LINUX ADOPTION BY FIRMS

by

Zheshi Peng

A thesis submitted to the Faculty of Graduate Studies and Research

in partial fulfilment of the requirements for the degree of

Master of Engineering in Telecommunications Technology Management

Department of Systems and Computer Engineering

Carleton University

Ottawa, Canada, K1S 5B6

May 3, 2004

The undersigned hereby recommend to

the Faculty of Graduate Studies and Research

acceptance of the thesis

LINUX ADOPTION BY FIRMS

submitted by

Zheshi Peng

in partial fulfilment of the requirements for the degree of

Master of Engineering in Telecommunications Technology Management

____

Rafik Goubran, Department Chair

____

A.J. Bailetti, Thesis Supervisor

Carleton University

May 3, 2004

ABSTRACT

The objective of this study is to examine the evolution of the market for Linux based products for the 1993-2003 period. Using data on 317 Linux suppliers available online, the differences in firms' size and in their first products were explored across the adoption stages of the Linux life cycle. Then two temporal patterns of the Linux-market were identified: changes in the entry rate of new Linux suppliers and changes in product diversity. Finally, the attributes of the partnerships formed by four major Linux distributors were examined. The study determined whether the number of partnerships formed by Linux distributors was related to the number of new entrants, whether the motives for partnerships formed by Linux distributors varied over adoption stages, and whether the type of partner selected by Linux distributors was a function of partnership motive. This study builds on the literature on open source software and traditional theories of technology adoption to make three important contributions. First, it develops a method to identify the stages of the life of a new technology. Secondly, it provides a way to measure the temporal patterns of the evolution of a new market. Finally, it validates the density-dependence model using data on open source.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor Professor Tony Bailetti for his guidance, understanding and patience throughout this study. I appreciate the manner and style with which he supervised this study. Professor Bailetti's approach allowed me the freedom to undertake this challenging topic, and made the process an enjoyable and educational one. Acknowledgement also goes to my parents and sister for their continuous support and endless love. Finally, I want to thank Maomi, my dear friend, for being such a wonderful source of encouragement and relief.

Table of Contents

List of Tables

List of Figures

List of Appendixes

1 INTRODUCTION

Linux is an open source software product that is freely available over the Internet. It is an operating system that runs on a number of hardware platforms including PCs and Macintoshes. Because Linux is an operating system, it is a key piece of technology that plays a role in connecting software applications to the hardware that it runs on.

Linux is a clone of the Unix operating system, written from scratch by Linus Torvalds with assistance from a large group of developers across the Internet. Developers share source code in order to refine the program and develop new features.

The Linux project is perhaps one of the most successful open source projects to date. Linux now holds 23.1% share of the server operating system market (Gonsalves, 2003). In 1995, Linux accounted for less than half of 1% of this market (Di Carlo, 2002). The success of Linux has attracted a significant amount of attention from researchers in diverse fields. Researchers examined various aspects of Linux development such as: 1) the merits of the open source model (Cubranic and Booth, 2000; Raymond, 1999; Von Hippel and Von Krogh, 2003); 2) software engineering issues relevant to the development of open source software (Godfrey and Tu, 2000); 3) the culture of open source developers (Hertel et al, 2003; Zeitlyn, 2003), and; 4) the economic issues around competitive firms' open source initiatives (Bonaccorsi and Rossi, 2003; Lerner and Tirole, 2002; Pal and Madanmohan, 2003; West and Dedrick, 2001; and West, 2003).

A careful review of the literature revealed very limited research on the adoption and diffusion of open source software. The study of the adoption of open source software such as Linux needs to be addressed from a life-cycle perspective and with a quantitative approach.

1.1 Relevance

This research is attractive because of the rapid adoption of Linux by commercial firms, and its open source nature. Since its introduction as a commercial product in 1993, Linux is now well beyond the steps of attempting to become a legitimate option for enterprise computing. It has become a legitimate option (Kusnetzky and Gillen, 2001).

As an operating system, Linux sits at the junction of hardware and software applications and has the potential to affect companies in both markets (Berquist et al, 2003). For example, independent software vendors (ISVs) need to write code using the Application Programming Interfaces (APIs) specific to the Linux operating system in order to access its functionality. Before Linux can run on a new hardware platform, vendors need to rewrite the particular part of the Linux kernel where the code is architecture dependent[1]. Similarly, vendors of peripheral devices need to develop specific driver programs for their hardware before it can be integrated into a Linux-base system.

Linux has received substantial support from software and hardware vendors. Linux-based software includes software that was specifically developed to run atop Linux and all the commercial Linux distributions. Linux-based hardware products are the hardware platforms and peripherals designed to run or support Linux.

1.2 Research Objectives

There are three objectives of this research. The first objective is to compare two attributes of Linux suppliers (i.e., companies that supply Linux based products) across adoption stages: firm size and product type.

To make comparisons of attributes across the stages, a method is developed to specify adoption stages in the Linux context. The researcher then tests whether the size of new entrants to the Linux market and the product types introduced by Linux suppliers are a function of adoption stage.

The second objective is to identify the temporal patterns in the evolution of the market for Linux products from 1993 to 2003. The purpose is to identify how the entry rate of new Linux suppliers (new entrants), and the diversity of products changed from 1993 to 2003.

The third objective is to examine the attributes of the partnerships formed by major Linux suppliers. The purpose is to identify: whether the number of partnerships formed by Linux suppliers was directly related to the number of new entrants; whether the motives for partnerships formed by Linux suppliers varied over adoption stages, and; whether the

---

[1] The *arch* subsystem of Linux source code contains the kernel code that is specific to particular hardware

type of partner selected by Linux suppliers was a function of the motive of the partnership.

1.3 Organization

This thesis is organized into seven chapters. Chapter 1 is the introduction. Chapter 2 reviews the relevant literature. Chapter 3 develops research hypotheses. Chapter 4 explains how this research is carried out in terms of research method, specification of research models, data gathering and data analysis. Chapter 5 provides the results. Chapter 6 discusses the results, and Chapter 7 draws conclusions, presents the limitations of this study and identifies opportunities for future research.

---

architectures/CPUs, including support for memory management and libraries.

2 LITERATURE REVIEW

This chapter is organized into five sections. The first section reviews the literature on Linux and open source software. The second reviews the literature on technology adoption theories. The third section reviews the literature on partnerships that is relevant to this research. The fourth section summarises the software taxonomy provided by the International Data Corporation (IDC). Finally, section five provides the lessons learned from the literature review.

2.1 Research on Linux and open source software

Linux is open source software. According to the trademarked definition, open source software is a product for which the source code is distributed or accessible via the Internet without charge or limitations on modifications and future distribution by a third party. As exemplified in Table 1, four areas of study are identified in the literature on open source software research. Among them, the most relevant one concerns economic issues around firms' open source initiatives.

Table 1: Existing literature on the research of open source software

Streams	Examples of Research Paper
Motivation and the culture of open source software developers; Protection of open source software	Zeitlyn (2003) • Gift giving and kinship amity Hertel, Niedner, and Herrmann (2003) • Individual's motivations O'Mahony (2003) • Practices to allow open source software to be governable and publicly available
The engineering issues of open source software	Godfrey and Tu (2000) • Examine the evolution of the Linux kernel code base at the system level and within the major subsystems.
Making sense of the open source innovation process and addressing the issues around the management of open source projects	Cubranic and Booth (2000) • Communication, Co-ordination and Version management in open source project Raymond, E. (1999) • First article that compares the open source development process with that of the commercial software. Von Hippel and von Krogh (2003) • Open source software development is an exemplar of a compound "Private-collective" model of innovation Von Krogh et al (2003) • The evolution of software architecture relates to the specialization of newcomers in a project
The economic issues around open source software	Pal and Madanmohan (2003) • Propose some strategies and practices for competing on Open source West and Dedrick (2001) • Open source standardization: The rise of Linux in the network era West (2003) • Hybrid strategies for computer platforms that include open source software Bonaccorsi and Rossi (2003) • Diffusion of open source software in a market dominated by commercial SW Lerner and Tirole (2002) • A preliminary exploration of the economics of open source software

2.1.1 Reasons why firms adopt Linux and the open source initiatives

Improvements in open source software are not appropriable by firms. Thus, a number of authors have examined the motivation firms have for adopting and backing open source software. Researchers have addressed the interesting question on how firms can profit from open source initiatives.

Lerner and Tirole (2002) found that firms who have adopted open source initiatives benefit indirectly in a complementary proprietary segment. Red Hat and VA Linux for example profit from providing complementary services and products that are not supplied efficiently by the open source community. Companies such as Hewlett-Packard release the source code of some of their existing proprietary products. This effort has helped the firm boost its profits on complementary market segments, while successfully offsetting their loss incurred due to the release of the source code. Similarly, Hawkins (2002) suggested that competitive firms release source code they are entitled to keep private so that it will become part of a maintained public code base, resulting in lower costs involved in maintaining it independently.

West and Dedrick (2001) found that the reasons for widespread support for Linux by established firms of hardware, software and service were that Linux responds to customer demands and reduces software development costs. There are strategic reasons for supporting Linux as well. Companies like Sun, IBM, and Oracle use Linux as a tool for challenging Microsoft in the server market. Firms like HP and IBM used Linux to catch up to Sun in the fast-growing hardware market for Internet servers. Other hardware vendors used Linux to become independent from Microsoft.

2.1.2 Open source software diffusion

Other papers of interest are those that consider the diffusion of open source software. West and Dedrick (2001) used the adoption patterns of Linux to illustrate the tension between control of a standard and the imperative for adoption. Linux takes an approach that has encouraged the widest possible adoption options through a liberal licensing regime, which encourages copying, and distribution.

Bonaccorsi and Rossi (2003) built a simulation model to evaluate the relevant factors in the diffusion of open source software in a market dominated by commercial ones. They found that under many plausible situations, commercial and open source software are likely to coexist in the future, given a distribution of beliefs biased towards open source and, in the absence of incumbent advantages, open source software could end up dominating the market.

In summary, previous studies contribute to our understanding of firms' adoption of Linux and open source initiatives, in terms of motives and means.

2.2 Theories of technology adoption and diffusion

Among the numerous theories and models of technology diffusion and adoption, three have been identified in the formation of a theoretical foundation for this research. These models are: Rogers's Innovation diffusion model, Moore's Technology Adoption Life Cycle, and the Density-dependence model. They have been chosen for their life cycle perspective on technology adoption as well as the impressive structural synopsis offered in describing the evolution of a market anchored around a new technology from birth to maturity.

2.2.1 Rogers' Innovation diffusion model

The innovation diffusion model by Rogers (1983) considers the manner and rapidity with which an individual or other unit of adoption responds to the offer of an innovation. Rogers categorised the adopters into five groups. Rogers observed that adopter distributions closely approach normality and he utilised the two statistics, mean (*X*) and standard deviation (*SD*) of a normal distribution of adopters to perform the adopter categorization.

Figure 1 shows that the area lying to the left of the mean time of adoption minus two standard deviations includes the first 2.5 percent of the individuals to adopt an innovation. These individuals are referred to as *innovators*. Similarly, the other four categories of adopters were identified and referred to as *early adopters*, *early majority*, *late majority*, and *laggards*.

Figure 1 also shows the two important curves of Rogers' innovation diffusion model: 1) the bell-shaped curve that represents the adoption rate of an innovation over time, and 2) the S-shaped curve, which represents the cumulative number of adopters.

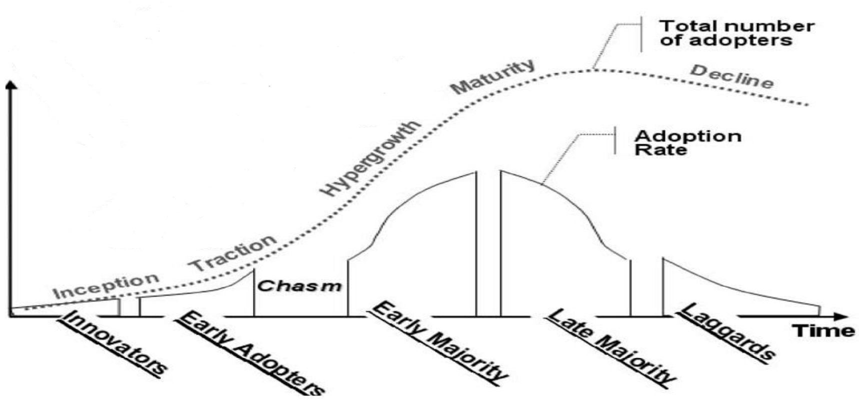
The characteristics of the five categories of adopters as defined by Rogers are as follows:

1. ***Innovators*** are eager to try new ideas. They embrace the new technology on its first appearance, in large part just to explore its properties to determine if it is "cool". In the Linux case, the innovators are the creators and early developers of Linux: the technology enthusiasts and advocates of the operating system.
2. ***Early adopters*** have the vision to adopt a new technology because of business opportunities or technology needs. They adopt it as a means for capturing a dramatic advantage over competitors who do not adopt it.
3. ***Early majority*** adopt new ideas just before the average number of a social system[2] (Rogers, 1983). They prefer to avoid the associated risk by staying away from the bleeding edge technology. They are quick to adopt the technology however, when the early adopters demonstrate benefits.
4. ***Late majority*** adopt new ideas just after the average number of a social system. They do not adopt a new technology until a majority of others in their systems have done so. The weight of public opinion must definitely favour the innovation before the late majority are convinced.
5. ***Laggards*** are last to adopt an innovation. They tend to be openly suspicious of innovations, innovators and change agents. When laggards finally adopt an

[2] A social system is defined as a population of individuals who are functionally differentiated and engaged in collective problem-solving behaviour (Rogers, 1983, p. 14).

innovation, it may already be superseded by a more recent idea that the innovators are using.

Figure 1: Technology Adoption Life Cycle



X – 2SD *X – SD* *X* *X + 2SD*

2.2.2 Moore's Technology Adoption Life Cycle (TALC) model

From Rogers' innovation diffusion model, it is evident that there are certain differences in the characteristics of adopters that influence them in their decision to try innovations within various time periods. Moore's Technology Adoption Life cycle (Moore 1991), which is derived from Rogers' model, clearly illustrates the evolution of a technology-enabled market that develops in a characteristic pattern. This is due to the aggregate effects of a certain population distributing its choices in the proportions outlined by the S-curve.

Moore argues that there is a “market chasm” between early adopters and the early majority. Moore believes that the needs of the early adopters are radically different from those of the pragmatic early majority that constitute the mass market. Moore has observed that many new products failed simply because they were not able to “cross the chasm”, in terms of new product design and marketing strategy, from the early market to the mass market. Figure 1 also illustrates the market chasm of Moore’s model.

2.2.3 Temporal Patterns and Density-Dependence Model

Temporal patterns refer to how entry, exit, network structure, supplier and product diversity and innovation vary from the birth of a new market through to maturity.

Temporal patterns of a technology-enabled market provide insights into the evolution and diffusion of a particular technology around which that market has formed and evolved. Among the temporal patterns concerning the evolution of a new market, density (i.e. cumulative number of entrants) and entry rate (i.e. rate of entry of new firms) are the two factors that have received the most attention in previous studies. The relationship between these two factors is usually interpreted by the Density-dependence theory.

Density-dependence theory explains the S-curve of adoption life cycle using the twin forces of “legitimization” and “competition”. These two forces help establish new technologies and then ultimately limit their take-up (Geroski 2000). Built on the density-dependence model, Debackere et al. (1998) investigated the entry pattern of the emerging technological communities. His results validated previous findings on the Bell-curve relationship between density and entry rate. The explanation for the Bell-curve relationship is as follows: entry rate initially accelerates with an increase in density and then slows down when density reaches a certain critical level. Another study by Wade (1995), investigating the rate at which communities of microprocessor manufacturers attracted organizational support, also confirmed the Bell-curve relationship between entry rate and density during the evolution of a new market.

In summary, three models of technology adoption offer a life-cycle perspective and a theoretical basis for the study of Linux adoption by firms and the Linux based market that has evolved from its inception to date.

Finally, there is no existing literature that provides an applicable method to identify the adoption stages of a new technology that has yet to reach the end of its lifecycle. The stages of Linux adoption by firms and the differences between them are two questions that remain unanswered in the current literature.

2.3 Literature review on firm strategic partnership formation

2.3.1 Motives of partnership formation

Various perspectives are used to examine the motives behind partnerships formation. These perspectives include: transaction costs (Williamson, 1985; Hennart, 1991), resource dependency (Pfeffer and Nowak, 1976), organizational learning (Hamel, 1991; Grant, 1996), strategic positioning (Porter and Fuller, 1986), and institutional theory (DiMaggio and Powell, 1983; Meyer and Rowan, 1977). From the resource dependence theory perspective, partnership formation is conceptualized as an organization's response to environmental changes demanding improvement or change in its resources or understanding of rapidly changing markets (Kogut, 1988).

Dodourova (2003) put forward a framework to classify the motives behind firms' forming strategic partnerships. According to this framework, the motives of partnership formation can be classified into eight groups: market-related motives, product-related motives, industry/market structure modification-related motives, timing-related motives, cost-related motives, competencies-related motives, and technology-related motives. The definitions of the partnership motive types in Dodourova's classification framework are included in Appendix I.

2.4 A review on software product categorization

IDC is global market intelligence and advisory firm in the information technology industry and its software taxonomy represents a collectively exhaustive and mutually exclusive view of the worldwide software marketplace (Heiman and Byron, 2003). Table 2 shows the IDC software taxonomy. It provides a basis for this research's categorization of Linux-based software products.

Applications include the software designed to automate specific sets of business processes in industry or business functions. Application development and deployment tools include application design tools, information and data management software, application deployment platforms, and middleware and application life-cycle management software. System infrastructure software includes system management software, network management, security software, storage software, serverware, networking software, and system-level software.

In summary, the software taxonomy from IDC provides a starting point for the categorization of the Linux based software products. Because this taxonomy is a general categorization of packaged software products, for the study of a particular context, a more concrete and specific categorization becomes necessary. Therefore, the IDC software taxonomy will be redefined to fit the context in this study.

Table 2: Software Taxonomy by IDC (2003)

Categories	Sub-categories
Applications	• Consumer software • Collaborative application • Content applications • Back-office application • Engineering applications • CRM and Sales and Marketing applications
Application development and deployment tools	• Information and Data management software • Application Design and construction software • Application life-cycle management software • Application deployment platforms • Middleware • Other development tools • Information access and delivery
System infrastructure software	• System management software • Network management • Security software • Storage software • Serverware • Networking software • System-level software

2.5 Lessons learned from literature review

A comprehensive review of existing literature has advanced our understanding of Linux adoption by firms and the traditional models of technology adoption. This section provides the five key lessons learned from the literature review.

2.5.1 Technology adoption follows an S-curve

Traditional models of technology adoption explain the dominant stylized fact that the use of new technologies over time typically follows an S-curve (Geroski, 2000). The Density-dependence model is one of these models. It explains the diffusion of a new technology with the twin forces: legitimization and competition. Legitimization helps establish a new technology. Competition ultimately limits its take-up.

2.5.2 Adopters of new technology can be classified into five types

The second lesson learned from the literature review is that technology adopters can be classified into categories based on level of innovation. Rogers (1983) was able to classify the adopters into five categories. This means that comparisons can be made for the purpose of identifying differences between categories.

The model proposed by Moore (1991) is an extension of Rogers' model. It proposes the existence of a market "chasm" between the early adopters and early majority. These two models provide a theoretical basis that this writer builds on to classify adopters and subsequently identify the differences in attributes of firms that adopt Linux along the technology adoption life cycle.

2.5.3 Lack of a method to define stages

The third lesson from the literature review was that a method to define a stage before the technology life cycle ends does not exist. Researchers have not defined how to identify the stages of the adoption life cycle prior to its ending.

2.5.4 Strategic and operational reasons support Linux adoption

The fourth lesson relates to the reasons for Linux adoption. Large firms have strong strategic reasons to support the adoption of Linux. Thus, in addition to cost reduction and meeting customer needs, there are a variety of issues relating to competitive aggressiveness that drive Linux adoption.

2.5.5 Lack of empirical research on open source diffusion

Most of the studies on open source software are case-based. Very few studies have examined Linux adoption by firms with a life-cycle perspective, and there is no empirical research on the temporal patterns of the evolution of the market for Linux-based products.

3 HYPOTHESES

This section is comprised of two parts. In section 3.1 the hypotheses to be tested are developed. Section 3.2 provides a list of the hypotheses developed.

3.1 Development of hypotheses

This chapter develops testable hypotheses based on three research objectives.

3.1.1 Compare attributes of Linux suppliers across adoption stages

The first research objective is to compare Linux suppliers across adoption stages by two attributes: firm size and product type. For a firm to enter a market at a certain stage, it requires a certain level of competence that enables it to successfully compete at that stage. It is expected that firm characteristics such as size and type of product introduced have some level of association with the adoption stages.

Firm size is important because many variables that affect competitive action are a function of company size. Firm size is taken as a proxy for technical competence and financial capability in many empirical studies of technology diffusion (Geroski, 2000). This writer expects the size of the company entering the Linux market at a given stage to be a function of the stage. Therefore,

Hypothesis 1: The size of a new entrant is a function of adoption stage.

Product type is a firm level attribute of interest. According to Rogers (1983) and Moore (2002), the characteristics and needs of the adopters of a new technology differ at varied stages along the adoption life cycle. Firms with different goals and abilities are likely to adopt the new technology at different times (Geroski 2000). Early stage adopters such as electronic design automation firms may adopt Linux-based solutions to run design suites, while at later stages corporate users may adopt Linux as a platform for mission-critical databases and applications. The change in the volume of each type of products introduced by new entrants at each stage reflects the change of the buyers' needs across the stages. This writer expects that the needs of the buyers will be associated with adoption stage. Given that the needs across stages are different, it is anticipated that new entrants will introduce new product types at each stage. Therefore,

Hypothesis 2: Type of product introduced by new entrants is a function of adoption stage.

3.1.2 Identify temporal patterns of Linux market evolution

The second research objective is to discover the temporal patterns in the evolution of the market for Linux products. The temporal patterns of interest in this study include information on how the entry rate of new firms and the diversity Linux products changed during the 1993-2003 period.

The adoption of Linux during the past decade has created a variety of Linux-based products. Linux product types range from the single category Slackware running on PCs to today's many Linux applications. There are obvious differences between 1993 and today in terms of diversity of Linux-based products, yet it remains unclear what the change pattern of product diversity is over time. For example, the availability of a database product running on Linux can attract new entrants such as vendors who deliver business applications based on that database. In turn, the increasing number of applications available on Linux further attracts associated vendors to provide Linux-based infrastructure software and certified hardware products. This writer expects product diversity to increase very quickly during the early stages and then to level off during the latter stages, despite the fact that the number of cumulative firms is still growing.

Hypothesis 3: Product diversity increases rapidly during the early adoption stages and then levels off, increasing slightly over the latter stages.

The twin forces of legitimization and competition have been used to explain the relationship between entry rate and density in a new market. At the early stages, Linux only attracted the technical enthusiasts. Mainstream vendors and large corporate customers were not likely to invest in Linux because they were not certain that the technology was going to survive, help create revenue or reduce costs. As start-ups such as Caldera and Red Hat entered the marketplace with successful business models, other firms started to follow. As more applications were ported to Linux and Linux itself was ported to various hardware platforms, Linux attracted more adopters and started to gain legitimacy. During the legitimization period, a new market experiences a rapid growth in the number of entrants and there is a positive relationship between the number of new entrants per month and the cumulative number of firms that have adopted Linux.

The legitimization period does not continue indefinitely. As the cumulative number of firms (or density) grows, competition for a limited number of customers and resources becomes the prevalent environmental force. During the competition period, a negative relationship exists between the number of new entrants per month and the cumulative number of firms that have adopted Linux (Debackere and Clarysse, 1998; Wade, 1995).

Based on what has been published to date, this writer expects the new entrants per month to increase and then fall when the density reaches a certain critical level. Therefore,

Hypothesis 4: The number of new entrants has a bell shaped relationship with the cumulative number of Linux suppliers.

The potential effect of firm age was not taken into account in Hypothesis 4. The regularity of entry rate for established firms potentially differs from that of start-ups. The entry rate of all new firms may follow the bell-shaped trajectory, while that of established firms may differ from this profile, therefore it is necessary to examine the relationship with two separate sets of data.

Hypothesis 4a: The number of new entrants that are established firms has a bell shaped relationship with the cumulative number Linux suppliers.

Hypothesis 4b: The number of new entrants that are start-up firms has a bell shaped relationship with the cumulative number of Linux suppliers.

3.1.3 Examine attributes of partnerships formed by major Linux distributors

To better understand the temporal patterns in the evolution of the Linux market, it is necessary to examine the partnership structure anchored around major Linux distributors (e.g., Caldera, Red Hat, SuSE, and TurboLinux). Thus, the third objective is to examine the attributes of the partnerships formed by major Linux suppliers from 1993 to 2003. The purpose is to identify: whether the number of partnerships formed by Linux suppliers was related to the number of new entrants; whether the motives for partnerships formed by Linux suppliers varied over adoption stages and; whether the type of partner selected by Linux suppliers was a function of the motive of the partnership.

Vendors possessing complementary assets collaborate with each other in order to provide customers with an integrated solution or "whole product" (Moore, 1991). Red Hat, for example, the market leader of Linux distribution in North America, attributes its success to its partnerships with platform vendors such as Dell, HP and IBM and software vendors such as Oracle, BEA, and Veritas (Alex, 2003).

During the early stages of evolution of the Linux market, there was a lack of collaboration between the distributors and established software or platform vendors. As Linux evolved, there have been an increasing numbers of partnerships established and the network structure of Linux-based market has become better developed. Debackere et al. (1998) suggested that potential entrants often face considerable 'searching costs' and they will tend to evaluate the overall attractiveness of an industry against the prestige position of a limited number of organizations. Debackere et al. (1998) used the number of partnership established by an organization as a proxy to measure the organization's prestige. This writer expects that as the number of partnerships formed by the major Linux distributors increased, so did the prestige of Linux in the overall market, attracting a greater number of new firms to provide Linux-compatible products. This writer predicts a positive relationship between the number of new entrants and the number of new partnerships formed by the Linux distributors. Therefore,

Hypothesis 5: The number of new partnership with Linux distributors is positively related to the number of new entrants.

The potential effect of firm age was not taken into account in Hypothesis 5. The regularity of entry rate for established firms potentially differs from that of start-ups, While the number of new partnerships may be positively related to the entry rate of established firms, it may have no such relationship with the entry rate of start-ups. It is necessary to examine the relationship with two separate sets of data. Therefore,

Hypothesis 5a: The number of new partnership with Linux distributors is positively related to the number of new entrants that are established firms.

Hypothesis 5b: The number of new partnership with Linux distributors is positively related to the number of new entrants that are start-ups.

Dodourova (2003) reported on the relationship between a firm's motives in forming a partnership and its strategic responses to the industry evolution. One interesting question to consider is whether there is a relationship between partnership motives and adoption stage. This writer expects that a relationship exists. Therefore,

Hypothesis 6: The motives for new partnerships are a function of adoption stage.

Dodourova (2003) found that firms formed partnerships to collaborate with a complementary partner and maximize existing capabilities. Firms enter into collaborative relationships in order to stretch their boundaries and gain access to complementary assets for the purpose of achieving their strategic objectives. An interesting question to consider is whether there is an association between the type of partner a Linux distributor selects and the motives behind the partnership. Therefore,

Hypothesis 7: The type of partner selected by a Linux distributor is a function of the motives behind the new partnership.

3.2 List of Hypotheses

The following hypotheses are tested in this thesis:

Hypothesis 1: The size of a new entrant is a function of adoption stage.

Hypothesis 2: Type of product introduced by new entrants is a function of the adoption stage.

Hypothesis 3: Product diversity increases rapidly during the early adoption stages and then levels off increasing slightly over the latter stages.

Hypothesis 4: The number of new entrants has a bell shaped relationship with the cumulative number of Linux suppliers.

Hypothesis 4a: The number of new entrants that are established firms has a bell shaped relationship with the cumulative number of Linux suppliers.

Hypothesis 4b: The number of new entrants that are start-up firms has a bell shaped relationship with the cumulative number of Linux suppliers.

Hypothesis 5: The number of new partnership with Linux distributors is positively related to the number of new entrants.

Hypothesis 5a: The number of new partnership with Linux distributors is positively related to the number of new entrants that are established firms

Hypothesis 5b: The number of new partnership with Linux distributors is positively related to the number of new entrants that are start-ups

Hypothesis 6: The motives behind new partnerships are a function of adoption stage.

Hypothesis 7: The type of partner selected by a Linux distributor is a function of the motives behind the new partnership

4 RESEARCH METHOD

4.1 Unit of analysis

The unit of analysis is a firm who made a commitment to Linux during the study period. A firm made a commitment to Linux when it introduced a Linux based product for sale.

4.2 Study period

The study period is from June 1993, the introduction of the first Linux commercial product, to December 2003.

4.3 Sample

4.3.1 New entrants sample

The sample that was used to test hypotheses 1 to 4 was comprised of all the suppliers that introduced Linux-based products from June 1993 to December 2003. This sample was referred to as the new entrants sample.

4.3.2 Partnerships sample

The sample that was used to test hypotheses five, six and seven was comprised of all the partnerships established by the major Linux distributors from June 1993 to December 2003. This sample was referred to as the partnership sample.

4.3.3 Important buyers sample

The sample that was used to identify the stages of the Linux adoption life cycle was comprised of important buyers who purchased Linux products from June 1993 to December 2003. This was referred to as the important buyer sample

4.4 Sample Selection

4.4.1 Selection of the new entrants sample

Online searches of the database *Business Source Premier* were performed to identify the suppliers that introduced Linux products during the study period.

Business Source Premier is a widely used database for business research. It includes 3300 scholarly journals and business periodicals, such as InfoWorld, PC week, Computing Canada, Computer World, PC Magazine, and Byte.com.

This writer first entered the word “Linux” into the search utility of the database in *Business Source Premier* and then examined all the entries to identify those that referred to a supplier introducing a new Linux product for the first time.

This writer created a database that included the names of suppliers that introduced a Linux product during the study period and the date on which the first product was introduced.

4.4.2 Selection of the partnerships sample

The partnerships sample was selected in two stages. In the first stage, the major Linux distributors were identified. A major Linux distributor had to meet all the following criteria:

- was embraced by IBM as being a Linux distributor
- was identified by Gartner as being a Linux distributor
- was based in North America or Europe

In the second stage, the partnerships formed by each of the major Linux distributors were identified. To identify the partnerships formed by the major Linux distributors, the *Business Source Premier* was searched, utilizing keywords such as alliance, collaboration, co-operation, agreement and partnership as well as the names of the Linux distributors.

4.4.3 Selection of important buyers sample

An important buyer was defined as a company identified by articles in the *Business Source Premier* and IDC reports as being an important buyer of Linux products.

Typically, an important buyer was one of the first adopters of a new type of Linux product or a first Linux adopter in a market segment that did not previously used Linux. Telia AB is a good example of a buyer that was important because it was one of the first companies to run Linux on an IBM mainframe. Jay Jacobs Inc. was an important buyer because it was the first retailer to adopt Linux.

Buyers identified in articles stored in *Business Source Premier* and IDC reports comprised the sample of important buyers. IDC did not conduct a market survey on Linux users until June 1997. Thus, important buyers before June 1997 were identified solely from *Business Source Premier.* If two or more buyers within a six-month period deployed the same type of Linux based product, only one was selected to reduce redundancy.

4.5 Data Collection

4.5.1 Company information

Company information such as size, years of operation, and products were obtained from the individual company websites and/or *Business Source Premier*.

4.5.2 Partnership information

Partnership information such as date of partnership announcement, firms participating in the partnership, firm type and a description of the partnership were obtained from the individual websites of major Linux distributors and/or articles stored in *Business Source Premier*.

4.6 Classifying Linux products and the motives to form partnerships

To test the hypotheses this writer needed to classify Linux products and the motives behind partnership formation. This section discusses the taxonomies used to accomplish this.

4.6.1 Classification of Linux products

Linux-based products in this research include both hardware and software products. For simplicity, the Linux-based hardware products were categorised into hardware platforms and peripherals, while a more structured way will be employed for the categorization of Linux-based software. This is due to the much higher diversity of Linux-based software and the fact that our Linux context has a naturally closer association with the software products.

Each Linux product was classified into one of the five sub-categories shown in Table 3. The product taxonomy shown in Table 3 uses the software taxonomy from IDC developed by Heiman and Byron (2003). Hardware products are classified in one of two sub-categories: platforms or peripherals.

Table 3: Classification of Linux-based products

Category	Sub-category	Examples of products
Linux-based software	1. Business applications ❑ Consumer software ❑ Messaging and conferencing software ❑ Content management software ❑ Back-office application ❑ Engineering applications ❑ CRM and Sales and Marketing applications	SAP R/3 on Linux, Accpac financial software for Linux.
	2. Application development and deployment tools ❑ Information and Data management software ❑ Application Design and construction software ❑ Application life-cycle management software ❑ Application deployment platforms ❑ Middleware	C-Forge integrated development tool for Linux, Adabas relational database.
	3. System software ❑ Linux distribution ❑ System management software ❑ Linux system management software ❑ Clustering and high-availability software ❑ Security software ❑ Backup, archive and storage management software	Red Hat Linux Advanced server, CA Unicenter TNG for Linux.
Linux-based hardware	4. Hardware platforms Hardware platforms designed or certified to run Linux	Certified Dell PowerEdge server for Linux.
	5. Peripherals Hardware components designed to provide specific support for Linux operating system	Ultra160 SCSI card from Adaptec Inc.

4.6.2 Classification of motives to form partnerships

The framework provided by Dodourova (2003) was used to classify the motives of major distributors in forming partnerships. According to this framework, the motives behind partnership formation can be classified into the eight groups. The framework is shown in Appendix I.

If there is more than one participating firm in the formation of a partnership with Linux distributors, the given partnership will be analysed as ***k*** individual partnerships (where ***k*** is the number of participating firms partnering with the Linux distributor). And, for each of the ***k*** partnerships, motives of Linux distributor's partnership formation are identified respectively according to the classification framework.

4.7 Identifying adoption stages

To test Hypotheses 1, 2, 3 and 6, the stages of the Linux adoption life cycle needed to be identified. No academic study provides a description of the adoption stages of the Linux life cycle or a process that can be used to identify them.

The researcher had three options:

- identify stages based on time intervals of equal duration (i.e., divide 10 years into five stages and define stage one as comprised of the first two years, stage two of the next two years, and so on)
- identify stages based on Rogers' approach (Rogers, 1983) (i.e., first stage ends at a time equal to the mean time of adoption minus 2 standard deviations, second stage ends at a time equal to the mean time of adoption minus one standard deviation, an so on)
- identify stages based on changes in the diversity of the products introduced to the market place (i.e., first stage ends when a new product type is introduced, second age ends when another new product type is introduced, and so on)

The first approach, while convenient, is not supported by logic or theory. No study has claimed that adoption stages have equal duration. Defining stages based on equal time periods is not supportable.

The second approach requires knowledge of the mean time of adoption and the standard deviation. These statistics will not be known, however, until the end of the Linux adoption life cycle. The distribution of Linux adopters is not known at this time. While the Rogers model (Rogers 1983) supports this approach, it cannot be used for the purpose of this research or any other research where the life cycle of a technology has not ended.

The third approach came from the realization that over the adoption life cycle, the type of products enabled by a new technology tends to change. According to Rogers (1983), Moore (2002), and Geroski (2000), the needs of the adopters of a new technology differ at varied stages along the adoption life cycle. Based on this information, it was concluded that the diversity of the set of products enabled by a new technology changes over time. If product diversity can be measured over time, changes in product diversity could be identified and used to mark the change from one stage to the next.

This writer decided to use the third approach to identify when one adoption stage changes to the next. This approach requires clarification on:

- what is meant by product
- what is the sample of interest

- what are the criteria to decide when changes in product creation signal a change in adoption stage

4.7.1 Product

In the Linux context, a "whole product" is comprised of 1) an application type, 2) a running environment type, and 3) a version of the Linux kernel. Commercial interests affect the first two, but not the third one. Accordingly, the concept of product was anchored around two constructs: application type and running environment type.

An application type refers to the type of workload or application that runs using the Linux kernel. It focuses on how Linux is being used. For example, at the early stages of the Linux adoption life cycle, the applications that used Linux included e-mail and printing. Today, mission critical applications use Linux.

The following software applications identified by Kusnetzky and Gillen (2001) were used in this research:

- Internet and Intranet
- File and print
- Application development
- Messaging
- Personal use
- System management
- Data warehouse and mart

- Scientific, engineering and Computer Assisted Design
- Enterprise applications
- Real time control
- Firewall
- Other

A running environment type refers to the hardware platform that the Linux kernel runs on. It focuses on the hardware platforms available to run Linux. For example, at the early stages of the Linux life cycle, Linux ran on Intel 386+ processors. Today, the hardware platforms that run Linux include: Intel IA-32/64, RISC based systems, mainframes, clusters, and appliance and embedded devices.

The following types of running environments were used in this research:

- PC (e.g. Intel 386+, Pentium, etc.)
- Intel IA-32/64 based server
- RISC-based workstation or server (e.g. HP PA-RISC, Compaq Alpha etc.)
- Mainframe (e.g. IBM mainframe)
- Linux cluster
- Embedded appliances (e.g. cell phone, task-specific appliances)

4.7.2 Sample of interest

To determine when an application type or a running environment type has a new value, a sample of companies that can be observed is needed. For the purpose of this research, it was decided to use a sample comprised of "important buyers."

Focusing on important buyers offers quality information on what and when a new type of Linux-based product was adopted. Considering the fact that reports of important buyers often provide first-to-the-world information on Linux adoption, the use of important buyers' data can be an effective and efficient way to capture required information to identify changes in Linux adoption stages.

4.7.3 Criteria that mark a stage change

The criteria used to decide when a change in product marked a change in stage were required. The criteria comprised three factors. The first focused on changes in application types and changes in running environment types. A change in either of the two types marked a change in the diversity of products. The second focused on the length of time between stage changes. The third focused on the minimum percentage of firms classified in a stage. The last two criteria were used in order not to generate a large number of stages.

For the purpose of this research, a new stage begins when all the following criteria are met:

- a new type of application or running environment is deployed by a firm in the sample of interest (i.e. an important buyer)
- the last stage change occurred more than 12 months ago
- at least 5 percent of the firms in the sample are classified to be part of the stage.

4.8 Measurement of Variables

This section describes how each variable was measured.

Adoption stages

Adoption stages = 1, 2, 3, 4… etc., they are identified using the approach described in section 4.7.

Density

Density = the cumulative number of firms that enter the Linux-based market minus those that exit the related business.

Density squared (DENSQ)

Density squared = DENSQ = (Density * Density) / 1000

Entry rate

Entry rate = the number of new firms that enter the Linux-based market at the end of each month.

Established firms vs. start-up

Company age is used to distinguish the established firms from the start-ups. If the firm was more than two years old when it introduced a Linux product, it was categorised as an established firm. Otherwise it was a considered a start-up.

Firm size

This study utilises two ways of measuring firm size: employee number and annual revenue reported at a date closest to the day date in which the firm introduced a Linux product.

Number of partnerships

Number of partnership = the total number of new partnerships major Linux distributors formed at the end of each month during the observation period.

Product category

Product category includes five major types: Business applications, Application development and deployment tools, System software, Hardware platforms, and Peripherals.

Product diversity

Product diversity = $1- \Sigma (p_i)^2$, where p_i is the percentage product category i represents among the total products.

Product diversity was measured using data on the first product introduced by a firm instead of including all the available products in the market at the end of each month. The number of firms who introduced multiple types of products was small and would not greatly influence the measure of product diversity.

A small program written in C language was developed for the purpose of calculating product diversity.

4.9 Specification of the Density-dependence model

The entry rate of firms into the market for Linux based products is estimated in discrete time (i.e. event count analysis). In event count analysis, the observation period is divided into fixed disjoint time intervals that occur in series, and then the number of events that occur in every interval is counted (King 1988).

In this study, the new firms that commit to Linux in month t were counted (this was denoted by $\mathbf{Y_t}$). According to Amburgey and Carroll (1984), the probability that exactly **k** events occur ($\mathbf{Y_t} = \mathbf{k}$, indicating **k** new firms commit to Linux) in a given interval (i.e., a given month, t, in our observation period) follows the Poisson distribution with parameter λ:

$$\Pr(Y= k) = \lambda^{k}\exp(-\lambda)/k! \qquad (1)$$

The mean of the Poisson distribution of the number of events in a fixed interval of length **t** equals its variance (all equal to λ), i.e.:

$$E[Y] = \lambda \text{ and } Var[y] = \lambda. \quad (2)$$

As in a linear regression model, the conditional means function will be modelled using a linear combination $\alpha_t x_i$ of the explicative variables:

$$E[Y_t|x_i] = \lambda_t = \exp(\alpha_t x_i) \quad (3)$$

Equation (3) denotes that every Y_t (at time 0 to t) conditional on certain characteristic x_i follows a Poisson distribution with parameter λ_t. The use of the exponential function in (3) assures that the predicted entry rate will be non-negative[3] (Debackere 1998).

To test hypotheses 4, 4a, and 4b, two explicative variables have been identified for the Poisson model, Density and Density Squared (denoted DENSQ), each of which is known to have an influence on the dependent variable entry rate.

Specifically, the full model in our analyses is as follows:

$$E[Y_t|x_i] = \lambda_t = \exp(\alpha_1 DEN_t + \alpha_2 DENSQ_t + C) \quad (4)$$

Where α_1 and α_2 are the coefficients, and C is a constant, and Hypothesis 4, 4a and 4b requires that $\alpha_1>0$ and $\alpha_2<0$, in order to reject the Null hypothesis. As mentioned above, the assumption of Poisson regression model requires mean and variance all equals to parameter λ; however, in the case of a violation, the above analyses will cause inconsistency and hence invalidate the hypotheses test. To solve this problem, various existing literatures recommend the additional use of Negative Binomial model to secure the findings (Debackere and Clarysse, 1998; Wade, 1995). It is an extension of the Poisson regression model. It captures the degree of over dispersion in the entry rates. Both models will be used in this study to test the Hypotheses 4, 4a, and 4b.

4.10 Data Analysis

4.10.1 Descriptive statistics

The data collected were coded and calculations were performed in order to derive measures for the variables used in this study. Descriptive statistics were computed for all the variables including frequencies, means, standard deviations, and correlation coefficients.

4.10.2 Hypothesis 1: The size of new entrant is a function of adoption stage

To test Hypothesis 1, this writer would determine if the firm size data was normally distributed, then would check whether the variance is equal (or approximately equal) between the data when split across the different stages.

If the firm size data were normally distributed, a one-way ANOVA test (such as Duncan's multiple range Test) would be used. If the data were not normally distributed, the Kruskal-Wallis H Test and Mann-Whitney U Test will be used to perform the multiple comparison of firm size between stages.

---

[3] This is a desirable characteristic, as negative entry rates are meaningless.

For the one-way ANOVA Duncan's multiple range test, if the mean difference between a pair of stages is significant at $P < .05$, or for both Kruskal-Wallis H Test and Mann-Whitney U Test, if the mean rank difference is significant at $P < .10$, this writer can conclude that the size of new entrants is significantly different between the two stages (at $P < .05$ and $P < .10$ respectively).

4.10.3 Hypothesis 2: Type of product introduced by new entrants is a function of adoption stage

This writer first created a contingency table, filled the cell frequencies of product type at each stage and then used the SPSS software to calculate the Chi-square.

If the Chi-square test shows a result significant at $P < .05$, this writer can conclude that the type of product introduced by new entrant is significantly associated with adoption stage at $P < .05$.

4.10.4 Hypothesis 3: Product diversity increases rapidly to a critical level during the early adoption stages and then levels off, increasing slightly over the latter stages

Compute the product diversity using a formula for product diversity = $1 - \Sigma (p_i)^2$ based on all the Linux-based products available at the end of each month.

If the data were normally distributed, this writer would perform multiple comparisons of product diversity between every two adjacent stages using a one-way ANOVA test. If the data were not normally distributed, the Kruskal-Wallis H Test and Mann-Whitney U Test would be used.

For the one-way ANOVA test, if the mean difference between two adjacent stages is significant at $P < .05$, or for Kruskal-Wallis H Test and Mann-Whitney U Test, if the mean rank difference is significant at $P < .10$, it can be concluded that product diversity is significantly different between the two stages (at $P < .05$ and $P < .10$ respectively). Otherwise, product diversity is considered to level off over the two adjacent stages.

4.10.5 Hypothesis 4: The number of new entrants has a bell shaped relationship with the cumulative number of Linux suppliers

To test hypothesis 4, this writer first calculated the entry rate and density for each month during the observation period, then ran the Poisson regression test using the LIMDEP software provided by Greene (2000).

If the regression test result does not meet any of the two conditions 1) coefficients of density and density squared are in the hypothesised direction (i.e. $\alpha_1 > 0$ and $\alpha_2 < 0$ in the model specified in section 4.9) and 2) test result is significant at $P < .01$, this writer will reject Hypothesis 4, otherwise, Hypothesis 4 will be tested again using the Negative Binomial model.

When testing with the Negative Binomial model, if the test result corroborates that of the Poisson regression model, this writer will conclude that Hypothesis 4 is supported.

To test Hypotheses 4a and 4b, this writer will split the sample into established firms and start-ups and repeat the procedure described above.

4.10.6 Hypothesis 5: The number of new partnerships with Linux distributors is positively related to the number of new entrants

According to Wade (1996), industry quarterly data have a substantial impact on firms' strategic decision making. The writer surmised that there is a three-month delay between when a firm decides to enter the market due to a partnership with a Linux distributor and when that partnership is announced to the public. For this reason, to test Hypothesis 5, this writer measured the strength of the association between the number of new entrants at month k and the number of new partnerships with Linux distributors at month k-3.

The writer first checked the normality of the data, and if normally distributed, used Pearson correlation. Otherwise, Spearman correlation was used.

If the correlation coefficient in Pearson correlation test result was positive and significant at $P < .01$, or the correlation coefficient in Spearman correlation test result was positive and significant at $P < .05$, this writer concluded that the number of new partnerships with Linux distributors is positively correlated with the number of new entrants.

To test Hypotheses 5a and 5b, this writer will split the sample into established firms and start-ups and repeat the procedure described above.

4.10.7 Hypothesis 6: The motives for new partnerships are a function of adoption stage

To test Hypothesis 6, this writer first created a contingency table, filled the cell frequencies using the motives of partnerships formed by Linux distributors at each stage and then computed the Chi-square.

If the Chi-square test showed a result of association between rows and columns that was significant at $P < .05$, this writer concluded that the motives for new partnerships are a function of stage.

4.10.8 Hypothesis 7: The type of partner selected by a Linux distributor is a function of the motives for the new partnership

To test Hypothesis 7, this writer first created a contingency table, filled the cell frequencies using the types of partners and the types of partnership motives and then computed the Chi-square.

If the Chi-square test showed a result of association between rows and columns that was significant at $P < .05$, it was concluded that the type of partner selected by a Linux distributor is a function of the motives for the partnerships.

5 RESULTS

5.1 Sample

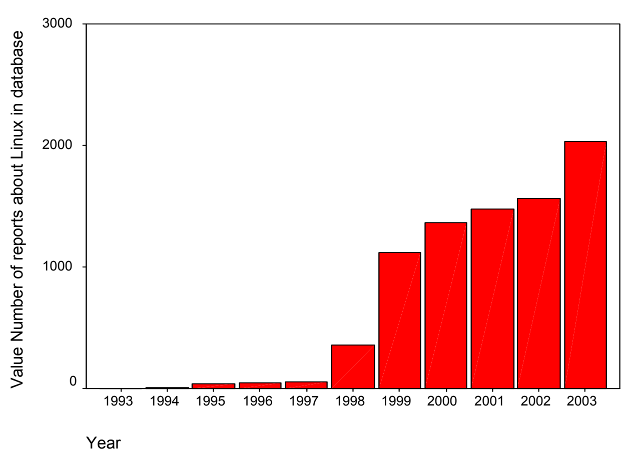
5.1.1 New Entrants

The new entrants sample was comprised of 317 suppliers of Linux products. Data on annual revenue was available for 259 of the 317 suppliers in the sample, data on number of employees was available for 212, and data on firm age was available for 258 suppliers.

Entering the word "Linux" into the search utility of *Business Source Premier* resulted in 8100 entries. For the study period, Figure 2 shows the number of entries by year.

From the 8100 entries 317 suppliers were identified that introduced Linux products from June 1993 to December 2003.

Figure 2: Yearly number of articles about Linux in business source premier



5.1.2 Partnerships

The partnership sample was comprised of 61 partnerships formed by four major Linux distributors from June 1993 to December 2003.

Four companies met the criteria for major Linux distributor: Red Hat Linux, Caldera, SuSE and TurboLinux. These four companies were

- embraced by IBM as being Linux distributors
- identified by Gartner as being Linux distributors
- based in North America or Europe

Searches using the names of these four companies and a set of keywords (alliance, collaboration, co-operation, agreement and partnership) resulted in 61 entries. Examining these entries 61 different partnerships were identified.

5.1.3 Important Buyers

The important buyer's sample comprised 21 companies.

Using the search utility in *Business Source Premier*, 56 online articles were found describing the purchase of Linux products. Two IDC reports that provided details about Linux purchases were found (Adelson, 2002; Kusnetzky and Gillen, 2001). From the 56 online articles and two IDC reports, 36 organizations that purchased Linux products were identified.

Of these 36 buying organizations identified, 15 were excluded. Two were excluded because they were government departments. Thirteen were excluded because they were buyers for which it was not clear what type of Linux products were bought and deployed.

5.2 Descriptive Statistics

5.2.1 Characteristics of the community comprised of new entrants over the study period

This study's observation period (June 1993 to December 2003) includes 127 months. Thus, there are 127 observations of the variables entry rate, exit rate, density, number of partnerships formed, and product diversity. Table 4 shows the descriptive statistics of these variables.

Table 4: Descriptive statistics for variables that describe the characteristics of the community of Linux suppliers during the 127 months of the study period

	N	Min.	Max.	Mean	Standard Deviation	Skew Std. Error = .22	Kurtosis Std. Error = .43
Firm entry rate (total)	127	0	17	2.50	3.18	2.03	5.17
Entry rate (established firms)	127	0	14	1.99	2.64	2.08	5.34
Entry rate (start-ups)	127	0	4	.50	.85	1.93	3.51
Firm exit rate	127	0	3	.28	.56	2.15	4.97
Density (cumulative number of firms)	127	1	282	97.91	100.16	.53	-1.43
Density squared/1000	127	.001	79.52	19.54	25.44	.94	-.56
Number of partnership with Linux distributors	127	0	4	.46	.82	1.97	3.74
Product diversity	127	.00	.92	.80	.15	-1.85	5.22

Figures 3 and 4 illustrate the entry rate of Linux suppliers (i.e., the number of suppliers that introduced a Linux product for the first time in a given month) and the density (i.e., end of month cumulative number of Linux suppliers) by month for the 127-month study period.

Figure 3: Entry rate of Linux suppliers

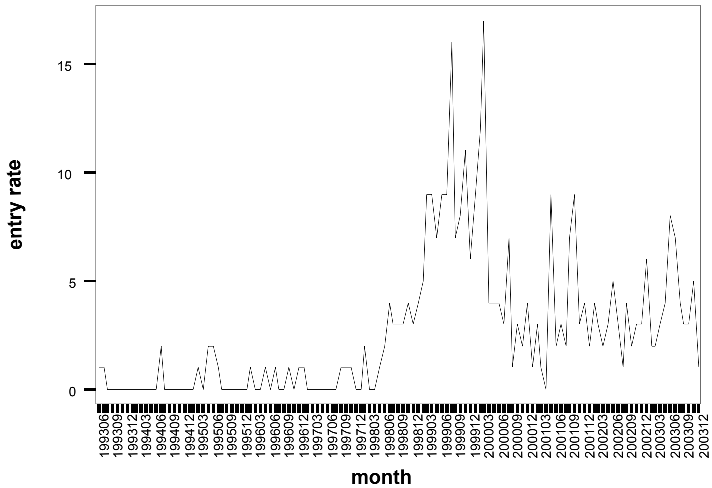


Figure 4: End-of-month cumulative number of Linux suppliers

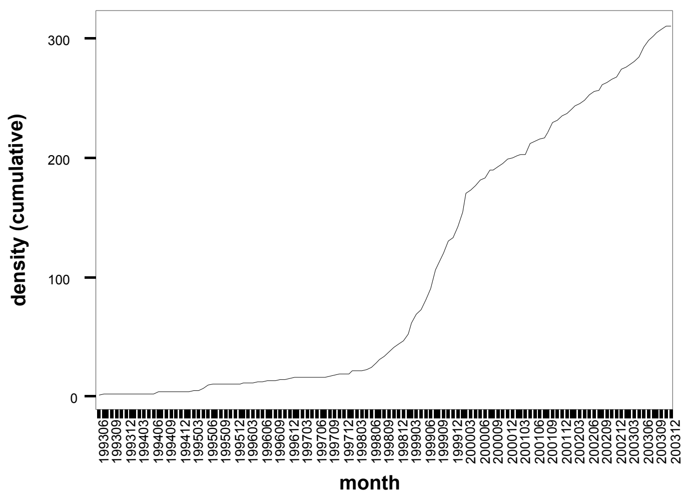
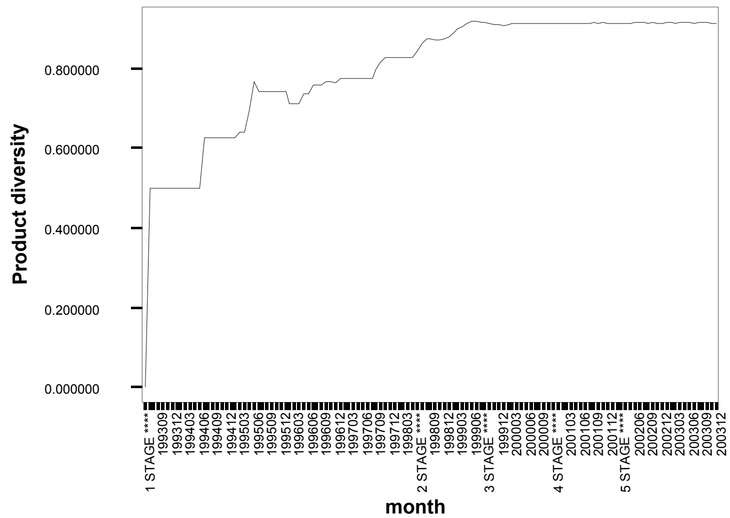


Figure 5 shows how product diversity changed over the study period. Product diversity increased very quickly at the very beginning. The appearance of the second product type results in product diversity increasing from 0 to 0.5. Product diversity increases rapidly during the period from June 1993 to the middle of 1999 and then levels off over time.

Figure 5: Product diversity



5.2.2 Annual revenue, employee number and age of new entrants at the time they introduced their first Linux products

Table 5 shows the descriptive statistics for the suppliers in the sample observed when they introduced their first Linux product. While the sample is comprised of 317 companies, annual revenue data was available for only 259 companies, employee number data for 212 companies, and company age data for 258 companies. Complete data for only 211 companies existed.

Table 5: Descriptive statistics for the new entrants' annual revenue, employee number, and age at the time they introduced their first Linux product

New entrant	**N**	**Min.**	**Max.**	**Mean**	**Standard Deviation**	**Skews**	**Skewness (Standard Error)**	**Kurtosis**	**Kurtosis (Standard Error)**
Annual revenue	259	.10	85337	2733.99	10313.08	5.28	.15	31.67	.30
Number of employees	212	4	426000	11985.29	44983.51	6.33	.17	47.56	.33
Age	258	5	105.0	15.35	18.61	2.85	.15	8.66	.30

Table 5 indicates significant levels of kurtosis and skewness in the data for annual revenues, employee number, and supplier's age. For each variable, the kurtosis is outside of the –2 to +2 range and the skewness is greater than twice the standard error. Figures 6a, 6b and 6c illustrate the distribution of Linux suppliers by annual revenue, number of employees and firm's age. The three figures illustrate positive skewness. There are far more Linux suppliers that have low annual revenue, fewer employees and are younger than Linux suppliers that have high annual revenue, large number of employees, and are established.

Figure 6a: Histogram for annual revenues

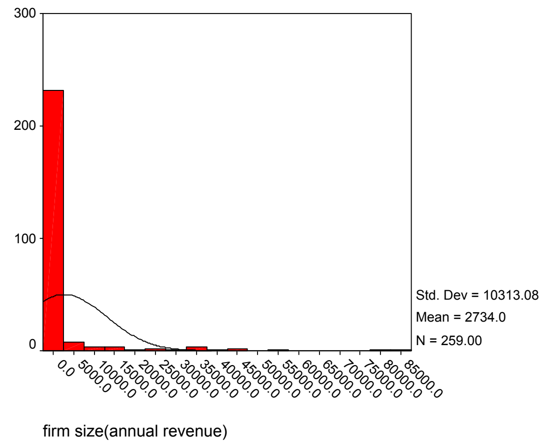


Figure 6b: Histogram for number of employees

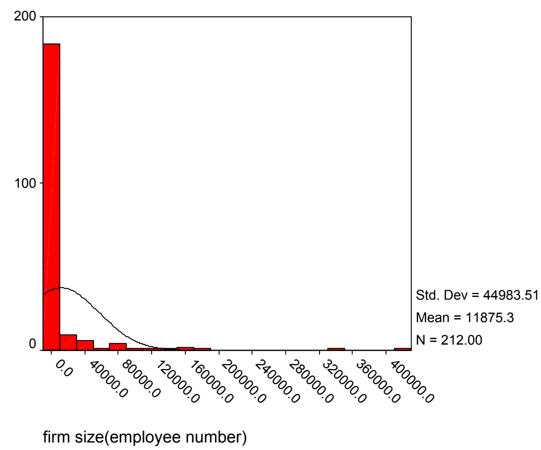


Figure 6c: Histogram for firm age

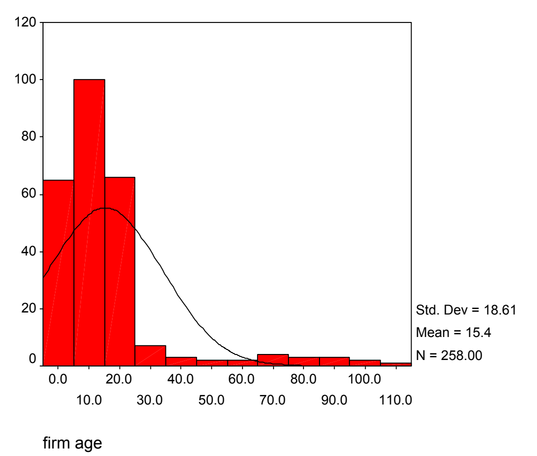


Table 6 provides the Spearman correlation coefficients for annual revenue, employee number, and firm age. The three variables are highly correlated at $P < .01$ (two-tailed).

Table 6: Spearman correlation coefficients for suppliers' annual revenue, number of employees and age

	Annual revenue	**Number of employees**
Number of employees	.959**	1.000
Age	.622**	.666**

**. Correlation is significant at the 01 level (2-tailed).

5.2.3 Characteristics of Linux partnerships

A total of 61 announcements of partnership agreements formed by the four major Linux distributors were identified. Examination of the 61 announcements led to the identification of 96 motives for partnering. Many announcements described more than one motive for the partnership. In only two partnership announcements, more than one partner was identified.

Table 7 organizes the motives for partnerships into the "partnership motive" types identified by Dodourova (2003).

Table 7: Motives for partnerships announced by Linux distributors organized by type

Partnership motive type	Count
Product/service related motive	39
Market-related motives	33
Technology-related motives	13
Competencies and skills-related motives	6
Industry/market structure modification-related motives	5
Total	96

5.3 Identification of adoption stages

The 21 companies in the important buyers sample are listed in the first column of Appendix II. The second and third columns of Appendix II show the application type and running environment type identified for each important buyer. The fourth column shows the date the buyer deployed the Linux products.

The changes in adoption stages were identified applying the criteria described in section 4.7.3. A new stage begins when an important buyer introduces a new type of application or running environment and the last stage occurred more than 12 months prior to this date. For the new stage to exist, it requires at least 5 percent of the firms in the new entrants' sample.

5.3.1 Changes in application and running environment types

Table 8 shows how potential adoption stage changes were identified. The start of this study, June 1993, marks the beginning of stage 1. At this time, Linux was used to run applications that were usually developed in-house where no third party applications were available.

The first stage change occurred in December 1996, when it was announced that small ISPs were using Linux for a new type of application: Linux was served as a platform for Internet and Intranet infrastructures that include Web server, File/Print server, DNS and Firewall.

The second stage change occurs in June 1998 when Digital Domain announced that it was using Linux for a new type of application: computing processing. The change in application type occurred more than 12 months after the previous change had occurred.

The next three changes in applications and running environments shown in Table 8 (use of RISC-based machines and Linux clusters as the running environments and use of Linux as the front end processor of business applications) did not result in a stage change because they occurred less than 12 months after the previous stage change.

The third stage change occurred in August 1999 with the announcement that Motorola was using Linux as a platform for telecommunication appliances. This announcement came more than 12 months after the second stage change (i.e., June 1998).

The next change in running environments announced by Kenwood in November 1999 did not result in a stage change because it occurred within three months after the third stage change. For similar reasons, the change in application type announced by 1stUP.com occurred in July 2000 did not trigger a stage change.

The fourth stage change occurred in December 2001, with the announcement that Telia AB was using Linux in IBM mainframes S/390.

The fifth stage change occurred in February 2002 with the announcement that Merrill Lynch was using Linux to support mission critical applications.

Table 8: Changes in applications and running environments identified from important buyers data

Name of construct that receives a new value	Value	Year and month	Sources	Stage transition
Running environment	PC	1993/06	Researchers, scientists	Start of Stage 1
Application type	Internet/Intranet Infrastructure (WEB, DNS, File/Print, Firewall)	1996/12	Small ISPs	YES
Application type	Computing	1998/06	Digital Domain	YES
Running environment	RISC –based machine	1998/06	Digital Domain	NO
Application type	Business application (front-end)	1999/02	Burlington Coat Factory warehouse corp.	NO
Running environment	Linux cluster	1999/05	Amerada Hess Corp	NO
Application type	Platform for telecommunication appliances	1999/08	Motorola	YES
Running environment	Embedded system	1999/08	Motorola	NO
Running environment	Intel IA-32/64 Server	1999/11	Kenwood	NO
Application type	Data warehousing	2000/07	1stUp.com	NO
Running environments	IBM mainframe S/390	2000/12	Telia AB	YES
Application type	Business application (mission-critical back-end)	2002/02	Merrill Lynch	YES

5.3.2 Each stage must include at least five percent of the sample

Based on the dates corresponding to the five stage changes shown in Table 8, six potential adoption stages were identified. Table 9 shows the results of classifying the 317

Linux suppliers in the sample into these six stages based on the date they introduced their first Linux product.

Table 9: New entrants in sample by adoption stage using six stages

	Year and month stage started	Frequency	Percent	Cumulative percent
Stage 1	1993/06	14	4.4	4.4
Stage 2	1996/12	8	2.5	6.9
Stage 3	1998/06	72	22.7	29.7
Stage 4	1999/08	107	33.8	63.4
Stage 5	2000/12	45	14.2	77.6
Stage 6	2002/02	71	22.4	100.0
Total		317	100.0	

Table 9 shows that the first stage includes less than five percent of the sample. Thus, the first stage in Table 9 does not meet the criteria for an adoption stage. For this reason, the first and the second stages were consolidated. Table 10 shows the results of classifying the 317 Linux suppliers into five adoption stages. The five stages shown in Table 10 were the ones used to test hypotheses 1, 2, 3 and 6.

Table 10: New entrants in sample by adoption stage using five stages

	Year and month stage started	Frequency	Percent	Cumulative percent
Stage 1	1993/06	22	6.9	6.9
Stage 2	1998/06	72	22.7	29.7
Stage 3	1999/08	107	33.8	63.4
Stage 4	2000/12	45	14.2	77.6
Stage 5	2002/02	71	22.4	100.0
Total		317	100.0	

5.4 Hypothesis 1: The company size of a new entrant is a function of adoption stage

Hypothesis 1 is to test the association between adoption stage and the company size of Linux suppliers at the time they introduce their first Linux product. Annual revenue was used to measure the size of new entrants. When annual revenue was used, the size of the sample went from 317 to 259.

For each of the five adoption stages identified in the previous section, Table 11 provides the numbers of new entrants for which annual revenues were available.

Table 11: Number of new entrants for which annual revenues were available by adoption stage

	Year and month stage starts	Frequency	Percent	Cumulative percent
Stage 1	1993/06	17	6.6	6.6
Stage 2	1998/06	60	23.2	29.7
Stage 3	1999/08	89	34.4	64.1
Stage 4	2000/12	37	14.3	78.4
Stage 5	2002/02	56	21.6	100.0
Total		259	100.0	

The Kruskal-Wallis test was used on the data in Table 11 to test for differences in means. The Kruskal-Wallis test is a popular nonparametric alternative to the standard one-way ANOVA. It was used because the annual revenue data was skewed.

The Kruskall-Wallis test resulted in a Chi-square equal to 12.915 with 4 degrees of freedom and a significance of $P < .012$. Thus, at $P < .05$ the result of the Kruskal-Wallis test indicated that the means for annual revenues are different when grouping is based on the five adoption stages.

Table 12 shows the results of using the Mann-Whitney U test that compared the annual revenues of the firms entering the Linux market at stage 1 and stage 2. The results indicate that the mean rank of firm size in stage 1 is significantly smaller than that of stage 2.

Table 12: Mann-Whitney U test comparing firm size between Stage 1 and Stage 2

	Adoption stage	N	Mean Rank	Sum of Ranks
firm size(annual revenue)	Stage 1	17	24.15	410.50
	Stage 2	60	43.21	2592.50
	Total	77		

Test Statistics[a]

	firm size(annual revenue)
Mann-Whitney U	257.500
Wilcoxon W	410.500
Z	-3.102
Asymp. Sig. (2-tailed)	.002

a. Grouping Variable: Adoption stage

Appendix III provides the Mann-Whitney U tests for the remaining pairs of stages. The results of all the Mann-Whitney U tests, indicate the following:

- Suppliers entering the Linux market at stage 1 are significantly smaller companies than those entering the Linux market at stage 2 ($p<0.002$).
- Suppliers entering the Linux market at stage 2 are significantly larger companies than those entering the Linux market at stage 3 ($p < 0.051$)
- There is no significant difference between the company size of the suppliers that entering the Linux market at stages 3, 4 and 5

The results support the acceptance of Hypothesis 1. The findings above suggest that there is an association between the company size of new entrants and the adoption stage; however, this association only exists in the early stages. As the Linux adoption process advances, this association diminishes.

5.5 Hypothesis 2: The type of product introduced by new entrants is a function of adoption stage

The Chi-square was used to test Hypothesis 2. Table 13 shows the contingency table and the results of the Chi-square test obtained using the SPSS software.

Table 13: Chi-square test for the association between product type and adoption stage

Contingency table (cells provide counts)

	Product type				
Adoption Stage	**Business applications**	**Application development and deployment tools**	**System software**	**Hardware platform and peripherals**	TOTAL
1	9	4	6	3	22
2	25	12	18	17	72
3	31	14	36	26	107
4	9	5	18	13	45
5	24	7	19	21	71
Total	98	42	97	80	317

Chi-square test results

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	9.754[a]	12	.638
Likelihood Ratio	10.043	12	.612
Linear-by-Linear Association	2.503	1	.114
N of Valid Cases	317		

a. 1 cells (5.0%) have expected count less than 5. The minimum expected count is 2.91.

Hypothesis 2 is not supported. The results shown in Table 13 suggest that there is no significant association between adoption stage and the types of product introduced by new entrants.

5.6 Hypothesis 3: Product diversity increases rapidly to a critical level during the early adoption stages and then levels off increasing slightly over the latter stages

Table 14 provides the descriptive statistics for product diversity grouped by adoption stage. For stage 2, 3, 4 and 5 product diversity is normally distributed. For these four stages, skewness values are less than twice the standard error and the kurtosis values are within the +2 to –2 range. Product diversity at stage 1 is not normally distributed.

Table 14: Descriptive statistics for product diversity grouped by stage

	N	Min.	Max.	Mean	Std. Deviation	Skew		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Product diversity (Stage 1)	60	.0000	.8264	.6778	.1432	-2.03	.31	7.03	.61
Product diversity (Stage 2)	14	.8438	.9170	.8843	.0224	.11	.60	-.83	1.15
Product diversity (Stage 3)	16	.9075	.9133	.9111	.0016	-.76	.56	.66	1.09
Product diversity (Stage 4)	14	.9109	.9138	.9120	.0009	.38	.59	-.55	1.15
Product diversity (Stage 5)	23	.9123	.9144	.9133	.0006	.09	.48	-1.02	.94

Because product diversity at stage 1 is not normally distributed, the Mann-Whitney test was used to compare product diversity of stage 1 and the later four stages. Table 15 provides the results of using the Mann-Whitney test to compare the product diversity at stage 1 with product diversity at stage 2. The results show that product diversity at stage 1 is lower than the product diversity at stage 2.

Table 15: Results of using the Mann-Whitney test to compare product diversity at stage 1 and product diversity at stage 2

Ranks

	STAGE	N	Mean Rank	Sum of Ranks
Product diversity	1	60	30.50	1830.00
	2	14	67.50	945.00
	Total	74		

Test Statistics[a]

	Product diversity
Mann-Whitney U	.000
Wilcoxon W	1830.000
Z	-5.819
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: STAGE

The results of using the Mann-Whitney test to compare product diversity at stage 1 with product diversity at stages 3, 4, and 5 respectively are shown in Appendix IV. These results show that product diversity at stage 1 is significantly lower than the product diversity at the other four stages.

To compare product diversity across stages 2 to 5, this writer first checked whether or not the variance of product diversity was homogeneous across these stages. The Levene statistic for the homogeneity of variance test was 48.076 with three degrees of freedom 1 and 63 degrees of freedom 2. The significance level was 0.000. The result indicates that the variances of product diversity are not equal across stages 2, 3, 4 and 5.

Since product diversity measures for stages 2 to 5 are normally distributed and have unequal variance, the one-way ANOVA Tamhane T2 test was used to compare the means. This test does not assume equal variances.

Table 16 shows the multiple comparison test results. The third column in Table 16 shows the mean differences between the stages. Table 16 shows that product diversity at stage 2 is significantly lower than product diversity at the other stages. The mean difference between product diversity at stage 2 and product diversity at stages 3, 4 and 5 are -.0268, -.0277, and -.0290 respectively.

The mean difference between product diversity at stages 2 and 3 (-.0268) is 30 times greater than the mean difference between product diversity between stages 3 and 4 (-.0009) and 20 times greater than the mean difference between product diversity at stages 4 and 5 (-.0013). This suggests that the mean differences diminish as the product diversity of the latter stages are compared.

Product diversity for stage 3 and stage 4 are not significantly different. The result suggests that product diversity levels off during stages 3 and 4.

Product diversity at stage 5 has the highest value. Product diversity at stage 5 is slightly greater than product diversities at stages 3 and 4 (by 0.0022 and 0.0013 respectively).

Table 16: Multiple comparison of product diversity at stages 2, 3, 4, and 5

Multiple Comparisons Dependent Variable: Product diversity Tamhane						
					95% Confidence Interval	
(I) STAGE	(J) STAGE	Mean Difference (I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
2	3	-.0268(*)	.00599	.004	-.0453	-.0083
	4	-.0277(*)	.00599	.003	-.0463	-.0092
	5	-.0290(*)	.00598	.002	-.0476	-.0105
3	2	.0268(*)	.00599	.004	.0083	.0453
	4	-.0009	.00046	.276	-.0022	.0003
	5	-.0022(*)	.00040	.000	-.0034	-.0010
4	2	.0277(*)	.00599	.003	.0092	.0463
	3	.0009	.00046	.276	-.0004	.0022
	5	-.0013(*)	.00027	.001	-.0021	-.0005
5	2	.0290(*)	.00598	.002	.0105	.0476
	3	.0022(*)	.00041	.000	.0010	.0034
	4	.0013(*)	.00027	.001	.0005	.0021
* The mean difference is significant at the .05 level.						

The results support acceptance of Hypothesis 3. The multiple comparison tests of product diversity led to the following conclusions:

- Product diversity increased rapidly during the first two adoption stages (period from June 1993 to middle 1999) and then levelled off increasing slightly over the latter stages
- Product diversity of stage 5, the latest stage, is greater than product diversity in the previous stages, although the difference is very slight in comparison with stages 3 and 4.

5.7 Hypothesis 4: The number of new entrants has a bell shaped relationship with the cumulative number of Linux suppliers

Hypotheses 4, 4a and 4b test the association between entry rate and density. To test Hypothesis 4a and Hypothesis 4b, the sample was divided into two groups: established firms and start-ups, according to firm age. The Poisson regression model implemented by the LIMDEP software provided by Greene (2000) was used to test this set of hypotheses. Table 17 provides the results.

For the entire sample and the sub-sample comprised of established firms, the results in Table 17 show that the coefficients of density and density squared are in the hypothesized direction and are statistically significant ($p = .0000$). For the sub-sample comprised of start-up firms, the results in Table 17 show that the coefficients of density and density squared are not significant at $P < .10$.

Table 17: Poisson regression tests for hypothesis 4, 4a and 4b

Hypothesis	Variables	Coefficient	Standard error	P-value	Log-likelihood
4 (entire sample)	DENSITY	0.2518	0.016371181	.0000	-253.60
	DENSQ	-0.8513	0.076888329	.0000	
4a (established firms)	DENSITY	0.2110	0.018059597	.0000	-233.16
	DENSQ	-0.7067	0.083694413	.0000	
4b (start-ups)	DENSITY	0.0083	0.03531165	.8124	-141.38
	DENSQ	-0.1259	0.15991441	.4311	

The Negative Binomial model was also used to test hypotheses 4, 4a and 4b. The results are shown in Appendix V. The inferences from Negative Binomial model accord with and reinforce the Poisson model that hypotheses 4 and 4a are supported and hypothesis 4b is rejected.

The acceptance of hypothesis 4 and 4b indicates that entry rate of new firms has a bell-shaped relationship with the density of Linux suppliers; that entry rate of established firms has a bell-shaped relationship with the density of Linux suppliers. This same relationship does not exist between the entry rate of start-ups and the cumulative number of firms supplying Linux products.

5.8 Hypothesis 5: The number of new partnerships with Linux distributors is positively related to the number of new entrants

Hypotheses 5, 5a and 5b aim to identify the relationships between the number of partnerships formed by Linux distributors and the rates at which all suppliers, established suppliers, and start-ups entered the Linux market. Spearman rank order correlation was used to test Hypotheses 5, 5a and 5b.

Table 18a shows the correlation coefficients between the number of partnerships formed by the four major Linux suppliers at month k - 3 and the entry rate of all Linux suppliers, the entry rate of established firms and the entry rate of start-ups at month k.

Table 18a: Spearman correlation between number of partnerships formed by Linux distributors and the entry rate of Linux suppliers three months afterwards

	Entry rate (total sample)	Entry rate (established firms)	Entry rate (start-ups)
Number of new partnerships	.297**	.363**	.066

**. Correlation is significant at the .01 level (2-tailed).
*. Correlation is significant at the .05 level (2-tailed).

The results in Table 18a seemed to indicate support for the acceptance of hypotheses 5 and 5a and the rejection of hypothesis 5b. There is a positive relationship between number of new partnership established by the four Linux distributors, and the rate at which established suppliers enter the Linux market, but not the rate at which start-ups enter the market.

Based on the insights obtained from testing the previous hypotheses, the writer decided to split the sample into two sub-samples and calculate Spearman correlations for these samples. The first sub-sample was referred to as the "legitimization period". It included entry rates for all the months in stages 1 and 2 (from the sixth month of 1993 to the eight-month of 1998). The second sub-sample was referred to as the "competition period". It included entry rates for the months in stages 3, 4, and 5.

Table 18b shows the correlation coefficients for the number of partnerships-entry rates pairs. The results indicate that:

- The number of new partnerships is positively associated with the entry rate of established firms during the legitimization period
- The number of new partnerships is not associated with the entry rate of established firms or start-ups during the competition period
- The number of new partnerships is not associated with the entry rate of start-ups during the legitimization or competition periods

Table 18b: Spearman correlation between number of partnerships formed by Linux distributors and the entry rate of Linux suppliers for the legitimization and competition periods

	Entry rate (total sample)	Entry rate (established firms)	Entry rate (start-ups)
Number of new partnerships during the legitimization period (months in stages 1 and 2)	.481*	.576**	.136
Number of new partnerships during the competition period (months in stages 3, 4 and 5)	.207	.268	.037
Number of new partnerships during the legitimization and the competition periods (months in stages 1, 2, 3, 4 and 5)	.297**	.363**	.066

**. Correlation is significant at the .01 level (2-tailed).
*. Correlation is significant at the .05 level (2-tailed).

5.9 Hypothesis 6: The motives for new partnerships are a function of adoption stage

Hypothesis 6 was designed to examine whether or not the Linux distributors' motives for forming partnerships changed over the adoption stages.

Table 19 provides a contingency table for "motives for forming partnerships" and "adoption stages." The numbers in the cell are counts of the motive type per stage. Dodourova's (2003) classification for partnership motives was used to organize the 96 motives identified in the 61 partnership announcements.

Table 19: Contingency table: Partnership motives vs. Adoption stage

	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Sub-Total
Product/service related motive	1	9	8	3	18	39
Market-related motives	0	13	5	4	11	33
Technology-related motives	1	5	3	2	2	13
Competencies and skills-related motives	0	2	1	1	2	6
Industry/market structure modification-related motives	0	1	0	0	4	5
Sub-Total	2	30	17	10	37	96

The Chi-square test was used to test the association between partnership motives and adoption stages. The Chi-square was 12.60 with 16 degrees of freedom. For significance at the .05 level, the Chi-square should be greater than or equal to 26.30.

Hypothesis 6 was rejected. The Chi-square results show that a significant association between adoption stages and the motives of Linux distributors when forming partnerships was not found.

5.10 Hypothesis 7: The type of partner selected by a Linux distributor is a function of the motive for the new partnership

Hypothesis 7 examined the relationship between "partner type" and "partnership motive type". The partners identified in the 61 partnership announcements were classified into four types: hardware platform firms, infrastructure software firms, business applications firms and firms in other industries.

Table 20 provides the contingency table for “partner type” and “partnership motive type” built by the researcher.

Table 20: Contingency table for “partnership motive type” and “partner type”

Partnership motive type	Hardware platform firms	Infrastructure software firms	Business applications firms	Firms in other industries	Sub-total
Product/service related motive	18	14	5	2	39
Market-related motives	16	13	2	2	33
Technology-related motives	0	8	0	5	13
Competencies and skills-related motives	4	1	0	1	6
Industry/market structure modification-related motives	1	4	0	0	5
Sub-total	39	40	7	10	96

The Chi-square test was used to examine the association between “motive type” and “partner type”. The Chi-Square for the contingency table shown in Table 20 is 27.51 with 12 degrees of freedom. The association is significant beyond the 0.007 level.

Hypothesis 7 was accepted. The Chi-square results show that a significant association exists between the partnership motive type and partner type.

6 DISCUSSION OF RESULTS

This chapter is organized into four sections. Section 6.1 presents a summary of the results. Section 6.2 and 6.3 discusses the key results linking them to the existing literature. Section 6.4 reviews the method used to define the Linux adoption stages.

6.1 Summary of results

Table 21 provides a summary of the results presented in chapter 5.

Table 21: Summary of results

Hypothesis	**Accepted or Rejected**	**Rationale**
Hypothesis 1: The size of a new entrant is a function of adoption stage.	Accepted	Results of Kruskall-Wallis test shows that the means of firm size vary across adoption stages Results of Mann-Whitney U tests show that: mean size of firms at stage 1 is significantly smaller than mean size of firms at other four stages; that mean size of firms at stage 2 is larger than mean size of firms at stage 1 and 3; that mean size of firms at stage 3, 4 and 5 do not differ.
Hypothesis 2: Type of product introduced by new entrants is a function of adoption stage.	Rejected	Results of Chi-square test indicate that the association between product type and stage is not significant at $P < .10$.
Hypothesis 3: Product diversity increases rapidly during the early adoption stages and then levels off increasing slightly over the latter stages.	Accepted	Results of Mann-Whitney test show that product diversity at stage 1 is lower than the product diversity at stage 2. Results of one-way ANOVA Tamhane T2 tests show that product diversity at stage 2 significantly lower than product diversity at the other three stages; that the mean difference

		between product diversity at stages 2 and 3 is 30 times greater than the mean difference between product diversity at stages 3 and 4 and 20 times greater than the mean difference between product diversity at stages 4 and 5; that product diversity for stage 3 and stage 4 are not significantly different; and that product diversity at stage 5 has the highest value, which is slightly higher than product diversities at stages 3 and 4.
Hypothesis 4: The number of new entrants has a bell shaped relationship with the cumulative number of Linux suppliers.	Accepted	Results of using the LIMDEP software on the entire sample to estimate the parameters of the Poisson regression model shows that the coefficient of the density variable is positive and significant and the coefficient of the density-squared variable is negative and significant at $P < .01$.
Hypothesis 4a: The number of new entrants that are established firms has a bell shaped relationship with the cumulative number of Linux suppliers.	Accepted	Results of using the LIMDEP software on the established firms' sub-sample to estimate the parameters of the Poisson regression model show that the coefficient of the density variable is positive and significant and the coefficient of the density-squared variable is negative and significant at $P < .01$.
Hypothesis 4b: The number of new entrants that are start-up firms has a bell shaped relationship with the cumulative number of Linux suppliers.	Rejected	Results of using the LIMDEP software on the start-up firms' sub-sample to estimate the parameters of the Poisson regression model show that the coefficient of the density variable and the coefficient of the density-squared variable are not significant at $P < .01$.
Hypothesis 5: The number of new partnership with Linux distributors is positively related to the number of new entrants.	Accepted	For the total period and the period comprised of stages 1 and 2, the Spearman rank order correlation coefficients between number of new partnerships and number of new entrants is positive and significant at $P < .01$ and $P < .05$ respectively. The corresponding coefficient for the period comprised of stages 3, 4 and 5 was not significant at $P < .05$.
Hypothesis 5a:	Accepted for	For the total period and the period comprised of

The number of new partnership with Linux distributors is positively related to the number of new entrants that are established firms.	the legitimization period and the total period, rejected for the competition period	stages 1 and 2, the Spearman rank order correlation coefficients between number of new partnerships and number of new entrants that are established firms are positive and significant at P < .01 (two tailed)
Hypothesis 5b: The number of new partnership with Linux distributors is positively related to the number of new entrants that are start-ups.	Rejected for the legitimization and competition periods as well as the total study period.	For the total period, the period comprised of stages 1 and 2, and the period comprised of stages 3, 4 and 5 the Spearman rank order correlation coefficients between number of new partnerships and number of new entrants that are start-ups is not significant at P < .05.
Hypothesis 6: The motives for new partnerships are a function of adoption stage.	Rejected	Chi-square for the contingency table for partnership motive type and adoption stage is not significant at P < .10.
Hypothesis 7: The type of partner selected by a Linux distributor is a function of the motives for the new partnership	Accepted	Chi-square for the contingency table for partner type and partnership motive type is significant at P < .10.

6.2 Key findings

There were three objectives of this research. The first was to compare two attributes of Linux suppliers across adoption stages: firm size and product type. It was found that firms entering the Linux market at stage 1 tend to be significantly smaller in size than those entering at latter stages. Firms entering the Linux market at stage 2 tend to be significantly larger than those entering at stage 1 and 3, the two adjacent stages. Size does not distinguish firms that enter the Linux market at stages 3, 4, and 5.

No significant association was found between the type of product introduced by the new entrants and adoption stages.

The second objective of this research was to identify how the entry rate of new firms and the diversity of products changed from 1993 to 2003.

This writer found that diversity of Linux product types increases rapidly up to the end of adoption stage 2 (October 1999) and then levels out over the last three stages.

It was found that firm entry rate increases initially and then begins to slow down when it reaches a critical level. Established firms exhibited the same pattern regularity, however start-ups did not.

The third objective was to examine the attributes of the partnerships formed by four major Linux distributors from 1993 to 2003. The purpose was to identify: whether the number of partnerships formed by Linux distributors was related to the number of new entrants; whether the motives for partnerships formed by Linux distributors varied over adoption stages; and, whether the type of partner selected by Linux distributors was a function of the motive of the partnership.

It was found that the number of new partnerships is positively associated with the entry rate of established firms during the legitimization period (stages 1 and 2), but not during the competition period (stages 3, 4 and 5). The number of new partnerships is not associated with the entry rate of established firms or start-ups during the competition period. The number of new partnerships is not associated with the entry rate of start-ups during the legitimization or competition periods

This study determined that the motives for partnership formation are not significantly associated with the adoption stages.

A significant association was found between partnership motives and the type of partner Linux distributors chose to collaborate.

6.3 Implications

6.3.1 Comparing firm size across adoption stages

Firm size, a proxy for financial capability and technical competence, has been a commonly explored variable in the empirical literature of technology diffusion (Geroski, 2000). Many studies report that firm size is positively related to speed of diffusion, while several others show negative or insignificant effects (Geroski, 2000; Mansfield, 1963; Oster, 1982).

In this study, it was found that smaller firms committed to Linux earlier than larger firms did. This may be explained by the fact that the open source nature of Linux reduces the entry barrier for a firm to pursue Linux related businesses. Therefore, firm size is not a deciding factor for a firm to adopt Linux at early adoption stage.

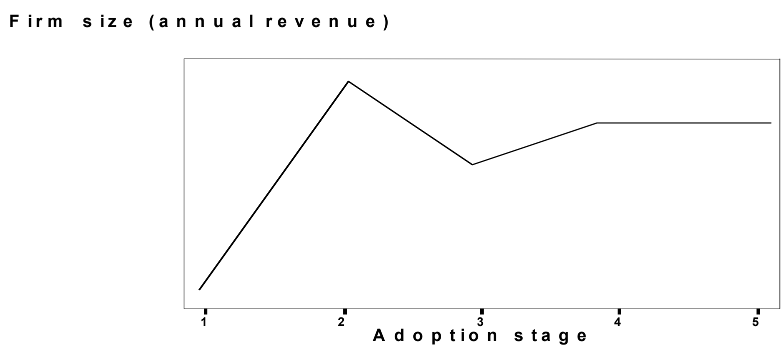
Schumpeter (1934) argued that small entrepreneurial firms are most likely to be the source of most innovation. This is true in the Linux context. Start-ups like Red Hat adopted Linux early on and leveraged its technical competence to effectively manage the technology evolution in the Linux community (Alex et al., 2003).

Figure 7 provides a graphic illustration of the results obtained when testing hypothesis 1. Figure 7 shows that at stage two, firms are significantly larger than those at stages 1 and 3. When examining the collected data, it was found that most of the prestigious firms of infrastructure software and hardware platform started to support Linux at stage 2.

If legitimization is the dominant force during the first two adoption stages, the following two characteristics can be observed in the Linux legitimization process:

- Small firms started the process of legitimizing Linux at stage 1, but the large firms ultimately legitimized it at stage 2
- Legitimization is a sequential process. No large firms initiated the legitimization process and the actions of many small firms would have not been able to institutionalize Linux without the adoption by large firms after the first wave of Linux adoption by small firms.

Figure 7: A graphic illustration of firm size across the stages



Results indicate that mean rank of size of firms at stage 3 is significantly smaller than that of stage 2 ($p<0.051$). To interpret this result, the first group the firms were grouped into four categories by firm annual revenue: 1) Those with an annual revenue < \$100 million, 2) annual revenue between \$100 and 999 million, 3) annual revenue between \$1000 and 9999 million and 4) annual revenue greater or equal to \$10000 million. Subsequently, the frequency of each category of firms at each stage was examined. The results are shown in Appendix VI. The results indicate the following: that firms with an annual revenue less than 100 million account for 71.9% of the total firms who commit to Linux at stage 3 (Stage 1 has the highest – 82.4%), whereas they represented only 58.3% at stage 2 and 54.1% and 62.5% respectively at stage 4 and 5. This supports the argument that the entry of prestigious firms such as IBM, Oracle, and Dell etc. at stage two has increased the attractiveness of the market and laid a foundation, based on which potentially new entrants who are relatively smaller VARs (value added resellers) can build to sell their own Linux-based products. With the entry of these relatively smaller firms, mean rank of firm size at stage 3 has decreased significantly as opposed to the previous stage.

Mean rank of size of firms that enter the market at stage 4 and 5 are found to have no significant difference between one another and between them and Stage 2 or 3. This suggests that similar proportions of large and small firms enter the market and coexist in later stages. This is corroborated by the frequency of the four categories of firms in Appendix VI.

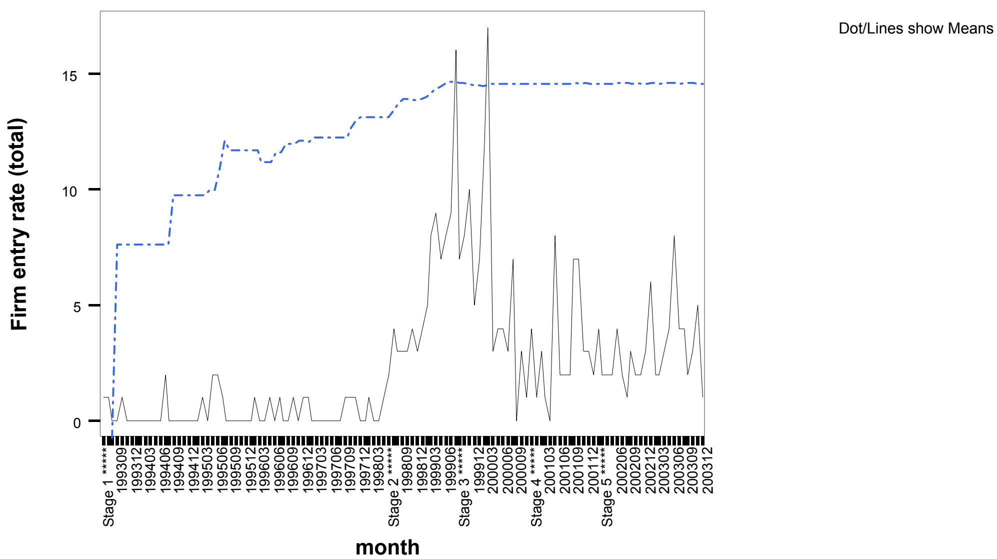
6.3.2 Comparing product diversity and entry rates across adoption stages

This writer constructed Hypothesis 4 based on the density-dependence model for the purpose of relating entry rate to the cumulative number of firms that have adopted Linux. Results confirmed the bell shaped relationship between the number of new entrants and the cumulative number of firms supplying Linux products (density). The twin forces of density-dependence model, legitimization and competition, function as predicted in the Linux context.

Figure 8 shows how the entry rate and product diversity changed over the five stages. It suggests that numerous changes in product diversity mark the process of legitimization while low numbers of changes in product diversity mark the process of competition. Low numbers of changes in entry rate followed by numerous changes in entry rate mark the process of legitimization while average changes in entry rate mark the process of competition.

Figure 8 shows that the Linux technology life cycle began in June 1993. The entry rate was low for the first five years. The beginning of stage 2 is marked by an abrupt rise in the entry of firms. Results show that many of these firms were large established firms.

Figure 8: Entry rate (continuous line) and product diversity (dotted line) over time



According to the density-dependence model, the competition force becomes the prevalent one when a cumulative number of firms reach a critical level. Results of this study provide evidence that entry rate began to slow down after reaching the peak at stage 3 (around March 2000). After stage 1, small and large firms tend to coexist in order to contribute to the process of competition.

Product diversity increases steadily and rapidly during the first two stages while Linux is being legitimized, with the support from various types of vendors who provide different types of Linux products. From stage three on, product diversity becomes saturated with the availability of products in almost every major category. At that point, new entrants to the market are required to compete with the first-movers who already provide similar Linux-based products. Not surprisingly, Figure 8 suggests that the increased competition causes firm entry rate to slow down when it reaches the peak at around the critical time when product diversity becomes saturated.

6.3.3 Number of partnerships and number of new entrants

A positive correlation was found between number of partnerships with Linux distributors and established firm entry rates during the legitimization period.

The four Linux distributors in our study were relatively small firms with less diversified product and services relative to the established firms in the computer industry. In the open source environment, users of open source software are not beholden to any producer or distributor, thus, reputation, reach and customer loyalty are much more crucial for vendors than for those in the closed source environment (Zander, 2002). There is no better way for a fledgling company like a Linux distributor to gain these attributes than to associate with a larger and respected industry leader like IBM. The partnerships formed by a Linux distributor with an established firm therefore, can facilitate the legitimization process of Linux. This may be the reason for the positive correlation between established firm entry rate and number of partnerships.

6.3.4 Motives for partnerships and adoption stages

Results of this study show no significant association between adoption stages and Linux distributors' partnership motives. This suggests that Linux distributors' partnership motives do not change over time and all their efforts support Linux adoption in the market.

6.3.5 Types of partners and motives for partnerships

Three interesting findings have been found about the relationship between partnerships motives and partner types. First, Linux distributors tend to partner with hardware platform and infrastructure software vendors to accelerate Linux adoption by gaining access to their partners' existing customer base.

Secondly, when Linux distributors partner with firms in other industries such as Ericsson in telecommunications, they are more likely to have a technology-related motive, such as conducting experiments to explore the application of Linux in 3G handsets. Linux distributors on the other hand, collaborate with vendors of applications (e.g. Hummingbird) with the sole purpose of making business applications available on Linux, instead of addressing other strategic concerns.

Thirdly, Linux distributors tend to partner with platform vendors to overcome their lack of certain competencies or skills. By forming a single source with HP for all Linux hardware and software support for customers, Red Hat leverages the resources from a prestigious platform vendor who has access to a global service team and traditionally possesses extensive experience in an Unix-like operating system.

In summary, the significant association between partnership motives and types of partners implies that seeking complementary assets are the main purpose for the Linux distributors' partnership formations. This confirms the importance of complementary skills and findings on partnership motives by previous studies (Dodourova, 2003).

6.3.6 Comparing product type across adoption stages

Results of testing hypothesis 2 suggest that there is no significant association between adoption stages and types of products introduced by new entrants. This result was not expected.

The reason for the lack of association between product type and adoption stages may be due to the nature of open source. Open source enables the development of various types of products concurrently, one product type complementing other product types. Teece (1986) suggests that innovations such as computer systems require the provision of complementary assets to commercialize the innovation. When additional investment is required to co-specialize the asset to be useful with a given innovation, the successful adoption of the innovation and the related assets are mutually reinforcing, providing a positive feedback cycle.

The results suggest that the adoption of Linux has provided a positive feedback across product types. This reinforces the concurrent introduction of products at different layers of a "whole product": the business application layer, application development & deployment tool layer, system level software layer and hardware platform layer.

6.4 Comments on stage identifications

This section comments on the method that has been used to identify the stages of Linux adoption life cycle in this study.

Because classifying the adopters only is a simplification that aids in understanding of the technology adoption phenomenon (Rogers 1983), this writer believes that the resulting stages delimitation cannot be judged as right or wrong. As long as insights can be gained from the process of stages delimitation and the subsequent comparisons between stages, this proposed method would prove to be useful.

First of all, figure 8 shows that the five-stage delimitation captures the dynamics of entry rate and product diversity over the 1993-2003 period. Secondly, the timing of stage 2 and 4 are coincidentally paralleled by the two important events in Linux history: 1) Linux received the first wave of established software firms' adoption in July 1998 (e.g. Oracle, Informix, Sybase) and, 2) IBM's announcement of $1 billion investment in Linux made in December 2000. This coincidence corroborates the fitness of our stages identification. Finally, the insights gained from the stage-based discussions in section 6.2 and 6.3 also strongly validate the effectiveness of our stages identification.

7 CONCLUSIONS, LIMITATIONS AND SUGGESTIONS FOR FUTURE RESEARCH

7.1 Conclusions

The results of this study lead to the following conclusions:

- The density-dependence model is valid in the Linux domain, indicating that legitimization and competition shape market evolution. Small firms initiated the process of legitimizing Linux, but large firms were the ones that ultimately legitimized it. Small and large firms coexist successfully to compete in the marketplace.
- Numerous changes in product diversity mark the progress of legitimization (with rapid increase in entry rate) while a number of changes in product diversity mark the process of competition (indicated by the fact that entry rate starts to slow down).
- The method used to define the adoption stages provides a framework that can be used for the research of the adoption of other technologies from a life-cycle perspective.

7.2 Limitation of the research

The limitation of this research is that this writer only focused on data for the suppliers of Linux-based products. The reason for this is that data for buyers were not available. The only buyer data that were available pertained to important buyers.

7.3 Suggestions for future research

I make five suggestions for future research.

First, future research can examine the differences in other firm-level attributes across stages. This study compared firm size and product type across adoption stages. In order to gain additional insights into Linux adoption, researchers may wish to examine the differences in firms' strategic concerns and top management team's reasons to adopt Linux between stages.

The second suggestion for future research is testing the validity of the method used to define the stages. A method to identify adoption stages was proposed and then used in the thesis. The method relies on data from a sample of important buyers and criteria to operationally identify the month in which a stage change occurs. Researchers can further examine the approach used in this thesis to identify adoption stages.

The third suggestion for future research is to compare the number of partnerships Linux distributors form at each stage with the number of partnerships Linux suppliers form at each stage. This study only examines the attributes of partnerships formed by four major Linux distributors. Future research could further examine the partnerships formed by other individual companies that have also played an important role in the market for Linux. Identifying the attributes in those partnerships and comparing them with what was found about Linux distributors' partnerships would help our understanding of the change of partnership structure during the Linux market evolution.

The fourth suggestion for future research is replicating this study using buyer data from 1993 to 2003. Buyer data can also be used to measure product diversity and company size. It will be interesting to compare the results obtained using buyer data with those obtained in this study using supplier data.

The fifth suggestion for future research is to examine and compare Linux adoption in each specific market segment. This researcher examined the aggregate Linux-based market. Future researchers could examine Linux adoption by market segment: server, desktop and embedded appliances. Although competitive dynamics and the state of market adoption are vastly different among the three specific market segments for Linux all three markets were treated collectively in this research study.

APPENDIX I: Dodourova's classification framework for partnership motives

Motives for alliance formation	Description
Market-related motives	Gain access to new markets, create and experiment with new markets, circumvent legal, regulatory, or other barriers, defend/enhance market position in present markets
Product-related motives	Create and experiment with new products, broaden/fill gaps in present product line, differentiate/enter new product domains, add value to a product, property rights protection
Industry/market structure modification-related motives	Reduce (potential) competition, raise entry barriers, alter the technological base of competition
Timing-related motives	Experiment and/or exploit new product/market/technology opportunities faster by accelerating pace of R&D or market entry
Cost-related motives	Lower R&D/manufacturing/marketing/organizational costs, achieve economies of scale
Risk reduction-related motives	Lower risk in the face of large R&D costs required, technological/market or other uncertainties
Competencies and skills-related motives	Overcome shortcomings in internal competencies; Internalize competencies, enhance knowledge-generation activities, learn new skills, enhance present skills
Technology-related motives	Need to experiment, gain access to technology, gain access to assets that are essential for efficient commercialization of a technology, exploit the potential for broad application of technology, need for technical coordination/develop technological standards, observe technological change but remain flexible, collective protection of technological advances

Source: Dodourova, 2003.

APPENDIX II: Important buyers

Important Buyer	Application type	Running environment type	Date
ISPs	Web/File/print	PC	1996/12
Digital Domain (created visual effect for Titanic)	Computing	Alpha workstation	1998/06
Jay Jacobs Inc	Front-end business application	PC	1998/11
Burlington Coat Factory warehouse corp.	Front-end business application	PC	1999/02
Canadian National Railway Co	Web/File/print	PC	1999/05
Amerada Hess Corp	Computing	Linux cluster	1999/05
B.F. Goodrich	Web/File/print	PC	1999/06
Cendant Corp	Business application	PC	1999/07
Motorola	Telecom product platform	Embedded system	1999/08
International Transportation Specialist (ITS)	Firewall	PC	1999/08
Cisco	Print server	PC	1999/11
Kenwood	Business application	Intel based Dell PowerEdge server	1999/11
Google	Computing (for searching algorithm)	PC	2000/06
1stUp.com	Data warehousing	Intel based server	2000/07
Amazon	Web	Intel based servers	2000/09
Telia AB	Web	IBM mainframe	2000/12
Banco Mercantil	Print/file/Firewall/DNS	IBM mainframe	2001/05
Amazon	Business application	Intel based servers	2001/06
Southern Illinois Power Cooperative (SIPC)	File/Print	PC	2001/07
Merrill Lynch	Mission critical applications	IBM Mainframe	2002/02
Reuters	Mission critical market data platform	Intel-based servers	2002/05

APPENDIX III: Testing Hypothesis 1 - Mann-Whitney Test

Five tests show significant differences in the size of firms between the two compared stages.

Mann-Whitney Test (Comparing Stage 1 and 2)

Ranks

	Adoption stage	N	Mean Rank	Sum of Ranks
firm size(annual revenue)	Stage 1	17	24.15	410.50
	Stage 2	60	43.21	2592.50
	Total	77		

Test Statistics[a]

	firm size(annual revenue)
Mann-Whitney U	257.500
Wilcoxon W	410.500
Z	-3.102
Asymp. Sig. (2-tailed)	.002

a. Grouping Variable: Adoption stage

Mann-Whitney Test (Comparing Stage 1 and 3)

Ranks

	Adoption stage	N	Mean Rank	Sum of Ranks
firm size(annual revenue)	Stage 1	17	39.03	663.50
	Stage 3	89	56.26	5007.50
	Total	106		

Test Statistics[a]

	firm size(annual revenue)
Mann-Whitney U	510.500
Wilcoxon W	663.500
Z	-2.119
Asymp. Sig. (2-tailed)	.034

a. Grouping Variable: Adoption stage

Mann-Whitney Test (Comparing Stage 1 and 4)

Ranks

	Adoption stage	N	Mean Rank	Sum of Ranks
firm size(annual revenue)	Stage 1	17	19.97	339.50
	Stage 4	37	30.96	1145.50
	Total	54		

Test Statistics[a]

	firm size(annual revenue)
Mann-Whitney U	186.500
Wilcoxon W	339.500
Z	-2.384
Asymp. Sig. (2-tailed)	.017

a. Grouping Variable: Adoption stage

Mann-Whitney Test (Comparing Stage 1 and 5)

Ranks

	Adoption stage	N	Mean Rank	Sum of Ranks
firm size(annual revenue)	Stage 1	17	23.76	404.00
	Stage 5	56	41.02	2297.00
	Total	73		

Test Statistics[a]

	firm size(annual revenue)
Mann-Whitney U	251.000
Wilcoxon W	404.000
Z	-2.937
Asymp. Sig. (2-tailed)	.003

a. Grouping Variable: Adoption stage

Mann-Whitney Test (Comparing Stage 2 and 3)

Ranks

	Adoption stage	N	Mean Rank	Sum of Ranks
firm size(annual revenue)	Stage 2	60	83.41	5004.50
	Stage 3	89	69.33	6170.50
	Total	149		

Test Statistics[a]

	firm size(annual revenue)
Mann-Whitney U	2165.500
Wilcoxon W	6170.500
Z	-1.953
Asymp. Sig. (2-tailed)	.051

a. Grouping Variable: Adoption stage

Mann-Whitney Test (Comparing Stage 2 and 4)

Ranks

	Adoption stage	N	Mean Rank	Sum of Ranks
firm size(annual revenue)	Stage 2	60	49.65	2979.00
	Stage 4	37	47.95	1774.00
	Total	97		

Test Statistics[a]

	firm size(annual revenue)
Mann-Whitney U	1071.000
Wilcoxon W	1774.000
Z	-.290
Asymp. Sig. (2-tailed)	.772

a. Grouping Variable: Adoption stage

Mann-Whitney Test (Comparing Stage 2 and 5)

Ranks

	Adoption stage	N	Mean Rank	Sum of Ranks
firm size(annual revenue)	Stage 2	60	59.40	3564.00
	Stage 5	56	57.54	3222.00
	Total	116		

Test Statistics[a]

	firm size(annual revenue)
Mann-Whitney U	1626.000
Wilcoxon W	3222.000
Z	-.298
Asymp. Sig. (2-tailed)	.765

a. Grouping Variable: Adoption stage

Mann-Whitney Test (Comparing Stage 3 and 4)

Ranks

	Adoption stage	N	Mean Rank	Sum of Ranks
firm size(annual revenue)	Stage 3	89	61.15	5442.00
	Stage 4	37	69.16	2559.00
	Total	126		

Test Statistics[a]

	firm size(annual revenue)
Mann-Whitney U	1437.000
Wilcoxon W	5442.000
Z	-1.122
Asymp. Sig. (2-tailed)	.262

a. Grouping Variable: Adoption stage

Mann-Whitney Test (Comparing Stage 3 and 5)

Ranks

	Adoption stage	N	Mean Rank	Sum of Ranks
firm size(annual revenue)	Stage 3	89	68.22	6072.00
	Stage 5	56	80.59	4513.00
	Total	145		

Test Statistics[a]

	firm size(annual revenue)
Mann-Whitney U	2067.000
Wilcoxon W	6072.000
Z	-1.726
Asymp. Sig. (2-tailed)	.084

a. Grouping Variable: Adoption stage

Mann-Whitney Test (Comparing Stage 4 and 5)

Ranks

	Adoption stage	N	Mean Rank	Sum of Ranks
firm size(annual revenue)	Stage 4	37	46.80	1731.50
	Stage 5	56	47.13	2639.50
	Total	93		

Test Statistics[a]

	firm size(annual revenue)
Mann-Whitney U	1028.500
Wilcoxon W	1731.500
Z	-.059
Asymp. Sig. (2-tailed)	.953

a. Grouping Variable: Adoption stage

APPENDIX IV: Comparing product diversity between stage 1 and 2, 3, 4 and 5

Mann-Whitney Test (comparing stage 1 and 2)

Ranks

	STAGE	N	Mean Rank	Sum of Ranks
Product diversity	1	60	30.50	1830.00
	2	14	67.50	945.00
	Total	74		

Test Statistics[a]

	Product diversity
Mann-Whitney U	.000
Wilcoxon W	1830.000
Z	-5.819
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: STAGE

Mann-Whitney Test (comparing stage 1 and 3)

Ranks

	STAGE	N	Mean Rank	Sum of Ranks
Product diversity	1	60	30.50	1830.00
	3	16	68.50	1096.00
	Total	76		

Test Statistics[a]

	Product diversity
Mann-Whitney U	.000
Wilcoxon W	1830.000
Z	-6.138
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: STAGE

Mann-Whitney Test (comparing stage 1 and 4)

Ranks

	STAGE	N	Mean Rank	Sum of Ranks
Product diversity	1	60	30.50	1830.00
	4	14	67.50	945.00
	Total	74		

Test Statistics[a]

	Product diversity
Mann-Whitney U	.000
Wilcoxon W	1830.000
Z	-5.819
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: STAGE

Mann-Whitney Test (comparing stage 1 and 5)

Ranks

	STAGE	N	Mean Rank	Sum of Ranks
Product diversity	1	60	30.50	1830.00
	5	23	72.00	1656.00
	Total	83		

Test Statistics[a]

	Product diversity
Mann-Whitney U	.000
Wilcoxon W	1830.000
Z	-7.040
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: STAGE

APPENDIX V: Test result with Negative Binomial model (Hypothesis 4/4a/4b)

Negative Binomial Regression	
Maximum Likelihood Estimates	
Dependent variable	ENTRY
Weighting variable	ONE
Number of observations	127
Iterations completed	8
Log likelihood function	-231.9001
Restricted log likelihood	-253.6021
Chi-squared	43.40383
Degrees of freedom	1
Significance level	.0000000

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
DENSITY	.2331341810E-01	.23830873E-02	9.783	.0000	97.905512
DENSQ	-.7522401357E-01	.10621101E-01	-7.083	.0000	19.538220
	Dispersion parameter for count data model				
Alpha	.4362050258	.13618525	3.203	.0014	

Negative Binomial Regression	
Maximum Likelihood Estimates	
Dependent variable	ENTRYES
Weighting variable	ONE
Number of observations	127
Iterations completed	8
Log likelihood function	-215.2309
Restricted log likelihood	-233.1688
Chi-squared	35.87577
Degrees of freedom	1
Significance level	.0000000

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
DENSITY	.1939927079E-01	.27026992E-02	7.178	.0000	97.905512
DENSQ	-.6204754822E-01	.12341806E-01	-5.027	.0000	19.538220
	Dispersion parameter for count data model				
Alpha	.5066102198	.16577680	3.056	.0022	

```
+--------------------------------------------+
| Negative Binomial Regression               |
| Maximum Likelihood Estimates               |
| Dependent variable               ENTRYST   |
| Weighting variable                   ONE   |
| Number of observations               127   |
| Iterations completed                   5   |
| Log likelihood function      -137.7660     |
| Restricted log likelihood    -141.3800     |
| Chi-squared                   7.227942     |
| Degrees of freedom                     1   |
| Significance level           .7177747E-02 |
+--------------------------------------------+
+---------+-------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+-------------+----------------+--------+---------+----------+
DENSITY    .1006329536E-02  .31958464E-02     .315   .7528  97.905512
DENSQ     -.1347327498E-01  .15372752E-01    -.876   .3808  19.538220
          Dispersion parameter for count data model
Alpha       .1640991423            .22410083   .732       .4640
```

APPENDIX VI: A comparison of firm size between stages

FIRMSIZE [a]

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	<$100M	14	82.4	82.4	82.4
	$100M-$999M	2	11.8	11.8	94.1
	$1000M-$9999M	1	5.9	5.9	100.0
	Total	17	100.0	100.0	

a. Adoption stage = Stage 1

FIRMSIZE [a]

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	<$100M	35	58.3	58.3	58.3
	$100M-$999M	14	23.3	23.3	81.7
	$1000M-$9999M	6	10.0	10.0	91.7
	>=$10000M	5	8.3	8.3	100.0
	Total	60	100.0	100.0	

a. Adoption stage = Stage 2

FIRMSIZE [a]

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	<$100M	64	71.9	71.9	71.9
	$100M-$999M	12	13.5	13.5	85.4
	$1000M-$9999M	9	10.1	10.1	95.5
	>=$10000M	4	4.5	4.5	100.0
	Total	89	100.0	100.0	

a. Adoption stage = Stage 3

FIRMSIZE [a]

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	<$100M	20	54.1	54.1	54.1
	$100M-$999M	8	21.6	21.6	75.7
	$1000M-$9999M	5	13.5	13.5	89.2
	>=$10000M	4	10.8	10.8	100.0
	Total	37	100.0	100.0	

a. Adoption stage = Stage 4

FIRMSIZE [a]

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	<$100M	35	62.5	62.5	62.5
	$100M-$999M	13	23.2	23.2	85.7
	$1000M-$9999M	5	8.9	8.9	94.6
	>=$10000M	3	5.4	5.4	100.0
	Total	56	100.0	100.0	

a. Adoption stage = Stage 5

REFERENCES

Adeslon, Andrew. 2002. Amazon.com: Migration from Unix to Red Hat Linux. International Data Corporation.

Alex, Joseph, Joanna Ng, Jean-Bernard Rolland and Ed Zane. 2003. Red Hat: Innovative Capabilities Audit. Graduate School of Business. Stanford University.

Amburgey, T.A., Carroll, G.R. 1984. Time series models for event counts. Social Science Research. 13(1): 38-54.

Berquist, Thomas P., Sarah Friar, Derek Bingham, Peter Yang. 2003. Fear the penguin. Global Strategy Research. Goldman Sachs.

Bonaccorsi, Andrea, Cristina Rossi. 2003. Why Open Source Software Can Succeed. Research Policy. 32: 1243-1258.

Cubranic, Davor, Kellogg S. Booth. 1999. Coordinating Open-source Software Development. Enabling Technologies: Infrastructure for Collaborative Enterprises, 1999. (WET ICE '99) Proceedings. IEEE 8th International Workshops, 16-18 June 1999: 61 – 66.

Debackere, Koenraad, Bart Clarysse. 1998. Advanced Bibliometric Methods to Model the Relationship between Entry Behaviour and Networking in Emerging Technological Communities. Journal of American Society for Information Science. 49(1): 49-58.

Decrem, Bart. July 10, 2003. Desktop Linux technology & market overview. Open Source Applications Foundation. http://www.osafoundation.org/desktop-linux-overview.pdf (accessed 10 February 2004).

Di Carlo, Lisa. July 15, 2002. Linux not just for geeks anymore. Forbes.com. http://www.forbes.com/2002/07/15/0715linux.html (accessed 10 February 2004).

DiMaggio, P.J., Powell, W.W. 1983. The iron cage revisited: institutional isomorphism and collective rationality in organizational fields. American Sociological Review. 48: 147-60.

Dodourova, Mariana. 2003. Industry dynamics and strategic positioning in the wireless telecommunications industry: the case of Vodafone Group plc. Management Decision. 41(9): 859-870.

Easingwood, C., S. Harrington. 2002. Launching and re-launching high technology products. Technovation. 22: 657-666.

Eisenhardt, K. M., C. B. Schoonhoven. 1996. Resource-based view of strategic alliance formation: Strategic and social effects in entrepreneurial firms. Organization Science. 7: 136–150.

Geroski, P.A. 2000. Models of technology diffusion. Research Policy. 29: 603-625.

Godfrey, Michael W., Qiang Tu. 2000. Evolution in Open Source Software: A case study. Software Maintenance. Proceedings. International Conference, 11-14 Oct. 2000: 131–142.

Gonsalves, Antone. 2003. Windows Gains Market Share, Despite Linux Threat. http://www.techweb.com/wire/story/TWB20031008S0013 (accessed 10 February 2004).

Grant, R.M. 1996. Toward a knowledge-based theory of the firm. Strategic Management Journal. 17. Special issue. Winter: 109-22.

Greene, William H. 2000. Econometric Analysis, fourth edition, Prentice Hall, New Jersey.

Hamel, G. 1991. Competition for competence and inter-partner learning within international strategic alliances. Strategic Management Journal. Special Issue. 12. Summer: 83-103.

Heiman, Richard V., Dennis Byron. 2003. IDC's Software Taxonomy 2003. Industry Developments and Models. International Data Corporation.

Hennart, J. 1991. The transaction cost theory of alliances: an empirical study of Japanese subsidiaries in the United States. Management Science. 37(4): 483-97.

Hertel, Guido, Sven Niedner, Stefanie Herrmann. 2003. Motivation of software developers in open source projects: an Internet-based survey of contributors to the Linux kernel. Research Policy. 32: 1159-1177.

King, G. 1988. Statistical models for political science event counts: Bias in conventional procedures and evidence for the exponential Poisson regression model. American Journal of Political Science. 32 (3): 838 –863.

Klepper, Steven. 1996. Entry, Exit, Growth, and Innovation over the Product Life Cycle. American Economic Review. 86(3): 562-583.

Kogut, B. 1988. Joint ventures: theoretical and empirical perspectives. Strategic Management Journal. 9: 319-32.

Kusnetzky, Dan, Al Gillen. 2001. Linux: A journey into the enterprise. IDC White Paper. International Data Corporation.

Lerner, Josh, Jean Tirole. 2002. Some Simple Economics of Open Source. The Journal of Industrial Economics. L (2).

Mansfield, E. 1963. The speed of response of firms to new technologies. Quarterly Journal of Economics. 77: 290-311.

Meyer, J.W., Rowan, B. 1977. Institutionalized organizations: formal structure as myth and ceremony. American Journal of Sociology. 83: 340-63.

Mockus, Audris, Roy T Fielding, James D Herbsleb. 2002. Two case studies of Open source software development: Apache and Mozilla. ACM Transactions on Software Engineering and Methodology. 11(3): 309-346.

Moore, Geoffrey A. 1991. Crossing the chasm: marketing and selling high-tech products to mainstream customers. Harper Business, New York.

Moore, Geoffrey A. 2002. Living on the fault line: managing for shareholder value in any economy. Harper Business, New York.

Newbold, P. 1991. Statistics for Business and Economics, 3rd Ed. Prentice-Hall International, 1991.

O'Mahony, Siobhan. 2003. Guarding the commons: how community managed software projects protect their work. Research Policy. 32: 1179-1198.

Oster, S. 1982. The diffusion of innovations among steel firms: the basic oxygen furnace. Bell Journal of Economics. 13:45-56.

Pal, Nilendu, T.R. Madanmohan. 2003. Competing on Open Source: Strategies and Practice. http://opensource.mit.edu/papers/madanmohan.pdf (accessed 10 February 2004).

Pettigrew, Andrew. 1990. Longitudinal Field research on change: Theory and Practice, Organization Science. 1(3): 267-291

Pfeffer, J., Nowak, P. 1976. Alliances and inter-organizational interdependence. Administrative Science Quarterly. 21: 398-418.

Porter, M., Fuller, M. 1986. Coalitions and global strategy. Harvard Business School Press, Boston, MA.

Raymond, E. 1999. The Cathedral and the Bazaar. O'Reilly, Sebastopol, CA.

Rogers, Everett M. 1983. Diffusion of Innovations, third edition, Free Press, New York.

Sawhney, Mohanbir, Prandelli, Emanuela. 2000. Communities of Creation: Managing distributed innovation in turbulent markets. California Management Review. Summer. 42(4).

Schumpeter, J. 1934. The Theory of Economic Development. Harvard University Press, Cambridge, MA.

Teece, David J. 1986. Profiting from technological innovation: implications for integration, collaboration, licensing and public policy. Research Policy. 15(6): 285-305.

Von Hippel, Eric, Georg von Krogh. 2003. Open Source Software and the "Private-Collective" Innovation Model: Issues for Organization Science. Organization Science. 14 (2): 209-223.

Wade, J. 1995. Dynamics of Organizational communities and technological bandwagons: An empirical investigation of community evolution in the microprocessor market. Strategic Management Journal, 16: 111-133.

West, Joel. 2003. How open is open enough? Melding proprietary and open source platform strategies. Research Policy. 32: 1259-1285.

West, Joel, Jason Dedrick. 2001. Open source standardization: The rise of Linux in the network era. Knowledge, Technology, & Policy. 14(2): 88-112.

Williamson, O. 1985. The Economic Institutions of Capitalism, Free Press, New York.

Zander, Eric. 2002. Big Blue Gets A Red Hat: An analysis of the alliance between IBM and Red Hat. http://www.ericzander.com/Publications/RedHatAllianceWithIBM.pdf (accessed 10 February 2004).

Zeitlyn, David. 2003. Gift economies in the development of open source software: anthropological reflections. Research Policy. 32: 1287-1291.