Formal Training in Nascent Small U.S. Firms and Its Impact on Organization Level Performance

A Dissertation
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Christopher J. Hanson

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Adviser:
Alexandre Ardichvili, Ph.D.
Professor, University of Minnesota

May 2014

Acknowledgements

Like a small business that begins with a founder's vision but requires the help of many to become reality, this work is the result of many people who have helped me along the way.

First and foremost I would like to acknowledge and thank my advisor, Professor Alexandre Ardichvili. Your wisdom, foresight, and patience have been instrumental in helping me accomplish what at times seemed like an insurmountable task. Thank you Sasha.

I would also like to acknowledge my dissertation committee members, Professor Karen Seashore (committee chair), Dr. Catherine Twohig, and Dr. Jeremy Hernandez. Thank you all for your thoughtful assistance in helping me complete this paper and my degree.

Thank you to Dr. Alicia Robb and the team at the Kauffman Foundation for conducting the brilliant and much needed Kauffman Firm Survey. This longitudinal data set has had, and will continue to have, a dramatic impact on our understanding of small and nascent businesses. Their support of this research was critical to gaining access to the data used in this analysis. Thanks also to Athena Diaconis for her assistance with the NORC enclave data at the University of Chicago.

Much thanks and appreciation goes to my statistical consultant, Lindsey Dietz. Without your help I would have never been able to conduct the sophisticated statistical analysis required by this research.

And finally to my dear wife Anne who has been both my best friend and co-student in this journey. It's been a long road with many turns, stops and detours, but we've finally arrived at this journey's end. I'm looking forward to the next. Get er done!

Dedication

This work is dedicated the women and men who own or have founded any of the millions of small business that are the economic heart and soul of our cities, towns, and communities across America. As one of you, I understand the passion that drives you to make your vision a reality, and the huge hurdles you have to overcome to make that reality a business success. My hope is that this research, and the growing amount of practical research on small businesses, will provide insights to help entrepreneurs and small business owners be more successful.

Abstract

Small business firms play an increasingly important role in the economic and social well-being of the global community. In the US over 98% of all businesses employ fewer than 100 people (Heneman & Berkley, 1999). In the United Kingdom (UK), 99.8% of all businesses have fewer than 50 employees (BIS, 2011). While small firms have a significant impact on our economic and social well-being, there is a high failure rate. Finding ways to increase the survival and success rate of small businesses will have a significant positive impact on our economy and communities.

We know that formal training in large businesses contributes to increased productivity, and ultimately increased organizational performance, but there is little research on the use of, or impact of formal training in small businesses. The lack of longitudinal data on the use of formal training in small businesses is one of the main stated reasons for the lack of research and understanding in this area.

This research uses data from the Kauffman Firm Survey, the largest longitudinal survey of nascent small businesses ever conducted, and multivariate linear mixed modeling to look at the impact over time of formal training on organization level performance in those businesses. The results show a significant impact in profits, asset levels, and profit margin for those small businesses that did formal training. The results also show significant differences in the number of employees, asset levels and additional human resource activities between those firms doing formal training and those not doing formal training. The two major conclusions are: Formal training has a significant positive impact on organizational performance, and not all small businesses are the same.

Table of Contents

List of Tables

List of Figures

CHAPTER I

INTRODUCTION

Background

Small business firms play an increasingly important role in the economic and social well-being of the global community. In the US over 98% of all businesses employ fewer than 100 people (Heneman & Berkley, 1999). These small businesses account for over 50% of all nonfarm related gross domestic product and have accounted for 60-80% of all job growth in the last decade (Kobe, 2007). In the European Community (EU) small enterprises play an even more significant role. For example, in the United Kingdom (UK), 99.8% of all businesses have fewer than 50 employees (BIS, 2011).

While small firms have a significant impact on our economic and social well-being, there is a high failure rate. Some estimates put the failure rate at up to 80% within the first year of operation. Dunn and Bradstreet (2012) estimates that failure rates for small businesses increased 40% between 2007 and 2010. Being able to increase the success rate of new small businesses would speed up economic recovery for communities hard hit by economic recession and contribute to their long-term well-being.

While the importance of small businesses has been shown and well discussed, identifying practical ways to increase their success rate has been limited. Most of the discussion on small business success has been about access to capital and founder characteristics (Castrogiovanni, 2011).

Research on formal training and development practices in large firms has shown that such practices can have a significant positive impact on the overall well-being and success of those organizations (Lawler, Mohrman, & Lawford, 1998). However, there is an identified gap of research into the use of formal training in small firms and the impact such training has on firm performance (Hill, 2004; Kitching, 2008; Stewart & Beaver, 2004).

Research into the use of human resource development in small businesses has been limited (Burke, 2011). Most of the research to date has been in the form of case studies. While these studies provide useful data for understanding the contextual behaviors of the firms studied they do not provide the necessary data for forming general theories about formal training in small businesses (Hill, 2004). In fact, most of the research to date has focused on informal learning as the primary process for employee development in small firms (Hill, 2004).

There are numerous reasons that small firms rely on informal learning as the primary process for employee development. Resource constraints, employee turnover, rapidly changing dynamics, and founder ignorance are all reason cited for the use of informal learning small businesses (Barry & Milner, 2002; Birchall & Giambona, 2007; Bishop, 2009; Hoque & Bacon, 2006; Jayawarna, Macpherson & Wilson, 2007; Kitching & Blackburn, 2002; Taylor, Shaw, & Thorpe, 2004). However, given the high failure rate of small businesses, there is growing interest by governments, financial institutions and development professionals, to identify formal employee development practices that would increase small business success rates (Stone, 2010).

Over the last twenty years, the *Resource-Based View* (RBV) or *Resource-Based Theory* (RBT) of the firm has been developed to explain how an organization's resources can give it a competitive advantage in the marketplace (Acedo, Barroso & Galan, 2006; Barney, 1991; Barney, Wright & Ketchen, 2001; Coff, 1999; Crook, Todd, Combs, Woehr & Ketchen, Jr., 2001). RBV argues that human capital is one of the primary resources that is difficult to copy and therefore can be a significant competitive advantage for a firm if properly developed and utilized (Coff, 1997; Grant, 1991, 1996; Kogut & Zander, 1992).

While it has been shown that a broad array of human resource programs and policies can strategically contribute to the overall efficiency and performance of an organization (Becker & Gerhart, 1996), more recent studies have focused on human resource development practices and specifically on training and its impact on organizational performance (Aguinis & Kraiger, 2009; Aragon-Sanchez, Barba-Aragon & Sanz-Valle, 2003; Jones, Beynon, Pickernell, & Packman, 2013; Katou, 2009).

There is growing empirical evidence that formal training in the workplace has a positive impact on the overall efficiency and performance of an organization (Aragon-Sanchez, et al., 2003; Becker & Huselid, 2006; Bowen & Ostroff, 2004; Combs, Liu, Hall & Ketchen, 2006; Crook, Ketchen, Combs & Todd, 2008; Huselid, 1995; Subramony, Krause, Norton & Burns, 2008). However, there are conflicting results as to the areas of an organization's performance that training impacts (Tharenou, Saks & Moore, 2007).

There are several reasons for the lack of conclusive data on the effect of training on organizational performance. One of the reasons is that very few firms spend the time and resources to measure this within their own organization (Swanson, 2001). A second reason is the lack of available data to conduct causal types of analysis (Aragon-Sanchez et al., 2003; Henri d'Arcimoles, 1997; Jones et al., 2013; Katou, 2009).

Most of the research to date in this area has used either case studies or cross-sectional analysis (Martin-Alcazar, Fernandez, & Gardey, 2005; Baldwin & Danielson, 2002; Gerhart, 2005). While the use of such data can lead to better understanding of the relationship between training and organizational performance it cannot provide the type of analysis needed to assess a direct causal relationship. Such an analysis requires the use of longitudinal data so the relationship can be analyzed over time (Crook et al., 2001; Henri-d'Arcimoles, 1997).

Problem Statement

Small businesses play an important part in the economic and social well-being of our communities, and yet up to 80% of small businesses fail within the first year of operation. Understanding small businesses' use of formal training and what impact that training has on firm performance would significantly aid in the development and adoption of formal training programs designed to increase the success rate of small businesses.

Research Questions

1. What are the characteristics of small businesses that use formal training for employee development?

2. What impact over time does formal training have on the organization-level performance of small businesses?
3. Are there characteristics of small businesses that moderate the effect that training has on organizational performance?

Objectives of the Study

The major objective of this study is to better understand the characteristics of small businesses that use formal training and the impact that this training has on overall organization-level performance. The following questions will be addressed.

Question 1: Are there characteristics of small businesses that indicate their use of formal training?

Question 2: Are there significant variances in organization level performance for firms that do formal training versus those that do not?

Question 3: Is there a positive relationship between formal training and organization level performance?

Question 4: Are there organizational variables that moderate the effect that formal training has on organization level performance?

Significance of the Study

The resource-based view (RBV) of the firm suggests that investment in human capital within an organization can positively impact the organization's overall efficiency and performance. This research will provide quantifiable evidence to support this view. In addition, while there is a growing body of research looking at this phenomenon in large organizations, there is very little research that has looked at it within small

organizations. Because small organizations lack many of the capital resources (financial, operational, etc.) of larger firms, their dependence on human capital is even more critical to their success. As such, being able to show that the use of formal training in small organizations positively contributes to their overall performance will provide needed evidence to the value of the investment in training for such firms.

Conceptual Framework

The Resource-Based View of the firm has been one of the dominant theoretical models of competitive advantage in strategic management studies for the last twenty years (Becker & Gerhart, 1996; Crook et. al., 2011; Katou, 2009). Barney's (1991) seminal paper on the resource-based view of sustained competitive advantage provides the foundational framework for how investment in human capital can provide a sustainable competitive advantage to a firm by increasing its efficiency and performance over the competition. According to Barney, a firm's resources fall into three general categories: physical capital, human capital and organizational capital.

Barney's (1991) initial model, as shown below in Figure 1, illustrates how resources that are heterogeneous and immobile lead to sustained competitive advantage. While there have been many additions and challenges to Barney's original model (this will be discussed in Chapter 2) the theoretical foundations upon which he based his model remain solid. And, while Barney identifies three general categories of capital resources that can lead to competitive advantage, it is human capital that is seen as having the most potential for sustained competitive advantage as long as that capital is both heterogeneous and immobile.

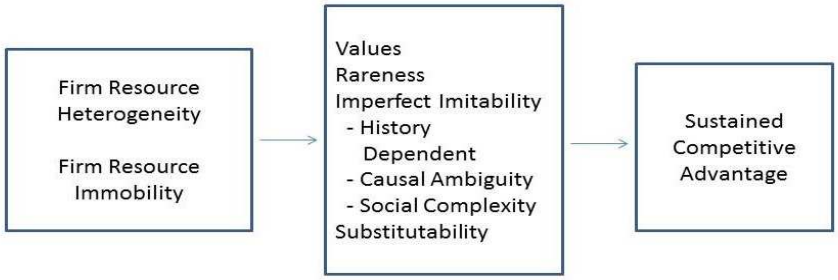


Figure 1: The Relationship between Resource Heterogeneity and Immobility, Value, Rareness, Imperfect Imitability, and Substitutability, and Sustained Competitive Advantage (Barney, 1991, p. 112)

Katou (2009) developed his causal pathway, shown below, based on RBV to show specifically how the development of human resources leads to increases in both employee performance and ultimately organizational performance.

Resourcing → Development → Skills → Attitudes → Behavior → Performance

Figure 2: Basic Causal Pathway between HRD and Organizational Performance (Katou, 2009).

Almost all theory development and research around RBV has been focused on large organizations. However, as we have mentioned, small businesses play an increasingly important role in our economic vitality. Applying RBV to small nascent firms reveals some important considerations. While access to financial capital has been widely researched and recognized as critical to the success of nascent firms, there is growing recognition of the vital role that human capital, beyond the founders, plays in

their survival and success. In fact, an argument can be made that human capital is even more critical to the survival and success of nascent firms than any other capital resource.

Summary of Methodology

This research will use longitudinal data collected by the Kauffman Firm Survey (KFS) to better understand the use of formal training in small firms and analyze what impact formal training has on organization level performance. The KFS survey is the first data set of its kind in the U.S. to systematically survey small nascent firms from startup through seven years of operation. It is currently the only large scale, longitudinal data set of small businesses that provides data on both the use of formal training and organizational performance metrics.

Hierarchical Linear Modeling (HLM) will be the primary statistical technique used in this analysis. HLM (also referred to as random effects models, mixed liner models, or random coefficient models) is specifically designed to examine nested data and allow for the simultaneous investigation of relationships within the same level as well as relationship across levels (Gentry & Martineau, 2010; Hoffman, 1997; Short, Ketchen, Bennett & du Toit, 2006).

By using HLM and longitudinal panel data with more than three years or waves of responses we hope to show conclusive links between formal training and organizational performance. Such an analysis will provide clear direction for the development and adaption of formal training programs for small businesses.

Limitation of the Research

While the goal of this research is to find significant correlations between formal training and various metrics of organizational level performance in the survey population it needs to be noted that there are many other factors which could impact an organization's performance beyond the ones used in this analysis or available from the KFS. And, while the statistical models used in this analysis, factor that into the analysis, this does limit the ability to generalize the results beyond the survey population.

Definition of Terms

Nascent Small Firms – For the purposes of this study all of the firms taking part in the Kauffman Firm Survey will be classified as Nascent Small Firms. All of the firms surveyed began operation in 2004, the first year of the survey. In the 2006 survey, over 90% of the firms had 10 or fewer employees.

Resource-Based View (RBV) – RBV states that sustained competitive advantage for an organization can be attained through strategic use of its resources and capabilities "that are neither perfectly imitable nor substitutable without great effort" (Barney, 1991, p. 117).

Training – "Training is defined as the systematic acquisition and development of the knowledge, skills and attitudes required by employees to adequately perform a task or job or to improve in the job environment" (Tharenou, Saxs & Moore, 2007, p. 252).

Organization Level Performance – Organization Level Performance is measured by various metrics that are taken by the organization as a complete unit. These metrics

would include: revenue, expenses, profit/loss, number of employees, and revenue per employee.

Multilevel Analysis – Multilevel analysis is a methodology used in the analysis of complex patterns of variability that exists when data resides at multiple levels of analysis. For example: students (level 1), classroom (level 2), and school (level 3). Multilevel analysis is used to look at the causal effects of relationships that exist across these levels.

Hierarchical Linear Modeling (HLM) – HLM is a multilevel statistical model with linear parameters used to examine nested data structures.

Assumptions

Small Business - For the purposes of this study it is assumed that all of the firms in the Kauffman Firm Survey are small businesses. While there are some firms that might have grown beyond what would normally be considered small, the average size of firms in the study is less than 10 employees.

Formal Training – It is assumed for the purposes of this study that any firm which states they have spent money on training is doing formal training.

Summary

Small businesses play an increasingly important role in the economic and social well-being of our local, national and global communities. A high number of new small businesses fail resulting in significant economic and social costs to those communities. Being able to identify tangible ways to increase the success rate of nascent small businesses would contribute to the overall well-being of our communities.

The Resource-Based View (RBV) of the firm provides a model to show how investment in human capital can increase an organization's performance. More specifically, formal training has been shown to lead to increased organization level performance.

Hierarchical Linear Modeling (HLM) can be used to analyze the relationship of data across multiple levels of analysis, such as the relationship of individual level training on organization level performance over time. This research will use longitudinal panel data from the Kaufmann Firm Survey to look at the impact of formal training in small nascent firms on their organization level performance over time.

The hope is that such analysis will provide needed conformation of the critical role that investments in formal training by small nascent firms has their organizational performance, contributing to higher levels of survival and success.

CHAPTER II

REVIEW OF LITERATURE

Introduction

Problem

The impact of formal worker training on individual or team level performance has been well researched (Aguinis & Kraiger, 2009). However there is a significant gap in the research of what impact training has on overall organizational performance (Swanson, 2001; Tharenou, Saks, & Moore, 2007).

For the purpose of this paper training is defined as "the systematic acquisition of and development of the knowledge, skills and attitudes required by employees to adequately perform a task or job or to improve performance in the job environment" (Tharenou et al., 2007, p. 252).

While much of the focus on performance outcomes of training has been at the individual or team level (Kozlowski, Brown, Weisbein, Cannon-Bowers, & Salas, 2000), there is significant interest, given the amount of money spent every year on training, and theoretical rationale, for researching its impact on organization-level performance (Alliger et al., 1997; Kozlowski et al., 2000).

Part of the reason for the lack of research on organization-level outcomes has been a lack of good theoretical models that span the gap between micro and macro levels of analysis (Wright & McMahan, 1992). Tharenou et al. (2007) discuss several theoretical models which could be used in the study of strategic human resource management. Three of the models, *resource-based view of the firm, behavioral*

perspective, and the cybernetic systems model, provide ways to look at training from an organization-level perspective.

Small businesses play an important part in the economic and social well-being of our communities, and yet up to 80% of small businesses fail within the first year of operation. Understanding small business use of formal training and what impact that training has on firm performance would significantly aid in the development and adoption of formal training programs designed to increase the success rate of small businesses.

This research will use the theoretical foundation of the Resource-Based View of the firm to look at the impact of a specific human resource development practice (formal training) on the organization level performance of small nascent firms over time. The research will attempt to answer the following questions.

Research Questions

1. What are the characteristics of small businesses that use formal training for employee development?
2. What impact over time does formal training have on the organization-level performance of small businesses?
3. Are there characteristics of small businesses that moderate the effect that training has on organizational performance?

By answering these questions we hope to provide needed insight into both the characteristics of small firms that use formal training and the effect that such practice has on their overall performance. The following provides the theoretical foundation for this

analysis and an overview of the specific functions of the organization which will be tested.

Theoretical Foundations

There has been significant discussion over the last twenty-five years within the field of human resource management on the theoretical foundations linking human resource management to the strategy of the organization (Dyer, 1985; Mahoney & Deckop, 1986; Swanson & Holton, 2009; Wright & McMahan, 1992). The emergence of Strategic Human Resource Management (SHRM) was in large part an attempt to link human resource practices to strategic management practices with an organizational context (Wright & McMahan, 1992).

One reason such a theoretical foundation is necessary is to provide the framework for how practices at the individual level within an organization can impact the overall effectiveness of the organization (Dubin, 1976). Without a theoretical framework it is difficult, if not impossible, to predict how actions at one level impact actions at another level. Being able to provide this predictive linkage is especially important in looking at the applied nature of human resource development practices such as formal training.

Both HRD practitioners and researchers benefit from having a sound theoretical foundation (Wright & McMahan, 1992). For the HRD practitioner, being able to understand how various development initiatives will impact the overall organization is critical to helping them decide which initiatives, or combination of initiatives, to implement. Further, such a theoretical foundation gives them an understanding of what

other variables within the organization would be affected by the initiative and therefore what they should be looking at to assess the effectiveness of their intervention.

For the HRD researcher, having a theoretical foundation is critical for them to be able to build models and tests to determine the accuracy of those models (Wright & McMahan, 1992), Without the theoretical framework such testing and analysis cannot provide predictive or generalizable results.

Classical Theories

Since Adam Smith's (1776) *division of labor* in *Wealth of Nations* there have been numerous theories of the organization. The industrial age introduced several new "classical" theories based on Smith's original concepts (Docherty, Surles, & Donovan, 2001). Classical theories of the organization focused the management and structure of the organization as the driving forces in increasing organizational efficiencies.

Taylor's (1947) *scientific management approach*, focused on improving worker productivity through efficiency, standardization, specialization and simplification. For Taylor, management's role was to scientifically analyze each part of the production process for maximum efficiency, select workers through some scientific process, and then expertly train them to be as productive in their specific work role as possible. This required a high level of assessment of both processes and employees. Taylor also believed that high levels of efficiency are only attainable through cooperation between management and workers.

Weber's (1947) *bureaucratic approach* viewed the organization in the larger context of society with clear lines of authority. Decisions are made at the management

level and disseminated through a clear chain of command. According to Weber, an organization's structure should reflect the maximization of this chain of command, jobs should be specialized to maximize productivity and rules and regulations should be predictable and stable so as not to create confusion or uncertainty.

Fayol's (1949) *administrative theory* also focused on the centralized role of management in planning, organizing, training, commanding, and coordinating functions within the organization.

While each of these classical theories has unique nuances, they are all based on Smith's notion of division of labor, and its four components: 1) hierarchy of authority, 2) span of control, 3) centralization versus decentralization, and 4) specialization of function or task, to increase organizational productivity (Gortner, Mahler, & Nicholson, 1987). And, while these theories may serve to help increase productivity in organizations that do mass production, they do not work well in organizations that are in industries of continual change, or in economies of continual changes such as the modern global economy.

These classical theories also shared the belief that the worker is basically lazy and needs order, structure and incentive in order to become more productive. Therefore, external forces could be applied to the worker to change their behavior.

Neoclassical Theories

The emergence of the human relations movement (Mayo, 1933), and the growth of behavioral psychology after WWII, gave rise to the neoclassical view of organizations. The fundamental shift from classical theory was the focus on the individual rather than the organization (Docherty et al., 2001).

The human relations movement focused on social factors such as the treatment of workers by management and the relationships between workers as the driving or determinant factor of worker's behavior within the organization.

Much of Mayo's (1933) work was based on the famous Hawthorne studies of 1927 (Roethlisberger & Dickson, 1939). The Hawthorne researchers worked with Western Electric to conduct a series of experiments focused on improving worker productivity. The conclusion of the research was factors such as employee relationships and employee motivation contributed more to their level of productivity than extrinsic work conditions. Even though subsequent analysis of the Hawthorne research revealed significant flaws in their analysis, the shift in focus from the efficiency models of classical theory to a focus on human behavior around employee development, motivation and satisfaction was significant.

Douglas McGregor (1960) contributed significantly to this shift in his discussion of *Theory X* and *Theory Y*. For McGregor, *Theory X*, representative of the classical approach of human nature as lazy and unambitious, is not reflective of the true nature of people. By contrast, McGregor's *Theory Y* purports that individuals naturally seek out responsibility and personal success. The role of managers in *Theory X* is one of control through policies, penalties and incentives, while in *Theory Y* management's responsibility is to create work conditions that are optimal for employee development and satisfaction. As such, one of the key elements of neoclassical organizational theory is decentralization of decision making.

While the shift in neoclassical theory from organizational efficiency to employee satisfaction was significant and led to many developments of the fields of human resource management and ultimately human resource development, it still lacked the multi-level perspective of the organization needed to understand how change at the individual level affects the organization. As we've discussed before, the lack of such theory makes it difficult, if not impossible, for both the HRD practitioner and the HRD researcher to assess and test changes in employee behaviors as they relate to changes in organizational outputs.

Contemporary Theories

The last twenty years have seen a dramatic shift in management research from an "operational" view of efficiency and routine to a "strategic" view where the organization is seen as part of a larger ecosystem (Hicks & Gullet, 1975; Hofer & Schendel, 1978; Miles & Snow, 1984; Porter, 1980). Accordingly, the field of Human Resource Management shifted to Strategic Human Resource Management (SHRM) (Miles & Snow, 1984; Schuler & Jackson, 1987; Wils & Dyer, 1984).

At the core of this shift to the strategic perspective is the notion that organizations do not exist in isolation, but rather are part of a larger ecosystem which is continually impacting the organization (Albretch, 1983). Management must *strategically* respond to outside influences to best utilize the organization's resources.

The foundation of Strategic Management is *Systems Theory* and what is commonly referred as the *systems approach* to the organization (Albretch, 1983). The shift to a *systems approach* of the organization provided a fundamentally different

perspective of how one thinks of an organization's resources. Instead of continually trying to maximize the efficiency of those resources, management began to focus on how to strategically use and develop those resources in response to the external environment. Now, instead of looking at human resource policies and programs as ways to reward or punish employee to behave is more productive ways, SHRM looked at human resources as integral to the overall strategy of the organization and began to look at the links between individuals within the organization on how their behavior impacts organization level outcomes (Dyer, 1984a, 1984b; Fombrun, Tichy, & Devanna, 1984; Jackson & Schuler, 1995; Jackson, Schuler, & Rivero, 1989; Schuler & Jackson, 1987a, 1987b; Schuler & MacMillan, 1984).

While *Systems Theory* and the *systems approach* provided researchers with a framework to look at the organization within the larger context of its environment, and it provided for the shift from operational management to strategic management, it still lacked a theoretic perspective of the organization to provide for a way to link changes at the individual level to changes at the organizational level (Bacharach, 1989; Dyer, 1985; Mahoney & Deckop, 1986).

Wright and McMahan (1992) looked at six theoretical models (*Resource-Based View, Behavioral Approach, Cybernetic Model, Agency/Transaction Cost Model, Resource Dependence, and Institutional)* which could provide the framework needed to better understand the determinants of HR practices and their impact on micro-macro relationships in the organization (Dyer, 1985; Guest, 1989; Lengnick-Hall & Lengnick-Hall, 1988).

While each of these proposed frameworks provides ways to view the link between human resources and organizational outcomes, for the purposes of this research, the *Resource-Based View* (RBV), because of its focus on sustained competitive advantage, provides the most robust framework for looking at the impact of formal training and organization level performance.

Resource-Based View of the Firm

RBV has been one of the dominant theoretical models of competitive advantage in strategic management studies for the last twenty years (Becker & Gerhart, 1996; Crook et al., 2011; Katou, 2009). Barney's (1991) seminal theoretical paper on the resource-based view of sustained competitive advantage provides the foundational framework for how investment in human capital can provide a sustainable competitive advantage to a firm by increasing its efficiency and performance over the competition. According to Barney (1991), a firm's resources fall into three general categories: physical capital, human capital and organizational capital. Of these three, human capital appears to be the best resource for creating sustained competitive advantage.

The resource-based view (RBV) of the firm suggests that investment in human capital within an organization can positively impact the organization's overall efficiency and performance (Barney, 1991; Conner, 1991; Wernerfelt, 1984; Wright & McMahan, 1992). Because small organizations lack many of the capital resources (financial, operational, and so on) of larger firms, their dependence on human capital is even more critical to their success. As such, being able to show that the use of formal training in

small organizations positively contributes to their overall performance will provide needed evidence to the value of the investment in training for such firms.

At the forefront of RBV are the notions of competitive advantage and sustained competitive advantage (Wright & McMahan, 1992). For Barney (1991) a competitive advantage exists "when a firm is implementing a value creating strategy not simultaneously being implemented by any current or potential competitors" (p. 102). Resources must exhibit two characteristics for them to be a competitive advantage, 1) they must be heterogeneous, and 2) they must be immobile (Wright & McMahan, 1992). Barney's initial model, shown below, shows how resources that are heterogeneous and immobile lead to competitive advantage. Resource heterogeneity refers to the differences of resources (physical, human, operational) between firms. Resource immobility refers to the inability of competing firms to obtain those resources from the other firms.

While resource heterogeneity and resource immobility are required for a competitive advantage to exist, a sustained competitive advantage requires additional criteria (Barney, 1991). First, the resource must add positive value to the firm. Second, the resource must be unique to the organization or rare to its competitors. Third, the resource must be imperfectly imitable. Forth, the resource cannot be substituted by another resource by a competitor. The figure below shows the various aspects of Barney's model.

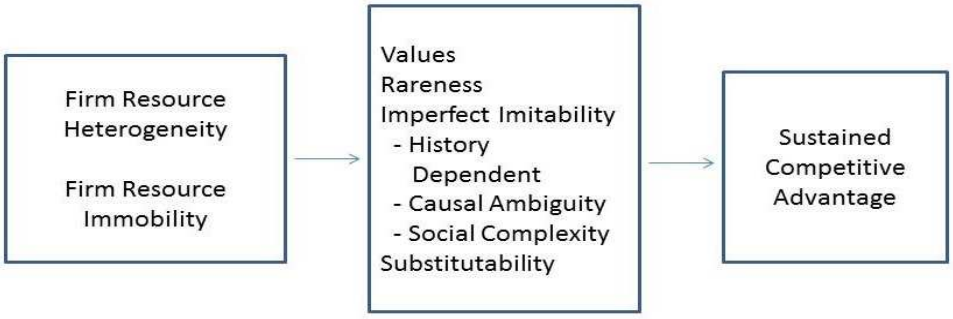


Figure 3: The Relationship between Resource Heterogeneity and Immobility, Value, Rareness, Imperfect Imitability, and Substitutability, and Sustained Competitive Advantage (Barney, 1991, p. 112).

While Barney identifies three general categories of capital resources that can lead to competitive advantage, it is human capital that is seen as having the most potential for sustained competitive advantage as long as that capital meets the criteria established by Barney (Wright & McMahan, 1992; Wright, McMahan, & McWilliams, 1994).

Following Barney's model, Wright et al. (1994) lay out the ways that human resources can meet the criteria for sustained competitive advantage.

First, human resources provide value for the firm when the demand for and supply of labor is heterogeneous. In other words, human resources provide value to the firm if a firm requires unique knowledge, skills, and abilities (KSA) and individuals differ in their knowledge, skills and abilities.

Second, based on the notion of normal distribution of knowledge, skills, and ability, human resources with high levels of each are rare. Therefore, organizations that

are selective in obtaining, or are actively developing, human resources with high levels of KSA, fit the criteria of rarity.

Third, Wright et al. (1992) use the concepts of historical conditions, casual ambiguity, and social complexity to demonstrate that well developed human resources are not likely to be imitated by a competitor.

Fourth, when looking at the notion of a resource not being substitutable, Wright et al. (1992) argue that human resources are the most difficult to substitute with regards to sustaining a competitive advantage. For example, if a competitor obtains or develops new technology which increases productivity, surpassing that of the firm which had a productivity advantage because of its human resources, once that firm obtained technology equal to the other, its human resources would again produce a competitive advantage.

While RBV offers a framework for looking at why investments in human resource development could lead to great organizational performance (Mayo, 2000; Mabey & Ramirez, 2005), it does not provide, on its own, a direct causal linkage, or pathway between HRD and increased organizational performance (Bartlett, 2001; Katou, 2009). Nor does it address the many individual, organizational and external variables that could mediate the linkage between these levels of analysis (Katou & Budhwar, 2006; Wright & Gardner, 2002; Wright, Gardner & Moynihan, 2003). This is sometimes referred to as the "black box" (Wright et al., 2003, p. 21) problem. To address these shortcomings researchers have suggested several alternative or supplementary perspectives. A look at three of these perspectives will help us to address these concerns.

Universalistic perspective – One group of researchers have focused on the *universalistic perspective* to address the issue of the moderating effects of human resource practices on other levels of analysis (Delaney, Lewin, & Ichniowski, 1989, Huselid, 1995; Osterman, 1994; Pfeffer, 1994; Terpstra & Rozell, 1993). The basic premise of this perspective is that some human resource practices are always better than others and that any organization would be better off adopting these *high performance work practices* or *best practices* (Osterman, 1994; Pfeffer, 1994). Pfeffer (1994) highlighted 16 management practices that would lead any organization to greater productivity and profit. Among these 16 are: incentive pay, employment security, promotion from within, and training and skill development. From this perspective the link between HR practices and organizational outcomes is direct in the sense that, given the same independent variable and dependent variable, we should see the same effect across all organizations (Delery & Doty, 1996).

Contingency perspective – Another group of researchers have focused on the *contingency perspective* to talk about how HR policies and practices can moderate organizational outcomes (Delaney & Huselid, 1996; Dewar & Werbel, 1979; Lengnick-Hall & Lengnick-Hall, 1988; Lopez, Peon, & Ordas, 2005). This perspective asserts that in order for HRD policies and practices to be effective they must be aligned with other processes and practices in the organization. The primary contingency which can moderate the effectiveness of HRD efforts is an organization's strategy. Thus, in order to assess the moderating impact that an organization's strategy has on HRD practices and organizational performance, one has to first identify the type or model of strategy the

organization is using (Delrey & Doty, 1996). This additional interacting variable complicates the relationship and therefore the analysis (Schoonhoven, 1981; Drazin & Van de Ven, 1985).

Configurational perspective – The final perspective to be noted is concerned with how a collection or "pattern" of independent variables relate to a dependent variable (Doty, Glick & Huber, 1993; Meyer, Tsui, & Hinings, 1993; Venkatraman & Prescott, 1990). Configurational approaches are complex and difficult, if not impossible, to assess without holistic or mixed methods analysis (Doty et al., 1993; Venkatraman & Prescott, 1990). The key for this perspective is to identify configurations of HR policies and practices that provide the best horizontal and vertical fit. Horizontal fit referring to the consistency of HR policies and practices across the organization. Vertical fit referring to the alignment of these HR policies and practices with other aspects of the organization such as strategy. Those configurations with the best fit in both directions will maximize the impact of those HR efforts (Delery & Doty, 1996).

Given the diversity of perspectives from which one can look at the linkages between HRD and organization level performance it's difficult to present a model from which all of these perspectives can be addressed. Each of these perspectives can be viewed independently or within the context of a larger theoretic framework such as RBV. In fact, all of these views may be necessary in order to fully understand the relationships or linkages between HR policies and practices and organization level outcomes. For the purpose of this research, primarily dictated by the available data, we will be most concerned with observing the direct link between formal training and organizational

performance, and then look at contingency and configurational factors which might moderate the impact of those links. Within this context it is helpful to examine a causal pathway model which takes into account different perspectives.

Katou (2009) developed his causal pathway, as shown below in Figure 3, based on RBV to show specifically how the development of human resources leads to increases in both employee performance and ultimately organizational performance.

Resourcing → Development → Skills → Attitudes → Behavior → Performance

Figure 4: Basic Causal Pathway between HRD and Organizational Performance (Katou, 2009).

Katou's (2009) causal pathway is based on the premise that improving employees knowledge, skills and attitudes will ultimately result in improved organizational performance. In other words, showing how HRD can impact organizational performance.

The starting point for the pathway is making sure that a firm's recruiting and hiring practices (resourcing) provide the correct pool of employees and then that those employees are given the correct development programs to increase their knowledge, skills and attitudes (KSA) (Raghuram. 1994). Starting with the correct employee pool, and then developing their KSAs, will result in higher motivation, increased productivity and higher retention (Bartlett, 2001; Fey, Bjorkman, & Pavlovskaya, 2000; Harde, 2003). The resulting change in employee behavior will result in changes in organization level

outcomes (Aguinis & Kraiger, 2009; Aragon-Sanchez et al., 2003; Becker & Gerhart, 1996; Crook et al., 2011; Tharneou et al., 2007; Wright, Gardner, & Moynihan, 2003;).

Some of the aspects of Katou's (2009) causal pathway that are important to note are that it contains both *indirect linkage* and *direct linkage*. The indirect linkage is between development and performance which is moderated by the outcomes skills, attitudes and behavior. The direct linkage is between the resourcing and development and the outcomes of skills, attitudes and behavior. In addition, this causal pathway can be moderated by other contextual variables and contingencies internal and external to the organization (Delaney & Huselid, 1996; Lopez et al., 2005).

Almost all theory development and research around RBV has been focused on large organizations. However, as we have mentioned, small businesses play an increasingly important role in our economic vitality. Applying RBV to small nascent firms reveals some important considerations. While access to financial capital has been widely researched and recognized as critical to the success of nascent firms, there is growing recognition of the vital role that human capital, beyond the founders, plays in their survival and success (The European Centre for the Development of Vocational Training, CEDEFOP, 2011; Gibb, 1997; Hill, 2004; Kotey & Slade, 2005; Sadler-Smith, Down, & Field, 1999). In fact, an argument can be made that human capital is even more critical to the survival and success of nascent firms than any other capital resources.

Now that we've looked at the theoretical framework from which we can model the causal pathway between human resource development and organizational

performance, let's turn our attention to the area of human resource development and specifically the role of training and organizational performance.

Human Resource Development

Human Resource Development is a field of study and practice focused on improving organizational effectiveness through the development of individuals and organizational systems and process. Swanson and Holton (2009) provide the following synthesized definition of HRD:

> "Human resource development is a process of developing and unleashing expertise for the purpose of improving individual, team, work process, and organizational system performance." (p. 4).

One can tell from this definition that the field of HRD encompasses all levels within an organization, from individuals to the entire organization. Because of the complexity of looking at all of the components of HRD at once, HRD researchers and practitioners typically focus on one of two primary areas: The training and development of employees, or development of organization level systems. The follow figure provides a detailed overview of the many parts of HRD.

Figure 5: Human Resource Development: Definitions, Components, Applications, and Contexts (Swanson, 2008, In Swanson & Holton, 2009).

While separating *training and development* from *organization development* allows researchers and practitioners to focus on specific aspects of HRD, the separation can create the impression that parts of the process can be viewed independently of, or operate in isolation from, other parts (Bierema, 2000; Swanson & Arnold, 1996). Systems theory of course tells us otherwise. And, while Swanson's (2008) model of HRD separates the view based on the focus of the activity (individual or organization), Gilley, Dean and Bierema (2001) provide a practitioner based perspective which separates HRD based on philosophical orientation, learning, performance, and change.

Organizational Learning

Practitioners and researchers that focus on organizational learning believe that individual learning is the primary purpose of HRD (Gilley et al., 2001). As Bierema

(2000) states, “HRD is about development, not profit, and HRD practitioners need to carefully consider how their work impacts the human growth, not just the corporate wallet” (p. 292). For those focused on organizational learning, learning is a prerequisite to performance improvement and change (Gilley et al., 2001). The organizational learning view comprises aspects of learning cultures, learning organizations and continuous employee learning (Senge, 1990; Redding & Catalanello, 1994), along with action learning (Marquardt, 2000), critical reflection (Argyris & Schon, 1978, 1996), and transformative learning (Mezirow, 1991). Organizational learning is based on five orientations: liberalism, progressivism, behaviorism, humanism and radical adult education (Zinn, 1983). For those that view HRD from the organizational learning perspective, improving organizational performance must be accomplished through formal design, development and implementation of learning interventions (Knowles, 1970). In other words, organizational change has to begin with change at the individual level. This perspective would align with the *“neoclassical”* view of the firm discussed earlier.

Organizational Performance

A second philosophical orientation of HRD focuses on performance and states that performance improvement and management are the essential components of HRD (Swanson, 1995; Rummler & Brache, 1995). Organizational performance focuses on systems theory, psychology and knowledge management (Brethower, 1999). While training can be a part of improving organizational performance, training is a secondary focus (Silber, 1992). From this perspective, management action is the best approach for improving organizational performance (Stolovitch & Keeps, 1999). One can assess and

improve an organization's performance through detailed analysis of the organization's structure, culture, employment practices, and motivational environment and, by improving an organization's human systems, one can improve organizational effectiveness (Fuller & Farrington, 1999; Rosenberg, 1998). For the organizational performance researcher or practitioner, the focus is on analysis of an organization's current systems, identifying inefficiencies, and developing new systems to increase efficiencies. The success of these initiatives is dependent on management's embrace and oversight. This perspective would align with the "*classical"* notion of the firm discussed earlier.

Organizational Change

The organizational change perspective focuses on an organization's culture, structure, mission and strategy (Burke & Litwin, 1992; Gilley & Gilley, 1998; Kissler, 1991). Changing the fundamental structure and culture of an organization will result in changes in organizational learning and performance (Nadler, 1998). At the foundation of organizational change is the notion that organizations are in a continual state of change. The goal then for the organizational change practitioner, or OD consultant, is to implement large scale interventions that foster structural processes and culture norms that easily adapt to change (Ulrich, 1998). A key for the organizational change effort is to develop *resilient employees* who can not only adapt to change and uncertainty, but actually thrive in such an environment (Conner, 1991; Patterson, West, Lawthom, & Nickel, 1997).

While each of these philosophical orientations of HRD offers unique perspectives, they all focus on the same goal, organizational effectiveness. Now that we have looked at some general frameworks for how to view HRD, let's turn our attention to the focus of this research, formal training and its impact on organizational performance.

Training and Performance

The impact of formal worker training on individual or team level performance has been well researched (Aguinis & Kraiger, 2009; Aragon-Sanchez et al., 2003; Jones et al., 2013; Tharenou et al., 2007; Wright et al., 2003). However there is a significant gap in the research on what impact training has on overall organizational performance (Swanson, 2001; Tharenou, et al., 2007).

For the purpose of this paper "formal training" will be defined as "the systematic acquisition of and development of the knowledge, skills and attitudes required by employees to adequately perform a task or job or to improve performance in the job environment" (Tharenou et al., 2007, p. 252).

Much of the research to date on the effects of formal training have focused on the individual or team level (Kozlowski et al.,, 2000). Researchers have shown clear links between formal training and increases in employee job performance, job satisfaction, and motivation (Arthur, Bennett, Edens, & Bell, 2003; Kraiger, 2002; Satterfield & Hughes, 2007). There is also a clear link between training and employee's knowledge, skills and attitudes (Barber, 2004; Davis & Yi, 2004; Kozlowski et al., 2000). The research seems clear that formal training has a direct impact an individual's knowledge, skills and attitudes which results in changes in employee performance.

While much of the focus on performance outcomes of training has been at the individual or team level (Kozlowski et al., 2000), there is significant interest, given the amount of money spent every year on training, and theoretic rational for researching its impact on organization-level performance (Alliger et al., 1997; Kozlowski et al., 2000).

Part of the reason for the lack of research on organization-level outcomes has been a lack of good theoretic models that span the gap between micro-macro levels of analysis. Tharenou et al. (2007) discuss several theoretical models which could be used in the study of strategic human resource management. Three of the models, *resource-based view of the firm, behavioral perspective, and the cybernetic systems model,* provide ways to look at training from an organization-level perspective.

Kozlowski et al. (2000) proposed a model designed to bridge the micro-macro gap by linking the transfer of individual level job performance to overall organizational effectiveness. This vertical transfer process happens when individual contributions work together because they involve the same content (composition), or when individual contributions are unique but work in collaboration with other individual contributions to increase overall organizational effectiveness (compilation).

While the models above suggest a direct linear relationship between training and organizational-level outcomes, other perspectives imply that other relational factors need to be taken into account (Delery & Doty, 1996; Ostroff & Bowen, 2000).

The *universalistic perspective* is the one most frequently used in studies. Its basic premise is that the more an organization uses training the greater the training's impact will be on an organization's overall performance. From this perspective there is a

positive linear relationship from training to individual level job performance to organizational effectiveness and productivity, which will result in positive financial outcomes for the organization.

The *contingency perspective* states that there are contextual factors in the organization that can contribute to or inhibit the impact of training on organization-level performance. HR practices that are well aligned with the organization's overall strategy will have a more positive effect on organization-level performance.

The *configurational perspective* states that training must be used in conjunction with other HR practices that are complimentary in order for the training to be effective. These other HR practices could include screening of employees to receive training, promotional practices, and incentive systems (Baron & Kreps, 1999; Lepak & Snell, 1999).

Any of these models could be used to determine the impact of training on organizational level outcomes. However, because of the contextual complexity of most of the models, existing research on training and organization-level outcomes is highly diverse and fragmented.

Tharenou et al. (2007) provide a review of 67 studies that looked at the relationship between training and organization-level outcomes. This review shows a lack of consistency in how the studies were conducted, what variables were measured and how the data was analyzed.

Their review showed that the studies viewed and measured training from four main perspectives: *absolute measures, proportional measures, content measures,* and

emphasis measures. Further, within each of the perspectives, there was variance in what was being measured and the number of training variables being measured.

In addition to the lack of consistent training variables, the studies also showed inconsistencies in what organizational-level outcomes were measured and how they were measured. These variables were categorized according to Dyer and Reeves' (1995) four-category definition of organizational effectiveness for evaluating effects of HR practices. The four categories are: *HR outcomes, organizational performance, financial or accounting performance,* and *stock market outcomes*. None of the studies they reviewed used stock market outcomes.

Training and Organizational Performance

Most of the organization-level research to date has focused on organizational performance outcomes of training. Measures of organizational performance can be classified into the following categories:

1. Productivity measures – productivity/value added per employee, productivity growth/gains, labor efficiency, and export growth.
2. Sales measure – sales per employee, sales output, sales growth, and new sales.
3. Quality measures – waste/defect levels, accuracy, customer satisfaction and other customer responses at the organizational level.
4. General performance measures – these are derived from a single measure or combination of measures from such factors as productivity, quality, customer satisfaction and growth.

5. Perceptual measures of organizational performance based on managers' and executives' perceptions.

Tharenou et al. (2007) analysis of research which used organizational performance measures showed an overall small positive relationship between increased use of training and organizational-level performance. They did find a stronger positive relationship in those studies that used longitudinal measures of organizational performance. For those studies that used sales measures, 14 of the 30 relationships studied were significant and positive, 3 were negative and the rest were not significant.

The results from the analysis of organizational performance measures indicate that, while training is shown to increase organizational performance, there are likely other moderating factors which influence the level of effect that training has on organizational performance.

Financial Outcomes

Twenty eight of the 67 studies reviewed by Tharenou et al. (2007) researched the relationships between training and financial outcomes. Financial outcomes were classified into the following categories:

1. Profit/profitability measures – gross or net profits, profit margin, and increases in profit margin.
2. Return measures – return on equity (ROE), return on assets (ROA), return on investments (ROI), or return on capital.
3. General financial measures – cash flow, total assets and liquidity.
4. Perceptual measures of financial outcomes by managers and executives.

Overall the analysis of financial outcomes showed little to no relationship between training and financial performance. The only significant relationship was found in managers' and executives' perceptions of the effect of training on financial outcomes.

As mentioned before, most of the studies looking at the relationship between training and organizational-level performance have focused on the direct effect of training on performance outcomes. Little research has looked at the mediating impact of other variables which might affect the relationship. Some research has found a positive relationship between training and organizational performance and organizational performance and financial outcomes (Guerrero & Barraud-Didier, 2004; Paul & Anantharaman, 2003). In addition, Faems, Sels, DeWinne, & Maes (2005), found that the relationship between training and financial performance was mediated by organizational productivity.

So, while the direct effect of training on organizational performance is small but significant, and the effect of training on financial performance appears not significant, a look at the mediating effects of other organizational variables, or the mediating effect of organizational performance variables on financial outcomes, could yield important information on the overall impact of training on organization-level outcomes.

There are two additional gaps in research which this study will attempt to address. First, there is little longitudinal analysis on the effects of training and organizational performance (Matlay, 1999). In addition, there is almost no research that has used panel data to track training and performance metrics in the same unit over time intervals. Second, because of the lack of available data, little quantitative research has been done on

training and organizational performance in small businesses (Rigg & Trehan, 2004; Storey & Westhead, 1997).

Training and Development in Small Business

Most of the thinking on training and development in small business has been done by using models derived from large organizations and applying them to small businesses (Harrison, 1997; Hill, 2004; Stewart & Beaver, 2004). There are many possible reasons why this is the case. One is the difficulty in accessing research populations. Lack of resources and expertise limits the desirability and feasibility of formal training and OD in most small businesses (Barry & Miller, 2002; Birchall & Giambona, 2007; Jayawarna et al., 2007; Kitching & Blackburn, 2002). In addition, the lack of formal structure and processes, typical in small business, presents barriers for development of theory that can be generalized.

Because of the lack of access to large populations of small businesses, most of the research on training and development in small business has been based on qualitative case studies. While such research can provide valuable contextual insight on processes happening at the individual level within the organization, it doesn't provide the type of organization-level data needed for theory development (Hill, 2004; Yin, 1994).

Whatever the real or perceived barriers to formal training and development in small businesses, current thinking about training and development in these organizations is that is it typically informal and takes place mostly within the context of an individual doing their job (Rigg & Trehan, 2004; Taylor, Shaw & Thorpe, 2004).

There are a number of stated reasons why training and development in small business is typically informal. Many studies show that management's lack of experience with formal training, and the perceived belief that such training doesn't provide a significant return on the investment, creates an organizational culture that relies on normal workday activities for employee development (Down, 1999; Taylor et al., 2004). The lack of dedicated HR staff is another reason stated for the lack of formal training.

Some researchers have argued that because of all of the complexities and development issues unique to small firms, there is no point in trying to assess the value of formal training within these organizations through quantitative analysis (Anderson & Skinner, 1999; Curran, West, & Finch, 1996; Lee, 2001; Vikerstaff & Parker, 1995).

Such a perspective may seem necessary given the lack of generalizable data on formal training in small business and its impact on organization-level outcomes. However, I would argue that such a perspective is more one of convenience rather than rigor. The lack of access to data in the past is not a sufficient excuse for not trying to collect the needed data and for not striving for new levels of understanding. In fact, there have been some studies showing that formal training can have significant impact on organizational performance in small firms (Bryan, 2006).

Other research suggests a growing need for training and development programs that help small firms deal with the increasing rapid pace of change in today business environment, but the difficulty for such programs is adoption (Stone, 2010).

The reality is that many small firms are doing formal training. In the UK, 20% of employees received training in firms with 1-9 staff. In firms with 10-49 staff over 60%

had received training (Department for Education and Employment [DfEE], 2000). Data from the Kauffman Firm Survey shows that 21% of the 2,329 firms surveyed had done training activities.

Many countries in EU, as well as Australia, and New Zealand have government sponsored training programs targeted at small to mid-size firms. While there is anecdotal evidence that some of these programs have been successful, there is little quantifiable evidence showing a positive return on the investment. And adoption rates for these programs are generally low.

Being able to show a significant positive relationship between formal training and organization-level performance in small businesses would be beneficial to increasing the discussion around this topic and ultimately increasing the development and adoption of training programs in small firms.

The preceding discussion provides us with a theoretical framework (RBV) from which we can view relationships between independent variables at the individual level and dependent variables at the organizational level. It also shows how other variables internal and external to the organization can moderate the impact of those relationships. We also looked at a causal pathway linking individual level development to changes in organization level performance through changes in a person's knowledge, skills and attitudes. We then looked at the role of HRD in facilitating these changes and specifically the impact of formal training on organizational performance. Finally, we talked about the role of formal training in small businesses and difficulty of doing quantitative analysis in such a setting, therefore making it difficult to develop generalizable theory. There is one

more area for us to discuss which is critical for doing research which looks at multiple level of analysis.

Multilevel Analysis

Foundational research in HRD comes from multiple perspectives looking at various levels of analysis. Each perspective gives insight into their respective level of analysis with varying degrees of cross-over into the other levels. It seems natural in any discipline that develops out of multiple perspectives that there is often a lack of flow across the various perspectives.

We know that multilevel theories can lead to better understanding and better transfer of theory to practice (Klein, Tosi, & Cannella, 1999). However there are significant hurdles in building multilevel theory.

While there are substantial barriers to doing this work, the benefits to having good multilevel theories are also substantial. The most obvious of these benefits is the ability for such theory to bridge the gap between individual (micro) and organization (macro) levels of analysis. In bridging this gap we are able to see both how the organization influences individual action and perception, and how individual action and perception impacts the organization.

From a practical standpoint, multilevel theories have the potential to provide one with the ability to understand how individual actions, and collective actions of individuals impact overall organizational performance. The extension of this benefit is that by strategically changing, or influencing a change in, behaviors, attitudes and

perceptions of individuals in an organization we can change the organization's performance in a desired way (Beer, 1980; Lawler, 2006).

Systems theory has developed from the efficient causal model first suggested by Aristotle, to current thinking about chaos and complexity where there are dependent, interdependent and independent parts that act in seemingly random patterns as the system transforms continuously (Burnes, 2006; Jones & Bos, 2007).

By extending the notion of the system out beyond the organization to the communal or even global context, the sheer magnitude of interdependencies becomes overwhelming. And yet, being able to better understand how the parts of a system impact each other can provide valuable information on how we can create systems which are beneficial to everyone (Beer, 1980; Stevenson, 2003; Tichy, 1980).

Research has shown that changes in an individual's knowledge, attitudes and skill affects their individual performance on the job (Lawler, 2006; Rigg & Trehan, 2004). It also seems clear that individual performance within a workplace impacts the overall performance of the organization (Nadler & Gerstein, 2006). It has also been shown that communities which have high performing organizations contribute to the overall well-being of the community (Coleman, 1990, Kanter, 2006). And finally, as we have seen in the recent financial crisis, the overall performance of organizations can impact national and international well-being. That these parts of our larger communal society are interconnected seems clear. How they are connected and in what ways they affect each other are areas that will require extensive study. Part of that work will be to build multilevel models, frameworks and theories.

There are barriers in building multilevel theories which makes it difficult. One of these barriers is the need for theorists to draw from multiple disciplines in building multilevel theories. Because of this, the volume of available research can be overwhelming. Individual level analysis requires informing from psychology, biology and others, while organization and community level analysis pulls in sociology, anthropology, economics, political science, and so on.

A second barrier is endemic to the academic community in that researchers are trained to look at things either from a micro or macro perspective. Because of this, it is difficult for them to see beyond their learned perspective. Being able to look at things from a holistic systems perspective will be necessary to move beyond this limitation.

One way for HRD researchers to move out of this boundary is to practice more *engaged scholarship* (Van de Ven, 2007). By actively moving outside of the theoretical realm and being actively "engaged" in the practical world of organizations as they go through processes of change or transformation, practitioner and researchers can become active co-creators of HRD research and theory that has real practical value. The reason I mention this here is that by doing active research within the organization, it is difficult for the HRD researcher to isolate themselves in only one level of analysis or one theoretic perspective.

One of the realities we may be faced with in looking at the multilevel perspective of HRD is that, because of the complexity, we may never be able to fully understand "how" or "why" change takes place within the system. Instead, we may have to be satisfied with understanding "what" is taking place while change occurs. This parallels

the shift that has occurred from earlier positivistic research to post-modern concept. This is one of the fundamental struggles within the current world of HRD research and practice.

Summary

Small businesses play a vital role in the well-being of our communities. Across the globe they provide the majority of jobs and account for a significant amount of economic production. And yet the survival rate of small business remains low and the lack of research focused on how to increase the survival rate of new small businesses is noticeably lacking. Understanding how to increase small business performance and therefore their survival rate would significantly help local, regional and national economies improve and sustain.

Formal training has been shown to have a significant positive impact on organizational performance. However the research is inconclusive as to why that might be the case and, with regards to small business, there is little evidence that they even do formal training, let alone how that might impact their performance.

The resource-based view of the firm (RBV) provides a theoretical framework from which we can view the relationship of individual level variables and organizational level variables. These relationships are not typically direct but are moderated through contingent or configurational factors. Katou's (2009) causal pathway shows us how development efforts at the individual level impact performance at the organizational level through changes in the individual's knowledge, skills and attitudes.

One of the primary reasons for the lack of research showing the causal relationship between formal training and organizational performance, especially in small firms, is lack of longitudinal data from which one can do this analysis. The next section will talk about the Kauffman Firm Survey and how using the statistical method of Hierarchical Linear Modeling (HLM) will allow us to address these issues.

CHAPTER III

METHODOLOGY

Background

As we have discussed there are many variables and levels of analysis that can impact the effect of formal training within an organization. One of the significant challenges in trying to understand the relationship of formal training and performance at the organizational level is estimating correlations between variables at the individual level of analysis and variables at the organizational level of analysis.

Most of the research to date has focused on the individual level or team level of analysis in part because of the complexity of analysis (Tharenou et al., 2007; Wright et al., 2003). But the primary reason is lack of data to properly conduct reliable and significant analysis (d'Arcimoles, 1997; Katou, 2009).

While research on the relationship between training and organization-level performance to date has shown marginal significance (Aguinis & Kraiger, 2009; Aragon-Sanchez et al., 2003), there are gaps in the research which leave one to question the reliability of these findings (Swanson, 2001; Tharenou et al., 2007)

First, most of the research in this area has looked at the direct relationship between training and firm performance, not taking into account mediating variables, including time, which could significantly influence the relationship, therefore limiting the reliability and significance of findings from these studies (Katou, 2009).

Second, most of the research studies had been done with relatively small sample sizes or have used cross-sectional analysis to view the relationships at a given point in

time or over two disparate points in time (d'Arcimoles, 1997; Tharenou et al., 2007; Wright et al., 2003). The use of such methods does not lend itself to understanding multivariate relationships and therefore cannot predict change across levels of analysis (d'Arcimoles, 1997; Katou, 2009).

This research will use data and statistical models specifically designed to overcome these barriers to doing multilevel analysis. The following discusses the two primary aspects of this research that differentiates it from previous research in this area: Panel Data and Multivariate Mixed Model Analysis.

Panel (Longitudinal) Data

Panel data contains measures of the same variables from the same units repeated over time. This data has a distinct advantage over "repeated cross-section" or "trend" data when trying to analyze causal interrelationships among variables (Finkel, 1995). For a causal relationship to exist between variables they must meet certain conditions: (a) they must co-vary; (b) one must precede the other in time, and (c) they must not be "spurious" or produced by their joint association with a third variable or set of variables (Menard, 1991).

With cross-sectional data it is difficult to establish temporal order therefore making it impossible to establish that co-variation between *X* and *Y* is produced by *Y* causing *X* (Finkel, 1995). In other words, without the ability to establish order of occurrence between *X* and *Y* it is not possible to establish a causal or even reciprocal relationship between them.

In contrast, because panel data takes the same unit of measurements of *X* and *Y* through time, it is possible to specify certain models that necessarily establish time precedence where prior values of each variable may affect later values of the other. It is important to note that any analysis of panel data is multilevel in nature as *time* constitutes a level of analysis. So if one were looking at the change of student test scores over a three year period, the student is the *Level 1* analysis and the test score measurements over time is the *Level 2* analysis.

In instances of reciprocal causality, panel data analysis can estimate nonrecursive models with feedback effects between variables with fewer restrictive assumptions than cross-sectional data. And panel data is also better for controlling the effects of outside variables that may result in the relationship between X and Y being either partially of fully spurious (Finkel, 1995).

Panel data satisfies three important criteria for analysis of change across multiple levels of analysis (Singer & Willet, 2003).

1. *Three or more waves of data* - The first criteria are that there are multiple waves of data. As mentioned above, these multiple waves of measuring the same unit at various times provides the researcher with the ability to establish temporal order in identified causal relationships.

   Cross-sectional analysis uses measurements collected at the same time with different samples. While it may be true that data from this type of analysis may represent causal relationships, there is no way to eliminate other explanations for the results.

Two waves of longitudinal data may suffice to show an increment of change but it is insufficient to show the process of change or trends in change. Therefore it is not possible to predict change based on two waves of data. Using only two waves of data is also subject to misinterpreting measurement errors as change (Rogosa, Brandt & Zimowski, 1982).

With three waves of data it is possible to begin to establish the shape or trajectory of change that has happened over time and therefore predict the trajectory of future change. Three waves of data is the minimum requirement for investigating change over time, however more waves of data allow for less restrictive assumptions. With only three waves of data the researcher is limited to using simpler models with more restrictive assumptions. For example: the measurement of change over time needs to be linear. More waves of data allow for changed to be viewed as nonlinear or less linear.

2. *A sensible metric for time* – Since time is the fundamental predictor in the study of change, its measurement must be reliable, valid and sensible (Gentry & Martineau, 2010). The selection of the best metric for time is best determined by the unit of analysis and what makes the most sense for the desired outcome. For an analysis of financial change of companies over time, the best time to measure would be when companies produce detailed financial reports. This could be monthly, quarterly or annually. The best metric for measuring would be financial measures that are consistent across all organizations: revenue, expenses, profit, loss, assets, liabilities and the like.

It is not necessary to have equal spaces of time between waves of measurement although equally spaced waves do provide a level of symmetry and balance that is appealing for consistency.

Another issue is whether everyone should be assessed at the same time or on the same schedule. If everyone is assessed on the same schedule the data set is *time-structured.* If data is collected at different times across subjects the data set is *time-unstructured*. Either can be used to show change over time, but the *time-structured* data set is simpler to use.

Finally, with regards to time metrics, it is not necessary for every subject in the data set to have the same number of waves. In other words, the data set does not need to be balanced. Attrition of participants is common in longitudinal studies. While non-random attrition can make drawing inferences difficult, balanced data is not required for individual growth modeling.

3. *A continuous outcome that changes systematically over time* – The specific *content* of what is being measured does not have statistical significance. However, *how* a given construct is measured does have statistical significance and not all variables are equally good or valid. Individual growth models used to examine the unique trajectories of individuals and groups in repeated measures of data require continuous outcomes whose values change systematically over time (Singer & Willett, 2003). Having this allows the researcher to formulate growth trajectories for individuals that can be compared to other individuals with the same type of data.

Continuous outcomes allow for mathematical manipulations (addition, subtraction, multiplication, division) so differences between scores can be scaled and have identical meanings. Outcomes must also be derived from instruments that ensure validity and precision.

In addition to the above requirements, longitudinal data also require that the measure, validity, and precision of the outcome be preserved over time. In other words, an outcome measure at one point in time (x) must be equitable with the same outcome measure at another point in time (x_{-1}).

Provided that panel data has been properly collected with a valid instrument, and at least three waves of data are available, it provides an excellent data source for the analysis of causal relationships across multiple levels of analysis.

Multilevel Modeling

As was previously discussed, having multilevel models for understanding causal relationships is needed in order to better understand the relationship between training and organization level outcomes. However such multilevel models contain additional complexity that require different statistical methods (Hoffman, 1997).

Starting at the individual level, human behavior is enormously complex and contextually affected. Individuals can and do behave differently in the same context. Adding additional levels of analysis such as individuals working together as a team, and being able to associate variables across levels of analysis becomes even more complex. This is why different research and analysis methods are required for multilevel modeling.

In its simplest form, multilevel modeling is needed because much of what we do as organizational researchers is multilevel. We study individuals in an organizational context. It seems natural then that we should use multilevel theories and analytic techniques specific to multilevel research. Without these, researchers can run into difficulties that can seriously impact the reliability and validity of their research. Some of the potential hazards include:

- Ecological fallacies where observations of relationship within a group erroneously assumed to hold true for individuals (Freedman, 2006).
- Atomistic fallacies where individual level measures are erroneously applied to the group (Hox, 2010).

By definition then, multilevel models are designed to analyze relationships that exist between different levels of analysis. The result of such analysis is to specify or define the links between these relationships. These links between phenomena at different levels can be top-down or bottom-up (Kozlowski & Klein, 2000). Researchers will often use both in their analysis.

In *top-down* processes one looks at contextual factors at the upper level that influence units at the lower level. These influences can have direct impact such as a company implementing a new employment policy that all employees must follow. Or the influences can be indirect such as the design of a work space that might encourage employee collaboration. These types of processes would align with the *organizational change* perspective previously discussed.

In *bottom-up* processes one looks at lower level phenomena that result in higher level "collective" phenomena. For example, an organization's culture is viewed as the collective result of its member's characteristics, behaviors and values. This process has become commonly referred to as *emergence,* meaning that the higher level phenomena emerge as a result of some *composition* or *compilation* of phenomena at lower levels (Cowan, Pines, & Meltzer, 1994; Kozlowski & Klein, 2000). These processes would align with the *organizational learning* perspective previously discussed.

Being able to describe and model the types of relationships that the researcher is studying is critical to determining the appropriate method(s) for analysis. In addition to answering how these relationships flow, one must also define where they exist. While this might seem like a rather simple task, the reality is that even small organizations have many horizontal and vertical levels of analysis. There are individuals. These individuals might work in a large team for some tasks. At other times they might work closely with one co-worker. Collectively all of the employees work for an organization. The organization operates within a community. And so on. In addition to formal structures in an organization, there are informal relationships that individuals will form on their own (Katz & Kahn, 1966).

The purpose of building a multilevel model then is to avoid these common problems of misalignment and misspecification (Kozlowski & Klein, 2000). The first step in building a multilevel model is specifying the levels of construct. In organizational research there are individual level and unit level constructs. In emergent processes the

researcher must show where the construct begins, this is usually at the individual level, and where the result is manifest at the unit level.

Defining the individual level construct is rather straightforward; however defining unit level constructs is problematic. Part of the problem is that the unit level construct may in fact be a result of lower level phenomena. Kozlowski and Klein (2000) have attempted to resolve this problem by distinguishing three basic types of unit level constructs: global, shared and configural.

- Global unit properties are single-level phenomena that are typically dictated by structure or function. A marketing department in an organization is a global unit that is not dependent on lower level units for its existence.
- Shared unit properties are those that are collectively shared by the members of the unit. Organizational culture and group efficacy are examples of shared units.
- Configural unit properties are constructs that are similar to shared units in that they originate at the individual level, but differ in that they are not seen as properties that coalesce or emerge from the members of the unit. Diversity (racial, gender, age, etc.) is one example of a configural unit property that, while it originates at the individual level, it is not shared by the entire unit, but is something that impacts the unit level's overall culture.

Properly defining the unit-level constructs is necessary for determining the proper method for researching and analyzing the multilevel phenomena. The next step is determining the levels of measurement.

Determining which level of measurement to use is dependent on what is the construct you're trying to assess. Individual level constructs are always measured at the individual level. Unit level constructs can be assessed by individual or unit level measurements. For example: a unit level construct could be assessed by a unit level measurement or could be assessed by an aggregation of individual level measurements that are then applied to the unit. There are potential problems with each of these unit level approaches.

Rousseau (1985) recommends that it is preferable to use unit level measures for assessing unit level constructs because they are more clearly linked and therefore less prone to ambiguity. Klein, Dansereau, & Hall (1994) argue that one should only use global measures in assessing unit level construct when the construct is "certain [or] beyond questions" (p. 210). Otherwise, using global measures assumes homogeneity at the individual level and provides no way to test for variance at the individual level. What seems clear is that the construct one is researching determines the level or levels of measurement.

Types of multilevel models

There are several distinct types of multilevel models. Determining which model to use is critical to determining the proper techniques for researching the desired constructs.

Single-level models are specific to relationships between constructs at a single level of analysis. Constructs at the individual level, such as those often used in psychology, are the simplest or most straight forward. While simple, they ignore

contextual influences at the unit level and therefore are inherently biased (Kozlowski & Klein, 2000).

Single-level models at the unit level are more complex and can differ significantly depending on the unit level construct being used. Group level models that compare global constructs are the simplest but suffer from the same bias issues mentioned above. Single level models using shared or configural constructs can be very complex and may require more advanced mixed-method statistical analysis and requires the researcher to not only link independent and dependent variables but also account for processes resulting out of the relationships between these variables (shared constructs).

Cross-level models describe relationships between independent and dependent variables at different levels of analysis (Rousseau, 1985). Initially, cross-level models were used to examine top-down phenomena where variables at the unit level moderate variables at the individual level. Research looking at how employee benefits impact employee morale is an example of this type of top-down model. It's only been the last 15 years or so that statistical processes has emerged that allow for accurate analysis of bottom-up cross-level models which look at emergent phenomena where individual level variables aggregate to moderate unit level constructs.

Cross-level models that include data sampled across time, such as panel data, require some special considerations. One such consideration is differential time scales across levels of analysis. Change happens much quicker at the individual level than the unit level. Therefore if one is looking at a top-down phenomenon, where a unit level variable is causing change at the individual level, the pace of change will happen

relatively quickly and the necessary time to observe change is short. However, if one is looking at emergent phenomena, where change is happening at the unit level, more time is needed to observe the change. This is why longitudinal series data is needed to analyze cross-level emergent phenomena.

A second issue with data sampled over time is entrainment. Entrainment is concerned with phenomena that are cyclical in nature where the strength of links between cross-level phenomena changes over time. The concern here is whether the timing of data sampling happens during periods of entrainment. The use of at least three waves of longitudinal data mitigates the potential effects of entrainment.

It is clear that multilevel models present significant hurdles which must be addressed in the research design. However, the benefit of being able to show links between phenomena within and across levels of analysis makes such effort worthwhile.

Research Design

This research utilizes data collected by the Kauffman Firm Survey to answer the stated research questions. The use of existing data that have been collected through a tested survey helps to eliminate issues around instrument bias and validity. The research design incorporates multilevel models and Hierarchical Linear Modeling techniques to look at the relationships of various phenomena within the data. The research looks at phenomena both within units of analysis and across levels of analysis in order to test the hypotheses.

Data

Kauffman Firm Survey

The Kauffman Firm Survey (KFS) is a panel study of 4,928 businesses founded in 2004 and tracked over their early years of operation. KFS is currently the largest longitudinal data set of small businesses ever produced. KFS is the only data set currently available that has the formal training data, organizational performance data, and the multi-waves (longitudinal) of responses needed to do this analysis. The KFS survey focused on the nature of new business formation activity; characteristics of the strategy, offerings, and employment patterns of new businesses; the nature of the financial and organizational arrangements of these businesses; and the characteristics of their founders. As of February 2014, there are seven follow-up waves of survey results beyond the base-line data. Having completed its final year of data collection, the KFS is becoming an invaluable resource for understanding the development of small firms.

Starting in 2008, the KFS began asking respondents specific questions about training activities including whether they had expenditures for training activities and the dollars spent on those activities. With the release of data from the seventh and final follow-up survey done in 2011, we now have four years of training data upon which to draw. Having three or more years (waves) of panel data can significantly increase the ability to estimate causal relationships (Finkel, 1995). Even two years of panel data will yield better results than cross-sectional data.

Participants in the survey were randomly selected from a D&B database of over 75,000 businesses started in 2004. Firms identified as “high-tech” were oversampled.

With over 2,000 variables there is a significant amount of data available for research. The data falls mainly into these general areas:

- Firm characteristics: Industry, Legal Form, # of Owners, # of Employees (PT/FT), Types of Customers, Location
- Firm strategy and innovation: Product/Service Offerings, Intellectual Property, Licensing In and Licensing Out, R&D
- Detailed financial information: Equity & Debt Financing, Income Statement Info (Revenue, Expenses, Profits), Balance Sheet Info (Assets, Liabilities, Equity)
- Employees: Types of Benefits Offered, Task/Work Structure
- Owner characteristics and work behaviors (Information on up to 10 owners): Education, Age, Race, Ethnicity, Gender, Citizenship, Immigrant Status, Hours Worked, Previous Years of Work Experience, Previous Start-up Experience (same/different industry as this firm)

In addition to the base-line data, new survey questions have been added in follow-up surveys. In 2008 (the 4th follow-up), questions were added asking participants about expenditures on intangibles assets. One of the intangible assets surveyed was employee training activities. The questions about expenditures on intangible assets have subsequently been repeated on the 5th, 6th, and 7th follow-up surveys.

A public-use microdata file of KFS data is available through download from the KFS website http://www1.kauffman.org/kfs/KFSWiki/Data-Dictionary.aspx. This dataset contains limited baseline data and a limited set of data from the first four follow-up

waves. The complete microdata set, from the baseline survey and all subsequent waves, are available through the NORC Data Enclave at the University of Chicago.

Researchers wishing to access the Enclave data must request permission from either the Kaufmann Foundation for a Kaufmann-sponsored seat or request permission directly from NORC. To receive one of the Kaufmann-sponsored seats the researcher submitted a 3 page summary of this proposed research and was approved for a 1 year sponsored seat with access to the Enclave data. Seats granted directly through NORC are charged $100 per week for access with a six month minimum duration, a $2,600 minimum cost. So receiving a sponsored seat made accessing the data affordable.

Given the large number of variables available it is necessary to succinctly identify which variables are specific to the research questions in this study. I will address these question by question.

Question 1: What are the characteristics of small businesses that use formal training for employee development?

Dependent variable – Expenditure on training (Y/N)

Independent variables – Research suggests that there are several variables which might impact the use of training by a firm (Aguinis & Kraiger, 2009). Understanding characteristics of the firm unique to those that do formal training will help us to predict what types of firms are most likely to use formal training. KFS has data in the following areas:

1. Firm size – we would expect that larger firms would be more likely to expend money on employee training.

2. Productivity rate – we would expect that more productive firms would be more likely to use formal training.
3. Funding – we would expect firms with more capital resources would be more likely to use formal training.
4. Industry sector – we would expect firms in industry sectors such as high technology, medical, and manufacturing, which require highly skilled employees to be more likely to use formal training.
5. Executive or Management education level – past research indicates there is a significant correlation between the education level of a firm's management and its use of formal training. The higher the education level of management, the more likely this will utilize formal training for employees.
6. Geographic location – we would expect firms located in areas where there is ready access to external training resources to be more likely to use formal training.
7. Asset levels – as with #3, we would expect firms with higher levels of assets to be more likely to use formal training.
8. Additional HR activities – the *configurational resource-based* view of the firm suggests that the more a firm invests in various human resource activities the more likely it is they would use formal training.

Question 2: What impact does formal training have over time on the organization level performance of small businesses?

Research has looked at many different performance variables to assess the effect of training (Tharenou et al, 2007). Dyer and Reeves (1995) outline four different categories of organizational effectives for evaluating the effects of HR practices. They are: a) HR outcomes such as absenteeism, turnover, motivation and job satisfaction; b) organizational performance outcomes such as productivity, quality and level of service; c) financial outcomes such as profit, ROI, and return on assets (ROA); d) if a public company, stock market metrics. KFS has data for assessing aspects of organizational performance and financial outcomes as follows:

Dependent Variables (Performance Metrics): Sales Growth, Profit, Employment levels, Productivity, Assets, Revenue per Employee

Independent Variables: Expenditures on training (Y/N)

Question 3: Are there characteristics of small businesses that moderate the effect that training has on organizational performance?

This question will address the contingency perspective that there is not a direct link between formal training and organizational performance by looking at other variables at multiple levels of the organization which might impact the effect that training has on organizational performance.

Dependent Variables: Performance metrics identified in Question 2 that showed a significant positive relationship to training.

Independent variables: Variables identified in Question 1 that showed a significant positive relationship to the use of training.

Human Subject Protection and IRB Issues

According to the University of Minnesota Institutional Review Board this research does not entail issues around human subject protection and does not require IRB review because:

1. It uses existing data that is anonymized, and
2. Data used in this research is from survey questions asked about organization level issues not individuals.

Data Analysis

Hierarchical Linear Modeling (HLM) will be the primary statistical technique used in this analysis. HLM (also referred to as random effects models, linear mixed models, or random coefficient models) is specifically designed to examine nested data and allow for the simultaneous investigation of relationships within the same level as well as relationship across levels (Gentry & Martineau, 2010; Hoffman, 1997; Short et al., 2006).

HLM is particularly well suited to analyzing change over time (Gentry & Martineau, 2010; Short et al., 2006). HLM was first used in education to look at student performance over time and test how classroom or school level variables might moderate individual level performance.

As was previously mentioned, there are significant hurdles to analyzing data measured over time and identify causal relationships with variables at different levels of analysis. HLM effectively overcomes these hurdles by allowing one to simultaneously estimate two models. First, HLM models relationships within lower level units, and

second, it models how these relationships within the units vary between the units (Bryk & Raudenbush, 1992).

The logic of HLM involves a two-stage procedure (Kozlowski & Klein, 2000). The Level 1 analysis is used to estimate with-in unit intercepts and slopes. Significant variance in unit intercepts and/or slopes across the units are then used as outcomes in the Level 2 analysis. The Level 2 analysis then models the effects of unit-level predictors on unit intercepts and slopes so that effects on intercepts indicate direct cross-level relationships and effects on slopes indicate cross-level moderation.

HLM relies on Generalized Least Squares (GLS) regression analysis procedure to estimate fixed parameters and the EM algorithm to generate maximum likelihood estimates of variances components (Raudenbush & Bryk, 2002). This allows for a more complete analysis of repeated measures over time because it is able to systematically investigate individual change patterns over time (Hoffman, Jacobs, & Baratta, 1993).

For this research HLM will provide for the estimation of both static and longitudinal performance parameters (as represented by an intercept and slope terms for each firm) and enable the analysis of change patterns within firm performance and between firm performance. In addition, HLM allows for the inclusion of both categorical and continuous independent variable at all levels of analysis (Raudenbush & Bryk, 2002).

Specific to firm performance, HLM allows us to examine both the overall effects of time on firm performance and address questions specific to understanding variance in performance within each firm and across firms. This is what will allow us to examine variance in organization level performance both within firms doing formal training and

across all firms. Specific analytic procedures to answer each of the questions have been developed in consultation with a statistician consultant. The appendix contains a "Consultant Technical Document" which provides an overview of the various methods used to model and answer the questions. The following outlines the proposed analytic models to be used in answering each question.

Question Modeling

Question 1: *Are there characteristics of small businesses that indicate their use of formal training?*

We will test within unit variance of predictors such as: Firm size, Growth rate, Funding, Employee turn-over, Industry sector, Executive or Management education level, Geographic location, Asset levels and Additional HR activities. Formal training will be tested as a dependent variable to determine if there exist any significant variance in the predictor variable between those firms doing formal training and those not doing formal training.

To answer this question the research will use Independent Sample t-Tests to look at sample mean variance of predictors between the group of firms doing formal training and the group of firms not doing formal training.

The first step is stating the hypotheses for Independent Sample t-Tests

Null Hypothesis: H0: $\mu 1 = \mu 2$ where $\mu 1$ standsfor the mean for the first group and $\mu 2$ stands for the mean for the second group.

Alternate Hypothesis: H0: $\mu 1 \neq \mu 2$

Sample size (n), sample mean, and sample variance have to be known for each sample. Then the next step is to calculate the t-Test as represented by the following:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(s_1^2 / n_1 + s_2^2 / n_2)}}$$

The results will show whether there is any significant variance in the means for each sample. This test will be computed for each of the identified predictors to identify any variables which can be used to predict the likelihood of a firm using formal training.

Question 2: *Are there significant variances in organization level performance over time for firms that do formal training versus those that do not?*

This test will use HLM for testing. The Level 1 test will use time as a nested level within the organization. The following is a Level 1 sample equation which is representative of what the actual equation might contain. The researcher will work with a consultant to develop the specific equation for this test.

Level 1: $Y_{Performance}j = ß0j + ß1j\ (Time) + rj$

Y is the organization level performance metric for group j. ß0j is the intercept value for group j and ß1j is the regression slope for group j. The relationship between performance variables and time intervals are estimated separately for each unit. The result will be a Level 1 intercept term and a slope term representing the estimated relationship between performance and time for each firm.

The resulting intercept value ß0j and regression slope ß1j from Level 1 are then used as outcomes for the Level 2 stage. The following equation is representative of what this might look like.

Level 2: $ß0j = Y_{00} + Y_{01}\ (\text{Training j}) + U0j$

$$ß1j = Y_{10} + Y_{11} (\text{Training } j) + U1j$$

ß0j and ß1j have already been defined above. Y_{00} and Y_{10} are Level 2 intercept terms. Y_{01} and Y_{11} are Level 2 slope terms. U0j and U1j are Level 2 residuals of variance in intercepts and slope respectively. The first equation investigates whether Training is related to the between firm variance in performance after controlling for time. The second equation assesses the degree to which Training moderates the relationship between time and performance.

Question 3: *Is there a positive relationship between formal training and organization level performance?*

The results of testing for Question 3 will reveal if formal training has any significance on organization level performance and if so, which direction is that relationship, positive or negative.

HLM provides a t-Test test for the two residual variances (U0j =T_{00} and U1j =T_{11}). This will test whether the variance in intercepts (T_{00}) and slopes (T_{11}) across groups is significantly different than zero and if so, whether the direction is positive or negative. This random-coefficient regression model provides a significance test for the mean of Level 2 coefficients and for the variance of Level 2 regression coefficient. Because ß0j and ß1j in Level 2 contain the predictor (Training) the resulting significance will be representative of this predictor.

Question 4: *Are there organizational variables that moderate the effect that formal training has on organization level performance?*

This test will look only at the subset of firms doing formal training and, at Level 1, model growth trajectories for those firms. The Level 2 model will use outcomes from Level 1 to test variables that might account for variance in Level 1 trajectories. The equations for answering this question will be similar to those in Question 2 except that we will be looking at within group variance instead of across group variance.

Level 1: $Y_{Performance}j = ß0j + ß1j\ (Time) + rj$

Level 2: $ß0j = Y_{00} + Y_{01}\ (Wj) + U0j$
$ß1j = Y_{10} + Y_{11}\ (Wj) + U1j$

Wj will be the predictor variable which might moderate the relationship that training has on organization level performance. As with Question 3, a t-Test will show if there is significant variance and if so whether the relationship is positive or negative.

Data Variables

The Kaufmann Firm Survey contains over 2,000 variables. Determining the correct variables for analysis is critical for producing significant and reliable results. KFS provides various ways to identify and search through this large data set. In addition to detailed data dictionaries starting with the base-line survey and each subsequent survey wave, the KFS website provides an online data dictionary, http://www1.kauffman.org/kfs/KFSWiki/Data-Dictionary.aspx, which was used to identify the key variables used to test for each question. The following details the specific variables (KFS data variable name) which were used for modeling and the subsequent testing for each question.

Question 1*: *Are there characteristics of small businesses that indicate their use of formal training?

Dependent variable: Formal training (y/n): (**f19b_e_worker_training)**

Independent variables:

1. Firm Size - Number of employees (#): (**c5_num_employees)**

2. Productivity rate – Look at productivity rates to see if there is a difference between firms that do formal training versus those that do not. Productivity rate = revenue: (**f16a_rev_200._amt) /** # of employees: (**c5_num_employees)**.

3. Funding – There are several funding variables. A question is do we aggregate all of them into one total funding amount or look at each independently of the other. Funding variables are:

Amount invested by angels: (**f4_eq_amt_angels_allyrs)**

Amount invested by other companies: (**f4_eq_amt_companies_allyrs)**

Amount invested by government: (**f4_eq_amt_govt_allyrs)**

Amount invested by others: (**f4_eq_amt_other_allyrs)**

Amount invested by venture capital firms: (**f4_eq_amt_vent_cap_allyrs)**

4. Industry Sector (NAICS code): (**c1a_naics_verification)**

5. Owner education level: (**g9_education_owner_01**)

6. Geographic location: (**a8_addr_ver)**

7. Asset level – There are several asset variables. For our purposes we will total all assets into one number “Total Assets”. However, it would be interesting to see if there was any variance in correlation among the various asset categories.

Value of accounts receivable: (**f29_assetval_acctrec)**

Value of cash: (**f29_assetval_cash)**

Value of equipment: (**f29_assetval_equip)**

Value of product inventory: (**f29_assetval_inv)**

Value of land and buildings: (**f29_assetval_landbuild)**

Value of other business properties: (**f29_assetval_othbusprop)**

Value of other assets: (**f29_assetval_other)**

Value of vehicles: (**f29_assetval_vehicles)**

8. Additional Human Resource activities – There are a number of HR variables that if possible, could be aggregated to see if there is a correlation between the number of other HR activities and the use of formal training.

Number of employees responsible for human resources: (**e1_a_num_human_res)**

Did the company offer a bonus plan for full-time employees?:

(**e2a_ft_emp_bonus_plan)**

Did the company offer alternative work schedules for full-time employees?:

(**e2a_ft_emp_flex_time)**

Did the company offer health insurance plan for full-time employees?:

(**e2a_ft_emp_hlth_plan**)

Did the company offer paid sick days for full-time employees?

(**e2a_ft_emp_paid_sick**)

Did the company offer paid vacation for full-time employees?

(**e2a_ft_emp_paid_vaca**)

Did the company offer a retirement plan for full-time employees?

(**e2a_ft_emp_retire_plan**)

Did the company offer stock options for full-time employees?

(**e2a_ft_emp_stock_own**)

Did the company offer tuition reimbursement for full-time employees?

(**e2a_ft_emp_tuit_reim**)

Question 2: *Are there significant variances in organizational performance levels for firms that do formal training versus those that do not?* See Question 3 for variables.

Question 3: *Is there a positive relationship between formal training and organization level performance?*

Dependent variables: Organization level performance metrics (Since the independent variable used in this test is only available starting with the 4th follow-up survey in 2008, we will use data from the 4-7th follow-up surveys for these tests)

1. Sales growth: Do firms that use formal training have higher revenue growth rates than those that do not? Total revenue for each year: (**f16a_rev_200._amt**)

2. Profit: Do firms that use formal training have higher profit margins than those that do not? Profit amount for year: (**f24_profit_amt_4**)

3. Employment growth: Do firms that use formal training have higher employment growth rates than firms that do not? Total number of employees: (**c5_num_employees**)

4. Assets: Do firms that use formal training have higher levels of asset growth than firms that do not? This would use the Total Assets variable calculated in Q1.

5. Productivity: Do firms that use formal training have higher level of productivity than those that do not? There are several productivity measures than can be calculated from KFS data.

Sales per employee = revenue/# of employees:

(**f16a_rev_200._amt / c5_num_employees**)

Profit margin = operation profit/revenue: (**f24_profit_amt_4 / f16a_rev_200._amt)**

Sales per $ of capital = revenue/assets: (**f16a_rev_200._amt**)/Total Assets

Capital intensity = assets/# of employees: Total Assets/ (**c5_num_employees**)

Question 4: *Are there organizational variables that moderate the effect that formal training has on organization level performance?*

Dependent variables: Performance metrics identified in testing Q3 that showed a positive relationship to formal training.

Independent variables: Variables identified in testing Q1 that showed a positive relationship to the use of formal training.

As mentioned before, HLM is a complex statistical process which can provide significant and reliable results when looking at multilevel causal relationships. In order to insure that the analysis models are properly formed and executed, the researcher has worked with a statistics consultant to help formulate the equation models, do data analysis, and verify results. A requirement for access to the NORC Enclave data is all analysis of data must be done within the secure Enclave environment. The Enclave provides access to SAS software which will be used for the analysis. The appendix contains a “Consulting Technical Document” provided by the statistics consultant,

Lindsey Dietz, detailing the specific procedures that will be used in the analysis. Ms. Dietz is a 3rd year Ph.D. student in the School of Statistics at the University of Minnesota.

Research Outcomes

The desired outcomes of this research were:

1. A better understanding of the characteristics of small firms that do formal training. This will aid in the identification of small businesses that might be receptive to implementing formal training in their organizations. It will also provide valuable information in the development of training programs specific to those organizations.
2. A better understanding the impact that formal training has on organization-level performance. One of the major identified barriers to adoption of formal training in small business is a perception by owners or management that training does not increase firm performance. Have credible data showing that formal training does have a significant positive impact on firm performance would be very beneficial in overcoming this barrier.
3. Formulation of multilevel models for investigating relationships of within unit and across unit variables in small businesses. Such models will be helpful in future research using the KFS data set.

Summary

The Kauffman Firm Survey (KFS) is the largest longitudinal study on new businesses ever done. The data provides researchers with a wealth of information about new business formation and development during their first eight years of operation. By

including a training variable in its fourth through eight waves, KFS provides us with the first large, longitudinal data set from which we can analyze the impact of formal training in nascent small firms on various organizational performance metrics. With four waves of data on formal training, KFS provides us with the type of data needed to do predictive statistical analysis showing causal linkages between formal training and organizational performance.

Hierarchical Linear Models (HLM) comprise a family of statistical models (linear mixed model, generalized linear mixed model, logit-normal mixed model, multivariate linear mixed model) that is used to do multilevel analysis of nested data. By using HLM we will be able to show the correlations between independent variables at the individual level of analysis with dependent variables at the organization level of analysis. The multiple waves of data allow us to determine temporal order of variable activity and therefore show indications of causal order. This *bottom-up* analysis will provide clear indications as to the impact of formal training on organizational performance in nascent small businesses. Now let us look at the results.

CHAPTER IV

RESULTS

Having set up the theoretical framework in Chapter II, and statistical framework for doing the analysis in Chapter III, the next step was to actually work with the data, refine the statistical models, and do the analysis. The following details the results of that analysis.

Question 1

Are there characteristics of small businesses that indicate their use of formal training?

The goal of this analysis is to identify characteristics of firms that do formal training versus those that do not. One of the identified problems in providing training to small businesses is the low level of adoption by those businesses when they are offered training. Understanding additional aspects of businesses that use formal training will help to target training initiatives to those firms most likely to use it, therefore increasing adoption rates. This information will also be used in Q4 to see if the variables identified as significant with regards to formal training in some way moderate the impact of formal training on organizational performance.

Statistical Model – Because formal training is a binary variable (Yes =1, No = 0) we could not use standard Linear Mixed Model (LMM) for this analysis. Instead we used a Logit-Normal Mixed Model (LNMM) which allows us to model the logit transform of the probability of a company saying yes (see the Consulting Technical Document in the

appendix for a full discussion of this). Following is the statistical equation and variables used for answering this question.

$$\text{Level 1: } f(y_{ij}|\eta_i) = \exp\{y_{ij}\eta_i - \log[1+e^{\eta_i}]\}$$
$$\eta_i = x_{ij1}\beta_1 + \ldots + x_{ijp}\beta_p + \alpha_1 + \ldots + \alpha_r$$
$$\text{Level 2: } \alpha_i \overset{ind.}{\sim} N(0, \sigma_i^2) \text{ for all } i$$

Final Model: survey variable is a categorical variable which is 4,5,6, or 7 stating where we are in time

Logit(*f19b_e_worker_training*)=

FIXED COEFFICIENTS

survey_variable+
number_employees+
NAICS_first_category+
sales_per_emp_category +
owner_education +
assets_category +
HR_emp_category +
hr_score+

survey_variable*number_employees+
survey_variable*NAICS_first_category+
survey_variable*sales_per_emp_category +
survey_variable*owner_education+
survey_variable*assets_category +
survey_variable*HR_emp_category+
survey_variable*hr_score

RANDOM VARIANCE COMPONENTS

Intercept by MPRID (company identifier) with AR(1) correlation structure over time

Population Size – The full data file going into this analysis was 4,928 records. The set was then reduced to include only those records which had responded to question f19b

(f19b_e_worker_training) in at least one of the four years (2008, 2009, 2010, 2011) that the question was asked. This brought the number of records to 2,445.

Dependent variable: Formal training (y/n): (**f19b_e_worker_training)**

> *Literal Question: Investments in intangible assets are expenditures expected to produce long-term benefits for businesses. I'm going to read you some types of intangible assets. When thinking about each category, please consider the cost of in-house activities in these areas including the time of the business owner(s), as well as services or license fees from outside providers. Did [NAME BUSINESS] have expenditures in [ITEM/Worker training] in calendar year 20..?*

Figure 6: Number of Firms Doing Formal Training in 2008 Survey

Responses 2008: No = 1764, Yes = 451, Null Response = 230

Figure 7: Number of Firms Doing Formal Training in 2009 Survey

Responses 2009: No = 1628, Yes = 402, Null Response = 415

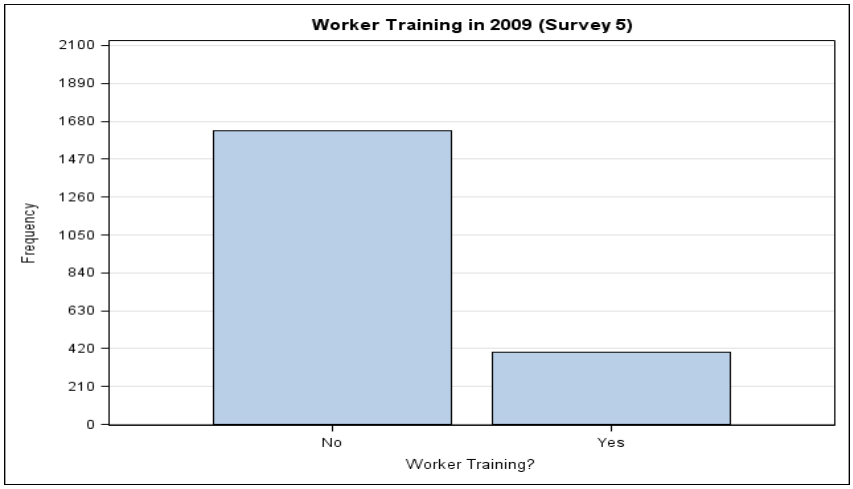


Figure 8: Number of Firms Doing Formal Training in 2010 Survey

Responses 2010: No = 1453, Yes = 349, Null Response = 643

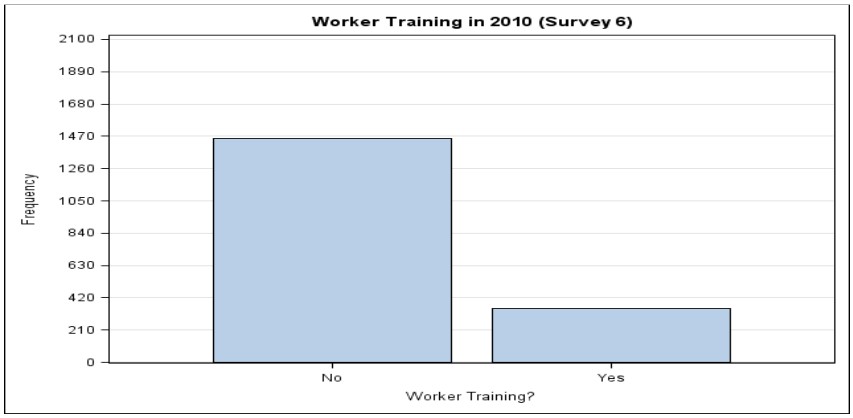
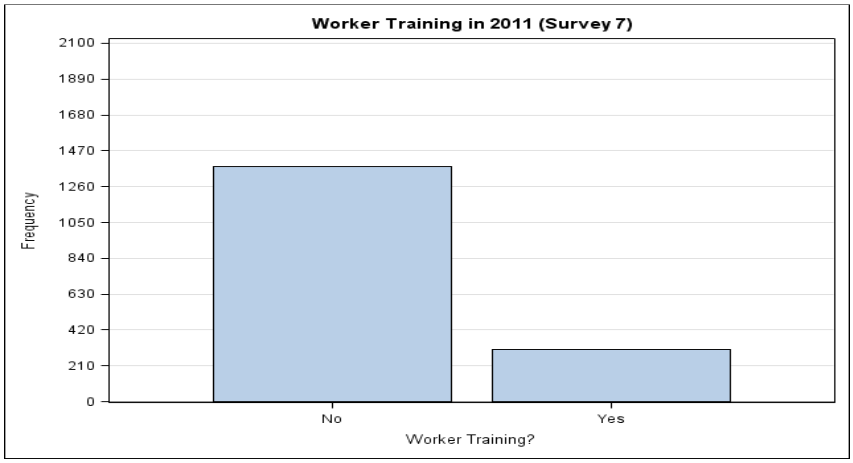


Figure 9: Number of Firms Doing Formal Training in 2011 Survey

Responses 2011: No = 1379, Yes = 307, Null Response = 759



A couple of items to note about these numbers: First, the growth in the number of null responses from Survey 4 to Survey 7 is reflective of the decline in overall survey responses for the full survey, and second, the percentage of firms indicating they expended money on worker training remains relatively fixed (20.36% - 18.21%) throughout the four surveys. Now that we have established the baseline of respondents for the dependent variable, let us look at which independent variables are significantly correlated to those firms doing formal worker training.

Table 1: Significance of Independent Variables to Worker Training

Independent Variable	p-value (Pr>F)	survey_variable* p-value (Pr>F)
num_emp	0.0020***	0.7733
sales_per_emp	0.1167	0.0157**
naics_first	0.1471	0.0444**
education	0.2693	0.0704*
assets	0.0144**	0.0219**
hr_score	<0.0001***	0.0558*

Notes
***: $p \leq 0.01$; **: $p \leq 0.05$; * $p \leq 0.1$.
survey_variable* = interaction of time

Table 1 shows us that there is significant correlation between formal training and: number of employees, assets, and HR score (this is calculated score based on the number of HR activities), when factored without the interaction of time. When adding the interaction of time there is significance with growth rate, industry sector, education, assets and HR score. We will look at each of these below.

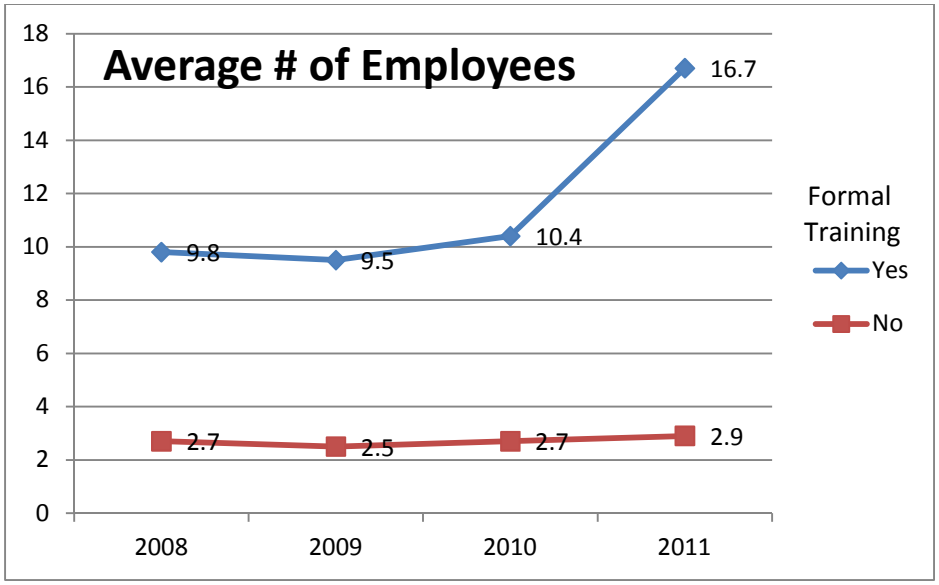
Independent variables:

1. Firm Size - Number of employees (#): (**c5_num_employees)**

   Literal Question -Not counting owner(s), on December 31, 20.., how many people worked for [NAME BUSINESS]? Please include all full- and part-time employees, but exclude contract workers who work for the business either full- or part-time but are not on the business' official payroll.

When looking at the relationship between the number of employees and the use of formal training, the results show there is a significant correlation (0.0020***) between the number of employees and the use of formal training without factoring the interaction of time. This tells us that firms with more employees are more likely to do formal training. This would support prior research that states that firm size positively relates to the use of formal training. The results also show that the interaction of time does not change the relationship between the number of employees and the use of formal training.

Figure 10: Average Number of Employees

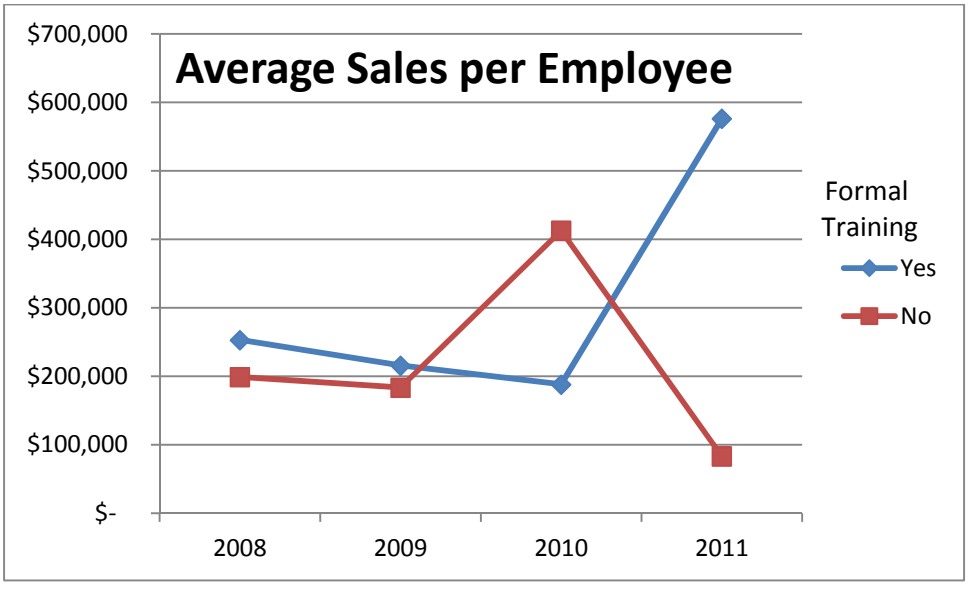


We can see from Figure 10 that the average size of firms doing formal training starts at 9.8 employees and grows to 16.7 over the four years. Firms not doing formal training are significantly smaller (2.7) and remain so throughout the four years of analysis.

2. Sales per Employee – Look at sales per employee to see if there is a difference between firms that do formal training versus those that do not. Sales per Employee = amount of revenue (**f16a_rev_200._amt)** / # of employees (**c5_num_employees).**

Sales per employee, as calculated by revenue divided by number of employees, was shown to have no significance without the interaction of time. However, when factoring the interaction of time, there is a significant correlation (0.0157**) between the use of formal training and a firm's productivity. This indicates a time latency in effect of formal training and a firm's productivity which would support the literature which talks about the need for longitudinal data to better understand the effect of formal training on firm performance over time. This also supports the emergent nature of *bottom-up* processes which have time latency.

Figure 11: Average Sales per Employee



Looking at Figure 11 we do not see any significant difference until 2011 when the average sales per employee for firms doing formal training have a significant increase. This again supports the time latency effect of bottom-up emergent processes.

3. Funding – There are several variables in the KFS data that reflect different types of funding. In looking at the amount of available data it was determined that was not enough data across the survey population to provide reliably significant results and so this analysis was dropped.

4. Industry Sector (NAICS code): (**c1a_naics_verification)**

Literal Question: Our records indicate the principal activity of the business was [D&B NAICS CODE DESCRIPTION]. Was that still the principal activity of the business as of December 31, 20..?

Factoring in a firm's NAICS code to identify if there were industry sectors more likely to use formal training was problematic. Using the full six digit code presented too much granularity to do the modeling for this research. However, it was decided to use the first digit of the code (1-9) to see what results this would reveal. Using this code there was no significance without the interaction of time. These was significance when factoring time (0.0444**). This would indicate that there is some moderation over time between the relationship of formal training and a firm's industry sector. It should be noted that NAICS code 5, which comprises multiple industries (Information, Finance, Insurance, Real estate, Professional technical and scientific services, Management companies, and Administrative support services) represented 46.38% of all firms answering this question and so this category was over sampled in the results. As

mentioned above, using the first digit of the NAICS code meant grouping several distinct industry sectors together so, while there was a significant correlation when factoring the interaction of time, there is not enough data in this analysis to identify particular industries that are more likely to use formal training. A list of two digit NAICS codes can be found in the appendices.

5. Owner education level: (**g9_education_owner_01**)

Literal Question: What is the highest level of education (you/[OWNER B-J]) (have/has) completed so far? Would you say . . .

This question was answered on a 1-10 scale as follows:

Value	Category
1	Less than 9th grade
2	Some high school, but no diploma
3	High school graduate (diploma or equivalent diploma
4	Technical, trade or vocational degree
5	Some college, but no degree
6	Associate's degree
7	Bachelor's degree
8	Some graduate school but no degree
9	Master's degree
10	Professional school or doctorate

The analysis showed no significant relationship between the owners(s) level of education and the use of formal training without factoring the interaction of time. There was significance (0.0704*) at the $p \leq 0.1$ level when factoring the interaction of time. This result is counter to research that suggests that an owner or manager's level of education positively relates to their use of formal training.

6. Geographic location: (**a8_addr_ver) –** This data point did not provide us with the level of analysis we had hoped for. Because of confidentiality concerns, KFS only asked

firms to identify there general geographic location by 1,2,3, 4, or 5 to represent census regions Northeast, Midwest, South, West, and unknown. It was determined by the researcher that there was little value in doing this analysis since there was not enough data to make it meaningful.

7. Asset level – There were eight asset variables that were totaled and then categorized as firms having <$100,000 in assets and those having ≥$100,000 in assets. The results show that assets are significantly related to a firm's use of formal training both without (0.0144**) and with (0.0219**) the interaction of time. In other words, the larger a firm's assets, the more likely there are to use formal training.

Figure 12: Average Asset Level

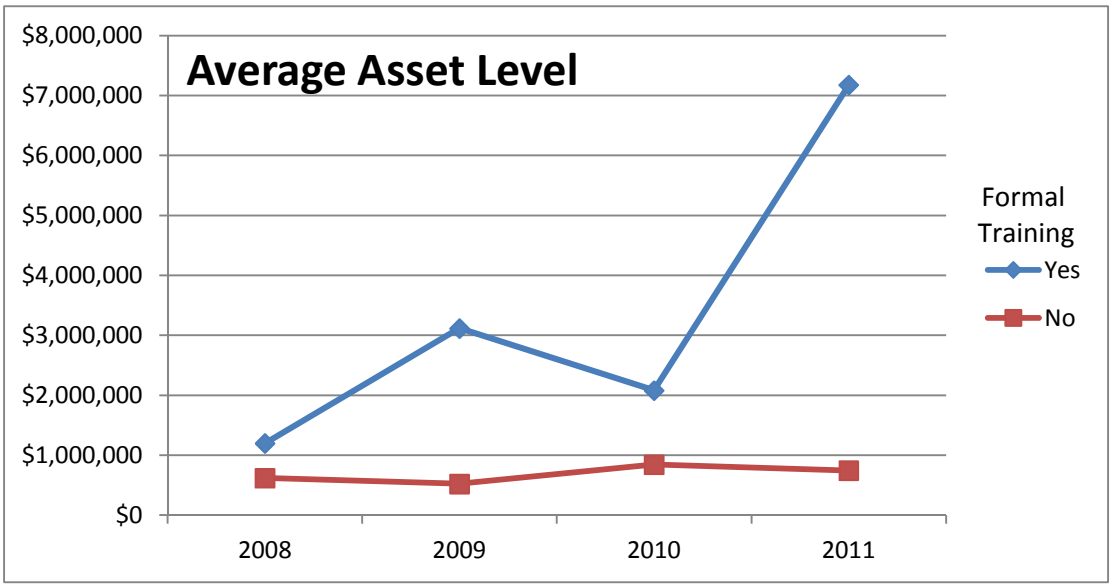
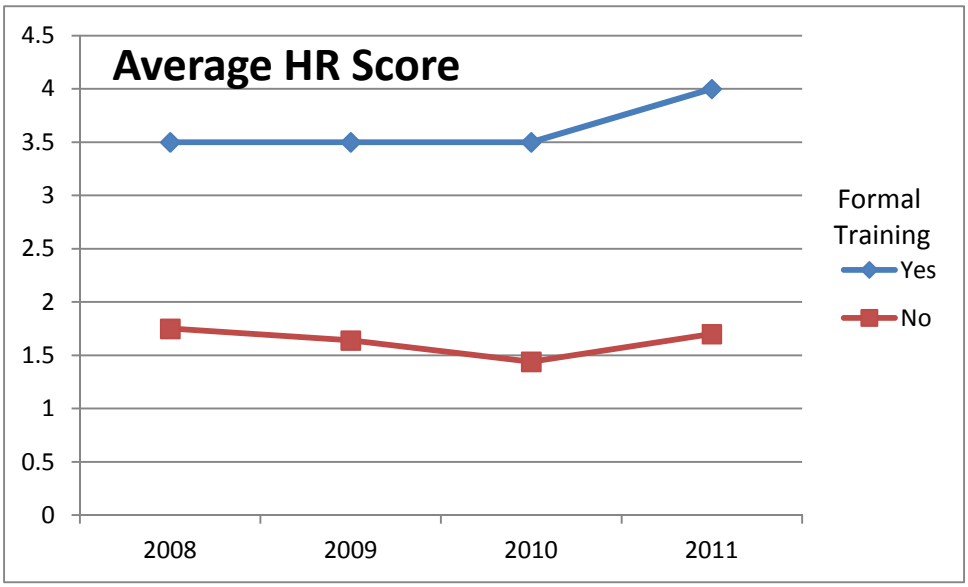


Figure 12 visually shows the significant difference between average asset levels of firms doing formal training and those that do not. This difference starts small in 2008 and continues to grow, with a significant increase in 2011. This parallels the 2011

increase we saw in average sale per employee, again supporting the time latent, emergent bottom-up nature of formal training initiatives.

8. Additional Human Resource activities – There were nine questions in KFS relating to additional human resource activities. Each firm's responses to these questions were combined to create a continuous variable of 0-8. The resulting analysis showed a high level of significance (<0.0001***) without factoring in the interaction of time. There was also significant correlation (0.0558*) with the interaction of time. This result strongly supports the configurational notion that various HR programs should be combined with others for optimal effect.

Figure 13: Average HR Score



As we can see with Figure 13, the average number of HR resources and programs for firms doing formal training starts and remains consistently higher than for firms not doing formal training.

The results for Q1 give us a sense of the characteristics of nascent small firms that are likely to do formal training. These characteristics include in order of significance:

1. HR Score – Firms with more HR resources and programs are more likely to do formal training.
2. Number of employees. The more employees a firm has the more likely they are to do formal training.
3. Assets – The more assets a firm has, the more likely they are to use formal training.
4. Sales per employee, Industry sector (NAICS), and Education Level of Owners, all have a significant correlation to formal training when factoring in the interaction of time.

Question 2

Are there significant variances in organizational performance levels for firms that do formal training versus those that do not? (See Question 3 for models and variables)

Based on the results from Q3 we know that there are significant variances in organizational performance metrics for firms doing formal training versus those that do not. These differences were significant in the following performance areas: profit, assets, sales per employee, profit margin, sales/capital and capital intensity. We've already seen, in Figure 10 and Figure 11, the differences in sales per employee and asset levels. The following figures illustrate differences in the other performance metrics.

Figure 14: Average Profit

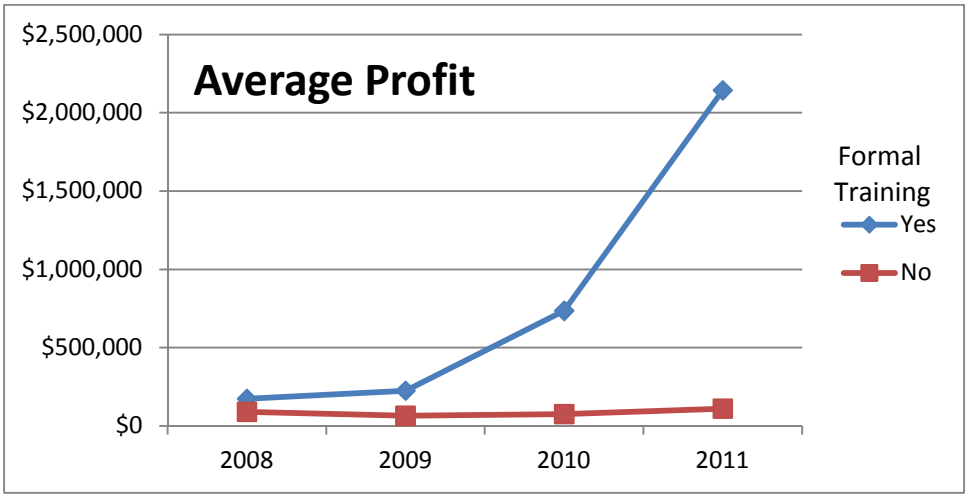
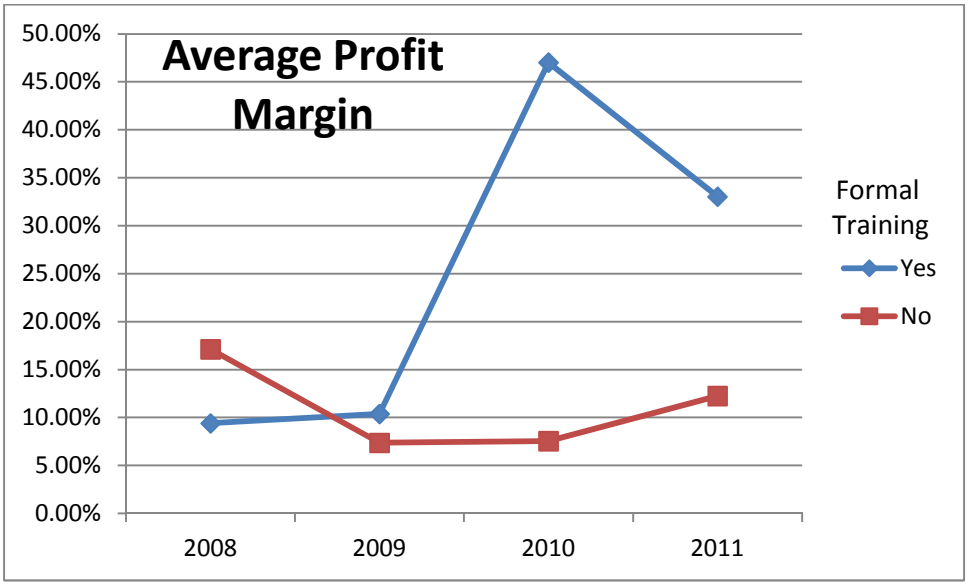


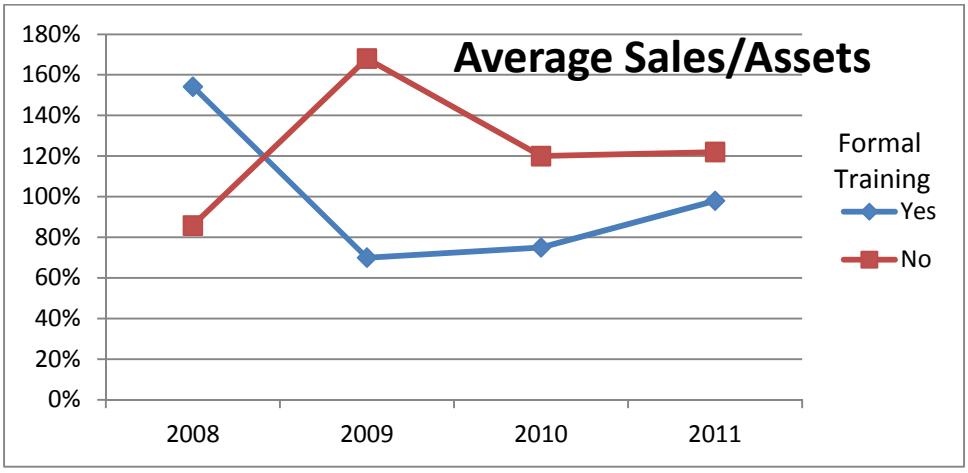
Figure 14 shows us the very significant difference between profit at firms that do formal training versus firms that do not. The difference increases over time.

Figure 15: Average Profit Margin



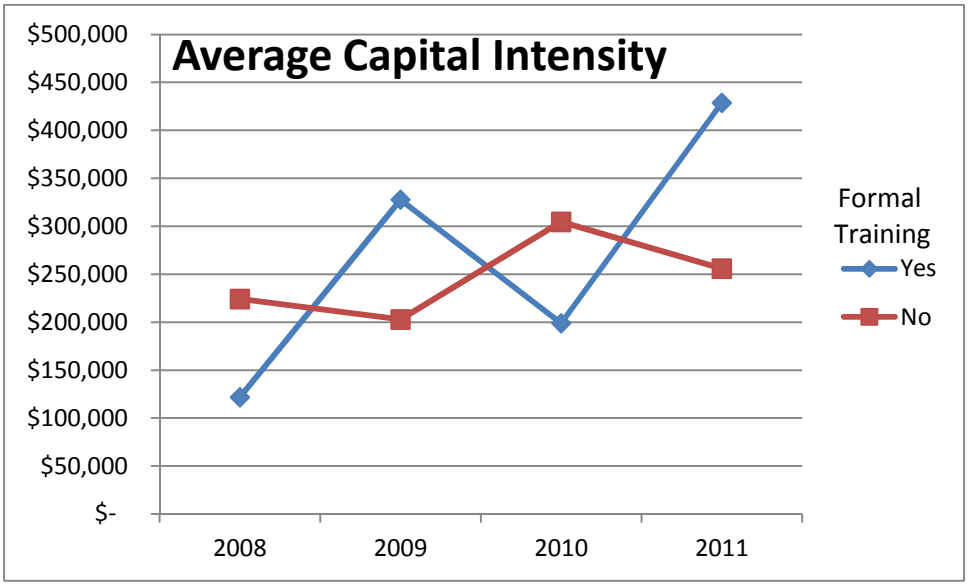
Profit margins begin at the same relative level but increases significantly at firms doing formal training. As we have seen with all of the data, these differences tend to increase over time.

Figure 16: Average Sales / Assets



Average sales divided by assets (total asset turnover ratio) provides a performance efficiency measure relating to the firms' efficient use of assets. While Figure 16 shows lower efficiency rations in years 2-4 for firms doing formal training that could be attributed to a number of factors including higher expenditures on employee development reducing assets in one year relative to revenue. What we see in 2010 and 2011 is an upward trend in efficiency for firms doing formal training, as revenues for those firms increase at a higher rate than firms not doing formal training.

Figure 17: Average Capital Intensity (Assets/# of Employees)



As with most of the charts we see in Figure 17 an increasing level of assets per employee in firms doing formal training. This indicates increased efficiency in use of employee resources.

Overall the results strongly show there are significant variances in organization level performance for firms that do formal training versus those that do not. This is especially true when factoring in the interaction of time as we see consistent increases in organizational performance in years 2010 and 2011.

Question 3

Is there a positive relationship between formal training and organization level performance?

Testing for this question required two different models of analysis. First a univariate analysis was done with each dependent variable. Then multivariate analyses were done with different combinations of dependent variables to test for overall impact

on wealth growth (profits and assets), organizational growth (sales growth, employee growth, and sales per employee), and over-all organizational performance (all five dependent variables).

Statistical Models – Testing for this question used Linear Mixed Modeling (LMM) to test for continuous response variables and Multivariate Linear Mixed Modeling (MLMM) to test for multiple continuous response variables. Following are the statistical equations and variables used for answering this question.

Linear Mixed Model (LMM)

$$Y_i = X_i\beta + Z_i\gamma_i + \epsilon_i$$
$$\gamma_i \sim N_{p_{random}}(0, \Psi)$$
$$\epsilon_i \sim N_{n_i}(0, \sigma^2\Lambda_i)$$

Multivariate Linear Mixed Model

$$[Y_i] = X_i[\beta] + Z_i[\gamma_i] + [\epsilon_i]$$
$$vec[\gamma_i] \sim N_{p_{random}}(0, \Psi)$$
$$\text{Each row of } [\epsilon_i] \sim N_{n_{responses}}(0, \Sigma)$$

Final Models:
survey_index is a categorical variable which is 4,5,6, or 7 stating where we are in time.
response_index is a categorical variable which is 1,2,3,4,5a,5b,5c,5d representing which response we have in the line of the data set.

<u>Univariate Models</u>
Sales growth: response= (r1)
Profits: response= (r2)
Employment Growth: response= (r3)
Assets: response= (r4)
Productivity (a): response= (r5a)
Productivity (b): response= (r5b)
Productivity (c): response= (r5c)
Productivity (d): response= (r5d)

Fixed formula-
response=response_index **+**
response_index*survey_index +
response_index*worker_training +
response_index*survey_index*worker_training

Random formula-
random intercept/subject=MPRID
repeated survey_index/subject=MPRID type=ar(1)

Bivariate Model
Profits and Assets: response= (r2, r4)
Fixed formula-
response=response_index +
response_index*survey_index +
response_index*worker_training +
response_index*survey_index*worker_training

Random formula-
repeated survey_index/subject=MPRID type=un group=response_index

Trivariate Models
Sales growth, Employment Growth and Productivity (a): response= (r1, r3, r5a)
Sales growth, Employment Growth and Productivity (b): response= (r1, r3, r5b)
Sales growth, Employment Growth and Productivity (c): response= (r1, r3, r5c)
Sales growth, Employment Growth and Productivity (d): response= (r1, r3, r5d)

Fixed formula-
response=response_index +
response_index*survey_index +
response_index*worker_training +
response_index*survey_index*worker_training

Random formula-
repeated survey_index/subject=MPRID group=response_index

Pentavariate Models
Sales growth, Profits, Employment Growth,Assets and Productivity (a): response= (r1,r2, r3,r4, r5a)
Sales growth, Profits, Employment Growth,Assets and Productivity (b): response= (r1,r2, r3,r4, r5b)
Sales growth, Profits, Employment Growth,Assets and Productivity (c): response= (r1,r2, r3,r4, r5c)

Sales growth, Profits, Employment Growth,Assets and Productivity (d): response= (r1,r2, r3,r4, r5d)

Fixed formula-
response=response_index +
response_index*survey_index +
response_index*worker_training +
response_index*survey_index*worker_training

Random formula-
repeated survey_index/subject=MPRID group=response_index

The results are shown in Table 2.

Table 2: Significance of Formal Training on Organizational Performance

FIXED COEFFICIENTS								
p-values Univariate Response	Sales Growth	Profit	Employment Growth	Assets	Sales/ Employee	Profit Margin	Sales/ Capital	Capital Intensity
worker_training	0.9316	**0.0033*** **	**0.0031*** **	**<0.0001*** **	0.297	**0.0032** **	0.733	0.4956
survey_index*worker_training	0.5534	**0.0119*** **	0.3699	**0.0008*** **	0.2617	0.7703	0.7207	0.8456
p-values Bivariate Response (2 and 4)		Profit		Assets				
response_index*worker_training		**<0.0001*** **						
response_index*survey_index*worker_training		**0.0002*** **						
p-values Trivariate Response (1,3, and 5)	Sales Growth		Employment Growth		Sales/ Employee	Profit Margin	Sales/ Capital	Capital Intensity
response_index*worker_training					**0.015* **	**<0.0001*** **	**0.02** **	**0.0163* **
response_index*survey_index*worker_training					0.2098	0.4762	0.395	0.4521
p-values Pentavariate Response (1,2,3,4, and 5)	Sales Growth	Profit	Employment Growth	Assets	Sales/ Employee	Profit Margin	Sales/ Capital	Capital Intensity
response_index*worker_training					**<0.0001*** **	**<0.0001*** **	**<0.0001*** **	**<0.0001*** **
survey_index*worker_training					**0.0006*** **	**0.0019*** **	**0.0014*** **	**0.0017*** **

Notes
***: $p \leq 0.01$; **: $p \leq 0.05$; * $p \leq 0.1$.
survey_index* = interaction of time
Results for “response=response_index” and “response_index*survey_index” have been left off this table as they are needed for the model but are not relevant to the discussion.

Dependent variables: Organization level performance metrics (Since the independent variable used in this test is only available starting with the 4th follow-up survey in 2008, we used data from the 4-7th follow-up surveys for these tests).

1. Sales growth: Do firms that use formal training have higher revenue growth rates than those that do not? Total revenue for each year: (**f16a_rev_200._amt**)

The results from Table 2 show that formal training does not have any significant correlation to increased sale growth. The interaction of time does not change this. These results do not show a positive relationship.

2. Profit: Do firms that use formal training have higher profit margins than those that do not? Profit amount for year: (**f24_profit_amt_4**)

The results from Table 2 show a very significant correlation in the univariate analysis between formal training and profit (0.0033***), and when factoring the interaction of time (0.0119***). We see even a higher level of significance in the bivariate analysis when adding the “asset” variable. These results show a significant positive relationship.

3. Employment growth: Do firms that use formal training have higher employment growth rates than firms that do not? Total number of employees: (**c5_num_employees**)

The results show a very significant correlation (0.0031***) between formal training and employment growth without the interaction of time, but no correlation (0.3699) with the interaction of time. This result suggests that employment growth is an antecedent to formal training rather than a result of formal training.

4. Assets: Do firms that use formal training have higher levels of asset growth than firms that do not? This would use the Total Assets variable calculated in Q1.

The results show a significant correlation between formal training and a firm's asset level (<0.0001***) which could indicate that firms with greater assets are more likely to do formal training. However, the significant correlation with the interaction of time (0.0008***) indicates that formal training also leads to increased assets. Both results strongly support a positive relationship.

5. Productivity: Do firms that use formal training have higher level of productivity than those that do not? There are several productivity measures than can be calculated from KFS data.

a. Sales per employee = revenue/# of employees:

(**f16a_rev_200._amt / c5_num_employees**)

The univariate results do not show a correlation between formal training and increased sales per employee. This is not surprising given the lack of significance between formal training and sales growth, and the lack of significant correlation between formal training and employee growth when factoring the interaction of time.

The trivariate results, which included sales growth and employment growth variables, showed a significant correlation without the interaction of time but no significance with the interaction of time. This indicates that sales growth and employee growth have more to do with sales per employee than formal training.

The pentavariate results, (adding the sales growth, profit, employment growth, and asset variables to the model) show very significant results both independent of time

and with the interaction of time. This indicates that adding profit and assets has a significant impact on sales per employee. Given the strong significant correlation between formal training and profit and assets, we can deduce that formal training does have an impact on sales per employee through its relationship with profit and assets, therefore supporting a positive, though indirect, relationship. In addition, the strong level of significance in this analysis when including the interaction of time indicates that there is an indirect cause-effect relationship.

b. Profit margin = operation profit/revenue:

(**f24_profit_amt_4 / f16a_rev_200._amt)**

The results from the univariate analysis show a significant correlation between formal training and profit margin (0.0032**) without the interaction of time, and no significant correlation with the interaction of time. This indicates that profit margin might be antecedent to formal training rather than a result of formal training.

Factoring in sales growth and employment growth in the trvariate analysis shows an even more significant correlation (<0.0001***) without the interaction of time and again no significant correlation with the interaction of time. This suggests that sales growth and employee growth have more significance to profit margin than formal training.

However, like sales per employee above, when we add profit and assets to the analysis in the pentavariate model, we see high levels of correlation for both levels of analysis. And again, because we have a very significant correlation between formal

training and profit and assets, these results indicate that formal training has an indirect impact on profit margin.

c. Sales per $ of capital = revenue/assets: (**f16a_rev_200._amt**)/Total Assets

The results from the univariate analysis show no significant correlation with or without the interaction of time. When we add the variables of sales growth and employment growth in the trivariate analysis, we see a significant correlation (0.02**) without the interaction of time. This is most likely the result adding the employment growth variable which showed a significant correlation in the univariate to formal training.

In the pentavariate analysis, we have significant correlations both with (0.0014***) and without (<0.0001***) the interaction of time, for the same reasons as mentioned with the other performance metrics.

d. Capital intensity = assets/# of employees: Total Assets/ (**c5_num_employees**)

The results from the univariate analysis show no significant correlation with or without the interaction of time. When we add the variables of sales growth and employment growth in the trivariate analysis, we see a significant correlation (0.0163*) without the interaction of time. This is most likely the result adding the employment growth variable which showed a significant correlation in the univariate analysis to formal training.

In the pentavariate analysis, we have significant correlations both with (0.0014***), and without (<0.0001***) the interaction of time, for the same reasons as mentioned with the other performance metrics.

The results show strong positive relationships between formal training and several organizational performance metrics. The most significant direct correlations between formal training and organizational performance are in profit and assets. These significant correlations exist with or without the interaction of time indicating that they are both antecedent to formal training and a result of formal training.

When we run the analysis including all performance metrics to see results on calculated productivity measures we see significant correlations with and without the interaction of time. This indicates that formal training, through its direct impact on profit and assets, has an indirect impact on these productivity metrics.

Question 4

Are there variables that moderate the effect that formal training has on organization level performance?

Testing for this question used the same MLMM process used in Q3. What we were testing for here was to see if any of the independent variables identified in Q1 as having significance to formal training have moderating effects on performance variables identified in Q3.

Dependent variables: Performance metrics identified in testing Q3 that showed a positive relationship to formal training. These are: Profit, Employment Growth, Assets, and Profit Margin.

Independent variables: Variables identified in testing Q1 that showed a positive relationship to the use of formal training. These are: Number of Employees, Sales per Employee, NAICS, Owner's Education Level, Assets, and HR Score.

Table 3: Moderating Effect of Variables on Organizational Performance

Independent Variable	Profit p-value (Pr>F)	Emp. Growth p-value (Pr>F)	Assets p-value (Pr>F)	Profit Margin p-value (Pr>F)
num_emp	<0.0001***	<0.0049***	0.0059***	0.1549
sales_per_emp	<0.0001***	<0.0001***	0.0889*	<0.0001***
naics_first	0.8920	0.0391**	0.5832	0.6609
owner_ed	0.5341	0.7662	0.6556	0.5143
assets	0.0035**	0.2304	n/a	0.9261
hr_score	0.5195	0.0184**	0.1545	0.7490
survey_index* num_employee	0.5195	0.7250	<0.0001***	0.0254**
survey_index* naics_first	0.1216	0.0171**	0.1842	0.8300
survey_index* owner_ed	0.8999	0.2371	0.1936	0.9648
survey_index* sale_per_empl	0.7636	0.4991	0.0809*	0.1009
survey_index* assets	0.5650	0.7842	n/a	0.7488

Notes
***: $p \leq 0.01$; **: $p \leq 0.05$; * $p \leq 0.1$.
survey_index* = interaction of time

Table 3 shows us that several of the variables identified in Q1 as having a positive relationship to an organization's use of formal training also have positive relationships to an organization's performance metrics identified in Q3. Let us look at each of the performance metrics and the effects that the independent variables have on them.

1. Profit – As identified in Q3, formal training has a very significant positive relationship to an organization's profit. This test shows us that a firm's number of

employees, their productivity as measured by “sales per employee”, and asset level all have a significant positive moderating effects on the impact of formal training on a firm’s profit independent of the interaction of time. There appears to be no moderating effects by a firm’s location, the owner(s) education level, their composite HR score, or any of the variables when including the interaction of time.

2. Employee Growth – Q3 showed us that formal training has a significant correlation to a firm’s employee growth without factoring the interaction of time. The Q4 analysis shows that additional variables which moderate that effect include number of employees, sales per employee, HR score, and a firm’s location. Location also has an effect when factoring the interaction of time.
3. Assets – Formal training has a very significant correlation to assets with or without the interaction of time. This analysis shows that the number of employees and sales per employee, both with and without the interaction of time, have moderating effects on a firm’s asset level.
4. Profit Margin – Sales per employee without the interaction of time and number of employee with the interaction of time are the only two other variables shown to have moderating effects on profit margin.

The test for Q4 shows that there are additional independent variables that have a moderating effect on the organizational performance metrics identified in Q3. Number of employees and sales per employee appear to have the strongest moderating impact. Firm location, assets, and HR score also have some moderating effect.

Summary

The results from this analysis show that nascent small firms doing formal training have more employees, more assets, and more HR resource and programs than firms not doing formal training. These differences are significant given the relative size of the firms. We also see that these differences increase over time. When factoring the interaction of time we also see differences with sales per employee, firm location and owner's education level.

We also found strong results for Q2 and Q3. Formal training has a significant impact on a firm's profit, employee growth, assets and profit margin without the interaction of time, and a significant impact on profit and assets when factoring the interaction of time. We also found a significant indirect impact of formal training on a firm's calculated productivity metrics when factoring the additional variables of sales growth, profit, employment growth and assets.

And finally, we found that the number of employees and sales per employee have significant moderating effects on a firm's performance metrics in addition to formal training. Assets and HR score also have some effect.

Now that we know what organizational characteristics indicate a firm's use of formal training, the impact formal training has on the firm's performance, and other independent variables that have moderating effects on that performance, we can move onto the discussion of our findings' significance.

CHAPTER V

DISCUSSION OF FINDINGS

Introduction

We began this research by pointing out the important role that small businesses play in our economy, and the high level of failure rates of those businesses. Increasing the success rate for small businesses would contribute to the overall well-being of local, national and international economies.

Formal training has been shown to improve individual and team level performance, and there is evidence that formal training also leads to increased organizational performance (Aguinis & Kraiger, 2009; Aragon-Sanchez, et al., 2003; Lawler, Mohrman, & Lawford, 1998). Most studies in this area have used large businesses for their analysis and there are questions whether small business would even use formal training and if that training would have the same effects as that in large businesses (Hill, 2004; Kitching, 2008; Stewart & Beaver, 2004).

There are several reasons why we do not have good evidence that formal training has a positive impact on organization level success in small businesses. First, there has been a lack of theoretical frameworks for showing how changes in behavior at the individual level of an organization can lead to change at the organizational level (Tharenou et al., 2007; Wright & McMahan, 1992). Second, given such a theoretical framework, we need good statistical models which can estimate correlations across multiple levels of analysis (Katou, 2009). And third, we need access to longitudinal data from small businesses that allows us to do multilevel analysis (Crook et. al, 2011;

d'Arcimoles, 1997). This research was designed to address those issues and answer the following questions.

Research Questions

1. What are the characteristics of small businesses that use formal training for employee development?
2. What impact over time does formal training have on the organization-level performance of small businesses?
3. Are there characteristics of small businesses that moderate the effect that training has on organizational performance?

The hope was that by answering these questions we could provide evidence that formal training can lead to improved performance and success for small businesses. The following discussion talks about the results of our research; what that might mean for human resource development researchers, practitioners, and small business owners; and future directions for research and practice in this area.

Discussion of Results

What are the characteristics of small businesses that use formal training for employee development?

The goal in answering our first research question was to give us a better understanding of the types of small businesses that are using formal training in their operations. A better understanding of what types of businesses do formal training will give us insight into what types of formal training we should be developing and who to

target those programs to. This would be especially beneficial to the many government and other publicly developed and funded training efforts targeting small businesses.

The results of our analysis provide us with valuable insights into the characteristics of small businesses that have successfully used formal training to increase their performance.

"Small" is a Relative Term

Defining what exactly is meant by the term "small business" is really relative to what one is trying to understand. Most practical and research definitions use various metrics of number of employees, revenue, and assets to define a small business (Longenecker, Petty, Palich, & Hoy, 2013). The U.S. Small Business Administration defines a small manufacturing business as one having less than 500-1000 employees depending on their specific industry, and non-manufacturing businesses as having less than $7M in revenue. Other countries and regions, such as Australia (15 employees) and the EU (50 employees) have much lower size thresholds (OECD, 2013). However, there really are no set standards in either practice or research as to what qualifies as a small business.

All of the firms in the Kaufmann Firm Survey were started in 2004, the first year of the survey. As such, they all started in the same relative position regarding number of employees and assets. Of course things changed as some firms received outside investment and others saw rapid increases in revenue growth. As a result, there were significant statistical differences in the size of firms in regards to number of employees, level of assets, and revenue by 2007, the first year of our analysis.

In the 2007 survey results, the average number of employees in firms not doing formal training was 2.7 fulltime equivalents (FTE). This would certainly qualify these firms as small businesses. For firms that were doing formal training in the 2007 survey, the average number of FTEs was 9.8, three times the size of those firms not doing formal training. However, 10 FTEs would still qualify as a "small" business under most any definition of small business (Longenecker, et al., 2013; Scott & Bruce, 1987).

While the difference between 3 employees and 10 might not seem significant, anyone who has started and run a small business will tell you the difference between having 3 employees and having 10 employees is both significant and substantial. With just 3 employees many businesses will not have a lot of formal processes or practices in place. Often the employees of firms this size will be made up entirely of founders (Scott & Bruce, 1987).

With 10 employees you start to have structures in place to facilitate communication, and help with understanding of job responsibilities. HR policies and programs are developed to attract, retain and develop talent. And it is possible there are multiple levels to the organizational structure. The point is that, while a 3 person firm and a 10 person firm are both "small" businesses, the structural and operational differences between them are significant and substantial (Ashton, Sung, & Raddon, 2005; Barrett & Mason, 2007).

In the same way that there is a significant difference between a firm with 3 employees and one with 10 employees, there is a difference between a firm with $600k in assets (the average for those not doing formal training in 2007) and one with $1.2M in

assets (the average for those doing formal training in 2007). After human capital, having access to financial capital is the most important resource for a small business (Chittenden, Hall, & Hutchison, 1996). Again, both asset levels would easily qualify as a "small" business. But having twice the level of asset resources, which in a small business would likely be in the form of operating capital, receivables, or hard assets, is a significant and substantial difference.

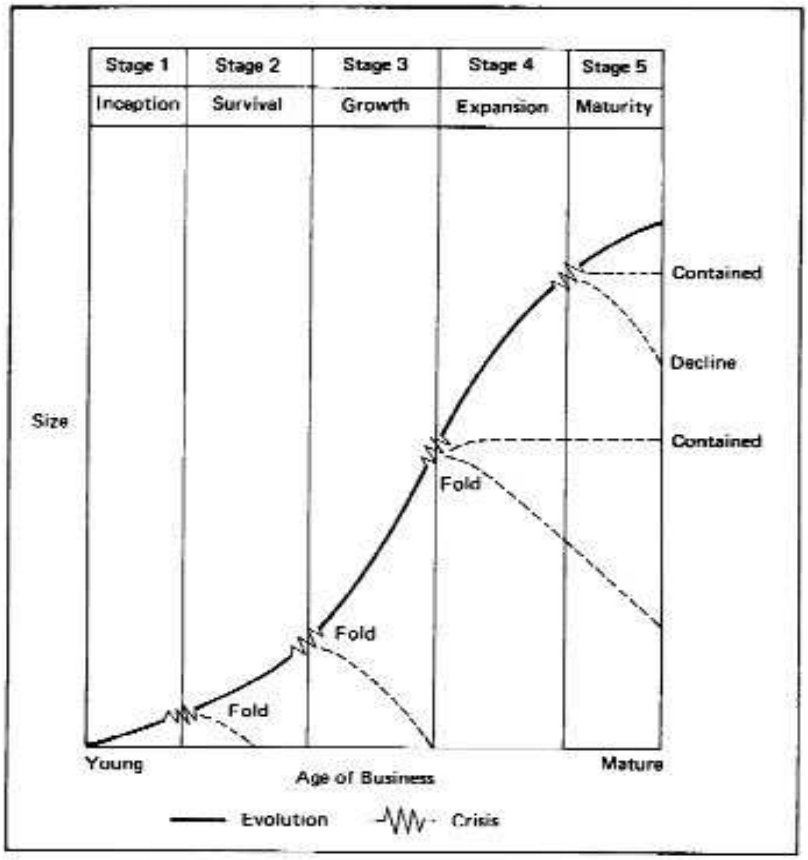
Higher asset levels allows a small business to make more investments in things like sales and marketing efforts to increase revenue, and employee development programs such a formal training. We see this in our analysis of the calculated HR Score were those firms doing formal training, with higher asset levels, are also doing multiple HR programs.

Because numeric differentiations among and between small businesses do not explain why there can be significantly different needs between firms of the same relative size, researchers have developed organizational, structural, and contextual models to help us better understand small business. Scott and Bruce (1987, p. 45-46), used a definition from the American Committee for Economic Development to define small business as follows:

1. Management is independent. Usually the managers are also owners.
2. Capital is supplied and ownership is held by an individual or small group.
3. Area of operations is mainly local. Workers and owners in one home community, but markets need not be local.

Another way to understand small business is by their stage of growth (Scott & Bruce, 1987). While size and age (young-mature) of the business are still defining variables, for Scott & Bruce, the main foci of the *stages of growth model* are specific contextual "crises" that most businesses face as they move from one stage of growth to another. This model can help both researchers and practitioners to better understand the specific issues and what HRD or OD interventions can help a small business as it grows. Scott and Bruce's model, derived from Greiner's five stages of growth model (Greiner, 1972), is shown below.

Figure 18: A Lifecycle Model for Small Business Growth (Scott and Bruce, 1987)



In this model a firm's size and age, relative to their stage of growth, will vary from business to business. Various internal and external factors also impact a firm's growth path. Subsequent researchers have developed growth models to focus on different aspects of the organization. Mount, Zinger & Forsyth (1993), focus their *phase of small business development* on changes in an organization's structure and processes as it transitions from on stage to another. Ardichvili, Harmon, Cardozo, Reynolds, and Williams (1998) discuss three phases of growth from an HRD perspective relative to changes in delegation of authority within a growing business.

All of these models contain certain points in a firm's growth where a change is needed within the organization. Scott and Bruce (1987) talk about "crisis points" that a small business must pass through in order to move to the next stage of growth. For Mount et al. (1993) businesses go through "transition phases" which are times of potential instability, as a business' leadership or management structure changes. In a similar way, Ardichivili et al. (1998) discuss changes in "delegation of functions" as points of transition in an organization.

Each of these "transition points" in a business requires new levels of knowledge, skills and abilities in employees, and new organizational processes and procedures in order for the business to move to the next stage of growth. These are the times in an organization's development where training and development initiatives can be beneficial in helping the company move onto its next phase of growth.

If we look at the numeric differences of firms doing formal training and those not doing such, we could hypothesize that a small business moving from 3 employees to 9

employees, might be going through one of these critical “transition points” where the use of formal training is beneficial to increasing their performance and accelerate growth. Of course testing this would require additional research beyond the scope of this paper and beyond the scope of the KFS data.

What’s clear from these various models is that small business organizations are complex systems that operate in larger complex cultural and economic systems. Each organization has its own internal and external factors which impact its operation and growth. While we can rightly say that all of the businesses in the KFS data fall under the general term of “small business”, there are numerous operational, structural, cultural and economic differences, beyond the numbers, that separate one small business from another and therefore affect both its use of formal training and the impact that training has on their performance.

This complexity presents significant challenges for small business owners, researchers, and HRD and OD practitioners when looking for ways to increase performance within a small business. While the results from our research indicates that formal training can and does result in increased organizational performance in small businesses, there are many other internal and external factors which must be taken into account when deciding whether formal training is an appropriate strategy for a particular small business. This is where the Resource Based View of the firm can help us to better understand how some of those other factors impact a small firm’s growth and performance.

Our discussion of the Resource Based View of the firm, and subsequent discussion of the contingency and configurational perspectives, tells us that having multiple human resource policies and programs that are aligned with the overall business strategy, will provide the optimal effect of those HR initiatives (Delery & Doty, 1996). While we do not have the data in KFS needed to analyze this from the contingency perspective, we do know that firms doing formal training had a significantly higher number of HR programs, supporting the configurational perspective that more HR programs will result in increased performance impact. This would indicate that small businesses that have already made investments in other HR activities are more likely to use and benefit from additional HR activities such as formal training. One must also keep in mind that, from a contingency perspective, all of those HR activities should be aligned with the firm's overall strategy for maximum benefit.

Other characteristics that prior research tells us we should look to when thinking about formal training did not provide the kind of results that we had anticipated. In some cases, such as business location, this was because of limits in the available data. We had hoped to see if there were specific geographic areas with higher rates of formal training but the location data provided in KFS was too general for any significant analysis.

We had also hoped to look at industry sectors, using NAICS codes, to determine if there are particular industries that are more likely to use formal training. However, the statistical models we used to do the multi-level analysis could not accommodate the large number of industry codes needed to do this analysis. There are likely other statistical methods that could be used to do this analysis. Being able to identify particular industries

where formal training is most likely to be used would be helpful in developing and marketing training specific to those industries.

And finally, we looked at the education level of the owners because prior research suggests that the higher an owner's level of education, the more likely they are to do formal training. In our analysis we did not see any significant correlation between the owner(s) level of education and the use of formal training, except for a low level of significance ($p \leq 0.01$) when factoring the interaction of time. These results are inconclusive about the effect of owner education on the use of formal training.

What we know from our analysis is that small businesses that do formal training are still very much "small" businesses. However, within the realm of small businesses, those that do formal training have three times the number of employees, they have twice the level of assets, and they have significantly more HR policies and programs in place.

From a research perspective, these results tell us that small businesses, doing formal training, have many of the same characteristics of larger firms, just at a relatively smaller scale. This challenges current thinking about small businesses not having enough resources or formal structures to facilitate the use of formal training, and then leverage that training into increased organizational performance (Cardon & Stevens, 2004; Kitching & Blackburn, 2002; Storey, 2004; Storey & Westhead, 1997). We know there are many small firms (about 20% in or data) doing formal training, and we also know that those businesses are seeing significantly increased organizational performance as a result.

From the HRD/OD practitioner, these results can help identify those small businesses that are large enough, from an employee, asset and HR perspective, to adopt

and benefit from formal training. Then, a simple survey or questionnaire could be used to understand other aspects of the business, such as their current growth or transition point, to help determine the correct training or development intervention.

For the small business owner, this information is helpful for them to gauge when their business reaches a point in its growth where formal training might be doable and beneficial in increasing their business performance. Now let us discuss the impact of formal training on organizational performance.

What impact over time does formal training have on the organization-level performance of small businesses?

A stated goal of this research was to show the significant positive impact that formal training has on organizational performance. By showing this we could provide evidence that doing formal training is one way small businesses can improve their performance and therefore their success rates.

The results of our analysis show that those small businesses doing formal training had significantly higher levels of organizational performance than those that did not do formal training.

Profit is the Difference

As the founder or co-founder of six small businesses, the primary researcher knows from experience that the single most important financial performance metric to the success of a small business is profit. Current small business research supports this idea (Gorgievski, Ascalon, & Stephan, 2011; Walker & Brown, 2004). In addition, Scott and Bruce (1987) discuss that at the early “Inception” stage of a small business, profitability

is the most likely crisis point for the business. Without profit, the small business will not be able to move to its next stage of growth.

While some research, and the media, might have us believe that large outside investments and high growth are what all small businesses are focused on, the reality is that very few businesses receive the outside funding that is typically needed to fuel high growth rates. In fact in our data, which was drawn from over 75,000 businesses that started in 2004, there was not enough outside investment data to provide a large enough sample for analysis. The Wall Street Journal estimated that in 2004 about 2,400 businesses received venture funding, or about 3% of all businesses started that year. We point this out because of what we believe to be an inordinate focus in small business research in two areas: founder characteristics and funding. The usual reason why this is the case is accessible research populations, but the fact that venture funded firms constitute about 3% of small businesses creates a false impression of what is small business.

What we see in our research are the results of actual internal business practices and their effect on business performance. In our case, the business practice is formal training and the result is profit. And the difference in profit is very significant.

Our analysis of the effect of formal training on profit showed a significant correlation (0.0033***) without the interaction of time. This tells us that there is a relationship, but not which direction that relationship flows. When we factor in the interaction of time, through multilevel analysis, we still see a significant level of correlation (0.0119***). But, because we have the benefit of observing that relationship

over time, we now see that formal training has a direct impact on a firm's profit over time. This, more than any other result of this research, is significant. If this research produced no other result than this we would consider the work a success.

If our goal as small business HRD researchers and practitioners is to help increase the success rate of those businesses, focusing on profit maximization is one clear way to do this (Barkham; Gudgin, Hart, & Hanvey, 1996; Hall & Fulshaw, 1993; Ibrahim & Goodwin, 1986; Jarvis, Curran, Kitching, & Lightfoot, 2000; Kelmer, 1994; Walker & Brown, 2004). When we factor this with our finding that formal training had a significant, positive, and direct impact on the profit levels of those firms, it stands to reason that formal training provides a proven way to increase the success rate of small businesses.

The other performance metric that had a high level of significance to formal training is assets. We already stated in discussing the first question that higher asset levels is one of the characteristics that differentiates firms doing formal training from others. So we know that higher asset levels are antecedent to doing formal training. When we look at the direct impact of formal training on assets over time, we also see that formal training has a direct positive impact on asset levels. So it is a two way relationship. Having assets helps a small business have the necessary resources to do formal training and doing formal training leads to increased assets.

In fact, the effect over time is dramatic. In the first year of our analysis, 2007, the difference between the average assets levels of firms doing formal training and those not

was approximately 2 to 1 ($1.2M to $620K). By 2011, the final year of analysis, the difference had grown to almost 10 to 1 ($7.2M to $750K).

A similar look at profits shows an even larger growth disparity over the four years of analysis. In 2004, the average profit for firms doing formal training was $174K, while the average for those not doing formal training was $91K. In 2011, that difference had increased to over 19 to 1, $2.14M to $111K.

As we saw when doing the analysis for Q4, there are other variables that can moderate these increases of profit and assets, but the direct level of correlation to formal training over time is positive and significant. We also found, in doing the univariate analysis, significant levels of correlation for employment growth and profit margin without the interaction of time. This would suggest that, while these variables are related to formal training, they are antecedent to the use of formal training.

When we combined all of the direct performance metrics (sales growth, profit, employment growth, and assets) with formal training to look at their impact on the calculated productivity metrics (sales per employee, profit margin, sales/capital, and capital intensity), we saw significant correlations both with and without the interaction of time. We did not do any further analysis to look at these particular results. However, when we compare the results of this analysis, which included profit and assets, with those from the trivariate analysis that only factored in sales and employment growth, we can deduce that profit and assets were the main contributors to the increased levels of correlation in the pentivariate analysis. And, with the addition of profit and assets in the analysis, we now see significant correlations over time suggesting a causal flow.

Since we have already seen that formal training has a direct and significant impact on an organization's profit and assets, and we now see that profit and assets have a significant impact on an organization's productivity metrics, we can deduce that formal training has an indirect, but significant impact on an organization's overall productivity.

This would fit with Katou's (2009) causal pathway we discussed in Chapter II.

Resourcing → Development → Skills → Attitudes → Behavior → Performance

The premise in Katou's pathway is that by first having HR practices that help insure the hiring of people best suited for the job (resourcing), and then strategically developing the knowledge, skills, and attitudes of those employees, they will behave and perform at a higher level. This equates to increased productivity and ultimately leads to increased organizational performance.

So in answer to this research question the data provides evidence that formal training in nascent small business has a significant, positive direct and indirect impact over time to a firm's profit, assets, and productivity. We will now discuss the final research question.

Are there characteristics of small businesses that moderate the effect that training has on organizational performance?

The configurational perspective of the organization suggests that no two variables in an organization have a relationship that is totally independent of influence from other variables (Delaney & Huselid, 1996). Systems theory provides the basis for this understanding (Brinkerhoff & Gill, 1994; Kauffman, Jr., 1980). Therefore, we decided to

see what other variables might have an impact on the effect of formal training on organizational performance. There were a few significant results.

An organization's number of employees, and subsequently their sales per employee, both had significant impacts on profit, employee growth and assets. Sales per employee also had a significant correlation with profit margin. On the face of it, it is not surprising that number of employees has an impact on a firm's profit and assets. This is especially the case in a small business where employee wages and benefits are typically the single largest expense category and often the most capital intensive part of the business (Brock & Evans, 1989).

When factoring in the interaction of time, the number of employees provides the only significant correlation of importance with regards to assets and profit margin. This would indicate that, in addition to formal training, the number of employees has the most significant impact on increases in assets and profit margin. This would align with the resource-based view of the firm which considers human resources to be the most valuable resource for sustained competitive advantage (Barney, 1991; OECD, 2013; Tuan & Ngoc, 2012; Wright & McMahan, 1992). Therefore increases in human resources, when aligned with properly configured HRD programs such as formal training, will result in better organizational performance. And, as we have discussed, these effects would be amplified in a small business where employees have a larger direct contribution to organizational expenses and performance than in large businesses.

Again, it is important to note that our analysis did not take into account such things as owner motivation and orientation, business strategy, industry sector,

competitive landscape, access to capital, and other internal and external factors that have been shown to moderate a firm's profit and assets (Walker & Brown, 2004). While these limitations do not reduce the significance of our findings, they do point out that there are other factors which must be taken into account when assessing the need for and type of training to initiate in a small business.

Our research has addressed several important areas. First we discussed the Resource Based View of the Firm (RBV), and the contingency and configurational perspectives, as a theoretical framework that gives us a way to look at the causal impact of formal training on organizational performance. Then we identified the statistical models in hierarchical linear modeling (HLM) which provided us with a way to show correlations of variables at multiple levels of analysis, including the interaction of time. And finally, we have the needed longitudinal data set (KFS) which provided us with the kind of data needed to do the analysis. The results are clear and significant.

Formal training has a direct and significant impact of the profit and asset levels of small businesses. Formal training also has an indirect, moderated through profit and assets, but significant impact on the productivity levels of those businesses. And, while those small businesses doing formal training are still small, they typically have three times the number of employees, twice the level of assets, and have multiple HR programs in place in addition to formal training.

With this understanding we can move beyond the questions of, do small businesses do formal training, and does formal training impact organizational

performance, to ones about when and how small businesses do formal training and what types of formal training they do.

Recommendations

The results of this research raise a number of questions which provide opportunities for further exploration beyond the scope of this work.

All Small Businesses Are Not The Same

We know from the data that there are significant differences in terms of size, assets, and HR activities between firm's doing formal training and those that do not. There are likely many other differences between these populations that we did not look at. As previously mentioned, we were not able to look at business location or industry sector. Both of these could likely affect a firm's use of formal training and the impact that training has on organizational performance.

We would expect that a firm located in an area with access to many training resources would be more likely to use those resources (OECD, 2013). We would also expect that firms in industries that are rapidly changing, like technology or medical, or firms in industries that require highly trained staff, like manufacturing, would be more likely to use formal training (Fernald, Jr., Solomon, & Bradley, 1999; OECD, 2013).

Also, as we previously mentioned, "small" can mean a business with one person or a business with 50 people. The difference between the average size of firms not doing formal training and those doing formal training was six employees. Not a lot. But that meant firms doing formal training were three times the size of those not doing training. The difference in communications paths between three employees and nine employees is

exponential. It's easy in research, and often necessary, to group populations together to better understand general shared characteristics. However, as we have seen from our investigation, there are many contingency and configurational aspects of small businesses that can moderate the impact of training on performance.

Previous researchers have attempted to address the size issue by grouping small businesses into categories by number of employees. For example, Aston, Sung, Raddon, & Riordan, (2008), discuss several different size categories: micro enterprise (1-10 employees), small enterprises (10-30 employees), small-medium enterprises (30-50), medium enterprises (40-249 employees), and large enterprises ((250+ employees). For Aston et al. a significant change takes place in an organization when they reach 30-50 employees requiring more formal skill acquisition processes. Others place this figure closer to 20 employees (Bishop, 2009; Kotey & Folker, 2007). However, as our research has shown, even within the smallest size grouping (micro enterprises) there are significant and substantial differences in organizational structures, practices, and resources that can impact a firm's use of formal training.

This points out some areas for consideration. First, attempts to classify and group small businesses by employee, revenue, or asset size with regards to HRD practices can lead to generalized miss-understandings of what is actually happening within those businesses. Because of this, HRD researchers and practitioners might be better off using growth stage or operational process definitions of small business that are not dependent on grouping firms together by size. Second, as much as researchers want to do research that generates generalizable results, doing so within the context of small businesses is

very difficult. The exponential growth curve of small nascent firms means that a firm with nine employees has significantly more potential communications pathways than one with three employees. A three person firm can easily rely on informal communications to share information across the organization. With nine employees an organization will likely have developed some formal structures and process for sharing information. So when we talk about small business in the context of HRD, we should not think of them as a homogenous group (Kitching, 2008; Rabemananjara & Parsley, 2006).

The more we can understand the unique characteristics of small businesses doing formal training the better we will be able to develop and deliver training to those firms. This is an opportunity for researchers and practitioners to work together.

All Training Programs Are Not the Same

The KFS data only contained a binary value for whether a firm had expenditures on training or not. As such, we were only able to test for firms that did formal training and those that did not. We know from previous research that certain types of training are better suited for certain human resource development needs (Jones et al., 2013). Leadership training can provide managers with better knowledge, skills and abilities to implement strategic policies in the workforce. On-the-job training can be beneficial for gaining competencies that require a high level of tacit knowledge. And so on.

Jones et al. (2013) also found that the way training was delivered made a difference in performance results. Training delivered by an outside expert but internally in the workplace had the biggest ROI. This was followed by workplace training delivered by internal staff.

Knowing that formal training can improve organizational performance is not enough to know what type of training is needed. For the HRD practitioner, there needs to be further investigation into the context of the situation, and other contingencies, in order to determine the correct training to deploy. However, having the knowledge that small businesses can and do benefit from formal training should make this task easier.

If You Build It They Will Not Come

In spite of numerous efforts by local, regional and national governments to develop and deliver training to small businesses, there is still a low level of adoption (OECD, 2013). In the KFS data this was around 20%. Other studies that have looked at SMEs and training in Europe and Asia saw similar levels of adoption, even when training was available for free (Gibb, 1997; OECD, 2013; Stone, 2010). Researchers have looked at a number of different reasons why there is such low adoption. The lack of funds, lack of education, lack of understand, lack of ROI, and many other "lacks" have been cited for why this might be the case (Bishop, 2009; Jones et al., 2013; Stone, 2010).

Our experience as a small business owner, and president of an association of small business owners, is that one of the biggest "lacks" is lack of time. Small business owners are so busy taking care of business that doing something which does not have immediate results for their business is extremely difficult to do. Even when they have the resources and the need, taking the time for training is seen as a distraction.

One possible way to overcome the "lack of time" resistance is to look at training delivery methods that reduce the amount of time staff need away from their job. This

could include more training delivered on-site or delivered on-line. Shorter more focused and frequent trainings might be more attractive.

In addition, the entrepreneurial persona contributes to the resistance of doing training, especially when the training comes from outside of the organization. Entrepreneurs tend to have a “can do” attitude that says they can do it all. It would be interesting to look at the KFS data and do an analysis if owner experience has an impact on a firm doing formal training. I would suspect that the more experienced a small business owner is, the more likely they would use formal training.

The Individual or the Organization

As we saw with Swanson’s model of HRD (2008), a division has developed in HRD separating individual level training and development from organization level programs and policies. One area of research and practice focuses on developing the person, the other focuses on developing the organization within which the person operates. This reminds us of the classic dualist separation of mind and body, where the mind operates and exists independent of the body. Philosophers have argued for centuries about this separation. Modern scientific research has effectively refuted the notion that the mind exists or operates independent of the body.

Of course we know, and we have seen in our results, that businesses operate as a system. While we can isolate parts of the system to gain better understanding of that individual part, no parts in a system operate independently. As such, it would be beneficial to take a more system wide approach to HRD research and practice. The very notion that practitioners and researchers specialize in training, career development,

leadership development, mentoring, organization development, and many other specializations, create impressions counter to the concept of the business operating as a system. While specialization is necessary for understanding, the knowledge gained must be tied to the greater understanding for us to fully appreciate it. This is why theory and practice working together matters. Researchers must be informed by practice and practitioners must be informed by research.

Conclusion

This research has provided us with several significant results. We know that small businesses are not only doing formal training, but also that those firms doing formal training have significantly improved profits, assets levels and productivity over those firms not doing formal training. Formal training in small business does result in improved organizational performance.

We also know that there are several characteristics such as number of employees and asset levels that are indicators of small businesses doing formal training. As we showed, the average number of employees in firms doing formal training was 9 versus 3 employees in firms not doing formal training. We also found that the asset level in firms doing formal training was double that of firms not doing training. However, we also know that relying on size variables to identify which small businesses might be good candidates for formal training is not enough and in fact can be misleading. Firms of the same relative size can be in very different stages of development. As a result, their need for employee development programs could be very different. In addition, there are many other internal and external factors which can impact a firm’s readiness for formal

training. Understanding these factors is needed to help us better focus training opportunities to the small businesses most likely to take advantage of the training.

The resource-based view of the firm provides us with a theoretical framework for understanding the multilevel impact of changes across levels of an organization. And the contingency and configurational perspectives help us understand that there are things which moderate the impact of formal training, Training, along with other HRD initiatives, should be aligned with a firm's strategic goals for maximum effectiveness, and there are configurations of HRD programs and policies that can optimize the impact of those initiatives.

Of all of the things we have seen through this research, the single biggest lesson to be learned is that, in small business, employees make all of the difference. This is not to say that there are no other critical and vital aspects for small businesses to consider. Obviously you need a good product or service and you need financial capital to grow. What we mean is that the single best way for small businesses to succeed is to invest in its employees. This is where small business can create and maintain competitive advantages that cannot be copied by others regardless of their size.

Investment in employees means hiring the right people for the right job. It means developing broad employee benefit and development programs that amply reward employees for their contributions and provide them with the necessary knowledge, skills and abilities to succeed in their work.

In our virtual economy, small businesses often find themselves competing against much larger, better funded, companies. Innovation can provide small businesses with

some advantage, but large companies buy innovation. Investing in people is the best way for small businesses to grow and succeed against the competition.

REFERENCES

Acedo, F. J., Barroso, C., & Galan, J. L. (2006). The resource-based theory: Dissemination and main trends. *Strategic Management Journal*, 27(7), 621-636.

Aguinis, H. & Kraiger, K. (2009). Benefits of Training and Development for Individuals and Teams, Organizations, and Society. *Annual Review of Psychology*, 60, 451-474.

Albrecht, K. (1983). *Organization development: A total systems approach to positive change in any business organization*. Englewood Cliffs, NJ: Prentice-Hall.

Alliger, G. M., Tannenbaum, S. I., Bennett, W., Traver, H., & Shortland, A. (1997). A meta-analysis on the relations among training criteria. *Personnel Psychology*, 50, 341−358.

Anderson, Valerie & Skinner, Denise. (1999). Organizational learning in practice: How do small businesses learn to operate internationally? *Human Resource Development International*, 2(3), 235-258.

Aragon-Sanchez, A., Barba-Aragón, I., & Sanz-Valle, R. (2003). Effects of training on business results. *The International Journal of Human Resource Management*, 14(6), 956-980.

Ardichvili, A., Harmon, B., Cardozo, R. N., Reynolds, P. D., & Williams, M. L. (1998). The new venture growth: Functional differentiation and the need for human resource development interventions. *Human Resource Development Quarterly*, 9(1), 55-70.

Argyris, C., & Schön, D. A. (1978). *Organizational learning: A theory of action perspective* (Vol. 173). Reading, MA: Addison-Wesley.

Argyris, C., & Schön, D. A. (1996). *Organizational learning. 2. Theory, method, and practice.* Reading, MA: Addison-Wesley.

Arthur Jr, W., Bennett Jr, W., Edens, P. S., & Bell, S. T. (2003). Effectiveness of training in organizations: A meta-analysis of design and evaluation features. *Journal of Applied Psychology*, 88(2), 234-245.

Ashton, D., Sung, J., & Raddon, A. (2005). A case where size matters: A preliminary investigation into the institutionalisation of skill formation and firm size. Research Paper No. 60, Autumn 2005, Warwick, UK: SKOPE Publications.

Ashton, D., Sung, J., Raddon, A. and Riordan, T. (2008). *Challenging the myths about learning and training in small and medium-sized enterprises: Implications for public policy*, ILO Employment Working Paper No. 1, Geneva: ILO.

Bacharach, S. B. (1989). Organizational theories: Some criteria for evaluation. *Academy of Management Review*, 14(4), 496-515.

Baldwin, T. T., & Danielson, C. C. (2002). Invited reaction: Linking learning with financial performance. *Human Resource Development Quarterly*, 13(1), 23-30.

Barber, J. (2004). Skill upgrading within informal training: Lessons from the Indian auto mechanic. *International Journal of Training and Development*, 8(2), 128-139.

Barkham, R., Gudgin, G., & Hart, M. (1996). *The Determinants of Small Firm Growth: An inter-regional study in the United Kingdom 1986-90*. UK: Routledge.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99-120.

Barney, J., Wright, M., & Ketchen, D. J. (2001). The resource-based view of the firm: Ten years after 1991. *Journal of Management*, 27(6), 625-641.

Baron, J. N., & Kreps, D. M. (1999). Consistent human resource practices. *California Management Review*, 41, 29−53.

Baron, J. N., & Kreps, D. M. (1999). *Strategic human resources: Frameworks for general managers (Vol. 149)*. New York: John Wiley.

Barrett, R., & Mayson, S. (2007). Human resource management in growing small firms. *Journal of Small Business and Enterprise Development*, 14(2), 307-320.

Barry, H., & Milner, B. (2002). SMEs and electronic commerce: a departure from the traditional prioritisation of training?. *Journal of European Industrial Training*, 26(7), 316-326.

Bartlett, K. R. (2001). The relationship between training and organizational commitment: A study in the health care field. *Human Resource Development Quarterly*, 12(4), 335-352.

Becker, B., & Gerhart, B. (1996). The impact of human resource management on organizational performance: Progress and prospects. *Academy of Management Journal*, 39(4), 779-801.

Becker, B. E., & Huselid, M. A. (2006). Strategic human resources management: Where do we go from here?. *Journal of Management*, 32(6), 898-925.

Beer, Michael. (1980). A Social System Model for Organization Development. In Thomas G Cummings (Ed.), *Systems Theory for Organization Development* (pp. 73-114). Chichester, UK: John Wiley & Sons.

Bierema, L. L. (2000). Moving beyond performance paradigms in human resource development. In Wilson, A.L. & Hayes, E.R. (Eds.), *Handbook of adult and continuing education*, (pp. 278-293). San Francisco, CA: Jossey-Bass.

Birchall, D., & Giambona, G. (2007). SME manager development in virtual learning communities and the role of trust: A conceptual study. *Human Resource Development International*, 10(2), 187-202.

BIS, (2011). *BIS small business survey*. Sheffield, UK: BIS Enterprise Directorate.

Bishop, Daniel, (2009). Learning in Small Firms: How is it different to learning in larger firms?
10th International Conference on Human Resource Development Research and Practice Across Europe, Working Paper.

Bowen, D. E., & Ostroff, C. (2004). Understanding HRM–firm performance linkages: The role of the "strength" of the HRM system. *Academy of Management Review*, 29(2), 203-221.

Brethower, D. M. (1999). General systems theory and behavioral psychology. In Stolovitch, H.D. & Keepr, E.J. (Eds.), *Handbook of human performance technology: Improving individual and organizational performance worldwide*, (pp. 67-81). San Fancisco, CA: Jossey-Bass.

Brinkerhoff, R. O., & Gill, S. J. (1994). *The Learning Alliance: Systems Thinking in Human Resource Development. The Jossey-Bass Management Series*. San Francisco, CA: Jossey-Bass

Brock, W. A., & Evans, D. S. (1989). Small business economics. *Small Business Economics*, 1(1), 7-20.

Bryan, J. (2006). Training and performance in small firms. *International Small Business Journal*, 24(6), 635-660.

Bryk, A. S., Raudenbush, S. W. (1992). *Hierarchical linear models.* Newbury Park, CA: Sage.

Burke, R. (2011) Human resource management in small and medium sized enterprises: Benefits and challenges. In Cooper, C & Burke, R. (Eds.) *Human Resource Management in Small Business: Achieving Peak Performance*, (pp. 10-70). Cheltenham, UK: Edward Elgar Publishing Limited.

Burke, W. W., & Litwin, G. H. (1992). A causal model of organizational performance and change. *Journal of Management*, 18(3), 523-545.

Burnes, Bernard. (2006). Kurt Lewin and the planned approach to change. In Joan V. Gallos (Ed.), *Organization Development,* (pp. 133-157). San Francisco, CA: Jossey-Bass.

Cardon, M. S., & Stevens, C. E. (2004). Managing human resources in small organizations: What do we know?. *Human Resource Management Review*, 14(3), 295-323.

Castrogiovanni, G. (2011). The role of human capital factors in small business success. In Cooper, C & Burke, R. (Eds.) *Human Resource Management in Small Business: Achieving Peak Performance*, (pp. 71-92). Cheltenham, UK: Edward Elgar Publishing Limited.

CEDEFOP (2011). *The impact of vocational education and training on company performance*, RP 19, Luxembourg: Publications Office of the European Union.

Chittenden, F., Hall, G., & Hutchinson, P. (1996). Small firm growth, access to capital markets and financial structure: Review of issues and an empirical investigation. *Small Business Economics*, 8(1), 59-67.

Coff, R. W. (1997). Human assets and management dilemmas: Coping with hazards on the road to resource-based theory. *Academy of Management Review*, 22(2), 374-402.

Coff, R. W. (1999). When competitive advantage doesn't lead to performance: The resource-based view and stakeholder bargaining power. *Organization Science*, 10(2), 119-133.

Coleman, James S. (1990). *Foundations of Social Theory*. Cambridge, MA: Harvard Press.

Combs, J., Liu, Y., Hall, A., & Ketchen, D. (2006). How much do high-performance work practices matter? A meta-analysis of their effects on organizational performance. *Personnel Psychology*, 59(3), 501-528.

Conner, K. R. (1991). A historical comparison of resource-based theory and five schools of thought within industrial organization economics: do we have a new theory of the firm?. *Journal of Management*, 17(1), 121-154.

Cowan, G., Pines, D. & Meltzer, D. (1994). *Complexity: Metaphors, models, and reality.* Reading, MA: Addison-Wesley.

Crook, T. R., Todd, S. Y., Combs, J. G., Woehr, D. J., & Ketchen Jr, D. J. (2011). Does human capital matter? A meta-analysis of the relationship between human capital and firm performance. *Journal of Applied Psychology*, 96(3), 443.

Crook, T. R., Ketchen, D. J., Combs, J. G., & Todd, S. Y. (2008). Strategic resources and performance: A meta-analysis. *Strategic Management Journal*, 29(11), 1141-1154.

Curran, Patrick J., West, Stephen G. & Finch, John F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1(1), 16-29.

d'Arcimoles, C. H. (1997). Human resource policies and company performance: A quantitative approach using longitudinal data. *Organization Studies*, 18(5), 857-874.

Davis, F. D., & Yi, M. Y. (2004). Improving computer skill training: Behavior modeling, symbolic mental rehearsal, and the role of knowledge structures. *Journal of Applied Psychology*, 89(3), 509-523.

Delaney, J. T., & Huselid, M. A. (1996). The impact of human resource management practices on perceptions of organizational performance. *Academy of Management Journal*, 39(4), 949-969.

Delaney, J. T., Lewin, D., & Ichniowski, C. (1989). *Human resource policies and practices in American firms*. US Dept. of Labor, Bureau of Labor-Management Relations and Cooperative Programs.

Delery, J. E., & Doty, D. H. (1996). Modes of theorizing in strategic human resource management. *Academy of Management Journal*, 39, 802−835.

Dewar, R., & Werbel, J. (1979). Universalistic and contingency predictions of employee satisfaction and conflict. *Administrative Science Quarterly*, 24(3), 426-448.

DfEE. (2000). *Smaller Firms Set for Healthy Future.* Skills and Enterprise Executive Issue. 4/2000. London: SEN Publication.

Docherty, J.P, Surles, R.C., & Donovan, C.M. (2001). Organizational Theory. In Talbott, J. A., & Hales, R. E. (Eds.). *Textbook of Administrative Psychiatry: New Concepts for a Changing Behavioral Health System*, (pp. 33-42). Washington D.C.: American Psychiatric Publishing, Inc.

Doty, D. H., Glick, W. H., & Huber, G. P. (1993). Fit, equifinality, and organizational effectiveness: A test of two configurational theories. *Academy of Management Journal*, 36(6), 1196-1250.

Down, S. (1999). Owner—manager learning in small firms, *Journal of Small Business and Enterprise Development*, 6(3), 267-280.

Drazin, R., & Van de Ven, A. H. (1985). Alternative forms of fit in contingency theory. *Administrative Science Quarterly*, 30(4), 514-539.

Dubin, R. (1976). Theory building in applied areas. In Anderson, N., Ones, D. S., & Kepir, H (Eds.), *Handbook of industrial and organizational psychology, Vol. 1* (pp. 17-39): Chicago: Rand McNally..

Dunn and Bradstreet, (2012). *Global Business Failures Report*. London: Dunn and Bradstreet Limited.

Dyer, L. (1984a). Linking human resources and business strategies. *Human Resource Planning*, 7(2): 79-84.

Dyer, L. (1984b). Studying human resource strategy: An approach and an agenda. Industrial Relations: *A Journal of Economy and Society*, 23(2), 156-169.

Dyer, L. (1985). Strategic human resources management and planning. *Research in Personnel and Human Resource Management*, 3(1): 1-30.

Dyer, L.,& Reeves, T. (1995). Human resource strategies and firm performance. *International Journal of Human Resource Management*, 6, 656−670.

Faems, D., Sels, L., DeWinne, S., & Maes, J. (2005). The effect of individual HR domains on financial performance. *International Journal of Human Resource Management*, 16, 676−700.

Fayol, H. (1949). *General and industrial administration*. NY: Pitman.

Fernald Jr, L., Solomon, G., & Bradley, D. (1999). Small business training and development in the United States. *Journal of Small Business and Enterprise Development*, 6(4), 310-325.

Fey, C. F., Björkman, I., & Pavlovskaya, A. (2000). The effect of human resource management practices on firm performance in Russia. *International Journal of Human Resource Management*, 11(1), 1-18.

Finkel, Steven, (1995). *Causal Analysis with Panel Data*. Thousand Oaks, CA. SAGE Publications, Inc.

Fombrun, C. J., Tichy, N. M., & Devanna, M. A. (1984). *Strategic human resource management*. New York: Wiley.

Freedman, D. A., (2006). 'On the so-called "Huber Sandwich estimator" and "robust standard errors"', *The American Statistician*, 60, 299-302.

Fuller, J., & Farrington, J. (1999). *From training to performance improvement: Navigating the transition.* San Francisco: Jossey-Bass/Pfeiffer.

Gentry, W. A. & Martineau, J. W. (2010). Hierarchical linear modeling as an example for measuring change over time in a leadership development evaluation context. *The Leadership Quarterly.* 21, 645-656.

Gerhart, B. (2005). Human resources and business performance: Findings, unanswered questions, and an alternative approach. *Management Revue*, 16(2), 174-185.

Gibb, A. A. (1997). Small firms' training and competitiveness. Building upon the small business as a learning organisation. *International Small Business Journal*, 15(3), 13-29.

Gilley, J. W., Dean, P. J., & Bierema, L. L. (2001). *Philosophy and Practice of Organizational Learning. Performance and Change*. Cambridge, MA: Perseus

Gilley, J. W., & Gilley, A. M. (1998). *Strategically integrated HRD: Partnering to maximize organizational performance*. Reading, MA: Addison-Wesley.

Gorgievski, M. J., Ascalon, M. E., & Stephan, U. (2011). Small business owners' success criteria, a values approach to personal differences. *Journal of Small Business Management*, 49(2), 207-232.

Gortner, H. F., Mahler, J., & Nicholson, J. B. (1987). *Organization theory: A public perspective*. Chicago, IL: Dorsey Press.

Grant, R. M. (1991). The resource-based theory of competitive advantage: Implications for strategy formulation. *California Management Review*, Summer: 114-135

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, 17, 109-122.

Greiner, L. E. (1972). Evolution and revolution as organizations grow. *Harvard Business Review*, July-August: 37-46.

Guerrero, S., & Barraud-Didier, V. (2004). High-involvement practices and performance of French firms. *International Journal of Human Resource Management,* 15, 1408−1423.

Guest, D. E. (1990). Human resource management and the American dream. *Journal of Management Studies*, 27(4), 377-397.

Hall, G. and Fulshaw, S. (1993) Factors associated with the relative perfomance amongst small firms in the British instrumentation sector, In H. Klandt (ed.) *Entrepreneurship and Business Development*, (pp. 227-237). Avebury: Aldershot.

Harde, P.L. (2003). Beyond two decades of motivation: A review of research and practice in instructional design and human performance technology. *Human Resource Development Review,* 2(1), 54081.

Harrison, R. (1997). *Employee Development*, London: Institute of Personnel and Development (IPD).

Heneman III, H.G. & Berkely, R.A. (1999). Application attraction practices and outcomes among small businesses. *Journal of Small Business Management*, 37(1), 53-74.

Hicks, G. H., & Gullet, C. R. (1975). *Organizational: Theory and Behavior*. New York: McGraw-Hill Education

Hill, Rosemary. (2004). Why HRD in small organizations may have become a neglected field of study. In Jim Stewart & Graham Beaver (Eds.). *HRD in Small Organizations: Research and practice* (pp. 8-25). New York: Routledge.

Hofer, C. W., & Schendel, D. (1978). *Strategy formulation: Analytical concepts (Vol. 1986).* St. Paul, MN: West Publishing Company.

Hofmann, David A. (1997). An overview of the logic and rationale of hierarchical linear models. *Journal of Management,* 23(6), 723-744.

Hoffman, D. A., Jacobs, R, & Baratta, J. (1993). Dynamic criteria and the measurement of change. *Journal of Applied Psychology, 77, 194-204.*

Hoque, K., & Bacon, N. (2006). The antecedents of training activity in British small and medium-sized enterprises. *Work, Employment & Society*, 20(3), 531-552.

Hox, J. J., (2010). *Multilevel analysis: Techniques and applications, 2nd edn*. Mahwah, NJ: Erlbaum.

Huselid, M. A. (1995). The impact of human resource management practices on turnover, productivity, and corporate financial performance. *Academy of Management Journal*, 38(3), 635-672.

Ibrahim, A. B., & Goodwin, J. R. (1986). Perceived causes of success in small business. *American Journal of Small Business*, 11(2), 41-50.

Jackson, S. E., & Schuler, R. S. (1995). Understanding human resource management in the context of organizations and their environments. *Strategic Human Resource Management*, 46, 237-264.

Jackson, S. E., Schuler, R. S., & Rivero, J. C. (1989). Organizational characteristics as predictors of personnel practices. *Personnel Psychology*, 42(4), 727-786.

Jarvis, R., Curran, J., Kitching, J., & Lightfoot, G. (2000). The use of quantitative and qualitative criteria in the measurement of performance in small firms. *Journal of Small Business and Enterprise Development*, 7(2), 123-134.

Jayawarna, D., Macpherson, A., & Wilson, A. (2007). Training commitment and performance in manufacturing SMEs: incidence, intensity and approaches. *Journal of Small Business and Enterprise Development*, 14(2), 321-338.

Jones, C. & Bos, René ten. (2007). Introduction. In Campbell Jones & René ten Bos (Eds.), *Philosophy and Organization* (pp. 1-17). London: Routledge.

Jones, P., Beynon, M. J., Pickernell, D., & Packham, G. (2013). Evaluating the impact of different training methods on SME business performance. *Environment and Planning C: Government and Policy*, 31(1), 56-81.

Kanter, Rosabeth Moss. (2006). From Cells to Communities: Deconstruction and reconstructing the organization. In Joan V. Gallos (Ed.), *Organization Development* (pp. 858-887). San Francisco, CA: Jossey-Bass.

Katou, A. A. (2009). The Impact of Human Resource Development on organisational performance: Test of a causal model. *Journal of Behavioral & Applied Management*, 10(3). 335-356.

Katou, A. A., & Budhwar, P. S. (2006). Human resource management systems and organizational performance: a test of a mediating model in the Greek manufacturing context. *The International Journal of Human Resource Management*, 17(7), 1223-1253.

Katz, D. & Kahn, R. L. (1966). *The social psychology of organizations.* New York: Wiley.

Kauffman, D. L. (1980). *Systems 1: An Introduction to Systems Thinking*. Minneapolis, MN: Future Systems, Inc.

Kelmar, J. H. (1994). Measurement of success and failure in small business: A two-factor approach. *Journal of Enterprising Culture*, 1(03n04), 421-436.

Kissler, G. D. (1991). *The change riders: Managing the power of change*. Reading, MA: Addison Wesley.

Kitching, J. (2008). Rethinking UK small employers' skills policies and the role of workplace learning. *International Journal of Training and Development*, 12(2), 100-120.

Kitching, J., & Blackburn, R. M. (2002). *The nature of training and motivation to train in small firms*. RR330, Small Business Research Centre, Kingston University.

Klein, K. J., Dansereau, F., & Hall, R. J. (1994). Levels issues in theory development, data collection, and analysis. *Academy of Management Review*, 19(2), 195-229.

Klein, K. J., & Kozlowski, S. W. (2000). *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*. San Francisco: Jossey-Bass.

Klein, K., Tosi, H. & Cannella, Jr., A. (1999). Multilevel theory building: Benefits, barriers, and new developments. *Academy of Management Review*, 24(2), 243-248.

Kobe, Kathryn. (2007). Advocacy: The voice of small business. SBA Fact Sheet. U.S. Department of Commerce, Bureau of the Census. Retrieved May 8, 2008 from www.sba.gov/advo/research/rs299tot.pdf.

Kogut, B., & Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization science*, 3(3), 383-397.

Kozlowski, S., Brown, K., Weissbein, D., Cannon-Bowers, J., & Salas, E. (2000). A multilevel approach to training effectiveness. In K. Klein, & S. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations*. San Francisco, CA: Jossey-Bass.

Kotey, B., & Folker, C. (2007). Employee training in SMEs: Effect of size and firm type—Family and nonfamily. *Journal of Small Business Management*, 45(2), 214-238.

Kotey, B., & Slade, P. (2005). Formal human resource management practices in small growing firms. *Journal of Small Business Management*, 43(1), 16-40.

Kraiger K. 2002. Decision-based evaluation. In Kraiger, K (Ed.*) Creating, implementing, and maintaining effective training and development: State-of-the-Art lessons for practice*, (pp. 331–375). San Francisco, CA: Jossey-Bass

Lawler, Edward E. (2006). What Makes People Effective?. In Joan V. Gallos (Ed.), *Organization Development* (pp. 634-655). San Francisco, CA: Jossey-Bass.

Lawler E.E., Mohrman, S.A, & Lawford, G.E.,(1998). *Strategies for high performance organizations: the CEO report: employee involvement, TQM, and reengineering programs in Fortune 1000 corporations.* San Francisco, Jossey-Bass Publishers.

Lee, M. (2001). A Refusal to define HRD. *Human Resource Development International*, 4(3), 383-390.

Lengnick-Hall, C. A., & Lengnick-Hall, M. L. (1988). Strategic human resources management: A review of the literature and a proposed typology. *Academy of Management Review*, 13(3), 454-470.

Lepak, D. P., & Snell, S. A. (1999). The human resource architecture. *Academy of Management Review*, 24, 31−48.

Longenecker, J., Petty, J., Palich, L., & Hoy, F. (2013). *Small business management*. Cengage Learning.

López, S. P., Peón, J. M. M., & Ordás, C. J. V. (2005). Organizational learning as a determining factor in business performance. *The Learning Organization*, 12(3), 227-245.

Mabey, C., & Ramirez, M. (2005). Does management development improve organizational productivity? A six-country analysis of European firms. *The International Journal of Human Resource Management*, 16(7), 1067-1082.

Mahoney, T. A., & Deckop, J. R. (1986). Evolution of concept and practice in personnel administration/human resource management (PA/HRM). *Journal of Management*, 12(2), 223-241.

Marquardt, M. J. (2000). Action learning and leadership. *The Learning Organization*, 7(5), 233-241.

Martin-Alcazar, F., Romero-Fernandez, P. M., & Sanchez-Gardey, G. (2005). Strategic human resource management: integrating the universalistic, contingent, configurational and contextual perspectives. *The International Journal of Human Resource Management*, 16(5), 633-659.

Matlay, H. (1999). Vocational education and training in Britain: a small business perspective. *Education and Training*, 41(1), 6-13.

Mayo, E. G. (1933). *The human problems of an industrial society*. New York: Macmillan.

Mayo, A. (2000). The role of employee development in the growth of intellectual capital. *Personnel Review*, 29(4), 521-533.

McGregor, D. (1960). *The human side of enterprise*. New York: McGraw Hill.

Menard, S. (1991). *Longitudinal Research*. Newbury Park: Sage Publications.

Meyer, A. D., Tsui, A. S., & Hinings, C. R. (1993). Configurational approaches to organizational analysis. *Academy of Management Journal*, 36(6), 1175-1195.

Mezirow, J. (1991). *Transformative dimensions of adult learning*. San Francisco: Jossey-Bass.

Miles, R. E., & Snow, C. C. (1984). Designing strategic human resources systems. *Organizational Dynamics*, 13(1), 36-52.

Mount, J., Zinger, J. T., & Forsyth, G. R. (1993). Organizing for development in the small business. *Long Range Planning*, 26(5), 111-120.

Nadler, D. A. (1998). Leading executive teams. In Nadler, D. & Spencer, J.L. (Eds.), *Executive teams*, (pp. 3-29). San Francisco: Jossey-Bass.

Nadler, David A. & Gerstein, Marc S. (2006). Designing High-Performance Work Systems: Organizing People, Work, Technology, and Information. In Joan V. Gallos (Ed.), *Organization Development*, (pp. 730-747). San Francisco: Jossey-Bass.

OECD (2013), *Skills Development and Training in SMEs, Local Economic and Employment Development* (LEED), OECD Publishing.

Osterman, P. (1994). How Common is workplace transformation and who adopts it*? Industrial and Labor Relations Review*, Vol. 47(2), 173-188.

Ostroff, C., & Bowen, D. E. (2000). Moving HR to a higher level: HR practices and organizational effectiveness. In K. J. Klein, & S. W. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations* (pp. 211−266). San Francisco: Jossey-Bass.

Patterson, M. G., West, M. A., Lawthom, R., & Nickel, S. (1997). *Impact of People Management on Business Performace*. London: Institute of Personell and Development.

Paul, A., & Anantharaman, R. (2003). Impact of people management practices on organizational performance. *International Journal of Human Resource Management*, 14, 1246−1266.

Pfeffer, M.E. (1994). *Competitive advantage through people*. Boston: Harvard Business School Press.

Porter, M. (1980). *Competitive Strategy*. New York: Free Press.

Rabemananjara, R. N., & Parsley, C. (2006). *Employee training decisions, business strategies and human resource management practices: A study by size of business*. Small Business Policy Branch, Ottawa: Industry Canada.

Raghuram, S. (1994). Linking staffing and training practices with business strategy: A theoretical perspective. *Human Resource Development Quarterly*, 5(3), 237-251.

Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods, 2nd edn*. Thousand Oaks, CA: Sage.

Redding, J. C., & Catalanello, R. F. (1994). Strategic readiness: The making of the learning organization. San Francisco: Jossey-Bass.

Rigg, Clare & Trehan, Kiran. (2004). Now you see it now you don't: comparing traditional and discourse readings of HRD in small organizations. In Jim Stewart& Graham Beaver (Eds.). *HRD in Small Organizations: Research and Practice* (pp. 48-73). London: Routledge.

Roethlisberger, F. J., & Dickson, W. J. (1939). *Management and the Worker. An Account of a Research Program conducted by the Western Electric Co., Hawthorne Works, Chicago*. Cambridge, MA: Harvard University Press.

Rogosa, D. R., Brandt, D., & Zimowski, M. (1982). A growth curve approach to measurement of change. *Psychological Bulletin*, 90, 726-748.

Rosenberg, E. L. (1998). Levels of analysis and the organization of affect. *Review of General Psychology*, 2(3), 247-270.

Rousseau, D. M. (1985). Issues of level in organizational research: Multilevel and cross-level perspectives. In L. L. Cummings & B. Staw (Eds.), *Research in organizational behavior*. Greenwich, CT: JAI Press.

Rummler, G. A., & Brache, A. P. (1995). *Improving Performance: How to Manage the White Space in the Organization Chart*, San Francisco: Jossey-Bass, Inc.

Sadler-Smith, E., Down, S., & Field, J. (1999). Adding value to HRD: Evaluation, investors in people and small firm training. *Human Resource Development International*, 2(4), 369-390.

Satterfield, J. M., & Hughes, E. (2007). Emotion skills training for medical students: a systematic review. *Medical Education*, 41(10), 935-941.

Schoonhoven, C. B. (1981). Problems with contingency theory: testing assumptions hidden within the language of contingency theory. *Administrative Science Quarterly*, 26(3), 349-377.

Schuler, R. S., & Jackson, S. E. (1987). Linking competitive strategies with human resource management practices. *The Academy of Management Executive*, 1(3), 207-219.

Schuler, R. S., & Jackson, S. E. (1987). Organizational strategy and organization level as determinants of human resource management practices. *Human Resource Planning*, 10(3), 125-141.

Schuler, R. S., & MacMillan, I. C. (1984). Gaining competitive advantage through human resource management practices. *Human Resource Management*, 23(3), 241-255.

Scott, M., & Bruce, R. (1987). Five stages of growth in small business. *Long range planning*, 20(3), 45-52.

Senge, P. (1990). *The fifth discipline: The art and practice of learning*. New York: Doubleday.

Short, J. C.; Ketchen, Jr., D. J.; Bennett, N. & du Toit, M. (2006). An examination of firm, industry, and time effects on performance using random coefficients modeling. *Organizational Research Methods,* 9, 259-284.

Silber, K. H. (1992). Intervening at different levels in organizations. In H.D. Stolovitch & E. Keeps, (Eds.), *Handbook of human performance technology: A comprehensive guide for analyzing and solving performance problems in organizations*. San Francisco: Jossey-Bass.

Singer, J. D. & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurance.* New York: Oxford Press.

Stevenson, W. B. (2003). Introduction: Planned organizational change and organizational networks. *The Journal of Applied Behavioral Science*, 39(3), 238-242.

Stewart, Jim & Beaver, Graham. (2004). Researching and practicing HRD in small organizations. In Jim Stewart& Graham Beaver (Eds.). *HRD in Small Organizations: Research and Practice* (pp. 1-4). London: Routledge.

Stolovitch, H. D., & Keeps, E. J. (1999). *Handbook of human performance technology*. Washington DC: National Society for Performance and Instruction.

Stone, I. (2010). *Encouraging small firms to invest in training : learning from overseas*. Wath-Upon-Dearne,Yorkshire: UK Commission for Employment and Skills. Praxis., 5.

Storey, D. J. (2004). Exploring the link, among small firms, between management training and firm performance: a comparison between the UK and other OECD countries. The *International Journal of Human Resource Management,* 15(1), 112-130.

Storey, D. & Westhead, P. (1997). Management training and small firm performance: Why is the link so weak? *International Small Business Journal,* 14(4), 13-24.

Subramony, M., Krause, N., Norton, J., & Burns, G. N. (2008). The relationship between human resource investments and organizational performance: A firm-level examination of equilibrium theory. *Journal of Applied Psychology*, 93(4), 778-788.

Swanson, R. A. (1995). Human resource development: Performance is the key. *Human Resource Development Quarterly*, 6(2), 207-213.

Swanson, R. (2001). *Foundations of Human Resource Development*, San Francisco, CA: Berrett-Koehler Publishers, Inc.

Swanson, R. A., & Arnold, D. E. (1996). Part one: What is the purpose of human resource development? The purpose of human resource development is to improve organizational performance. *New Directions for Adult and Continuing Education*, 1996(72), 13-19.

Swanson, R.A. & Holton, E.F. (2009). *Foundations of Human Resource Development 2nd Edition*. San Francisco: Berrett-Koehler.

Taylor, F. W. (1947). *Scientific management*. New York and London: Harper and Row.

Taylor, Scott; Shaw, Sue; & Thorpe, Richard. (2004). Neither market failure nor customer ignorance: the organizational limitations of employee training and development. In Jim Stewart & Graham Beaver (Eds.). *HRD in Small Organizations: Research and Practice* (pp. 1-4). London: Routledge.

Terpstra, D. E., & Rozell, E. J. (1993). The relationship of staffing practices to organizational level measures of performance. *Personnel Psychology*, 46(1), 27-48.

Tichy, Noel. (1980). A Social Network Perspective for Organization Development. In Thomas G Cummings (Ed.), *Systems Theory for Organization Development* (pp. 115-162). Chichester, UK: John Wiley & Sons.

Tharenou, P., Saks, A. M., & Moore, C. (2007). A review and critique of research on training and organizational-level outcomes. *Human Resource Management Review*, 17, 251-273.

Tuan, N.G. & Ngoc, N.M., (2012). Effects of human resource management on business performance of small and medium size manufacturers in Hanoi – Vietnam. *Australian Journal of Business and Management Research*, 2(6), 47-54.

Ulrich, D. (1998). A new mandate for human resources. *Harvard Business Review*, 76, 124-135.

Van de Ven, A. H. (2007). *Engaged Scholarship: A Guide for Organizational and Social Research*. Oxford, UK: Oxford University Press.

Venkatraman, N., & Prescott, J. E. (1990). Environment-strategy coalignment: An empirical test of its performance implications. *Strategic Management Journal*, 11(1), 1-23.

Vickerstaff, S. & Parker, K. (1995). Helping small firms: The contribution of TECs and LECs. *International Small Business Journal*, 13, 56-72.

Walker, E., & Brown, A. (2004). What success factors are important to small business owners?. *International Small Business Journal*, 22(6), 577-594.

Weber, M. (1947*). The theory of economic and social organization*. Trans. A.M. Henderson and Talcott Parsons. New York: Oxford University Press.

Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal*, 5(2), 171-180.

Wils, T., & Dyer, L. (1984). Relating business strategy to human resource strategy: Some preliminary evidence. In *44th annual meeting of the Academy of Management*, Boston.

Wright, P.M. and Gardner, T.M. (2002). Theoretical and empirical challenges in studying the HR practice - firm performance relationship. In D. Holman, T. Wall, C. Clegg, P. Sparrow and A. Howard, (Eds). *The New Workplace: A Guide to the Human Impact of Modern Working Practices*. (pp. 311-330). Chichester: John Wiley and Sons.

Wright, P. M., Gardner, T. M., & Moynihan, L. M. (2003). The impact of HR practices on the performance of business units. *Human Resource Management Journal*, 13(3), 21-36.

Wright, P. M., & McMahan, G. C. (1992). Theoretical perspectives for strategic human resource management. *Journal of Management*, 18, 295−320.

Wright, P. M., McMahan, G. C., & McWilliams, A. (1994). Human resources and sustained competitive advantage: a resource-based perspective. *International Journal of Human Resource Management*, 5(2), 301-326.

Yin, R. K. (1994). *Case Study Research Design and Methods, 2nd ed*., Newbury Park, CA: Sage.

Zinn, L. M. (1983). *The Philosophy of Adult Education Inventory (PAEI)*. Doctoral dissertation, The Florida State University. Dissertation Abstracts International, 44, 1667A–1668A.

APPENDICES

Consulting Technical Document

Lindsey Dietz

January 9, 2014

1 Linear Mixed Model

Hierarchical Linear Models (HLM) can cover a broad variety of response data. Linear Mixed Models (LMM)[1] require a continuous outcome that can be assumed to follow a normal distribution. The LMM is classically written for subject i as

$$Y_i = X_i\beta + Z_i\gamma_i + \epsilon_i$$
$$\gamma_i \sim N_{p_{random}}(0, \Psi)$$
$$\epsilon_i \sim N_{n_i}(0, \sigma^2\Lambda_i)$$

where

- n_i is the number of observations for subject i
- Y_i is a $n_i \times 1$ vector of responses for subject i
- X_i is a $n_i \times p_{fixed}$ design matrix for the fixed effects of subject i; characterizes the systematic part of the response
- β is a $p_{fixed} \times 1$ vector of fixed parameters (coefficients)
- Z_i is a $n_i \times p_{random}$ design matrix for the random effects of subject i; characterizes random variation in the response attributable to among-unit sources
- γ_i is a $p_{random} \times 1$ vector of the random effects (coefficients) that completes the characterization of among-unit variation.
- Ψ is the $p_{random} \times p_{random}$ covariance matrix for the random effects
- ϵ_i is a $n_i \times 1$ vector of random errors for subject i; these represent within-unit deviations characterizing variation due to sources like within-unit fluctuations and measurement error
- $\sigma^2\Lambda_i$ is the $n_i \times n_i$ covariance matrix for errors in subject i
  - Λ_i is the $n_i \times n_i$ correlation matrix for random errors among subject i's different time points

The assumptions of the model (besides normality) are:

1. ϵ_i is independent of γ_j for all i, j
2. $\gamma_1, ..., \gamma_n$ are independent
3. $\epsilon_1, ..., \epsilon_n$ are independent

[1] http://cran.r-project.org/doc/contrib/Fox-Companion/appendix-mixed-models.pdf for other standard ways to write the model

Note that it is allowed for there to be correlation within subject but that all subjects are independent. We can assume no within-group correlations, i.e. that $\Lambda_i = I_i$. Suppose we have 3 observed time points for subject i. A more reasonable correlation structure may be to use an autoregressive correlation matrix within each subject

$$\Lambda_{AR1} = \begin{bmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{bmatrix}$$

or a compound symmetry style correlation matrix

$$\Lambda_{CS} = \begin{bmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{bmatrix}$$

Λ_{CS} would represent the belief that a person's time point should be correlated the same regardless of how far apart the surveys were administered. Λ_{AR1} would indicate that time points farther apart will be less correlated (since $\rho^2 \leq \rho$).

1.1 LMM Example modified from Hypothesis 3

Let's say we just want to make a model with sales growth as our response. Assume we only have 3 companies total and each company has all 4 responses for the 4^{th}-7^{th} surveys. Also, let's just look at a couple of covariates, sales per employee and profit margin and again assume there are no missing values. We will write a model with a random intercept by company and will assume an AR(1) correlation structure for the error terms. Here is some fictitious data:

Company 1:

	Sales Growth	Sales/Emp	Profit Margin
S4	40,000	10,000	0.51
S5	50,000	15,000	0.48
S6	80,000	13,000	0.35
S7	70,000	14,000	0.25

Company 2:

	Sales Growth	Sales/Emp	Profit Margin
S4	500,000	50,000	0.35
S5	700,000	70,000	0.24
S6	800,000	40,000	0.27
S7	850,000	42,000	0.28

Company 3:

	Sales Growth	Sales/Emp	Profit Margin
S4	200,000	10,000	0.01
S5	100,000	5,000	0.02
S6	50,000	2,500	0.02
S7	40,000	2,000	0.01

Model for Company 1:
(Note that we add a row of ones to the X_i - matrix to represent having an intercept in the model.)

$$\begin{bmatrix} 40,000 \\ 50,000 \\ 80,000 \\ 70,000 \end{bmatrix} = \begin{bmatrix} 1 & 10,000 & 0.51 \\ 1 & 15,000 & 0.48 \\ 1 & 13,000 & 0.35 \\ 1 & 14,000 & 0.25 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \gamma_1 + \begin{bmatrix} \epsilon_{11} \\ \epsilon_{12} \\ \epsilon_{13} \\ \epsilon_{14} \end{bmatrix}$$

where

$$\gamma_1 \sim N(0, \sigma^2_{company}) \text{ and } \begin{bmatrix} \epsilon_{11} \\ \epsilon_{12} \\ \epsilon_{13} \\ \epsilon_{14} \end{bmatrix} \sim N_4(\mathbf{0}, \sigma^2_{error}\mathbf{\Lambda}_1) \text{ where } \mathbf{\Lambda}_1 = \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}$$

Model for Company 2:

$$\begin{bmatrix} 500,000 \\ 700,000 \\ 800,000 \\ 850,000 \end{bmatrix} = \begin{bmatrix} 1 & 50,000 & 0.35 \\ 1 & 70,000 & 0.24 \\ 1 & 40,000 & 0.27 \\ 1 & 42,000 & 0.28 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \gamma_2 + \begin{bmatrix} \epsilon_{21} \\ \epsilon_{22} \\ \epsilon_{23} \\ \epsilon_{24} \end{bmatrix}$$

where

$$\gamma_2 \sim N(0, \sigma^2_{company}) \text{ and } \begin{bmatrix} \epsilon_{21} \\ \epsilon_{22} \\ \epsilon_{23} \\ \epsilon_{24} \end{bmatrix} \sim N_4(\mathbf{0}, \sigma^2_{error}\mathbf{\Lambda}_2) \text{ where } \mathbf{\Lambda}_2 = \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}$$

Model for Company 3:

$$\begin{bmatrix} 200,000 \\ 100,000 \\ 50,000 \\ 40,000 \end{bmatrix} = \begin{bmatrix} 1 & 10,000 & 0.01 \\ 1 & 5,000 & 0.02 \\ 1 & 2,500 & 0.02 \\ 1 & 2,000 & 0.01 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \gamma_3 + \begin{bmatrix} \epsilon_{31} \\ \epsilon_{32} \\ \epsilon_{33} \\ \epsilon_{34} \end{bmatrix}$$

$$\gamma_3 \sim N(0, \sigma^2_{company}) \text{ and } \begin{bmatrix} \epsilon_{31} \\ \epsilon_{32} \\ \epsilon_{33} \\ \epsilon_{34} \end{bmatrix} \sim N_4(\mathbf{0}, \sigma^2_{error}\mathbf{\Lambda}_3) \text{ where } \mathbf{\Lambda}_3 = \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix}$$

The parameters that the model estimates are:

- β_0= Fixed Intercept term
- β_1= Fixed Coefficient of Sales per Employee
- β_2= Fixed Coefficient of Profit Margin
- $\sigma^2_{company}$= Variance of the Random Intercept by Company
- σ^2_{error}= Variance of the Random Error
- ρ= Correlation between Error terms within each Company

2 Generalized Linear Mixed Model

2.1 Exponential Families

There is still the ability to model binary and count type variables using Generalized Linear Mixed Models (GLMM). One can think of a GLMM as an extension of LMM in which a specific transformation of the mean function is linear. The basic setup of these models requires the response to have an exponential family conditional on the random effect. This implies the response variable has a probability density function of the form

$$f(y|\eta) = h(y)\exp\{\eta \cdot T(y) - c(\eta)\}$$

where $h(y)$ is a function of only the data Y, $T(y)$ is the canonical statistic, η is the canonical parameter, and $c(\eta)$ is the cumulant function. It is possible to write many popular probability densities in this form, including the normal, Bernoulli, Poisson, geometric, etc. This form provides a method for modeling the mean in terms of the canonical parameter η by finding an appropriate transformation. One of the most useful properties of an exponential family is the derivative of $c(\eta)$ is the mean of Y which implies that it is easy to find this transformation.

2.2 Logit-Normal Mixed Model

Let Y_{ij} be a Bernoulli response for $i = 1, ..., m$ independent subjects, with $j = 1, .., n_i$ (possibly correlated) measurements per subject. The density for a single observation is

$$\begin{aligned} f(y_{ij}|\theta_i) &= [\theta_i]^{y_{ij}}\,[1-\theta_i]^{1-y_{ij}} \\ &= \left[\frac{\theta_i}{1-\theta_i}\right]^{y_{ij}}[1-\theta_i] \\ &= \exp\left\{y_{ij}\log\left[\frac{\theta_i}{1-\theta_i}\right] + \log\left[1-\theta_i\right]\right\} \end{aligned}$$

We make a change of variable to get the canonical form

$$f(y_{ij}|\eta_i) = \exp\left\{y_{ij}\eta_i - \log\left[1+e^{\eta_i}\right]\right\}$$

In this case, the cumulant function is $\log[1+e^{\eta_i}]$. Thus, the derivative is $\frac{e^{\eta_i}}{1+e^{\eta_i}} = \theta_i$. This implies that the transformation need to model the mean is $\eta_i = \log\left[\frac{\theta_i}{1-\theta_i}\right]$ (also known as the logit transformation).

For the logit-normal GLMM, we model this mean transformation as a linear function of fixed and random effects. That is,

$$\eta_i = x_{ij1}\beta_1 + ... + x_{ijp}\beta_p + \alpha_1 + ... + \alpha_r$$

where $\beta_1, ..., \beta_p$ represent the fixed effects and $\alpha_1, ..., \alpha_r$ represent random effects.
Finally, we let each of the random effects be independent and $\alpha_i \sim N(0, \sigma_i^2)$. Thus, the parameters to be estimated in the model are fixed coefficients $\beta_1, ..., \beta_p$ and the variance components $\sigma_1, ..., \sigma_r$.

Level 1: $f(y_{ij}|\eta_i) = \exp\left\{y_{ij}\eta_i - \log\left[1+e^{\eta_i}\right]\right\}$
$\eta_i = x_{ij1}\beta_1 + ... + x_{ijp}\beta_p + \alpha_1 + ... + \alpha_r$

Level 2: $\alpha_i \overset{ind.}{\sim} N(0, \sigma_i^2)$ for all i

2.3 Example of Logit-Normal Mixed Model pertaining to Hypothesis 2

We cannot model Formal Training as a LMM because it is a binary variable (Yes=1 and No=0). However, we can model the logit transform of the probability of a company saying yes[2]. Let's just look at Firm Size and Funding as independent variables. Pretend we have 2 companies and we are only looking at 3 time points. Here is some fake data:
Company 1:

$$\begin{array}{c} \\ \text{S1} \\ \text{S2} \\ \text{S3} \end{array} \begin{array}{c} \begin{array}{c} \text{Formal Training} \end{array} \\ \left(\begin{array}{c} 0 \\ 1 \\ 1 \end{array}\right. \end{array} \begin{array}{|c} \begin{array}{c} \text{Firm Size} \end{array} \\ \begin{array}{r} 5 \\ 10 \\ 12 \end{array} \end{array} \begin{array}{c} \begin{array}{c} \text{Funding} \end{array} \\ \left.\begin{array}{r} 10,000 \\ 100,000 \\ 50,000 \end{array}\right) \end{array}$$

Company 2:

$$\begin{array}{c} \\ \text{S1} \\ \text{S2} \\ \text{S3} \end{array} \begin{array}{c} \begin{array}{c} \text{Formal Training} \end{array} \\ \left(\begin{array}{c} 0 \\ 0 \\ 1 \end{array}\right. \end{array} \begin{array}{|c} \begin{array}{c} \text{Firm Size} \end{array} \\ \begin{array}{r} 15 \\ 16 \\ 18 \end{array} \end{array} \begin{array}{c} \begin{array}{c} \text{Funding} \end{array} \\ \left.\begin{array}{r} 10,000 \\ 15,000 \\ 50,000 \end{array}\right) \end{array}$$

Let θ_i =Probability of a firm saying Yes to having formal training.

Company 1:

$$\log\left[\frac{\theta_1}{1-\theta_1}\right] = \begin{bmatrix} 1 & 5 & 10,000 \\ 1 & 10 & 100,000 \\ 1 & 12 & 50,000 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \alpha_1 \text{ where } \alpha_1 \sim N(0, \sigma^2_{company})$$

Estimates are then doing using maximum likelihood (a standard statistical technique).
Company 2:

$$\log\left[\frac{\theta_2}{1-\theta_2}\right] = \begin{bmatrix} 0 & 15 & 10,000 \\ 0 & 16 & 15,000 \\ 1 & 18 & 50,000 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \alpha_2 \text{ where } \alpha_2 \sim N(0, \sigma^2_{company})$$

Estimates are then doing using maximum likelihood (a standard statistical technique).
The parameters that the model estimates are:

- β_0= Fixed Intercept term
- β_1= Fixed Coefficient of Sales per Employee
- β_2= Fixed Coefficient of Profit Margin
- $\sigma^2_{company}$= Variance of the Random Intercept by Company

[2]This is sometimes called the log-odds.

After estimation is complete, one can back-transform the results to get an estimate of the probability of a yes. For instance,

$$\hat{\theta}_1 = \frac{\exp\left\{ \begin{bmatrix} 1 & 5 & 10,000 \\ 1 & 10 & 100,000 \\ 1 & 12 & 50,000 \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \hat{\alpha}_1 \right\}}{1 + \exp\left\{ \begin{bmatrix} 1 & 5 & 10,000 \\ 1 & 10 & 100,000 \\ 1 & 12 & 50,000 \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \hat{\alpha}_1 \right\}}$$

and

$$\hat{\theta}_2 = \frac{\exp\left\{ \begin{bmatrix} 1 & 15 & 10,000 \\ 1 & 16 & 15,000 \\ 1 & 18 & 50,000 \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \hat{\alpha}_2 \right\}}{1 + \exp\left\{ \begin{bmatrix} 1 & 15 & 10,000 \\ 1 & 16 & 15,000 \\ 1 & 18 & 50,000 \end{bmatrix} \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \hat{\alpha}_2 \right\}}$$

3 Multivariate Linear Mixed Model

The Multivariate Linear Mixed Model (MLMM) is an easy extension of the LMM framework to multiple continuous response variables. The MLMM can be written as an extension of LMM for subject i as

$$\begin{aligned} [Y_i] &= X_i[\beta] + Z_i[\gamma_i] + [\epsilon_i] \\ vec[\gamma_i] &\sim N_{p_{random}}(0, \Psi) \\ \text{Each row of } [\epsilon_i] &\sim N_{n_{responses}}(0, \Sigma) \end{aligned}$$

where[5]

- n_i is the number of observations for subject i
- $[Y_i]$ is a $n_i \times n_{responses}$ matrix of responses for subject i
- X_i is a $n_i \times p_{fixed}$ design matrix for the fixed effects of subject i; characterizes the systematic part of the response
- $[\beta]$ is a $p_{fixed} \times n_{responses}$ matrix of fixed parameters (coefficients)
- Z_i is a $n_i \times p_{random}$ design matrix for the random effects of subject i; characterizes random variation in the response attributable to among-unit sources
- $[\gamma_i]$ is a $p_{random} \times n_{responses}$ matrix of the random effects (coefficients) that completes the characterization of among-unit variation.

[5] $vec()$ vectorizes a matrix by stacking it's columns

- Ψ is the $p_{random} \times p_{random}$ covariance matrix for the random effects
- $[\epsilon_i]$ is a $n_i \times n_{responses}$ matrix of random errors for subject i; these represent within-unit deviations characterizing variation due to sources like within-unit fluctuations and measurement error
- Σ is the $n_{responses} \times n_{responses}$ covariance matrix for each row of errors in subject i

Without conditioning on the random effects, the model becomes:

$$vec[Y_i] \sim N(vec(X_i[\beta]), (I_r \otimes Z_i)\Psi(I_r \otimes Z_i)^T + (\Sigma \otimes I_{n_i}))$$

The assumptions (besides multivariate normality) are:

1. $[\epsilon_i]$ is independent of $[\gamma_j]$ for all i, j
2. rows of $[\epsilon_i]$ and $[\epsilon_1], ..., [\epsilon_n]$ are independent

3.1 MVLMM Example modified from Hypothesis 3

Let's say we just want to make a model with sales growth and number of employees as our responses. Assume we only have 2 companies total and each company has all 4 responses for the 4^{th}-7^{th} surveys. Also, let's just look at a couple of covariates, sales per employee and profit margin and again assume there are no missing values. We will write a model with a random intercept by company. Here is some fictitious data:
Company 1:

	Sales Growth	# of Employees	Sales/Emp	Profit Margin
S4	40,000	10	10,000	0.51
S5	50,000	15	15,000	0.48
S6	80,000	16	13,000	0.35
S7	70,000	17	14,000	0.25

Company 2:

	Sales Growth	# of Employees	Sales/Emp	Profit Margin
S4	500,000	5	50,000	0.35
S5	700,000	10	70,000	0.24
S6	800,000	15	40,000	0.27
S7	850,000	20	42,000	0.28

Model for Company 1:
(Note that we add a row of ones to the X_i - matrix to represent having an intercept in the model.)

$$\begin{bmatrix} 40,000 & 10 \\ 50,000 & 15 \\ 80,000 & 16 \\ 70,000 & 17 \end{bmatrix} = \begin{bmatrix} 1 & 10,000 & 0.51 \\ 1 & 15,000 & 0.48 \\ 1 & 13,000 & 0.35 \\ 1 & 14,000 & 0.25 \end{bmatrix} \begin{bmatrix} \beta_{01} & \beta_{02} \\ \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{12} & \gamma_{13} \end{bmatrix} + \begin{bmatrix} \epsilon_{111} & \epsilon_{112} \\ \epsilon_{121} & \epsilon_{122} \\ \epsilon_{131} & \epsilon_{132} \\ \epsilon_{141} & \epsilon_{142} \end{bmatrix}$$

where

$$\begin{bmatrix}\epsilon_{111} & \epsilon_{112}\end{bmatrix} \sim N_2(\mathbf{0}, \Sigma), \begin{bmatrix}\epsilon_{121} & \epsilon_{122}\end{bmatrix} \sim N_2(\mathbf{0}, \Sigma), \begin{bmatrix}\epsilon_{131} & \epsilon_{132}\end{bmatrix} \sim N_2(\mathbf{0}, \Sigma), \begin{bmatrix}\epsilon_{141} & \epsilon_{142}\end{bmatrix} \text{ and } \sim N_2(\mathbf{0}, \Sigma)$$

and

$$\begin{bmatrix}\gamma_{11}\\ \gamma_{12}\\ \gamma_{12}\\ \gamma_{13}\end{bmatrix} \sim N_4(\mathbf{0}, \Psi),$$

Model for Company 2:

$$\begin{bmatrix}500,000 & 5\\ 700,000 & 10\\ 800,000 & 15\\ 850,000 & 20\end{bmatrix} = \begin{bmatrix}1 & 50,000 & 0.35\\ 1 & 70,000 & 0.24\\ 1 & 40,000 & 0.27\\ 1 & 42,000 & 0.28\end{bmatrix} \begin{bmatrix}\beta_{01} & \beta_{02}\\ \beta_{11} & \beta_{12}\\ \beta_{21} & \beta_{22}\end{bmatrix} + \begin{bmatrix}1 & 1\\ 1 & 1\\ 1 & 1\\ 1 & 1\end{bmatrix} \begin{bmatrix}\gamma_{21} & \gamma_{22}\\ \gamma_{22} & \gamma_{23}\end{bmatrix} + \begin{bmatrix}\epsilon_{211} & \epsilon_{212}\\ \epsilon_{221} & \epsilon_{222}\\ \epsilon_{231} & \epsilon_{232}\\ \epsilon_{241} & \epsilon_{242}\end{bmatrix}$$

where

$$\begin{bmatrix}\epsilon_{211} & \epsilon_{212}\end{bmatrix} \sim N_2(\mathbf{0}, \Sigma), \begin{bmatrix}\epsilon_{221} & \epsilon_{222}\end{bmatrix} \sim N_2(\mathbf{0}, \Sigma), \begin{bmatrix}\epsilon_{231} & \epsilon_{232}\end{bmatrix} \sim N_2(\mathbf{0}, \Sigma), \begin{bmatrix}\epsilon_{241} & \epsilon_{242}\end{bmatrix} \text{ and } \sim N_2(\mathbf{0}, \Sigma)$$

and

$$\begin{bmatrix}\gamma_{21}\\ \gamma_{22}\\ \gamma_{22}\\ \gamma_{23}\end{bmatrix} \sim N_4(\mathbf{0}, \Psi),$$

The parameters that the model estimates are:

- β_{01},β_{11},β_{21}= Fixed Intercept,Coefficient of Sales per Employee, and Coefficient of Profit Margin for Sales Growth
- β_{02},β_{12},β_{22}= Fixed Intercept,Coefficient of Sales per Employee, and Coefficient of Profit Margin for # of Employees
- Ψ= Covariance matrix the Random Intercepts by Company in each response
- Σ= Covariance matrix of a row of Random Errors

4 KFS Sampling Scheme

The KFS used a stratified sampling methodology, which oversampled high-tech firms. The D&B database was partitioned into six sampling strata defined by industrial technology categories (based on industry designation) and gender of the business owner or CEO (based on the D&B data element and supplemented by including businesses whose owners had a feminine first name).

Cross-sectional and longitudinal weights are available in the data. The weight that will be used in the mixed models is the final longitudinal weight represented as `weight_final_f1234567_long_7`. This is defined as % of firms who survived the first to seventh year.

The final longitudinal weights in the KFS are assigned for businesses that:

- responded to the survey in every follow-up from the first follow-up to the seventh follow-up,
- responded to the survey in every follow-up from the first follow-up to the follow-up when they permanently stopped operations or sold or merged, and
- responded to the survey in every follow-up from the first follow-up to the seventh follow-up and have reported that they have temporarily stopped operations in the seventh follow-up.
- `http://www1.kauffman.org/kfs/KFSWiki/ResearchForum.aspx?forumid=17&threadid=106&tpage=2`

5 Modeling Decisions

SAS will be the software of choice for analysis, and in particular, the proc GLIMMIX command will be exclusively used. This procedure can do univariate and multivariate linear and generalized linear mixed models and has the ability to incorporate the longitudinal sampling weights in the data.

In proc GLIMMIX, the SUBJECT= option in the RANDOM statement identifies the clustering structure for the random effects. When the subjects of the multiple RANDOM statements are nested, the model is a multilevel model and each RANDOM statement corresponds to one level. Using a pseudo-maximum-likelihood approach, you can extend the multilevel model framework to accommodate weights at different levels. Several of the estimation procedures discussed below will be tested.

Based on the structure of the model, the GLIMMIX procedure selects the estimation technique for estimating the model parameters. If you fit a generalized linear mixed model, you can change the estimation technique with the METHOD= option in the PROC GLIMMIX statement. The defaults are determined as follows: generalized linear mixed models The default technique is METHOD=RSPL, corresponding to maximizing the residual log pseudo-likelihood with an expansion about the current solutions of the best linear unbiased predictors of the random effects. In models for normal data with identity link, METHOD=RSPL and METHOD=RMPL are equivalent to restricted maximum likelihood estimation, and METHOD=MSPL and METHOD=MMPL are equivalent to maximum likelihood estimation.

5.1 Pseudo-Likelihood

Estimation methods ending in PL are pseudo-likelihood techniques. Pseudo-likelihood methods for generalized linear mixed models can be cast in terms of Taylor series expansions of the GLMM. The expansion locus of the expansion is either the vector of random effects solutions (S) or the mean of the random effects (M). The expansions are also referred to as the Subject-specific and Marginal expansions. The abbreviation PL identifies the method as a pseudo-likelihood technique.

Residual methods account for the fixed effects in the construction of the objective function, which reduces the bias in covariance parameter estimates.

5.2 Maximum Likelihood with Laplace Approximation

If you choose METHOD=LAPLACE with a generalized linear mixed model, PROC GLIMMIX approximates the marginal likelihood by using Laplaces method. Twice the negative of the resulting log-likelihood approximation is the objective function that the procedure minimizes to determine

parameter estimates. Laplace estimates typically exhibit better asymptotic behavior and less small-sample bias than pseudo-likelihood estimators. On the other hand, the class of models for which a Laplace approximation of the marginal log likelihood is available is much smaller compared to the class of models to which PL estimation can be applied.

Because the marginal likelihood of the data is approximated numerically, certain features of the marginal distribution are not availablefor example, you cannot display a marginal variance-covariance matrix. Also, the procedure includes both the fixed-effects parameters and the covariance parameters in the optimization for Laplace estimation.

5.3 Maximum Likelihood with Adaptive Quadrature

If you choose METHOD=QUAD in a generalized linear mixed model, the GLIMMIX procedure approximates the marginal log likelihood with an adaptive Gauss-Hermite quadrature. Compared to METHOD=LAPLACE, the models for which parameters can be estimated by quadrature are further restricted. In addition to the conditional independence assumption and the absence of R-side covariance parameters, it is required that models suitable for METHOD=QUAD can be processed by subjects.

Summary Statistics for Question 1

13:59 Thursday, January 23, 2014 1

The FREQ Procedure

F19b_e_Worker_Training_4	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	1764	79.64	1764	79.64
1	451	20.36	2215	100.00

Frequency Missing = 230

F19b_e_Worker_Training_5	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	1628	80.20	1628	80.20
1	402	19.80	2030	100.00

Frequency Missing = 415

F19b_e_Worker_Training_6	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	1453	80.63	1453	80.63
1	349	19.37	1802	100.00

Frequency Missing = 643

F19b_e_Worker_Training_7	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	1379	81.79	1379	81.79
1	307	18.21	1686	100.00

Frequency Missing = 759

13:59 Thursday, January 23, 2014 2

The MEANS Procedure

Variable	N	N Miss	Mean	Std Dev	Lower Quartile	Median	Upper Quartile
C5_Num_Employees_4	2241	204	4.2016957	14.5483927	0	1.0000000	3.0000000
C5_Num_Employees_5	2050	395	3.9829268	12.3577786	0	1.0000000	3.0000000
C5_Num_Employees_6	1823	622	4.3143171	15.3983494	0	1.0000000	3.0000000
C5_Num_Employees_7	1718	727	5.3806752	28.6190922	0	1.0000000	4.0000000

13:59 Thursday, January 23, 2014 3

The FREQ Procedure

naics_first	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	27	1.10	27	1.10
2	162	6.63	189	7.73
3	407	16.65	596	24.38
4	322	13.17	918	37.55
5	1134	46.38	2052	83.93
6	61	2.49	2113	86.42
7	102	4.17	2215	90.59
8	230	9.41	2445	100.00

13:59 Thursday, January 23, 2014 4

The MEANS Procedure

Variable	N	N Miss	Mean	Std Dev	Lower Quartile	Median	Upper Quartile
sales_per_emp_4	937	1508	217111.04	717078.78	33333.33	89029.33	187500.00
sales_per_emp_5	956	1489	208290.15	700979.65	37188.10	93750.00	188235.29
sales_per_emp_6	854	1591	343315.73	4147648.85	37500.00	95285.71	181312.60
sales_per_emp_7	799	1646	327588.75	2100860.01	45000.00	100524.33	200000.00

The FREQ Procedure

G9_Education_Owner_01_4	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1-3	224	10.46	224	10.46
4-6	747	34.89	971	45.35
7-9	1002	46.80	1973	92.15
10-12	168	7.85	2141	100.00

Frequency Missing = 304

G9_Education_Owner_01_5	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1-3	188	9.66	188	9.66
4-6	675	34.69	863	44.35
7-9	931	47.84	1794	92.19
10-12	152	7.81	1946	100.00

Frequency Missing = 499

G9_Education_Owner_01_6	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1-3	167	9.73	167	9.73
4-6	580	33.80	747	43.53
7-9	835	48.66	1582	92.19
10-12	134	7.81	1716	100.00

Frequency Missing = 729

G9_Education_Owner_01_7	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1-3	145	9.02	145	9.02
4-6	562	34.97	707	44.00
7-9	771	47.98	1478	91.97
10-12	129	8.03	1607	100.00

Frequency Missing = 838

13:59 Thursday, January 23, 2014 6

The FREQ Procedure

C8_Primary_Loc_4	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	1142	50.91	1142	50.91
2	847	37.76	1989	88.68
3	158	7.04	2147	95.72
4	96	4.28	2243	100.00

Frequency Missing = 202

C8_Primary_Loc_5	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	1048	51.05	1048	51.05
2	754	36.73	1802	87.77
3	161	7.84	1963	95.62
4	90	4.38	2053	100.00

Frequency Missing = 392

C8_Primary_Loc_6	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	940	51.45	940	51.45
2	674	36.89	1614	88.34
3	140	7.66	1754	96.00
4	73	4.00	1827	100.00

Frequency Missing = 618

C8_Primary_Loc_7	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	874	50.90	874	50.90
2	635	36.98	1509	87.89
3	143	8.33	1652	96.21
4	65	3.79	1717	100.00

Frequency Missing = 728

13:59 Thursday, January 23, 2014 7

The MEANS Procedure

Variable	N	N Miss	Mean	Std Dev	Lower Quartile	Median	Upper Quartile
asset_level_4	2039	406	741119.61	8064866.90	11325.00	48000.00	196000.00
asset_level_5	1856	589	1048701.40	18082966.76	10000.00	43000.00	195125.00
asset_level_6	1654	791	1094257.72	14461992.46	10000.00	45000.00	209300.00
asset_level_7	1541	904	1972091.41	30388125.18	10000.00	50000.00	224000.00

13:59 Thursday, January 23, 2014 8

The MEANS Procedure

Variable	N	N Miss	Mean	Std Dev	Lower Quartile	Median	Upper Quartile
E1_a_Num_Human_Res_4	1522	923	0.7608410	2.7008346	0	1.0000000	1.0000000
E1_a_Num_Human_Res_5	1408	1037	0.6654830	0.8066191	0	1.0000000	1.0000000
E1_a_Num_Human_Res_6	1191	1254	0.6633081	0.9649255	0	1.0000000	1.0000000
E1_a_Num_Human_Res_7	1140	1305	0.6798246	1.0757671	0	1.0000000	1.0000000

13:59 Thursday, January 23, 2014 9

The FREQ Procedure

hr_4	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	526	34.51	526	34.51
1-2	373	24.48	899	58.99
3-4	340	22.31	1239	81.30
5+	285	18.70	1524	100.00

Frequency Missing = 921

hr_5	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	518	36.66	518	36.66
1-2	345	24.42	863	61.08
3-4	305	21.59	1168	82.66
5+	245	17.34	1413	100.00

Frequency Missing = 1032

hr_6	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	906	53.17	906	53.17
1-2	315	18.49	1221	71.65
3-4	256	15.02	1477	86.68
5+	227	13.32	1704	100.00

Frequency Missing = 741

hr_7	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	418	36.60	418	36.60
1-2	241	21.10	659	57.71
3-4	253	22.15	912	79.86
5+	230	20.14	1142	100.00

Frequency Missing = 1303

Summary Statistics for Question 3

09:44 Thursday, January 30, 2014 1

Worker Training in 2011 (Survey 7)

The MEANS Procedure

Variable	N	N Miss	Mean	Std Dev	Lower Quartile	Median	Upper Quartile
sales_growth_4	1537	1335	33.6476536	1243.05	0.7142857	1.0000000	1.3333333
sales_growth_5	1558	1314	3.0435929	40.6930107	0.6547619	0.9565217	1.2350664
sales_growth_6	1657	1215	13.2386142	361.5742119	0.7714286	1.0000000	1.3574468
sales_growth_7	1554	1318	2.8816653	19.6258995	0.7911111	1.0330833	1.4000000

09:44 Thursday, January 30, 2014 2

Worker Training in 2011 (Survey 7)

The MEANS Procedure

Variable	N	N Miss	Mean	Std Dev	Lower Quartile	Median	Upper Quartile
F24_Profit_Amt_4	1525	1347	109411.66	499936.33	5000.00	20000.00	60000.00
F24_Profit_Amt_5	1438	1434	100117.13	607043.60	3500.00	15000.00	50000.00
F24_Profit_Amt_6	1342	1530	219874.76	4711457.25	4560.00	19000.00	50000.00
F24_Profit_Amt_7	1320	1552	522030.66	13787732.42	5000.00	20500.00	70000.00

09:44 Thursday, January 30, 2014 3

Worker Training in 2011 (Survey 7)

The MEANS Procedure

Variable	N	N Miss	Mean	Std Dev	Lower Quartile	Median	Upper Quartile
growth4	1361	1511	0.9414271	1.3722454	0.5000000	1.0000000	1.0000000
growth5	1266	1606	0.9238854	1.1536000	0.5000000	1.0000000	1.0625000
growth6	1077	1795	1.1832230	7.2603907	0.6250000	1.0000000	1.2000000
growth7	983	1889	0.9783275	0.8381825	0.5833333	1.0000000	1.1666667

09:44 Thursday, January 30, 2014 4

Worker Training in 2011 (Survey 7)

The MEANS Procedure

Variable	N	N Miss	Mean	Std Dev	Lower Quartile	Median	Upper Quartile
asset_level_4	2393	479	703837.50	7521616.88	11000.00	47600.00	200000.00
asset_level_5	2186	686	1004081.53	16828572.32	10000.00	43600.00	200000.00
asset_level_6	1952	920	1447835.57	24324440.51	9400.00	45000.00	210000.00
asset_level_7	1818	1054	2357038.13	34092974.48	10000.00	51250.00	237000.00

09:44 Thursday, January 30, 2014 5

Worker Training in 2011 (Survey 7)

The MEANS Procedure

Variable	N	N Miss	Mean	Std Dev	Lower Quartile	Median	Upper Quartile
sales_per_emp_4	1103	1769	217637.11	677456.60	34144.00	91000.00	192830.77
sales_per_emp_5	1132	1740	200881.70	650183.75	36666.67	93848.00	188927.86
sales_per_emp_6	1008	1864	350572.73	3921824.17	35333.33	94875.25	182952.38
sales_per_emp_7	946	1926	325134.12	1969905.32	42000.00	100000.00	200000.00

09:44 Thursday, January 30, 2014 6

Worker Training in 2011 (Survey 7)

The MEANS Procedure

Variable	N	N Miss	Mean	Std Dev	Lower Quartile	Median	Upper Quartile
profit_margin_4	1269	1603	0.3123211	1.2452791	0.0551724	0.1579587	0.4000000
profit_margin_5	1354	1518	0.3323625	1.5552818	0.0533333	0.1632468	0.4000000
profit_margin_6	1294	1578	0.2711616	0.3021595	0.0555556	0.1666667	0.4000000
profit_margin_7	1272	1600	0.3006334	0.9468286	0.0575001	0.1692724	0.4000000

09:44 Thursday, January 30, 2014 7

Worker Training in 2011 (Survey 7)

The MEANS Procedure

Variable	N	N Miss	Mean	Std Dev	Lower Quartile	Median	Upper Quartile
sales_per_capital_4	1814	1058	557.7077843	23478.96	0.9658537	2.4643650	5.5147059
sales_per_capital_5	1978	894	77.5158161	3001.35	0.9090909	2.4713900	5.4761905
sales_per_capital_6	1767	1105	14.4838580	287.6590277	0.8725490	2.4980645	5.4687500
sales_per_capital_7	1658	1214	12.4552362	205.1502790	1.0000000	2.5560334	5.7142857

09:44 Thursday, January 30, 2014 8

Worker Training in 2011 (Survey 7)

The MEANS Procedure

Variable	N	N Miss	Mean	Std Dev	Lower Quartile	Median	Upper Quartile
capital_intensity_4	1336	1536	217210.18	3484021.37	11000.00	32975.00	85000.00
capital_intensity_5	1165	1707	183647.46	1708527.72	10375.00	32500.00	90000.00
capital_intensity_6	1006	1866	191894.44	1191944.16	11937.50	35450.00	89750.00
capital_intensity_7	963	1909	784258.69	12719891.32	11944.44	35333.33	90000.00

09:44 Thursday, January 30, 2014 9

Worker Training in 2011 (Survey 7)

The FREQ Procedure

F19b_e_Worker_Training_4	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	2057	79.15	2057	79.15
1	542	20.85	2599	100.00

Frequency Missing = 273

F19b_e_Worker_Training_5	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	1918	79.85	1918	79.85
1	484	20.15	2402	100.00

Frequency Missing = 470

F19b_e_Worker_Training_6	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	1706	80.36	1706	80.36
1	417	19.64	2123	100.00

Frequency Missing = 749

F19b_e_Worker_Training_7	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	1617	81.09	1617	81.09
1	377	18.91	1994	100.00

Frequency Missing = 878

Summary Statistics for Question 4

Profit

Fixed Effect Testing:

Categorical Version

Variable	Num DF	Den DF	p-value (Pr>F)
survey_index	3	1182	0.2812
hr_emp	2	1182	0.9044
naics_first	7	1182	0.8920
owner_ed	1	1182	0.5341
sales_per_emp	2	1182	<0.0001***
assets	1	1182	0.0035***
num_employees	1	1182	<0.0001***
hr_score	1	1182	0.5195
survey_index*hr_emp	6	1182	0.5109
survey_index*naics_first	21	1182	0.1216
survey_index*owner_ed	3	1182	0.8999
survey_index*sales_per_emp	6	1182	0.7636
survey_index*assets	3	1182	0.5650

Continuous Version

Variable	Num DF	Den DF	p-value (Pr>F)
survey_index	3	1190	0.5715
hr_emp	1	1190	0.9569
naics_first	7	1190	0.3226
owner_ed	1	1190	0.3871

sales_per_emp	1	1190	0.0176**
assets	1	1190	0.7278
num_employees	1	1190	<0.0001***
hr_score	1	1190	0.3238
survey_index*hr_emp	3	1190	0.4077
survey_index*naics_first	21	1190	0.2075
survey_index*owner_ed	3	1190	0.9005
survey_index*sales_per_emp	3	1190	<0.0001***
survey_index*assets	3	1190	<0.0001***

Random Effects Testing/Estimates:

Categorical

Variance Component	Subject	Estimate	p-value (Pr>Z)
Intercept	MPRID	0	.
AR(1) (error correlation)	MPRID	0.7123	<0.0001***
residual	MPRID	2.689*10^11	<0.0001***

Continuous

Variance Component	Subject	Estimate	p-value (Pr>Z)
Intercept	MPRID	0	.
AR(1) (error correlation)	MPRID	0.7264	<0.0001***
residual	MPRID	2.633*10^11	<0.0001***

Employment Growth
Fixed Effect Testing:
Categorical Version

Variable	Num DF	Den DF	p-value (Pr>F)
survey_index	3	1572	0.3263
hr_emp	2	1572	0.6658
naics_first	7	1572	0.0391**
owner_ed	1	1572	0.7662
sales_per_emp	2	1572	<0.0001***
assets	1	1572	0.2304
num_employees	1	1572	<0.0049***
hr_score	1	1572	0.0184**
survey_index*hr_emp	6	1572	0.7250
survey_index*naics_first	21	1572	0.0171**
survey_index*owner_ed	3	1572	0.2371
survey_index*sales_per_emp	6	1572	0.4991
survey_index*assets	3	1572	0.7842

Continuous Version

Variable	Num DF	Den DF	p-value (Pr>F)
survey_index	3	1580	0.5213
hr_emp	1	1580	0.4887
naics_first	7	1580	0.0480**
owner_ed	1	1580	0.8667
sales_per_emp	1	1580	0.1956

assets	1	1580	0.8654
num_employees	1	1580	<0.0005***
hr_score	1	1580	0.6153
survey_index*hr_emp	3	1580	0.2359
survey_index*naics_first	21	1580	0.0200**
survey_index*owner_ed	3	1580	0.3155
survey_index*sales_per_emp	3	1580	0.2695
survey_index*assets	3	1580	0.9174

Random Effects Testing/Estimates:

Categorical

Variance Component	Subject	Estimate	p-value (Pr>Z)
Intercept	MPRID	0.03286	0.0025***
AR(1) (error correlation)	MPRID	-0.2814	<0.0001***
residual	MPRID	0.4449	<0.0001***

Continuous

Variance Component	Subject	Estimate	p-value (Pr>Z)
Intercept	MPRID	0.02622	0.0103**
AR(1) (error correlation)	MPRID	-0.2827	<0.0001***
residual	MPRID	0.4582	<0.0001***

Assets
Fixed Effect Testing:
Categorical Version

Variable	Num DF	Den DF	p-value (Pr>F)
survey_index	3	1860	0.1663
hr_emp	2	1860	<0.0001***
naics_first	7	1860	0.5832
owner_ed	1	1860	0.6556
sales_per_emp	2	1860	0.0889*
~~assets~~	NOT INCLUDED	NOT INCLUDED	NOT INCLUDED
num_employees	1	1860	0.0059***
hr_score	1	1860	0.1545
survey_index*hr_emp	6	1860	<0.0001***
survey_index*naics_first	21	1860	0.1842
survey_index*owner_ed	3	1860	0.1936
survey_index*sales_per_emp	6	1860	0.0809*
~~survey_index*assets~~	NOT INCLUDED	NOT INCLUDED	NOT INCLUDED

Continuous Version

Variable	Num DF	Den DF	p-value (Pr>F)
survey_index	3	1868	0.1976
hr_emp	1	1868	0.0984*
naics_first	7	1868	0.6917
owner_ed	1	1868	0.6957

sales_per_emp	1	1868	<0.0001***
~~assets~~	NOT INCLUDED	NOT INCLUDED	NOT INCLUDED
num_employees	1	1868	<0.0001***
hr_score	1	1868	0.2883
survey_index*hr_emp	3	1868	0.0855*
survey_index*naics_first	21	1868	0.0899*
survey_index*owner_ed	3	1868	0.2397
survey_index*sales_per_emp	3	1868	<0.0001***
~~survey_index*assets~~	NOT INCLUDED	NOT INCLUDED	NOT INCLUDED

Random Effects Testing/Estimates:

Categorical

Variance Component	Subject	Estimate	p-value (Pr>Z)
Intercept	MPRID	6.499*10^12	0.3268
AR(1) (error correlation)	MPRID	0.05644	0.0781*
residual	MPRID	7.62*10^14	<0.0001***

Continuous

Variance Component	Subject	Estimate	p-value (Pr>Z)
Intercept	MPRID	4.114*10^12	0.3653
AR(1) (error correlation)	MPRID	0.07285	0.0258**
residual	MPRID	5.83*10^14	<0.0001***

Profit Margin
Fixed Effect Testing:
Categorical Version

Variable	Num DF	Den DF	p-value (Pr>F)
survey_index	3	1182	0.7125
hr_emp	2	1182	0.2869
naics_first	7	1182	0.6609
owner_ed	1	1182	0.5143
sales_per_emp	2	1182	<0.0001***
assets	1	1182	0.9261
num_employees	1	1182	0.1549
hr_score	1	1182	0.7490
survey_index*hr_emp	6	1182	0.0254**
survey_index*naics_first	21	1182	0.8300
survey_index*owner_ed	3	1182	0.9648
survey_index*sales_per_emp	6	1182	0.1009
survey_index*assets	3	1182	0.7488

Continuous Version

Variable	Num DF	Den DF	p-value (Pr>F)
survey_index	3	1190	0.9553
hr_emp	1	1190	0.6137
naics_first	7	1190	0.8892
owner_ed	1	1190	0.5495
sales_per_emp	1	1190	0.5664

assets	1	1190	0.9275
num_employees	1	1190	0.0906*
hr_score	1	1190	0.0069**
survey_index*hr_emp	3	1190	0.5706
survey_index*naics_first	21	1190	0.9960
survey_index*owner_ed	3	1190	0.9615
survey_index*sales_per_emp	3	1190	0.5102
survey_index*assets	3	1190	<0.0001***

Random Effects Testing/Estimates:

Categorical

Variance Component	Subject	Estimate	p-value (Pr>Z)
Intercept	MPRID	0.000920	0.4870
AR(1) (error correlation)	MPRID	0.01411	0.7599
residual	MPRID	0.9023	<0.0001***

Continuous

Variance Component	Subject	Estimate	p-value (Pr>Z)
Intercept	MPRID	0.008738	0.3772
AR(1) (error correlation)	MPRID	0.002711	0.9534
residual	MPRID	0.8864	<0.0001***