n_i, n_j) = \begin{cases} \max\{0, n_i - n_j\} & \text{if } n_i < 10 \\ 10 + \frac{n_i}{n_i + n_j} & \text{if } n_i \geq 10 \end{cases}.$$

It is routine to verify that this model satisfies (P1)-(P5); in particular, $S_i = \{0, 1, \ldots, 9\}$ since $2 + \max\{0, n_i - n_j\} < 11 + \frac{n_j}{n_i+n_j}$ for $0 \leq n_i \leq 9$ and $n_j \geq 10$, and $12 + \frac{n_i}{n_i+n_j} > 11 + \frac{n_j}{n_i+n_j}$ for all $n_i, n_j \in \mathbb{N}$, so Type i agents will always choose good i when $n_i \geq 10$. But the networks are not independent; in fact, they are strictly competitive.

(Strictly) competitive networks do not imply that each $S_i = \mathbb{N}$.

The above example suffices in this case as well, for the networks are strictly competitive (and therefore competitive), but $S_i = \{0, 1, \ldots, 9\}$.

A.2 The Simultaneous Move Game

Some (strictly) competitive networks have a split equilibrium.

Suppose $i \neq j$ and let $v_{i,i} = 5$, $v_{i,j} = 1$, and $N_A = N_B = 25$. Suppose the network functions are of the form

$$e_i(n_i, n_j) = \begin{cases} 0 & \text{if } n_i = n_j = 0 \\ \frac{n_i}{n_i + n_j} \cdot 50 & \text{otherwise} \end{cases}.$$

Clearly, this model satisfies (P1)-(P5) as well as (NT), and also has a split equilibrium because $v_{i,i} + e_i(N_i, N_j) = 30 > 27 = v_{i,j} + e_j(N_j + 1, N_i - 1)$.

Some competitive networks do not have a split equilibrium.

Consider the above example, but with $v_{i,i} = 1$ and $v_{i,j} = \frac{1}{2}$. Then there is no split equilibrium, because $v_{i,i} + e_i(N_i, N_j) = 26 < 26 + \frac{1}{2} = v_{i,j} + e_j(N_j + 1, N_i - 1)$.

Pareto efficiency of pure and split equilibria under (P1)-(P5) and (NT).

Consider the model where $v_{i,i} = 3$, $v_{i,j} = 1$, $N_A = 15$, and $N_B = 8$. Suppose the network effect functions are

$$e_A(n_A, n_B) = \begin{cases} 0 & \text{if } n_A < 10 \\ 10 & \text{if } n_A \geq 10 \end{cases} \quad \text{and} \quad e_i(n_i, n_j) = \begin{cases} 0 & \text{if } n_B < 10 \\ 12 & \text{if } n_B \geq 10 \end{cases}.$$

It is easy to verify that this model satisfies (P1)-(P5) and (NT), and that both pure equilibria and split equilibria exist. Under a pure A equilibrium, the payoff to each Type A agent is 13, and the payoff to each Type B agent is 11. In the split equilibrium, Type

A agents receive payoffs of 13 and Type B agents receive payoffs of 11. But if two Type A agents switch to good B, then all Type A agents (including those who defect) still receive payoffs of 13, while the Type B agents receive payoffs of 15. This profile dominates both the pure and split equilibria, so they are not Pareto efficient.

Pareto efficiency of pure equilibria with strictly increasing network functions.

To see that both pure equilibria need not be Pareto efficient, take $v_{i,i} = 2$, $v_{i,j} = 1$, $N_A = 5$, and $N_B = 4$. Suppose the network functions are of the form $e_i(n_i, n_j) = \frac{n_i}{n_i+1} M_i$, where $M_i > 0$ is some constant.

If $M_A = M_B = 10$, then it is easy to verify that both pure A and pure B equilibria exist. In the pure A equilibrium, Type A agents receive utility of 11, and Type B agents receive utility equal to 10. But these numbers are reversed in the pure B equilibrium, so neither pure equilibrium dominates the other. Combined with Lemma 4.1.5, this means both pure equilibria are Pareto efficient.

On the other hand, if $M_A = 10$ and $M_B = 14$, then both pure equilibria exist, but only the pure B outcome is efficient. This is because payoffs are once again 11 and 10 for Type A and Type B agents, respectively, in the pure A equilibrium, but these become 13.6 and 14.6 in the pure B equilibrium, so (B) Pareto dominates (A). Together with Lemma 4.1.5, this means that only (B) is Pareto efficient.

Pareto efficiency of split equilibria with strictly increasing network functions.

Consider a model with $v_{i,i} = 3$, $v_{i,j} = 1$, $N_A = N_B = 10$, and network functions of the form $e_i(n_i, n_j) = \frac{n_i}{n_i+1} 11$. Then a split equilibrium exists, and in this equilibrium Type A agents receive payoffs of 13 and Type B agents see payoffs of 11.08. However, there is also a pure A equilibrium in which these payoffs become 13.47 and 11.47, dominating the split outcome (similarly, the pure B equilibrium will also dominate the split outcome).

A.3 Proof of the Cascade Theorem for Strategic Agents

The proof is an "induction"-style argument, but is only applied to a finite subset $\{1, \ldots, N\}$ of the natural numbers, and builds up backwards, starting with agents N and $N-1$. The first lemma below is the "base case".

Lemma A.3.1. *For $N > 2$, there exists an integer $1 < k < N$ such that for every $k \leq i < N$, $\mathfrak{c}_i$ satisfies the Cascade Property.*

Proof. Since $N > 2$, $k = N - 1$ is between 1 and N, and the only integer i for which $k \leq i < N$ is k itself. So, it suffices to show that $\mathfrak{c}_{N-1}$ satisfies the Cascade Property; that is, it suffices to show that $\mathfrak{c}_{N-1}(i, A_{N-1}, B_{N-1}) = j \Rightarrow \mathfrak{c}_N(t_N, A_N, B_N) = j \neq i$. For brevity, this will only be demonstrated for the case of Type B agents selecting A, but symmetric arguments may be used to demonstrate it for the other case as well. So, assume agent $N-1$ is Type B and that he optimally chooses A.

First, observe that from agent $(N-1)$'s perspective, the probability of a Type A agent is $\frac{A_{N-1}}{N-1}$ and the probability of a Type B agent is $\frac{B_{N-1}+1}{N-1}$. Since there is only one agent remaining after $N-1$, the set of possible forms from agent $(N-1)$'s perspective is $\mathcal{F}_{N-1} = \{\langle A \rangle, \langle B \rangle\}$, with $P(\langle A \rangle) = \frac{A_{N-1}}{N-1}$ and $P(\langle B \rangle) = \frac{B_{N-1}+1}{N-1}$. Then agent $(N-1)$'s expected utility to choosing A is

$$\begin{aligned} EU_{N-1}(A) = v_{B,A} &+ P(\langle A \rangle) e_A(A_{N-1} + A(\langle A \rangle) + 1, B_{N-1} + B(\langle A \rangle)) \\ &+ P(\langle B \rangle) e_A(A_{N-1} + A(\langle B \rangle) + 1, B_{N-1} + B(\langle B \rangle)) \ , \end{aligned}$$

and his expected utility to choosing B is

$$\begin{aligned} EU_{N-1}(B) = v_{B,B} &+ P(\langle A \rangle) e_B(B_{N-1} + B(\langle A \rangle) + 1, A_{N-1} + A(\langle A \rangle)) \\ &+ P(\langle B \rangle) e_B(B_{N-1} + B(\langle B \rangle) + 1, A_{N-1} + A(\langle B \rangle)) \ . \end{aligned}$$

Now, agent $N-1$ optimally chooses A, which means $EU_{N-1}(A) > EU_{N-1}(B)$. It suffices to show that a Type B agent N optimally selects A (because then a Type A agent N

would also select A, which means $A(\langle A\rangle) = 1$ and $B(\langle A\rangle) = 0$); so, suppose to the contrary that agent N, of Type B, optimally selects B. Agent $N-1$ is aware that a Type B agent would do this, because agent $N-1$ knows the values of $\underline{A}_N$ and $\underline{B}_N$. So, agent $N-1$ knows that $A(\langle B\rangle) = 0$ and $B(\langle B\rangle) = 1$. Plugging these into the above expressions gives

$$EU_{N-1}(A) = v_{B,A} + P(\langle A\rangle)e_A(A_{N-1}+2, B_{N-1}) + P(\langle B\rangle)e_A(A_{N-1}+1, B_{N-1}+1)$$

and

$$EU_{N-1}(B) = v_{B,B} + P(\langle A\rangle)e_B(B_{N-1}+1, A_{N-1}+1) + P(\langle B\rangle)e_B(B_{N-1}+2, A_{N-1}) \ ,$$

where, of course, the first expression is greater than the second because $EU_{N-1}(A) > EU_{N-1}(B)$. Next, notice that $v_{B,A}+e_A(A_{N-1}+2, B_{N-1}) \geq EU_{N-1}(A)$ (because $e_A(A_{N-1}+2, B_{N-1}) \geq e_A(A_{N-1}+1, B_{N-1}+1)$ and $P(\langle A\rangle) + P(\langle A\rangle) = 1$). Similarly, $EU_{N-1}(B) \geq v_{B,B} + e_B(B_{N-1}+1, A_{N-1}+1)$ because $e_B(B_{N-1}+2, A_{N-1}) \geq e_B(B_{N-1}+1, A_{N-1}+1)$. Since $EU_{N-1}(A) > EU_{N-1}(B)$, this means that

$$v_{B,A} + e_A(A_{N-1}+2, B_{N-1}) > v_{B,B} + e_B(B_{N-1}+1, A_{N-1}+1) \ . \tag{A.1}$$

But this contradicts the fact that agent N selected B, because given that agent $N-1$ selected A, agent N will select B if and only if $v_{B,A} + e_A(A_{N-1}+2, B_{N-1}) \leq v_{B,B} + e_B(B_{N-1}+1, A_{N-1}+1)$. Therefore agent N, of Type B, will select A also. This implies that a Type A agent in position B would also select N, and the proof is complete. □

The next lemma establishes the following: if agent ℓ is Type i and optimally selects $j \neq i$, then, if every agent after ℓ satisfies the Cascade Property, agent ℓ will satisfy the Cascade Property as well.

Lemma A.3.2. *Suppose that every choice function* $\mathbf{c}_i$, $1 < k \leq i < N$, *satisfies the Cascade Property. Then* $\mathbf{c}_{k-1}$ *satisfies the Cascade Property also.*

Proof. To see that $\mathfrak{c}_{k-1}$ satisfies the Cascade Property, suppose (for simplicity) that agent $k-1$ is Type B and $\mathfrak{c}_{k-1}(B, A_{k-1}, B_{k-1}) = A$. It suffices to show that a Type B agent k will also choose A, because that would cause an A cascade, which means $A_{k-1}+2 \geq \underline{A}_{k+1}$; and a Type A agent in position k would also choose A after $k-1$ chooses A, so that $A_{k-1}+2$ is realized and a cascade occurs.

The first thing to note is that since a Type B agent in position k will cause an A cascade if he selects A, this means $EU_k(A) = v_{B,A} + e_A(A_k + 1 + N - k, k - A_k - 1)$. This is the best possible outcome for agent $k-1$, given that he selected A. Thus $EU_k(A) \geq EU_{k-1}(A)$. Since $k-1$ chooses A, this inequality may be combined with $EU_{k-1}(A) > EU_{k-1}(B)$ to yield $EU_k(A) \geq EU_{k-1}(A) > EU_{k-1}(B)$. Finally, it must also be the case that $EU_{k-1}(B) \geq EU_k(B)$, because relative to agent $k-1$, agent k has a lower subjective probability of Type B agents appearing because agent $k-1$ selected A. Thus, combining inequalities yields $EU_k(A) > EU_k(B)$, so that all agents after $k-1$ choose A, as required.

A similar argument may be used to show that if $k-1$ is Type A and chooses B, then all subsequent agents will also choose B. Thus, $\mathfrak{c}_{k-1}$ satisfies the Cascade Property. □

With these results in place, it is easy to prove the Cascade Theorem for strategic agents.

Proof of the Cascade Theorem for Strategic Agents. By repeated use Lemma A.3.2, it is clear that if there exists some agent $k > 1$ for which every agent i in $k \leq i < N$ satisfies the Cascade Property, then every agent $1 \leq i < N$ satisfies the Cascade Property (simply take the lemma to get that $k-1$ satisfies the Cascade Property; then reapply it to get that $k-2$ satisfies it also, until finally every agent satisfies it). Lemma A.3.1 establishes that at least one such agent k exists (namely, $k = N-1$), and so all strategic agents satisfy the Cascade Property. □

A.4 Proof of the Cascade Theorem for Naïve Agents

In order to prove this theorem, two basic facts about the binomial distribution must first be established. They are as follows:

Lemma A.4.1. *Let* $\mathbf{x} = (x_0, x_1, \ldots, x_n) \in \mathbb{R}^n$ *be a vector for which* $x_i \leq x_{i+1}$ *for every* $1 \leq i < n$. *Let* $\mathcal{B}$ *be the random variable formed by placing a binomial distribution on* $\mathbf{x}$ *with probability of success* $p \in (0,1)$*; let* $\mathcal{B}'$ *be the random variable formed by placing a binomial distribution on* $(x_1, \ldots, x_n)$ *with probability of success* $q \in (0,p)$. *Then* $E[\mathcal{B}'] \geq E[\mathcal{B}]$.

Proof. Step 1: There exists a minimal $0 \leq j^* \leq N$ for which $j \geq j^* \Rightarrow P(\mathcal{B}' = x_j) \geq P(\mathcal{B} = x_j)$. To see this, note that $P(\mathcal{B}' = x_j) = \binom{n-1}{j} p^j (1-p)^{n-1-j}$ and $P(\mathcal{B} = x_j) = \binom{n}{j+1} q^{j+1}(1-q)^{n-j-1}$. Then the statement $P(\mathcal{B}' = x_j) \geq P(\mathcal{B} = x_j)$ is algebraically equivalent to

$$(j+1)\left(\frac{p}{q}\right)^j \left(\frac{1-q}{1-p}\right)^j \geq nq\left(\frac{1-q}{1-p}\right)^{n-1}. \tag{A.2}$$

Notice that the right hand side is constant in j, and that since $p > q$, the left hand side is increasing in j. Notice also that if $j = n-1$, then the condition is equivalent to $p^{n-1} \geq q^n$, which is satisfied because $p > q$ and $p \in (0,1)$ imply that $p^n \geq q^n \geq q^n p$. Thus, at least one value of j exists for which the inequality is satisfied, which means there is a smallest value of j, j^*, which satisfies the inequality. Since the left hand side is increasing in j and the right hand side is constant, this means that the inequality is also satisfied for every $j \geq j^*$ (subject, of course, to the constraint that $j \leq n-1$).

Step 2: $E[\mathcal{B}'] \geq E[\mathcal{B}]$. Step 1 implies that $P(\mathcal{B}' = x_i) < P(\mathcal{B} = x_i)$ for every $0 \leq i < j^*$ and that $P(\mathcal{B}' = x_i) \geq P(\mathcal{B} = x_i)$ for every $j^* \leq i \leq n$. So, the difference between the two distributions is that relative to $\mathcal{B}$, $\mathcal{B}'$ has shifted probability weight to higher indices x_i; since the x_i are increasing in i, this means that relative to $\mathcal{B}$, $\mathcal{B}'$ places higher probability mass on greater values and lower mass on smaller values, so that $E[\mathcal{B}'] \geq E[\mathcal{B}]$. □

Lemma A.4.2. *Let* $\mathbf{x} = (x_0, x_1, \ldots, x_n) \in \mathbb{R}^n$ *be a vector for which* $x_i \leq x_{i+1}$ *for every* $1 \leq i < n$. *Let* $\mathcal{B}$ *be the random variable formed by placing a binomial distribution on* $\mathbf{x}$ *with probability of success* $p \in (0,1)$*; let* $\mathcal{B}'$ *be the random variable formed by placing a binomial distribution on* $(x_0, \ldots, x_{n-1})$ *with probability of success* $q \in (0,p)$. *Then* $E[\mathcal{B}] \geq E[\mathcal{B}']$.

Proof. The argument is similar to that in Lemma A.4.1 and therefore omitted. □

These lemmas allow two results about the choice functions $\mathfrak{c}_i$ to be deduced. For simplicity, the following results are given (and the main theorem is proved for) the case in which an A cascade occurs. Obviously, symmetric arguments may be used to establish the same results for the case of B cascades.

The first result states that if any agent (regardless of type) optimally selects A, then the next agent, if he is Type A, will also select A.

Lemma A.4.3. *For every model* $\mathcal{M} \in \mathbb{M}(N)$ *with naïve agents and every* $1 \leq k < N$, $\mathfrak{c}_k(t_k, A_k, B_k) = A \Rightarrow \mathfrak{c}_{k+1}(A, A_{k+1}, B_{k+1}) = A$.

Proof. Suppose $\mathfrak{c}_k(t_k, A_k, B_k) = A$. In particular, this means that a Type A agent in position k finds it optimal to choose A, so that

$$v_{A,A} + \sum_{j=0}^{N-k} \binom{N-k}{j} \left(\frac{A_k+1}{k}\right)^j \left(\frac{B_k}{k}\right)^{N-k-j} e_A(A_k+1+j, N-A_k-1-j) >$$
$$v_{A,B} + \sum_{j=0}^{N-k} \binom{N-k}{j} \left(\frac{B_k}{k}\right)^j \left(\frac{A_k+1}{k}\right)^{N-k-j} e_B(B_k+1+j, N-B_k-1-j).$$

Notice that the left hand side is the expectation of a binomial distribution, with probability $q = \frac{A_k+1}{k}$, over the values $x_j = v_{A,A} + e_A(A_k+1+j, N-A_k-1-j)$ for $0 \leq j \leq N-k$. Notice also that the expression

$$v_{A,A} + \sum_{j=0}^{N-k-1} \binom{N-k-1}{j} \left(\frac{A_k+2}{k+1}\right)^j \left(\frac{B_k}{k+1}\right)^{N-k-1-j} e_A(A_k+2+j, N-A_k-2-j)$$

is the expectation of a binomial distribution, with probability $p = \frac{A_k+2}{k+1} > \frac{A_k+1}{k} = q$, over the values $x_j = v_{A,A} + e_A(A_k + 1 + (j+1), N - A_k - 1 - (j+1))$ for $0 \leq j \leq N-k-1$. Thus, Lemma A.4.1 applies, and so

$$v_{A,A} + \sum_{j=0}^{N-k-1} \binom{N-k-1}{j} p^j (1-p)^{N-k-1-j} e_A(A_k + 2 + j, N - A_k - 2 - j) \geq$$
$$v_{A,A} + \sum_{j=0}^{N-k} \binom{N-k}{j} q^j (1-q)^{N-k-j} e_A(A_k + 1 + j, N - A_k - 1 - j).$$

In a similar fashion, letting $p = \frac{B_k}{k}$ and $q = \frac{B_k}{k+1}$, Lemma A.4.2 may be used to determine that

$$v_{A,B} + \sum_{j=0}^{N-k} \binom{N-k}{j} p^j (1-p)^{N-k-j} e_B(B_k + 1 + j, N - B_k - 1 - j) \geq$$
$$v_{A,B} + \sum_{j=0}^{N-k-1} \binom{N-k-1}{j} q^j (1-q)^{N-k-1-j} e_B(B_k + 1 + j, N - B_k - 1 - j).$$

Combining these inequalities yields

$$v_{A,A} + \sum_{j=0}^{N-k-1} \binom{N-k-1}{j} p^j (1-p)^{N-k-1-j} e_A(A_k + 2 + j, N - A_k - 2 - j) \geq$$
$$v_{A,B} + \sum_{j=0}^{N-k-1} \binom{N-k-1}{j} q^j (1-q)^{N-k-1-j} e_B(B_k + 1 + j, N - B_k - 1 - j),$$

where $p = \frac{A_k+2}{k+1}$ and $q = \frac{B_k}{k+1}$, which is exactly what is required to show that a Type A agent in position $k+1$ will select A. □

Repeated application of this result shows that if one agent selects A, then any subsequent chain of Type A agents will also select A. Of course, it is possible that a Type B agent will appear and select B, breaking the cascade. So, the next theorem demonstrates that if a Type B agent, k, optimally selects A, then a Type B agent in position $k+1$ will also select B (which, by the previous lemma, means that a Type A agent in position $k+1$ would also select A). Because of this, the following lemma is generalized to show that if agent k of Type B selects A, then a Type B agent in position $k+m$ ($m \geq 1$) will also select A, given that agents $k, k+1, ldots, k+m-1$ all selected A.

Lemma A.4.4. *For every model $\mathcal{M} \in \mathbb{M}(N)$ with naïve agents and every $1 \leq k < N$, $\mathfrak{c}_k(B, A_k, B_k) = A \Rightarrow \mathfrak{c}_{k+m}(B, A_k + m, B_k) = A$ for every $1 \leq m \leq N - k$.*

Proof. Let $1 \leq m \leq N - k$ and suppose $\mathfrak{c}_k(B, A_k, B_k) = A$. This means that

$$v_{B,A} + \sum_{j=0}^{N-k} \binom{N-k}{j} \left(\frac{A_k}{k}\right)^j \left(\frac{B_k+1}{k}\right)^{N-k-j} e_A(A_k + 1 + j, N - A_k - 1 - j) >$$

$$v_{B,B} + \sum_{j=0}^{N-k} \binom{N-k}{j} \left(\frac{B_k+1}{k}\right)^j \left(\frac{A_k}{k}\right)^{N-k-j} e_B(B_k + 1 + j, N - B_k - 1 - j).$$

Notice that the left hand side is the expectation of a binomial distribution, with probability $q = \frac{A_k}{k}$, over the values $x_j = v_{B,A} + e_A(A_k + 1 + j, N - A_k - 1 - j)$ for $0 \leq j \leq N - k$. Notice also that the expression

$$v_{B,A} + \sum_{j=0}^{N-k-m} \binom{N-k-m}{j} p^j (1-p)^{N-k-m-j} e_A(A_k + m + 1 + j, N - A_k - m - 1 - j)$$

is the expectation of a binomial distribution, where $p = \frac{A_k+m}{k+m} > \frac{A_k}{k} = q$, over the values $x_j = v_{B,A} + e_A(A_k + m + 1 + j, N - A_k - m - 1 - j)$ for $0 \leq j \leq N - k - m$. Repeated application of Lemma A.4.1 therefore implies that

$$v_{B,A} + \sum_{j=0}^{N-k-m} \binom{N-k-m}{j} p^j (1-p)^{N-k-m-j} e_A(A_k + m + 1 + j, N - A_k - m - 1 - j) \geq$$

$$v_{B,A} + \sum_{j=0}^{N-k} \binom{N-k}{j} q^j (1-q)^{N-k-j} e_A(A_k + 1 + j, N - A_k - 1 - j).$$

Similarly, repeated application of Lemma A.4.2 may be used to establish that

$$v_{B,B} + \sum_{j=0}^{N-k} \binom{N-k}{j} p^j (1-p)^{N-k-j} e_B(B_k + 1 + j, N - B_k - 1 - j) \geq$$

$$v_{B,B} + \sum_{j=0}^{N-k-m} \binom{N-k-m}{j} q^j (1-q)^{N-k-m-j} e_B(B_k + 1 + j, N - B_k - 1 - j),$$

where $p = \frac{B_k+1}{k} > \frac{B_k+1}{k+m}$. Combining these inequalities gives

$$v_{B,A} + \sum_{j=0}^{N-k-m} \binom{N-k-m}{j} p^j (1-p)^{N-k-m-j} e_A(A_k + m + 1 + j, N - A_k - m - 1 - j) \geq$$

$$v_{B,B} + \sum_{j=0}^{N-k-m} \binom{N-k-m}{j} q^j (1-q)^{N-k-m-j} e_B(B_k + 1 + j, N - B_k - 1 - j),$$

which is exactly what is required to show that $\mathfrak{c}_{k+m}(B, A_k + m, B_k) = A$. □

With these results in place, it is simple to prove the Cascade Theorem for Naïve Agents:

Proof of the Cascade Theorem for Naïve Agents. Let $1 \leq k < N$ and suppose $\mathfrak{c}_k(B, A_k, B_k) = A$. It suffices to show that $\mathfrak{c}_\ell(t_\ell, A_\ell, B_\ell) = A$ for every $k < \ell \leq N$. Lemma A.4.3, any consecutive sequence of Type A agents after agent k will choose A; and, because of this, Lemma A.4.4 guarantees that the first Type B agent after agent k will also select A. Applying the same argument, then, all agents after this Type B agent will also select A, resulting in an A cascade. □