ARTICLES:
The Danger of Human Capital Depletion in the Accounting Profession  
Kevin L. James

The Promising Use of Blogs in Graduate Education  
Delaney J. Kirk

An Analysis of Economic and Social Recession-Affected Constructs in the Kansas City Region (2000–2009)  
James Martin, Jennifer Ball, Janie Schrum

Collegiate Post-Graduation Venture Creation: What is the Impact of Entrepreneurship Courses?  
Benjamin J. Blackford, Terrence C. Sebora

Ethnic Diversity, Technical Efficiency, Democracy Measures, and Economic Growth in Sub-Saharan Africa  
Seymour Patterson

Hedge Fund Returns, Risk, and Fees and Systematic Equity Factors  
Rosemary Walker, Robert A. Weigand

A Comparison of Multiple Regression and Neural Networks for Forecasting Real Estate Values  
Daniel Friesen, Mike Patterson, Bob Harmel

Electronic Pedigrees and Counterfeit Pharmaceuticals: The U.S. Experience with RFID  
Bill Roach, Gene Wunder

BOOK REVIEWS:
Higher Education? How Colleges Are Wasting Our Money and Failing Our Kids—and What We Can Do About It  
Andrew Hacker and Claudia Dreifus  
Reviewed by Cindy Kenkel

Available online at www.nwmissouri.edu

Volume 30  
May 2011
CONTENTS

Articles:

The Danger of Human Capital Depletion in the Accounting Profession ..................... 1
Kevin L. James

The Promising Use of Blogs in Graduate Education ............................................. 11
Delaney J. Kirk

An Analysis of Economic and Social Recession-Affected Constructs in the Kansas City Region (2000–2009). ................................. 32
James Martin, Jennifer Ball, Janie Schrum

Collegiate Post-Graduation Venture Creation: What is the Impact of Entrepreneurship Courses?.................. 52
Benjamin J. Blackford, Terrence C. Sebora

Ethnic Diversity, Technical Efficiency, Democracy Measures, and Economic Growth in Sub-Saharan Africa....................... 78
Seymour Patterson

Hedge Fund Returns, Risk, and Fees and Systematic Equity Factors....................... 96
Rosemary Walker, Robert A. Weigand

A Comparison of Multiple Regression and Neural Networks for Forecasting Real Estate Values. .......................... 114
Daniel Friesen, Mike Patterson, Bob Harmel

Electronic Pedigrees and Counterfeit Pharmaceuticals: The U.S. Experience with RFID................................. 137
Bill Roach, Gene Wunder

Book Review:

Higher Education? How Colleges Are Wasting Our Money and Failing Our Kids—and What We Can Do About It .................. 163
By Andrew Hacker and Claudia Dreifus
Reviewed by Cindy Kenkel

Available online at www.nwmissouri.edu

Volume 30 May 2011
Manuscript Submissions
The Regional Business Review welcomes manuscripts in the fields of management, marketing, finance, economics, accounting, management information systems, and other business-related fields. Interested authors should contact the editor for manuscript submission guidelines:

Terry Coalter, JD, PhD, Editor
Booth College of Business and Professional Studies
Northwest Missouri State University
800 University Drive
Maryville, MO 64468-6001
coalter@nwmissouri.edu
RBRv@nwmissouri.edu

from Northwest Missouri State University:
Casey Abington
Benjamin Blackford
Ben Collier
Al Kelly
Gerald Kramer
Chi Lo Lim

from other institutions:
Larry Acker, Harris-Stowe State University
Stephen Allen, Truman State University
Donald Baack, Pittsburg State University
John Cherry, Southeast Missouri State University
Christopher Decker, University of Nebraska–Omaha
Bruce Domazlicky, Southeast Missouri State University
Mary Goebel-Lundholm, Peru State College
Rebekah Heath, Pittsburg State University
Han Kang, Illinois State University
John David Kerr, Harris-Stowe State University
Brad Kleindl, Park University
Larry Lawson, Missouri Western State University
Frank Markham, Mesa State College
Marcel Minutolo, Robert Morris University
Willie Redmond, Southeast Missouri State University
Victor E. Sower, Sam Houston State University
Clarence Von Bergen, Southeastern Oklahoma State University

The editor gratefully acknowledges assistance from the following reviewers

Steve Ludwig
Jody Strauch
Deborah Toomey
Frank Veeman
Tekle Wanorie
Mike Wilson
The Danger of Human Capital
Depletion in the Accounting Profession

Kevin L. James
North Carolina A & T State University

ABSTRACT
The last several years have brought challenging trends related to human capital that seriously threaten the future effectiveness of the accounting profession. If not adequately addressed, these problems could result in a lower level of knowledge and proficiency among company accountants and among auditors at accounting firms that audit public and private companies of all sizes. The impact on the nation’s financial markets could be substantial. The efficiency and effectiveness of the U.S. financial market system depends on providing quality financial information from companies to financial statement users who invest in those companies. A decrease in the overall quality of the pool of accounting graduates could significantly impair the profession’s ability to provide such information. The current article discusses three trends in human capital that seriously threaten the future quality of accounting graduates entering the profession: (1) a shortage of qualified PhD graduates in accounting (2) the use of outdated teaching materials in academic accounting programs, and (3) a lack of diversity in the accounting workforce. The author provides practical strategies for addressing each of these threats.

Keywords: PhD shortage, diversity, human capital
INTRODUCTION
The last several years have brought many challenges to the accounting profession. Most visible have been the challenges related to restoring the public’s confidence after numerous frauds early in the previous decade. Less visible, but perhaps as troubling, are trends related to human capital that seriously threaten the future effectiveness of the profession. If not adequately addressed, these problems could result in a lower level of knowledge and proficiency among company accountants and among auditors at accounting firms that audit public and private companies of all sizes. The impact on the nation’s financial markets could be substantial. The efficiency and effectiveness of the U.S. financial market system depends on providing quality financial information from public companies to financial statement users who invest in those companies. A decrease in the overall quality of the pool of accounting graduates could significantly impair the profession’s ability to provide such information. Stated succinctly, less proficient auditors can lead to a higher incident of audit failures like those seen at Enron, WorldCom, and others. Such failures can erode the confidence that investors and lenders need to allocate capital efficiently. The current article explores three issues that threaten the sustainability of a high quality accounting and auditing workforce and provides practical strategies for addressing these threats.

THREAT 1: THE SHORTAGE OF PHD GRADUATES
A shortage clearly exists in the number of accounting faculty graduating from PhD programs. According to the 2004 National Study of Postsecondary Faculty, the number of accounting faculty declined 13.3% between 1988 and 2004 (U.S. Department of Education, 2004). The greatest loss, 31.0%, occurred at four-year, non-doctoral-granting universities. These universities train a large percentage of the accounting workforce. During the same period, student enrollment in accounting programs increased 12.3%. As a result, the ratio of students per full-time faculty member increased from 20:1 to 28:1 (American Accounting Association, 2004). The result is larger class sizes and less individual attention to students.

Going forward, the shortage of accounting faculty will be exacerbated by the fact that current accounting faculty are older on average than 10 to 15 years ago. Hence, the shortage will likely get worse in future years. In November 2006, one-third of accounting doctoral faculty were estimated to be at least 60 years old. One-half were 55 years old or older (Hasselback, 2007). Plumlee, Kachelmeier, Madeo, Pratt, and Krall (2006) investigated demand for accounting faculty compared to the number of projected graduates from PhD programs and found that less than half of that demand could be met. The most severe shortages by discipline were in audit and tax where graduates were expected to meet 23% and 27% of demand, respectively.

The result of these trends is understaffed accounting departments, and this problem is expected to worsen. Many departments have suffered from multiple years
of being unable to hire for empty faculty positions due to short supply. Budget
cuts caused by the current economic crisis amplify hiring woes as departments
have less money to pursue graduates who command higher salaries. In addition
to smaller departmental budgets, economic problems have caused an influx of
students who are out of work and retooling for an economic recovery. When a
short-handed department faces such increased student enrollment, it must find
ways to educate more students with fewer faculty members. Therefore, depart-
ments must increase class sizes which decreases attention that can be given to
each individual student. Furthermore, faculty of such departments often must
教 more classes leaving less time to prepare for classes and less time to per-
form research that keeps them current on developments in the profession. Larger
class sizes can also impact the mix of requirements in a course such that the
number of beneficial but time-consuming assignments (e.g., cases and research
reports) may be minimized or eliminated. These factors are combining to de-
crease the overall quality of instruction delivered to accounting students.

Recommendations:
Addressing this trend depends on increasing the number of students choosing
to obtain PhDs in accounting. The most significant barrier to entering a PhD
program is cost, both actual cost and opportunity costs. Most PhD students are
recruited from professional jobs in accounting. As such, most of them are earn-
ing in excess of $60,000 per year. Choosing to enroll in a PhD program means
forgoing this income in exchange for a graduate assistantship stipend that gener-
ally pays less than $20,000 per year. Furthermore, over the last several years the
length of time required to complete a doctoral program has increased at many
institutions, sometimes taking six to seven years instead of the more traditional
four to five years. Also, the potential student is often not the only person be-
ing asked to make this sacrifice because many potential applicants have spouses
and children. For many the perceived benefits of such a decision simply do not
outweigh the costs. To increase PhD applicants, the profession must find ways to
mitigate the costs of such a decision, thereby altering the cost/benefit relation-
ship. Three possible ways to do this follow:

1. State accounting societies should consider following the lead of the
California Society of CPAs by instituting a program to recruit ac-
tounting PhD students. The California CPAs Doctoral Scholarship
provides $10,000 per year for up to three years to accounting PhD
students. In return, recipients must agree to join the faculty of a four-
year California university for at least three years. Recipients who
do not fulfill this three year obligation must repay $10,000 for each
unfulfilled year. Adopting such a program may be challenging in
tough economic times, but may be worth the effort as it will also give
adopting states an edge in closing the recruiting gap in their states
(California Society of CPAs, 2008).
2. Schools of accountancy and colleges of business may consider “grow your own” programs in which they recruit high-achieving students from their undergraduate and master’s programs in accounting. These students could receive financial assistance from the school or college in exchange for an agreement to serve at the institution for a prescribed number of years. Again, the award would convert to a loan in the event the commitment is unfulfilled. Current school budgets will make implementation of such a program unlikely without aggressively pursuing private funding. Therefore, implementing this strategy would depend on clearly articulating the problem to company recruiters and helping them understand the danger it poses to the quality of accounting education at the institution. Recruiters must be convinced of the value to their companies of investing in the program in such a way.

Accounting firms and other business organizations should support the American Institute of Certified Public Accountants’ (AICPA) Doctoral Scholars program. This new program aims to help reverse the shortage of PhD accounting faculty in colleges and universities by recruiting current employees with public accounting firms to doctoral programs in accounting. The program will provide an annual stipend of $30,000 for up to four years. The AICPA intends to provide funding for up to 30 new candidates each year for a total of 120 newly trained PhDs in audit and tax (Accounting Doctoral Scholars, 2008).

THREAT 2: OUTDATED TEACHING MATERIALS

Entrants into the accounting profession today are facing an increasingly complex financial environment. Companies continue to engage in more complex transactions than ever, utilizing new financial instruments and arrangements. Globalization of American business and new technologies have also added significant complexity. To be prepared to enter today’s business world, students need exposure to the newest standards, technologies, and transaction types. However, this new cutting-edge information is not being incorporated in textbooks in a timely manner.

Developing and revising teaching materials involves a significant investment of time and resources on the part of textbook publishers. As a result, producers of accounting textbooks conservatively weigh the risks and rewards before publishing new materials (U. S. Department of the Treasury, 2008). Conservative analyses cause accounting textbook manufacturers to lag behind current developments in accounting, creating out-dated teaching materials. Incorporating new developments in the classroom becomes the responsibility of the professor. However, the professor is often already burdened by large class sizes and increased teaching loads (see above) which make it difficult to find time to research and incorporate additional materials. Furthermore, many professors have limited opportunity to educate themselves on new developments due to a lack of ready
sources of information or lack of departmental funds to send faculty to appropriate training. As a result, students graduate less prepared for what they may encounter when they join the workforce. Current examples of topics to which students are underexposed include International Financial Reporting Standards and eXtensible Business Reporting Language (commonly known as XBRL).

The American Accounting Association and the AICPA has formed the Pathways Commission charged with establishing a national strategy for the future of accounting education. Their goal is to craft a model with structure and content that better responds to issues that pose a threat to the development of accounting professionals like those noted above (American Accounting Association, 2010). Some additional, near-term recommendations to address these issues follow:

**Recommendations:**

1. Public accounting firms and academic institutions should consider forming alliances to provide courses and/or seminars to students on current topics that would supplement materials in textbooks. Firms typically have several trained staff members with expertise and experience in current accounting topics. Forming instructional alliances would allow universities to take advantage of firm expertise to provide current information to students. Firms would benefit from having improved relationships with the university and its students which will benefit recruiting efforts. To be effective, such arrangements would go beyond the one-time class visit for an hour long presentation we often see now. These alliances would likely entail multiple coordinated visits to a course or to a campus for all sections of a course, or coordinated times when students could visit firm offices. An efficient alternative would be online or DVD materials produced by firms for classroom use. Such alliances would also help academics and practitioners be in better communication to make sure professors know how important certain topics are and when they need to be incorporated more fully into accounting curricula.

2. Firms and university accounting departments could partner to better inform professors on new and emerging accounting topics. For example, firms could invite selected faculty to firm training events on current topics. A more involved possibility is offering faculty internships where a faculty could receive practical experience dealing with issues they can take back to the classroom.

3. Faculty members and members of the accounting profession should make stronger efforts to make textbook publishers understand the importance of updating materials promptly to reflect the most recent developments in the profession. This entails shifting publishers perceptions of the risk/reward tradeoff. For example, faculty members
can make clear their intentions to adopt textbooks for classroom use when the content remains current over those textbooks that contain outdated material. At a minimum, academicians should make clearer their desire for such updated materials, at least in the form of addendums to the text which could be published online. This low-expense option may be appealing in the current economy and could form a competitive advantage to publishers who are able to accomplish this.

**THREAT 3: LACK OF A DIVERSE WORKFORCE**

Over the last 25 years, the United States has experienced profound demographic changes. American businesses have seen their workforces and customer bases become increasingly diverse, and these trends are expected to continue in upcoming years. In response, a perusal of CPA firm websites will reveal that many accounting firms have undertaken significant diversity initiatives. The chief purpose of these diversity initiatives has changed over time from providing access to the profession to recognition that diversity makes good business sense (Hammond & Paige, 1999). Accounting firms have come to realize a staff that is diverse with respect to gender and ethnicity can provide better services and solutions to their clients. A more diverse accounting firm can provide a better understanding of the cultures of diverse businesses and can make more effective decisions by bringing diverse backgrounds, experiences, and perspectives into the decision-making process.

Despite some efforts toward diversification over the years, gains in the profession have been meager. In 2007, 13% of accounting graduates from bachelor’s programs were from underrepresented ethnic populations (African American, Hispanic American, and Native American). Only 5% of CPA firm professionals were from these same populations. Moreover, only 1% of CPAs with public accounting firms were African American while 2% were Hispanic American. None were Native American (AICPA, 2008). Women are being hired by accounting firms at representative levels. Females represented 52% of accounting graduates in 2007 and 52% of new hires. However, current data shows that they are not being retained. Females represent 45% of current CPA firm employees, 37% of CPAs employed by accounting firms, and only 23% of firm partners (AICPA, 2008). As these women leave public accounting, the profession loses a wealth of skills and intellectual capital as well as unique perspectives they can bring to decision-making. The hiring and retention of women and ethnic minorities should be emphasized to enrich the pool of human capital. Some ways to do this follow. Each recommendation below would require a firm to incur costs in redesigning aspects of its recruiting strategy and redeploying recruitment personnel. Each firm must weigh the costs and benefits of making such a investment in current economic conditions.
Recommendations:

1. Firms could benefit from altering or expanding current recruiting plans. For example, many accounting firms do little or no recruiting at historically black colleges and universities (HBCUs). Yet, approximately 14% of all African American students enrolled in higher education are enrolled at HBCUs (U. S. Department of Education, 2009). Recruitment at these schools, or predominately Hispanic American or Native American schools, could focus recruiting efforts on high quality minority graduates. Accounting firms may also consider recruiting talented minority students from liberal arts degrees and providing scholarships for them to obtain graduate degrees in accounting. In this way, they may access talented individuals who never considered accounting as a career, but who may be looking for a career that provides more stability and income than careers for which they are currently eligible.

2. Firms can start outreach initiatives to minority high school students. This tactic could prove especially beneficial since many students make their choices of careers prior to college enrollment (AICPA, 2000). One way to reach these students is by partnering with organizations that currently serve them. For example, the mission of INROADS is to develop and place talented minority youth in business and industry. Thus, INROADS students are pre-screened to find the most talented individuals and those who are specifically interested in business. Firms can expose these students to accounting and to their firms by actively partnering with INROADS or similar organizations by offering workshops to their students. Some firms have experimented with their own programs to reach high-achieving minority high school students. For example, PricewaterhouseCoo- pers (PwC) has reached out to high school students by pairing them with professional mentors and providing workshops to give guidance on colleges and careers (PwC, 2008). As more firms undertake such efforts, the pipeline of high quality minority students interested in accounting should grow.

3. To better retain women and minorities, firms can develop visible programs that seek to address barriers to retention. A study by the nonprofit organization Catalyst indicated that one of the primary factors that lead to retention of women of color is management’s role in diversity efforts (“Color Barriers,” 2005). Thus, while the firm’s programs will develop firm-specific ways to retain talent, simply having a visible program aids retention by making staff feel the firm is committed to diversity and values their contribution. Studies also show that a lack of influential mentors is a primary barrier to retention for both minorities and women (“Women Still Underrepresented,” 2007). Therefore, firms should seek to establish strong mentoring
programs to enhance retention. Flexible work arrangements have proven to be an advantage in the retention of women and should be considered by firms which do not have such programs. However, research shows that some women who might benefit from such programs still choose to depart because they believe people taking advantage of the programs are underutilized, i.e., given jobs of lesser significance ("Women Are Not On Par(tner)," 2007). Efforts should be made to ensure people taking advantage of flexible work arrangements are treated equally with other staff.
REFERENCES


The Promising Use of Blogs in Graduate Education

Delaney J. Kirk
University of South Florida Sarasota–Manatee

ABSTRACT
Faculty and universities alike are recognizing the need to change the way we teach in order to build the skill sets and competencies needed for our graduates to succeed in today’s global society. Building on the concepts of social learning theory, social constructivism, and connectivism, we are embracing more learner-centered educational opportunities. This paper addresses ways in which blogs can be used to create, evaluate, and share knowledge. There were over 133 million blogs in existence in 2010 and bloggers tend to be more educated (43% have graduate degrees) and more affluent (25% have household income of over $100K) than the general population. In addition, 350 million people around the world read blogs on a regular basis. Due to the success and ubiquity of blogs, the author examines ways in which these could be used to enhance graduate education.

Keywords: blogs, educational technology, graduate education, pedagogy

“Exciting times lie ahead for educators as the oft-desired, but rarely-realized, dream of learner-centered education moves daily closer to reality. Driven by the development of social learning theory and the advancement of participatory web technologies, new opportunities are rapidly becoming apparent. Learning theories, such as constructivism, social constructivism, and more recently, connectivism form the theoretical shift from instructor or institution-controlled teaching to one of greater control by the learner” (Siemens, 2008).
The challenge being faced by faculty and universities today is teaching the competencies needed to succeed in a world that is increasingly characterized by high speed technology, global access, cultural diversity, professional specialization, and change (Holaday, Weaver, & Nilson, 2007; Russell, 1999). However, as Anderson (2010) observes, “Leadership training (and perhaps the MBA as a whole) has become stale. We have become complacent. We haven't demanded more innovation from our MBA programs . . . “ To address this need for change, educators and business leaders in the United States (Wendler, et al., 2010), Canada (Canadian Association for Graduate Studies, 2008), and Japan (Vuksko, 2010) have hosted forums to discuss the future of graduate education. The outcome has been a call for the development of “professional skills” such as creativity, communication, and teamwork which are “behaviors that can be learned, that can be improved with practice, that require reflection, and that benefit from ongoing coaching” (Canadian Association for Graduate Studies, 2008).

To answer this demand, master’s level faculty are redesigning courses to address the current and future needs of students, organizations, and communities. As teachers, we are moving toward becoming “facilitators of knowledge creation instead of disseminators of knowledge” (Bilimoria, 1997, p. 241) and creating “permeable classrooms” where students learn how to learn in the unpredictable environments they will face after graduation (Sandy, 1998). Emphasis goes beyond just knowledge acquisition to building competencies in intrapersonal skills such as adaptability and interpersonal skills such as networking (Boyatzis, Stubbs, & Taylor, 2002). The ability to make connections between the classroom and the “real world” requires that we “shift away from the traditional-centered class to one that allows students to take responsibility for their own learning” (Costello & Brunner, 2008, p. 63). It is anticipated that students coming into graduate programs will be even more demanding in their desire for a meaningful educational experience that will enable them to be successful in their careers and lives. According to a comprehensive study of 50 master’s programs across the United States, students expect their courses to provide intellectual challenge, cooperative learning in and beyond the classroom, and ways for them to share their experiences, insights, and knowledge with other students and faculty in order to learn from one another (Conrad, Haworth, & Millar, 1993).

The development and ease of use of participatory web technologies make these a natural way to move from an instructor-centered educational process to one that gives the learner more control over his/her own learning. One promising tool I have found for successful teaching and learning at the master’s level is the weblog (blog). My example is taken from four years of experimentation with MBA students at two universities. However, I believe that blogs can be useful for teaching graduate classes in most degree programs given the current demand for up-to-date information, cooperative and experiential learning, and formats in which learning more about one’s self and others is pertinent to maximizing the learning experience. Before I present details about the processes and outcomes of
using a class blog, I will briefly outline what blogs are and my choice to use these as a pedagogical tool.

**WHAT IS A BLOG?**
The term, blog (or weblog), is defined as a “frequently updated website, normally with dated entries and usually with the newest entries at the top” (“Blog,” 2008). Blogs differ from traditional websites in several distinct ways (Quibble, 2005). Websites tend to be static and do not change often; blogs are dynamic and are usually added to frequently by the author (or authors). Also, the readers of a blog can respond to the writer’s text by making comments that can then be read by other readers. In addition, blogs are easy to create and add content to because they do not require software programming expertise. As a result, their use has grown from 150 blogs in existence in the late 1990s (Trammell & Ferdig, 2004) to over 133 million by 2010 (Sobel, 2010), with tens of thousands of new ones being added each day (“About Technorati,” n.d.). Blogs are an effect of and a force behind the “most explosive outbreak in the information world since the Internet itself” (Baker & Green, 2008, p. 1). Because most information today is in digital format, it is readily posted to and accessed in the blogosphere. Almost 350 million people globally read blogs (Singer, 2009) and bloggers tend to be more affluent and educated than the general population. Forty-three percent of bloggers have graduate degrees and 25% have a household income of over $100,000 (Sobel, 2010). Thus, due to the success and ubiquity of blogs, I decided to examine how these could be used to enhance graduate education.

**CASE STUDY:**
**ONE PROFESSOR’S EXCELLENT ADVENTURE WITH BLOGS**
I have been teaching graduate level courses for almost three decades now. Teaching the course content is relatively easy after all this time, but I am still struggling with some of the intangibles I want to achieve such as:

- How do I get 100% participation from the students in discussions of the readings, cases, and other course materials?
- How do I cope with the “fadeout effect” between classes, especially when classes may meet only once a week or even one weekend a month?
- How do I show students that the theories and concepts they are learning in class apply to a variety of situations they encounter on a daily basis?
- How can I help the students make connections with each other and with experts in their field, thus learning interpersonal skills such as networking that will benefit them both in their graduate program and in their future careers?
How do I assign writing exercises that encourage intrapersonal skills such as self-reflection and growth?

I first decided to use a class blog as part of my course design in the summer of 2006. In addition to the objectives above, I wanted to become competent in the educational technologies that could be useful in teaching. I was intrigued by the article by Sherer and Shea (2002) which showed how various tools such as chat rooms and videoconferencing can provide more interaction between and among students and faculty, support different learning styles, and allow students to take more active roles in their learning. Liebowitz (2003) found an asynchronous bulletin board was useful in showing students how to apply concepts learned in class to specific organizational situations. As he notes, “What impressed me the most about this format was the thorough, insightful, analytical, and well-written paragraphs and pages that the students posted . . .” (p. 82).

Little, though, has been written about how to use blogs as an integral part of academic coursework (Lohnes, 2006). I found Du and Wagner’s (2007) study interesting as they observed that students using blogs were more active in seeking external knowledge sources and explicating information from these, and that one-third of their students did more work on the blog than was required by the syllabus. Higdon and Topaz (2009) have also had success using blogs in getting feedback from their students as to whether they understood the class readings and how these were related to previous concepts they had learned.

I did have some personal experience with blogs as I had developed one in April of 2006 to share my (and others’) tips on teaching effectiveness. I had found my teaching blog very useful as I was able to set up a learning community to share links to articles, videos, and other blogs as well as to network with faculty in my field from all over the world. I hoped the blogging technology would help my graduate students develop their own intrapersonal and interpersonal competencies in addition to helping them see connections between what they were learning and their own and others’ experiences.

My first decision was to decide where to host the blog. One possibility was to use the university’s course-management software as students are used to this format and perceive it favorably (Korchmaros & Gump, 2009). While this is certainly doable and I do use Blackboard to post grades, I saw several downsides to its use for a class blog. Access back to a course (and thus to the class blog) is limited after a semester is over. In fact, my university, like many others, deletes old course information after a couple of years. Blackboard also seems to be slow and even inaccessible at certain times of the semester. I wanted to have complete control over the blogs I would be developing, and believed an outside site would allow me to do so. However, the most important reason to me to use a site outside the university setting is that I am teaching students skills that they can use after graduation. Knowledge of Blackboard technology will not be useful in the “real world” as much as skills using ongoing Internet sites would be.
There are numerous Internet sites available that provide free or inexpensive blog platforms such as:

- Blogger (http://www.blogger.com) is Google’s tool for publishing.
- Posterous (http://posterous.com) is very simple to use.
- Tumblr (http://www.tumblr.com) has lots of applications and features.
- Edublogs (http://edublogs.org) is specifically set up for teachers and students.
- Edmodo (http://www.edmodo.com) is another site set up for teachers and students that emphasizes password protection of groups of blogs.
- Typepad (http://www.typepad.com) has a free trial period and then low cost per month. Also, includes tech support for users.

I decided to use Typepad as I was familiar with it through my teaching blog and liked the professional look and ease of use of this site. Typepad allows me unlimited blogs for about $180 a year which has been useful over the past four years. However, all of the sites listed previously are user-friendly and can be set up easily.

**INSTRUCTIONAL BLOGGING**

The next decision to make before introducing a class blog to my students was to determine the focus it should take to help in achieving my learning objectives. As noted by Edbauer (2004, p. 1), the typical response she got from colleagues when she showed them some of the blogs on the Internet was, “Cool! But, how in the world would I actually use this” in my class? How blogs are used depends on the professor’s teaching style, the course subject matter, and the comfort level the instructor has with giving up some of the control in the classroom.

Approaches to class blogs can vary from instructor-focused, featuring one-way communication from faculty to students; learner-focused, using two-way communication between faculty and students; or an interactive community-focused teaching tool between and among the students and the professor, as well as others identified as experts to create a virtual learning environment (Kirk & Johnson, 2010). In other words, blogs can be used as a more efficient method to communicate information, as a way to get increased participation from students, and/or as a method of creating a virtual learning community.
Instructor-focused
The simplest way to use a blog is as a one-stop source where the professor posts syllabi, class agendas, announcements, assignments, and links to articles and websites for the students to read. Faculty retain ownership of the site and students are expected to access the blog on a regular basis to obtain class information. The blogs allow the professor to be more efficient as he or she does not have to make copies of course materials and can point students to the blog for answers to questions on assignments. Sample papers or grading rubrics can be posted to guide students in understanding expectations. The blog is also a good way to hyperlink to visual aids used in class so the students can access these. An instructor-focused blog may be the best approach for classes that do not involve a lot of discussion. As noted by a colleague who teaches finance, “I deal with questions such as, ‘what’s the present value of $500 received five years from now if the interest rate is 8%?’ The one and only right answer: $340. Not much discussion needed here.”

Learner-focused
However, it is easy to move beyond “the student as passive recipient of content” stage to a more learner-focused approach by posting weekly questions or links to relevant articles and requiring the students to comment on these. Even the finance colleague mentioned earlier could link to current articles and have the students read and comment, thus allowing the professor to see if the students understood the concepts being taught.

Thus in this approach, the professor would expect the students to be more active participants so that learning can occur peer-to-peer in addition to teacher-to-student. Students could write their own posts or comment on what they had been learning in class, relating the course topics to work and personal experiences. Students could create their own portfolios of their writing and get feedback from their fellow classmates. They could also create learning journals, a technique strongly advocated by Ramsey (2002) except in this case these would be online and easy to access offsite. Faculty can have students create their own individual blogs on topics either chosen by the instructor or by the students themselves. In this instance, the student creates and owns the blog, and the professor (and potentially classmates) is the commenter, thus giving the student more control over his or her own learning experience.

Sparacio and Witonsky (2006) believe that “teaching can be significantly enhanced by the integration of blogging” (p. 2). They advocate a learner-focus by having their students develop their own blogs on a specific topic of their choice that they then write about during the course of the semester. According to them, the students tend to put more effort and care in their work as they got to choose their subject matter and they know that their peers in the class are reading their work.
Tom Nelson uses group blogs in his classes at the University of Texas. Each “blogger group” writes on subjects such as politics or technology. Nelson allows his student groups to decide on their own goals and guidelines for posting and commenting on the blogs and finds this helps in building a collaborative learning environment (Edbauer, 2004).

Community-focused
A third approach to using blogs is to involve participants from outside the class itself. Students could be required to find, read, and evaluate blogs from “experts” outside class on assigned topics and then to share this information with their classmates. They can also develop questions related to course topics and contact others outside of class who participate as “guest teachers.”

In a class I taught on human resource management, we were discussing how to manage people who have disabilities. The students had a number of questions—many of which I was unable to answer. I used the Internet to identify and invite two individuals—one who was blind and another whose legs were paralyzed—to participate and then posed the students’ questions to those individuals. The students were able to get first-hand knowledge to questions they had such as:

I’d like to know what the best technique is to offer help to a disabled person without making him/her uncomfortable or even getting on his/her nerves. This is especially tricky when the disabled person is an introvert and is uncomfortable with asking for help. Is there some middle ground where both parties can meet and be comfortable?

It would be helpful and interesting for me to know about past experiences (very good or bad) you have had in the workplace that you would not mind sharing. This would help me understand how I can make the work environment more comfortable.

Do people who are blind get scared when they hear noises from people or objects near them if they didn’t know they were there? How do they not get hurt on a daily basis?

Obviously, having someone who has a disability to answer the students’ questions contributed greatly to the learning process and will enable the students to become better managers after graduation. The students took ownership of the learning process by posting their own questions on the blog and then reading and commenting on the answers shared by the outside participants. In addition, they were able to share their own personal and work stories with their classmates. As noted by Gould (2002), “research evidence and common sense tell us that students learn more about a subject when they are emotionally involved with the
intellectual context being taught” (p. 23). Thus, designing the course to provoke curiosity and emotional reactions can bring about insights and learning.

In some of my MBA classes, I have contacted authors of books that the students were reading and asked them to contribute comments and feedback on the class blog. I find for the most part that authors are enthusiastic about being part of the learning process. In addition, the students were very excited to have meaningful interaction with the person whose book they were discussing. Opening up the learning process to include personal connections with book and journal authors, subject matter experts with their own blogs, and others including potential employers created a virtual learning community unlike anything I had ever done or even envisioned in the physical classroom. One student summed up the experience as follows:

I thought it was so great to have the author of one of the books in our class write comments on our class blog. I never realized that published authors would be interested in participating. It gave a whole new meaning to my reading of his book.

So which blog focus would I recommend? Obviously, the instructor-focused blog is the easiest to start with and allows the instructor the most control. Typically, I use this approach for my undergraduate classes. I post the class syllabus, schedule, and classroom expectations there and use it as the major source of communication between the students and me. Each class period I post the class agenda, an outline of the lecture module, and their writing assignments. I also put an email link on the blog to encourage the students to contact me with any questions.

However, given the objectives I had for my graduate classes, I wanted more interaction. As noted by Kris (2009), “One of the major attractions for teachers to using blogging is that its focus is not necessarily on the content of the blog but more on the process of constructing and evaluating knowledge helping us reach the sometimes illusive upper levels—analyzing, evaluating, and creating—of Bloom’s Taxonomy.” Thus, each week, I posted a link on the class blog to an article related to that week's class. A typical post then would have one or more questions for the students to answer that required them to use examples drawn from their own experience(s) with the topic. By focusing on student experiences rather than opinions, I was able to keep the conversation concrete and informal rather than abstract and pedantic. The responses then helped me develop a repertoire of shared examples to draw on for further elaboration either online or during class. The students were also instructed to read over their classmates’ comments before our next session together and to respond to these comments. (See Appendix for sample questions used). The following feedback was typical of the students’ response to the weekly assignment:
While at the time I honestly didn’t enjoy the weekly reflection where we had to talk about our personal experiences, I’m glad you made us do them. Having to write every week about me helped me get comfortable in my own skin and helped get me in the groove with having to do writing assignments after being out of school for nine years.

**DID HAVING A BLOG MEET MY CLASS OBJECTIVES?**
I have now used class blogs in my graduate classes for four years. The students’ responses have been overwhelmingly positive to both the specific use of blogs in the classroom and to the use of online learning in general. Although only a handful of the students had ever read or commented on a blog before their class with me, they were comfortable with using the Internet and found the format easy to use. As a side note, I found it interesting that not a single one of my graduate students had their own blog given the proliferation of blogs today. Below are my thoughts and the students’ feedback on how the class blogs helped meet the objectives I stated earlier.

**Participation/Avoiding the “fadeout effect”**
My first goal of getting 100% participation was reached by using a blog. Part of this may have been the fact that 20% of each students’ grade was allocated to the blog assignments. However, I did find that the students tended to take the assignments seriously and to spend more time and effort than required by me. According to their feedback, they felt more comfortable writing their thoughts than speaking up in class, and they felt that the format allowed for more perspectives than they would have heard in a classroom setting.

The blog structure . . . allows us to express our opinions and experiences when we may not otherwise be able to [and] we got to hear a lot of different perspectives on a lot of different subjects.

The level of interaction keeps the class engaged, makes it interesting . . . I love the blog mainly for this reason. It allows us to speak freely and you can really gain something from reading other people’s thoughts.

The blog also kept the classroom conversation alive between classes and helped to avoid the fade-out effect, thus encouraging the students to think about the class even when they were not physically present. Another plus is that it allowed students with different learning styles (for example, visual learners rather than auditory, or those students who would rather reflect before answering) to interact with the class.
I was actually thinking about it [the class] last night. I couldn’t
wait to blog about the experience and hear everyone’s opinions
on yesterday’s speaker.

Often after class ended, I would think of good points that I
could have added to the discussion. However, after the week’s
gap, those ideas are long forgotten. Also, some students are
not as outspoken as others and often their opinions do not get
heard. Having the ability to share these unspoken thoughts
benefits the class as a whole.

**Applying theory to practice**

Using the blog to link to articles, videos, and other blogs exposed the students to
vast amounts of information and in the process made them become more knowl-
edgeable on a topic. As noted by Trammell and Ferdig (2004) doing this on a
regular basis “creates a repetitive process where the blogger student can build on
what he learned last week and find more advanced information for the current
week,” thus increasing both explicit and tacit knowledge.

Reading what our peers have experienced allows us to gain a
better understanding of not only the subject matter, but also
how to handle ourselves in a situation that we could face in the
future.

I have enjoyed the class activities and reading material and
been able to apply the skills immediately in my career. The
personal reflection papers and blog assignments give time each
week to relate to the concepts and really reinforce the material
(p. 62).

**Community building and networking**

I found that opening up the classroom to include authors, subject matter experts,
and others really created an environment where we could all learn from each
other. Medley (2005) described his experience with using a discussion board to
invite virtual experts to class. (He also provides useful tips on how to find these
experts and prepare the students for discussions in his article). Using each other
as well as guest experts gave the students opportunities to learn and to make
connections that they would not have been able to do otherwise. The students
summed up the experience as follows:

The irony is that the use of this blog showed me how little I re-
ally do know about some of my classmates, and how little they
really know about me. I think that one of the biggest downfalls
of our [graduate program] has been a lack of communication and interaction outside of the classroom, and this blog has incorporated a great piece of technology that kept the discussion going past the classroom.

It was a pleasant surprise having the author participate in the blog, I would have never expected it. It added some real validity to discussions and shows that he takes a real interest in the study of his work. More interaction should be encouraged between students and authors.

**Self-reflection and growth**

In addition to helping the students develop their interpersonal skills, I wanted them to work on their intrapersonal competencies and learn more about themselves as part of their educational experience. I enjoyed seeing how they develop over the course of the semester in sharing their insights and stories.

The aspects that I learned the most from have been the in-class activities and the blog posts. All of these activities required a tremendous amount of soul searching and creative thinking that lead to a significant amount of enrichment in my life and learning.

I had an “a-ha” moment Wednesday when [the teacher] was asking questions about the blog. She wanted to know what we could use on the blog for jobs. I was reminded again just how unsatisfied that I am in my current position. I heard people saying what they did and found myself a bit envious that I wasn’t doing more. It was just another sad wake up call to me that only I control my destiny and that I need to take some time to reevaluate my current position and my future goals. I really need to figure out exactly where I am and why, and where I really want to be. From there I should make a choice on how to get there.

I liked the suggestion by Gaudry-Hudson and Yalda (2008) of having the students formalize their self-reflection by creating a personal folder of their own comments and responses to classmates and other participants. These could then be used to reflect on the learning and confidence level gained throughout the course. A final paper could also be assigned on lessons learned.

**WHAT DO YOU NEED TO KNOW IF YOU ARE CONSIDERING A CLASS BLOG?**

In my experiment with using blogs in my classes, I have learned a few tips to making these both an effective and efficient tool. These include budgeting time,
designing the assignment questions, deciding whether to password protect the site, and setting expectations as to the quality and quantity of writing expected. I also had to decide how I was going to grade the blog comments and posts. In addition, I decided it was important to discuss “netiquette” with my students as to issues of appropriateness and respect.

Budget time for the class blog
Finding the time to develop good questions and to read the comments made by students sometimes felt overwhelming. I found it helpful to spend more time reading and responding earlier in the semester in order to set expectations. I also found it best not to jump in and respond on the blog too soon but instead to wait until a number of students had commented before I did.

Finding the outside experts took time also. In future classes I am considering assigning this task to the students. They could do their own initial networking with subject matter experts and authors whom they are interested in following by researching books, blogs, and organizations, and by connecting through Facebook, Twitter, and LinkedIn.

Another issue I found was that I needed to set a deadline by which the students’ comments were posted. This then gave me time to read over these and organize my own thoughts as to how to use the stories shared in the next class. I found setting a deadline 24 hours before class met was sufficient and doable for both the students and me.

Think carefully about the questions posted
I found it important to really think through the weekly questions. Some of the ones I used in the first class experiment with blogs resulted in a “me too” response from students and thus did not provide the interaction I had hoped for. The best blog questions required the graduate students to relate the topic and readings to their own personal work experiences as this resulted in better discussions. Even so, some fine tuning might still be needed, as the student comments below indicate.

Due to the detailed explanations [posted by individual students], the blog seemed lengthy and broadened the spectrum for responses. This I feel resulted in less interaction within the blog space. I wonder if it would be better to narrow the topic and promote more interaction within the group?

I agree with Paul, some of the comments came out to be a little too lengthy. Blogs could be a tool to teach people to express their opinions in a concise simple way—teach students to use the KISS method. Maybe you could limit how many characters can be posted per comment.
Consider whether to allow students to initiate blogposts
In the instructor-focused approach, I was the only one who could initiate a post although the students could comment on these similar to a discussion board and they could respond to each other. However, in the learner- and community-focused approaches, faculty might want to add junior and guest authors to the blog, a task that is easy to do as the owner of the blog. Junior authors can write posts but these must be approved and then uploaded by the professor. Guest authors are allowed to publish posts without approval but cannot edit or delete posts written by others although others can comment on their posts. As the blog owner, the instructor retains ultimate control.

Consider whether to password protect the blog
Setting a password is easy and means that only those who know the password can access the information in the blog, even through the use of search engines. The decision to password protect may be dependent on the subject matter. I did password protect the blog in a class on managing diversity due to the controversial topics being discussed and my hope that by doing so, the students would feel comfortable sharing personal topics. In other classes, however, I wanted the students to realize that others outside the graduate program would be invited to participate and that they needed to be professional in their comments. Another possibility is to have the students use their first names only when posting or commenting on the blog.

Communicate expectations as to quality and quantity of writing
There are obviously different views on this. I believe that content is important but that presentation is also important (e.g., spelling and grammar). In other words, I am modeling the importance of impression management for my students. I have colleagues, though, that want their students to just focus on getting their ideas and thoughts written down. Obviously, this difference in teaching philosophies affects writing expectations that needs to be shared with the students.

Also, because of the novelty of using this technology in the classroom, I found it important to set appropriate expectations of student behavior and activity for blogging. I found some students would comment to an extreme, while others did not take the assignment to blog (or to comment on a class blog) seriously enough. In cases where this was a problem, I would email the student personally to give feedback on what I expected. For the most part, this was not a problem as there was a great deal of peer pressure in knowing that your classmates will be reading your comments also.

I like that the blog allows us to share our ideas outside of the classroom versus writing individual papers that only the professor and the individual student see.
Decide how to grade student comments/posts
Because I needed time to read the students’ comments before class, I decided to give no credit for a comment not completed on time. This was easy to determine as blogs have a date stamp when anything is saved to the blog. Other than that, full credit was usually given if I believed the question was answered in sufficient detail and that the student took the assignment seriously. At the beginning of the course, I would have to talk one-on-one to a couple of students who were not meeting expectations on the blog. However, for the most part this was not a problem as there was peer pressure to do well. Sample (2010) shares an example of a rubric that could be used to grade comments on a blog and Kirk and Johnson (2010) have a rubric for grading student blogs.

Teach “netiquette” to students
Occasionally, a student would write an inappropriate or unreasonable comment or blog post, either stemming from personal beliefs and values, or possibly as a test to see how others would react. In one of my classes on managing diversity, I posted an assignment requiring the students to read a current newspaper article on the firing of a city manager after he revealed he was planning to have a sex change operation. Most of the students were indignant and wrote that this should not have anything to do with how the man performed his job. However, one student wrote:

After reading the article on the city manager, I feel like the right thing to do is be politically correct and say that I feel an injustice has been committed. My true feelings get in the way because I think this guy is disgusting. I don’t understand what kind of mental state a man has to be in to decide that he wants to become a woman . . . . It may not affect the managerial skills that this thing has used for the past 14 years, but anyone that would have this procedure has serious psychological issues and shouldn’t be in a position of leadership.

Obviously, I then felt the need to make this a “teaching moment” as I was trying to promote the inclusion of people perceived as “different” from ourselves. This took both classroom management skills as well as an ongoing rapport with the students to achieve, but ended up being a worthwhile exercise that probably would not have occurred in the physical classroom. (For more information on handling inappropriate blog comments related to diversity, check out my article published in the Journal of Management Education, Kirk & Durant, 2010).

SO WHAT HAVE I LEARNED FROM USING BLOGS IN MY GRADUATE CLASSES?
Despite my positive experiences with class blogs, I agree that they are not for everyone. There is a definite learning curve to setting one up and maintaining it
even though blog sites are very user-friendly. Table 1 presents some of the advantages and disadvantages of using a class blog.

The majority of students are familiar with and comfortable with using the Internet so the blog format is easy for them to use. The most recent information posted is listed at the top of the blogpage and the students can even have updates forwarded to their email accounts. I found that I only needed to be “tech comfy,” not “tech savvy,” as all blog writing is done in a Microsoft Word document and does not require any programming skills.

I found blogs to be much more useful than just discussion boards in terms of ownership and ability to hyperlink to journal or newspaper articles and other blogs. A blog is controlled by the person who created it; others can comment but cannot create blogpost content unless given administrative access. As noted by Trammell and Ferdig, “On a discussion board anyone can start a thread of conversation and all contributors have the same editorial authority” (2004, p. 61). I agree with Asllani, Ettkin, and Somasundar (2008) that blogs appear more successful than discussion boards in communicating tacit knowledge.

**Blogging invites more students into the conversation**
Students who were unwilling to speak up in class indicated they felt more comfortable writing their thoughts and responding to the thoughts of others in a blog/comment format. In addition, using a blog allowed students with different learning styles to interact with the class.

I thought that the use of a blog as a teaching tool was a very creative way to get everyone involved. I have found that one of the greatest resources in this program is my classmates. Learning from their different experiences and listening to their different points of view helps to bring my learning experience full circle.

**Blogging extends the conversation**
Allowing students to participate in mini-conversations via blogs keeps the classroom conversation alive even when the students are not physically present. As universities create flexible class times and schedules such as night and weekend courses to meet the needs of the students, using blogs may be even more useful in preventing the fade-out effect.

I found the blog to be very effective as a learning tool, mainly because I am one of those more quiet people in class. Things don’t always come to my mind right away, but once I give them time to sink in and realize how I really feel, this is when I find
<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provides communication medium between in-classroom experiences</td>
<td>Many instructors lack a comfort level with blog technology</td>
</tr>
<tr>
<td>Allows “quiet” students a forum for expressing themselves</td>
<td>Can create more work for the instructor (developing content, reading student blogs and comments)</td>
</tr>
<tr>
<td>Format allows instructors to disseminate information quickly and accurately outside of the classroom</td>
<td>Students may be uncomfortable sharing information through this medium</td>
</tr>
<tr>
<td>Information is available electronically for students to access from home</td>
<td>Instructors may not understand how to use the blog effectively</td>
</tr>
<tr>
<td>Lower cost than making hard copies</td>
<td>Student writing might be more casual and sloppy compared to turning in hard copies</td>
</tr>
<tr>
<td>High comfort level with blogs among Generation Y students</td>
<td>May be seen as not as useful in classes that do not require discussion</td>
</tr>
<tr>
<td>Promotes learning community between and among students and gives them a feeling of ownership</td>
<td>Students may resist commenting if it is not made part of their grade</td>
</tr>
<tr>
<td>Students may put more effort toward their writing knowing it will be read by their peers as well as the professor</td>
<td>Students must have access to computer and Internet to participate</td>
</tr>
<tr>
<td>Gives students skills in obtaining and evaluating knowledge found on the Internet</td>
<td>May be difficult to keep students focused on topics they are to discuss</td>
</tr>
<tr>
<td>Allows for teachable moments that might not take place in physical classroom</td>
<td>Students may write what they think professor wants to see rather than their own thoughts and feelings</td>
</tr>
<tr>
<td>Students perceive a “cutting edge” factor in their classes from using blogs as a powerful communication tool</td>
<td></td>
</tr>
</tbody>
</table>

the blog to be quite useful. It also allows me to express myself more clearly and to allow others to become familiar with my point of view.

**Classroom blogging provides a “safe” mechanism for introducing students to social media**

Even with the proliferation of blogging, Facebook, and Twitter, there are some students who were only vaguely aware of social media and its possibilities. Once students see how blogs can be used, they begin to realize the business benefits of extending “conversations” to the online world. While some of the students may resist a classroom blog at first, most quickly embrace the medium and start seeing ways to use these in both their professional and social lives.

Occasionally, I would have a student ask why I was using a blog in his or her class rather than one of the standard university course management programs. My response was that while these educational software packages can be useful, these programs are not available outside the academic setting. Learning to use blogs is a skill that translates to the business world. Many organizations are either developing their own in-house blogs and/or keeping track of what is being said about their products and services in the blog community.

This blogging experience has been informative for me as well. The longer that I am in this MBA program, the more I realize that I am one of the few people that do not read blogs on a regular basis and after this class, I may have to rethink that. It has been fun.

**Blogging makes the students into subject matter experts**

This process of creating a blogpost or answering a well-crafted question requires students to search for, filter, and then share information found on the Internet. The process exposes students to vast amounts of information and thus they become more knowledgeable on a topic. Doing this on a regular basis helps in increasing both their explicit and tacit knowledge (Trammell & Ferdig, 2004, p. 62). In one of the classes on leadership, one student started thinking about applying his personal brand to his class blog and was able to transition into a new career for himself in a field he was passionate about.

**Blogging helps students take ownership of their own learning**

I have been pleasantly surprised by the depth and breadth of the students’ blog comments and posts. They tended to spend more time and effort than required by the assignments. When given a choice of using a blog or a more traditional
paper method of journaling, those with blogs appeared to be more creative as they were leveraging online resources, linking to other blogs, and finding meaningful graphics. The students who wrote in more traditional ways tended to stay strictly within the stated expectations of the assignment.

**CONCLUSION**

As noted by the Council of Graduate Schools, “Individuals who earn a master’s degree gain an edge in today’s increasingly competitive global marketplace . . . [and] can play a substantial leadership role in the evolving knowledge economy of the 21st century” (2009). It is projected that jobs requiring a master’s degree will increase at a rate of 18% between 2008 and 2018 (Bureau of Labor Statistics, 2010). To address employer demand, “many master’s programs have shifted from having an arts and science focus and [have] become professional programs that prepare students for careers in business, government, and nonprofit organizations” (Wendler, et al., 2010, p. 18). Building competencies in creativity, collaboration, communication, and how to create, evaluate, and share knowledge will be invaluable to these new graduate students. One way to do so is to take advantage of existing participatory web technologies.

As observed by Sherer and Shea (2002), while there are numerous ways of using technology in the classroom, it is “how faculty integrate technology into their course design and assignments that is the critical piece when using technology to improve learning” (p. 15). Based on my own personal experience, I believe blogs can be a useful tool for teaching master’s level students some of the skills and competencies they will need in today’s environment. The knowledge that other students and professionals in the field were reading their writing assignments kept the students on task. They were able to make connections with each other and with virtual experts and extend the conversation beyond the physical classroom. The stories they shared helped them to see how others apply the theories and concepts being studied to their own situations. Although some of the students initially resisted the blog, this changed as they grew more fluent with its use and started seeing ways to use their new skills in both their professional and personal lives.

As my first experience blogging it took a few weeks to become used to this media, as I am far more familiar with classic paper and e-mail based communications. This type of collaborative technology was a unique approach to class assignments, but a necessary experience as blogs are becoming more and more popular in the workplace.
REFERENCES


Appendix:
Sample Blog Questions for MBA Course on Leadership

Example #1
David Lorenzo in his book, *Career Intensity*, notes that, “We should all be dedicated to discovering new things that we can use to improve our careers. This discovery and learning is the first part of the cycle of continuous career improvement. From there, we can apply what we’ve learned to our careers in a way that makes us different from others in our field. This difference makes us valuable to our customers. Then, we can market ourselves to the world in a way that will help people understand why we are different and how these differences—our ability to generate value—will benefit them . . .” (16).

Read the classic article by Tom Peters published in 1997 titled, *The Brand Called You*, and reflect on this week’s reading in Lorenzo’s book. In your comments this week, answer the following questions:

- What are the “feature-benefits” that you offer in your own personal brand? (deliver work on time, anticipate and solve problems before they become crises, etc.)
- What do you do that you are most proud of?
- Who in your organization would you say is your raving fan?

Example #2
John Maxwell in his new book, *Leadership Gold*, gives advice on learning from your defining moments. Read chapter 3 of his book and think about the choices you’ve made and how these have defined you. As he notes, many times these are situations when you step up and make a risky decision that then can really pay off for you. Share one of your defining moments with us.
An Analysis of Economic and Social Recession-Affected Constructs in the Kansas City Region (2000–2009)

James Martin
Jennifer Ball
Janice Schrum
Washburn University

ABSTRACT
The economic and social response to recessions in a specific region will vary when compared to other regions, and when compared to the United States as a whole. Economic constructs examined in this paper are unemployment and retail sales. The social construct analyzed is the rate of divorce. The region examined is a 20-county region encircling Kansas City. The analysis of the 2008 recession found the region’s increase in unemployment to be less than the nation’s increase. The unemployment increases in the KC region also began later than the nation’s. Decreases in unemployment at the end of the recession also began later in the Kansas City region than the nation. Similarly, recession related decreases in Kansas City region retail sales were less precipitous and later starting than national decreases in retail sales. Retail sales recoveries also began later in the region than in the nation. Regional divorce rates reversed a steady downward trend prior to the recession. However, national divorce rates continued an established pre-recession trend of downward movement even during the recession.

Keywords: Kansas City region, recession, unemployment, retail sales, divorce.

Authors extend their gratitude to student research assistants Lisa Fahey and Jessica Winkel for their contributions to this paper.
INTRODUCTION
It is a common occurrence. This time, maybe it is the luncheon speaker at a suburban Johnson County Rotary Club meeting. Or maybe it is two retirees discussing the latest weather and economic forecasts at a diner in a small west Missouri town. But when the discussion turns to the most recent recession, the story goes something like this, “We in the Midwest aren’t affected the same way as the rest of the country by recessions. We don’t experience the full brunt of the economic downturn like the rest of the country. We likewise don’t enjoy as much of the benefit of economic upturns as the rest of the U.S. when the nation’s economy is expanding.” But is this true? That is the first hypothesis investigated by this paper.

A related corollary deals with the social response to recessions in the Midwest. Statements such as the following also sound plausible: “There will be an increase in certain negative social outcomes (e.g., dissolutions, divorces) during periods of economic recession. The magnitude of these negative social outcome increases will be related to the depth of the recession.” But is this true? This statement summarizes the second hypothesis investigated by this paper. This hypothesis has been divided into Hypothesis 2A and 2B.

In testing the first hypothesis, we will analyze two recession related economic constructs, unemployment rate changes and retail sales growth or reduction, during the period of 2000 through 2009. Increases in unemployment (or decreases in employment) and decreases in retail sales are generally associated with recessions. This period of time (2000–2009) bracketed two United States recessions and is fertile for analysis. The bursting of the dot-com bubble and the September 11th attacks are generally credited with fueling the recession that ran from March through November of 2001. The subprime mortgage crisis, the resulting end of the United States housing bubble, and the demise of companies like Lehman Brothers and Bear Stearns are at the core of the current recession which began in December 2007. The end date of the current recession is still under debate as this paper is completed.

The purpose of this study is to bridge two intersecting areas of research, social outcomes on the one hand and economic factors on the other, in order to investigate the impact of recession on family dissolution. This study will quantify and compare these economic constructs in a specific Midwest region to comparable constructs in the nation as a whole. Relationships between the regional social constructs and the regional economic factors will also be analyzed. The specific Midwest region selected is a 20-county area that surrounds Kansas City. Ten of the counties selected (Atchison, Douglas, Franklin, Jackson, Jefferson, Johnson, Leavenworth, Miami, Shawnee, and Wyandotte) are in Kansas. At the beginning of the study (2000), these counties included 49.4% of the selected region’s retail sales. The other ten counties selected (Bates, Buchanan, Caldwell, Cass, Clay, Clinton, Jackson, Lafayette, Platte, and Ray) are in Missouri. At the beginning of the study, these counties included 50.6% of the region’s retail sales. As such,
the study area is roughly split 50/50 between Kansas and Missouri. Interestingly, during the 10 years of the study, each state’s percentage of retail sales shifted such that at the end of the study period (2009), 50.9% of the region’s retail sales were from Kansas counties while 49.1% of the sales were from Missouri counties.

**Economic Literature Review**

Considerable research has been completed analyzing varying region-specific economic reactions to national economic events such as recessions or changes in monetary policy (national shocks). Research has also been done on the regional impact of regional economic events such as the effect of a catastrophic oil spill in the Gulf Coast states or a drought in the Midwest (regional shocks). Given that this paper’s hypotheses parallel research into whether a region’s economy is more sensitive to national economic shocks or to regional economic shocks, it was not surprising to find research implicitly supporting our first hypothesis and additional research implicitly rejecting our first hypothesis in various regions of the United States. However, no research specifically addressed the Kansas City region’s reactions to a national (or regional) shock such as a recession. This paper therefore adds to the current body of economic research.

As an example, in other research contrary to our first hypothesis, Michael Kouparitsas (2002a) found that the Plains region (including this paper’s 20-county region) was the most sensitive to (and implicitly likely to follow) the national business cycle. This conclusion was explained in part by the finding that manufacturing and agriculture (concentrated in the region) are very cyclically sensitive. Similarly, Gerald Carlino and Robert DeFina (1998) found that the Plains region’s economic reaction to national monetary policy was the most consistent of the United States regions to the United States average reaction. Reasons identified were the region’s industrial diversity and the concentration of small businesses in the Plains region. Finally, Ellen Rissman (1999) found that regional shocks play a lesser role (than in other regions of the United States) in regional employment growth in the West North Central region (including this paper’s 20-county region). The research implied national events, rather than regional events, tend to drive this region’s employment growth.

Supportive of our hypothesis is research completed by Carlino and Sill (1996) finding that, when compared to other region cyclical changes, the Plains region had considerably less volatility in its business cycles when reacting to national events. Kouparitsas (2002b) found that a relatively high percentage (30%) of business cycle variation in per capita income in the Plains region can be explained by region specific shocks. This is because the Plains derive a large portion of its income from commodities, which has fluctuations more tied to international events than to national events. Finally, Carlino and DeFina (1993) found that of all the United States regions, the Plains had the largest portion of its growth variance (59%) explained by shocks within its own region.
Recessions
Historically, the widely accepted layman’s definition of what constitutes a recession was a period of time, lasting at least two quarters that experienced a decline in gross domestic product (GDP). Refinements to this definition include those of the National Bureau of Economic Research (NBER; Leamer & Medberry, 2008). NBER suggests using four series of data when identifying recessions. These series are the following: 1) employment, 2) industrial production, 3) real personal income less transfer payments, and 4) the volume of sales of the manufacturing and trade sectors. By definition, recessions are protracted periods of economic decline normally marked by decreases in employment, production, personal income, and sales. Unfortunately, monthly data for industrial production and personal income are not available by county; however, monthly data related to employment (countywide unemployment rates and employment levels) and related to sales of the manufacturing and retail sectors (sales tax receipts) are readily available for the 20-county region. Comparable data are also available on a national level for unemployment and sales, facilitating this paper’s analysis.

Recessions and Employment
Monthly county specific unemployment rates have historically been compiled by state agencies in Missouri and Kansas. These data were obtained from the Missouri Economic Research and Information Center (2010) and the Kansas Labor Information Center (Kansas Department of Revenue, 2010) for all months in the period January 2000–December 2009. Additionally, for weighting purposes, estimated county population data for the same periods were obtained from the U.S. Census Bureau (2010). U.S. Department of Labor (2010b) data for a single month (unemployment rates) for the 20 counties were then weighted according to county populations to arrive at an average 20-county unemployment rate for a specific month. This process was repeated for the 20 counties for all months in the study period resulting in the calculation of 20-county monthly unemployment rates for 2000–2009. Comparable national monthly unemployment rates were also obtained for the same periods (U.S. Department of Labor, 2010a). Chart 1 graphically displays the data.

A general observation of Chart 1 does indicate lower unemployment in the region than at the national level in most recession months. However, it does not provide overwhelming support for the first hypothesis that the recession affects the 20-county region less dramatically than it affects the nation. Additional statistical analysis of monthly unemployment rates is completed below. However, in order to analyze the data without monthly cyclical fluctuations at both the regional and national levels, rolling 12-month average unemployment rates were calculated and displayed in Chart 2. Use of rolling 12-month average rates eliminates cyclical spikes in unemployment, more clearly reveals trends, and could support or refute this paper’s hypothesis.
The increase in unemployment during the brief 2001 recession for both the national and 20-county region was not significant. However, increases in national and regional unemployment rates during the most recent recession were noteworthy. In this case, the rolling 12-month pre-recession unemployment rates of the nation (4.59%) and 20-county region (4.84%) increased dramatically through the remaining study period for the nation (9.25%) and the region (8.26%). The fact that the nation’s unemployment rate grew more rapidly is supportive of the first hypothesis and is illustrated above. A final depiction of the unemployment data is included as Chart 3. This chart illustrates the monthly increase or decrease (calculated over the prior 12-month period) in the unemployment rate for the 20-county region and the nation.

Chart 3

Twelve Month Change in Unemployment Rate

Chart 3 depicts, on a month-by-month basis, the percentage increase/decrease in the unemployment rate over the same month, 12 months prior. During the second recession (December 2007–current), the 12-month percentage increase in unemployment for the nation exceeded the 20-county region increase for all months except the first month (December 2007). This elevated national rate of increasing unemployment supports hypothesis one.

While basic statistical analysis supports the claim that unemployment in the region varies less than national unemployment, the reason requires additional research. We looked at national and regional employment and labor force growth to possibly shed some light on this relationship. Monthly national and regional employment growth and labor force growth were calculated from the Bureau of Labor Statistics data (U.S. Department of Labor, 2010a for national; U.S. Depart-
Table 1
Descriptive Statistics
Monthly Observations January 2000 to December 2009

<table>
<thead>
<tr>
<th></th>
<th>Obs</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regional Unemployment</td>
<td>119</td>
<td>3.0100</td>
<td>8.7363</td>
<td>5.4522</td>
<td>1.2218</td>
</tr>
<tr>
<td>National Unemployment</td>
<td>119</td>
<td>3.8000</td>
<td>10.1000</td>
<td>5.5504</td>
<td>1.4449</td>
</tr>
<tr>
<td>Regional Emp Growth</td>
<td>118</td>
<td>-2.3936</td>
<td>1.4117</td>
<td>-0.0136</td>
<td>0.7423</td>
</tr>
<tr>
<td>National Emp Growth</td>
<td>118</td>
<td>-2.0750</td>
<td>0.9347</td>
<td>0.0137</td>
<td>0.5541</td>
</tr>
<tr>
<td>Regional Labor Force Growth</td>
<td>118</td>
<td>-1.7439</td>
<td>2.0930</td>
<td>0.0224</td>
<td>0.7433</td>
</tr>
<tr>
<td>National Labor Force Growth</td>
<td>118</td>
<td>-0.8898</td>
<td>1.4035</td>
<td>0.0844</td>
<td>0.4790</td>
</tr>
</tbody>
</table>

Note: All numbers (except observations) in percentages.

As shown, the mean unemployment rate for the region is slightly lower than for the nation, and the standard deviation is smaller. This would support the idea that unemployment varies less regionally than nationally. On the other hand, the means of regional employment and labor force growth are lower than their national counterparts, and the regional standard deviations are substantially larger. (None of these differences in means are statistically significant.) It is interesting to note that both the national and regional labor force growth outpace employment growth on average, but the difference between the regional labor force growth and employment growth is smaller than the difference for the national data. These observations point to the smaller changes in regional unemployment (as compared to national) as a result of people moving in and out of the regional labor force at a faster rate than the national labor force. This is confirmed by ordinary least squares regressions of the national data on the regional data. Table 2 illustrates the results of these models.

These estimates confirm earlier observations that the regional unemployment rate varies less than the national unemployment rate for these years. For a one percentage point change in the national unemployment rate, the regional unemployment rate changes by 0.774 of a percentage point on average. On the
other hand, regional employment changes marginally more than national; for a one percentage point change in the national growth rate, regional employment changes by an average of 1.055 percentage points. Finally, the regional labor force changes substantially more than the national; for a one percentage point change in the national rate, the regional labor force changes by an average of 1.314 percentage points. Again, this indicates the smaller changes in the regional unemployment rate are due to greater movements (in percentage terms) of people in and out of the regional labor force. It would be informative to also compare the percentage of the regional labor force that is discouraged or marginally attached to the labor force to the relevant percentage of the national labor force. Unfortunately, data on discouraged workers and those marginally attached to the labor force are not available by county. If the percentage of these workers in the region were generally higher than in the nation during economic downturns, it might indicate workers in the region did not willingly move in and out of the labor force in greater percentages, but that the duration of unemployment tends to be longer in the region than in the nation. However, if this were the case, we might expect to find retail sales varying more in the region than the nation, which is shown in the next section not to be the case. This suggests movements in and out of the labor force for workers in the region may be somewhat easier to manage than for workers in the nation.

As stated above, data on industrial production are not available by county. However, it is interesting to note that in both Kansas and Missouri, the contribution to GDP by the manufacturing sector was smaller than for the nation at large in 2008. Manufacturing accounted for approximately 20.0% of GDP in 2008 (Bureau of
Economic Analysis, 2008) whereas it only accounted for 15.2% in Kansas (Kansas Department of Labor, 2009) and 13.5% in Missouri (Missouri Department of Economic Development, 2009). As industrial output tends to vary more than total output, the smaller manufacturing sectors may also be part of the reason why the region’s unemployment rate varies less than the nation’s.

**Recessions and Retail Sales**

A common source of revenues for counties in Missouri and Kansas is the retail sales tax. Both Missouri and Kansas have statewide sales taxes and varying county/city sales taxes charged on the majority of retail sales. Monthly county-specific sales tax receipts have historically been compiled by state agencies in Missouri and Kansas. These data were obtained from The Missouri Department of Revenue (Missouri Department of Revenue, 2010) and The Kansas Department of Revenue (Kansas Labor Information Center, 2010) for all months in the period January 2000–December 2009. Additionally, in order to gross up sales tax receipts to yield retail sales, actual sales tax rates by month were obtained from the same state agencies. Data for a single month (sales tax receipts) for a specific county were divided by the actual sales tax rate charged in that county that month. This process was repeated for the 20 counties for all months in the study period resulting in calculated 20-county monthly retail sales amounts that were summed for 2000–2009. Comparable national retail sales were also obtained for the same periods (U.S. Census Bureau, 2010). Using these calculated amounts, rolling 12-month retail sales receipts were calculated for the 20-county region and nation for the period 2000–2009. In order to analyze the data without monthly cyclical fluctuations at both the regional and national levels, rolling 12-month average retail sales were calculated. The use of rolling 12-month average rates eliminates cyclical spikes in sales, more clearly reveals trends, and could support or refute this paper’s hypothesis.

Using the rolling 12-month national and regional retail sales, Chart 4 was plotted to depict the percentage change in each month’s rolling 12-month sales amount when compared to a rolling 12-month sales amount, one year prior. Vertical lines show the most recent recession’s approximate time period. (Rolling 12-month base year data were not available for the first recession in the decade.) It is clear from the chart that the retail sales growth in the region from early 2003 through late 2007 trailed the national amounts, in support of our hypothesis. Differences in regional/national inflation rates could explain part of this. (See the Conclusions section of this paper for areas of additional research.) Likewise the decrease in retail sales during the current recession was also more precipitous in the nation than the 20-county region.

Finally, Chart 5 was prepared using the same national and regional sales data as Chart 4. However, the sales and percentage changes were calculated on a rolling 3-month basis (quarterly) instead of 12-month basis as in Chart 4. The results
depicted in Chart 4 are similar to those in Chart 5. Chart 5 is interesting in that it shows a moderate recovery beginning in the most recent quarters for the nation and region. But again, the region’s sales growth lags behind the nation’s sales growth.
Again, basic statistical analysis confirms these findings. Table 3 shows the means for the national and regional monthly changes of both the 12-month and 3-month rolling averages of retail sales. As can be seen, the means of the changes for the region are smaller than they are for the nation, and have smaller standard deviations. (The differences in the national and regional means are statistically significant at levels of confidence greater than 99%.)

| Table 3 |
| Descriptive Statistics, Monthly Change in Difference from 12 Months Prior January 2002 to December 2009 (for 12-Month Rolling Average) |
| April 2001 to December 2009 (for 3-Month Rolling Average) |

<table>
<thead>
<tr>
<th></th>
<th>Obs</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>12-month rolling average of regional retail sales</td>
<td>96</td>
<td>-6.52</td>
<td>4.52</td>
<td>1.651</td>
<td>2.522</td>
</tr>
<tr>
<td>12-month rolling average of national retail sales</td>
<td>96</td>
<td>-8.90</td>
<td>6.89</td>
<td>3.002</td>
<td>3.952</td>
</tr>
<tr>
<td>3-month rolling average of regional retail sales</td>
<td>105</td>
<td>-8.37</td>
<td>6.15</td>
<td>1.270</td>
<td>3.447</td>
</tr>
<tr>
<td>3-month rolling average of national retail sales</td>
<td>105</td>
<td>-9.76</td>
<td>8.19</td>
<td>2.654</td>
<td>4.361</td>
</tr>
</tbody>
</table>

*Note.* All numbers (except observations) in percentages.

These findings are also confirmed by ordinary least squares regressions of the national data on the regional data. Table 4 illustrates the results of these models. For each one percentage point change in national sales from a year prior, regional sales changed by 0.563 of a percentage point when using a 12-month rolling average, and 0.664 of a percentage point when using a 3-month rolling average. This indicates that, similar to unemployment rates, regional sales change less from year to year than do national sales. It is possible the lower costs of living mentioned above allow families in the region to be more flexible in their labor force participation without having to change their spending habits as substantially as the average family in the nation.

As mentioned above, monthly data on personal income by county are not available. However, if income were to vary more regionally than nationally, we might expect to see retail sales varying more as well. Since this is not the case, and since regional unemployment varies less than national unemployment, it is unlikely that personal income varies more regionally than nationally; however, there is no clear evidence it varies less, either.
Social Construct Literature Review

“Recession has always been a factor raising divorce rates . . . couples that experience any sudden significant and unexpected change in income, positive or negative, are at risk for divorce” (Becker, 1977). During a recession, the expectation develops that the economy will slow and firms will reduce employment levels to save money. People fear that they will lose their jobs and the loss of a job is known to have a negative impact on family stability (Vaitilingam, 2009). The family stress theory referred to as the ABC-X model provides a framework for addressing the impact of recession on dissolution rates. According to the ABC-X stress model proposed by Hill (1949) and expanded later by McCubbin and Patterson (1982), economic factors resulting from a recession play a significant role in the decision to dissolve a marriage. “In this model, A represents a stressor event, B represents the family’s coping resources, C represents the family’s perception of the event and X represents the crisis” (Doiron & Mendolia, 2007). Based on this model, the following applies to our analysis:

A – Stressor event—the economic recession that might result in loss of employment and a decline of monetary income for the family;

B – Family’s coping resources—including the family’s financial resources, available employment opportunities, as well as social support systems;

C – Family’s perception of the event—including the family’s perception of whether they could cope with unemployment and with the loss of income;

### Table 4

<table>
<thead>
<tr>
<th>Model equation estimates</th>
<th>Obs</th>
<th>Adjusted $R^2$</th>
<th>$F$-test significance (model)</th>
<th>Coefficient significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>12-month rolling average of sales $\Delta$Regl sales = 0.563 * $\Delta$Natl sales</td>
<td>96</td>
<td>0.858</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>(0.023)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-month rolling average of sales $\Delta$Regl sales = 0.588 * $\Delta$Natl sales</td>
<td>105</td>
<td>0.664</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>(0.041)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Standard error of coefficient in parentheses.
X – The crisis—the family’s perception that their expenses exceeded their financial resources, and that they could not cope with unemployment or potential loss of employment; therefore the family might determine that they must dissolve the marriage in order to resolve the crisis.

Ideally, an analysis of the impact of economic conditions on divorce rates should consider a multitude of factors. For instance, in recent decades, family and marriage characteristics have dramatically changed; divorce rates have risen and marriage rates fallen (Doiron & Mendolia, 2007). In addition, cohabitation has emerged as an important institution, often as a substitute for marriage (Doiron & Mendolia, 2007). Other factors not addressed by this study but relevant include length of marriage, number of children, level of partners’ education, and whether one or both of the partners is employed. Previous research regarding the length of marriage and its relationship to divorce reveals that there appears to be strong evidence of duration dependence in marital stability in that the longer people have been married, the smaller the probability of dissolution (Doiron & Mendolia, 2007). Further, studies have demonstrated that the wife’s employment status has no effect on the probability of divorce (Doiron & Mendolia, 2007). Also, education appears to have a non-monotonic impact in that the highest probability of divorce is found for the lowest level (i.e., high school) of education for both husband and wife (Doiron & Mendolia, 2007). Finally, studies have also indicated that a couple’s number of children has no effect on dissolution (Doiron & Mendolia, 2007).

Previous studies have investigated the relationship between unemployment and divorce rates. Kraft (2001) performed a study to investigate the impact of unemployment on married couples’ decisions to divorce and found an increase in divorce rate following the first year of unemployment and that this impact increased with the duration of unemployment. Blekesaune (2008) found a significant increase in the probability of family dissolution after any form of unemployment experienced by either husbands or wives. However, Blekesaune (2008) did not distinguish between different causes of unemployment. Lastly, Doiron and Mendolia (2007) investigated the impact of involuntary job displacements on the probability of divorce using discrete duration models and found that couples where the husband experienced a job loss were more likely to divorce. However, the wife’s employment status and level of non-labor income had no significant effect (Doiron & Mendolia, 2007).

Other studies have investigated the relationship between unfavorable economic conditions and divorce. Becker’s (1981) work on the economics of the family has been used as the theoretical basis to demonstrate adverse family-level economic shocks (such as the job loss of a husband) increase the probability of divorce. Fischer and Liefbroer (2006) performed a study to investigate the effects of
These authors found support of their general hypothesis that unfavorable economic conditions increase union dissolution rates. This finding supports the relational stress argument that suggests that unfavorable economic conditions increase the stress within the spousal relationship.

**Data and Analysis**

The purpose of this study was to investigate the relationship between social outcomes and economic factors in the Midwest in order to investigate recession effects. The hypothesis, in general terms, is that the social outcome variable of dissolution rate will be positively correlated with the economic factor of unemployment and negatively correlated with retail sales change. A second hypothesis suggests that when an economic recession becomes deeper, then divorce rates will likely increase.

Regression analyses were performed to investigate whether there was a relationship between measures of economic activity (i.e., unemployment and retail sales change) and social outcomes (e.g., dissolution/divorce rates).

**Hypothesis 2A**

Hypothesis 2A states that the social outcome variable of dissolution rate (regional level) would be positively correlated with the economic factor of unemployment and negatively correlated with retail sales change.

**Results (Hypothesis 2A)**

Basic statistical analysis supports the hypothesis above. Table 5 includes descriptive statistics for regional divorce rates and the results of regressions to explore the relationships between divorce rates and unemployment. Table 6 reports divorce rates and retail sales change.

<table>
<thead>
<tr>
<th>Table 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Descriptive Statistics, Regional Divorce Rates</strong></td>
</tr>
<tr>
<td>(Divorces/1000 Population)</td>
</tr>
<tr>
<td>December 2000 to December 2009</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Obs</td>
</tr>
<tr>
<td>Regional divorce rate</td>
</tr>
</tbody>
</table>
As shown, the regional divorce rate is positively correlated with the regional unemployment rate. The coefficient is small but statistically significant at the 99% level of confidence. The estimate suggests for every one percentage point increase in the unemployment rate, the regional divorce rate increases by 0.152 per thousand (population). Therefore, these results would support the hypothesis 2A in that regional dissolution rates would increase as a result of the economic factor of increased unemployment. Also as expected, the regional divorce rate is negatively related to the 12-month change in the rolling average of regional retail sales. Again, the coefficient is small but statistically significant at the 99% level of confidence. The estimate suggests for every one percentage point increase in the 12-month change in the rolling average of regional retail sales, the divorce rate decreases by 0.113 per thousand (population). Therefore, these results support the second half of the hypothesis 2A in that regional dissolution rates would decrease due to the economic factor of retail sales increase.

**Hypothesis 2B**

Hypothesis 2B suggests that with a deeper and longer lasting recession comes more pronounced increases in divorce rates when compared to a shorter less economically disruptive recession. Therefore, it is hypothesized that regional dissolution rates will be greater for the deeper recession beginning in December 2007 than for the 2001 recession. This is because the most recent recession has been deeper and lasted longer than the previous recession.

**Results (Hypothesis 2B)**

A comparison of the two recession periods revealed a dissolution rate of 5.85 per thousand for the region in March 2001.

---

**Table 6**

**OLS Regressions, Regional Divorce Rates and Regional Economic Factors**

<table>
<thead>
<tr>
<th>Model equation estimates</th>
<th>Obs</th>
<th>Adjusted R²</th>
<th>F-test significance (model)</th>
<th>Coefficient significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regl divorce rate = 4.620 + 0.152 * Regl unemp rate</td>
<td>108</td>
<td>0.097</td>
<td>.001</td>
<td>.001</td>
</tr>
<tr>
<td>Regl divorce rate = 5.639–0.113 * Regl sales change</td>
<td>95</td>
<td>0.278</td>
<td>.000</td>
<td>.000</td>
</tr>
</tbody>
</table>

Note. Standard error of coefficient in parentheses.
decreasing to a rate of 5.31 by November 2001. However, from December 2007 to December 2009, the regional dissolution rate increased from 5.89 to 6.55 per thousand. These findings support the hypothesis 2B that the dissolution rate would be greater for the deeper recession period.

Although the findings indicate an increased dissolution rate for the Midwest region during a deep recession, these findings appear to run counter to the national divorce rate trend. Based on the national rolling 12-month divorce rate from 2000 through 2009, Chart 6 indicates a general downward trend. At the national level, fewer divorces have been granted even during a time when growing economic strain might have produced a spike in divorce. National trends do not appear to react to the economic environment in a noticeable fashion. According to St. George (2010), “With housing values depressed and jobs disappearing, divorce has become a luxury beyond the reach of some couples. There is often not enough money to pay for separate households or to hire lawyers, fight over children and go to court.” However, this is not the case for Midwestern couples.

**Chart 6**

**National Rolling 12-Month Divorce Rate/1000**

**Conclusions and Further Research**

Our analysis of regional/national economic and regional social outcomes related to post-2000 recessions yielded noteworthy results. When analyzing changes in national and regional unemployment rates during the recessions, findings were robust. In addition to higher nominal national unemployment rates, it was also noted that the rate of increase in national unemployment during the early periods of the recession outpaced the more modest rates of increase in the Midwest region's unemployment. However when the economy began to recover in 2009
and unemployment began to decline, the rate of decline in national unemploy-
ment exceeded the rate of decline in the region. These findings supported the
paper’s hypothesis that the Midwest is affected less dramatically than the nation
as a whole. The Midwest does not experience the full brunt of the economic
downturn as dramatically as the rest of the country. Likewise it does not enjoy as
rapid of a rebound as the rest of the United States when the nation’s economy is
expanding (recovering).

Similar to the conclusions made following the analysis of unemployment, results
related to changes in retail sales during recessionary times were equally robust.
As expected, retail sales dropped precipitously within the region at the outset of
the most recent recession. However, national retail sales declined earlier, more
sharply, and bottomed out at a relative lower level than the region. When retail
sales began to recover in 2009, the start of the recovery was later for the region
than the nation, and the rate of recovery for the region lagged behind the na-
tional level of sales growth. Again, these findings supported the previously stated
first hypothesis.

The effect of the recession on divorce and marital dissolution in the region
was also analyzed. As expected, regional divorce rates increased as economic
measures declined and the economy sagged into recession. It was also noted,
when comparing the two post-2000 recessions, the more profound and lengthy
recession beginning December 2007 resulted in a greater increase in the region's
divorce rate than the shorter 2001 recession. These findings were consistent with
the paper's second hypothesis of an increase in regional negative social outcomes
(e.g., dissolutions, divorces) during periods of economic recession. The findings
were also in line with the expectation that the magnitude of these negative social
outcome increases would be directly related to the depth of the recession.

As is in the case with most research, areas of additional research are identified. In
this case, the following were noted:

1. Variance between regional and national reactions to the recession
   was noted during analysis of unemployment and retail sales. Further
   research into the cause of these differences including a comparison
   of Midwestern trends to trends in other regions warrants additional
   study.

2. The impact of varying national and regional inflation rates on
   changes in retail sales was identified as a potential reason for differ-
   ing national and regional reactions. This should be considered as an
   area requiring additional research.
3. A comparison of the percentage of the regional labor force that is discouraged or marginally attached to the labor force to the relevant percentage of the national labor force would be a meaningful addition to the current body of knowledge. This too should be considered as an area requiring additional research.
REFERENCES


Collegiate Post-Graduation Venture Creation: What is the Impact of Entrepreneurship Courses?

Benjamin J. Blackford  
Northwest Missouri State University

Terrence C. Sebora  
University of Nebraska–Lincoln

ABSTRACT
This initial exploratory study empirically examines factors reflecting the three core stages in new venture creation—opportunity recognition (entrepreneurial self-efficacy), opportunity evaluation (risk acceptance), and opportunity exploitation (willingness to act). Data were obtained via survey from more than 400 graduates from the college of business of a major public university in the Midwestern United States. The results of a logistic regression analysis reaffirm that the factors associated with the three stages, as well as the number of collegiate entrepreneurship courses taken, are associated with the post graduation startup of a business.

Key Words: entrepreneurship, entrepreneurship education, small business, small business education

The authors wish to thank Ronda Smith and a blind reviewer for their helpful comments on various revisions of this paper.
INTRODUCTION
As entrepreneurship education has expanded, researchers have debated whether entrepreneurship can be taught and whether entrepreneurship education has any effect on the startup of new businesses (Chell & Allman, 2003). Is the increasing availability of entrepreneurship courses actually increasing the number of entrepreneurs? What is the impact of the significant resources being expended to offer more and more entrepreneurship courses and programs? Are students who take entrepreneurship courses more likely to start a new firm than those who do not? This study attempts to take steps toward answering these questions. The framework presented by Shane and Venkataraman (2000) and Brush et al., (2003), which distinguishes entrepreneurship by the recognition, evaluation, and exploitation of opportunities, provides a guiding framework for the research. Specifically, this research investigates the relationships between the self-confidence of college students in identifying opportunities, their willingness to accept risk when evaluating an identified opportunity for possible exploitation, their willingness to act to exploit an opportunity, and their entrepreneurship education with the post-graduation start of a new business.

ENTREPRENEURSHIP EDUCATION LITERATURE
Effects of Education on Entrepreneurs
What is the impact of education on entrepreneurs? Prior literature has provided a variety of insights into this question. Clark, Davis, and Harnish (1984) found support for a relationship between entrepreneurship courses and the creation of new ventures. Although their study could not provide an interpretation of cause and effect in the relationship, their work is important in that it provides confidence that there is a relationship worthy of further study. Charney and Libecap’s (2000) results showed that graduates from the entrepreneurship education program were personally more successful, having a 27% higher income than those who did not graduate from the entrepreneurship program. Clouse’s (1990) study sought to determine if taking an entrepreneurship course would actually affect the decisions and choices made by entrepreneurship students in a simulated situation. The study simulated decisions that may be made in the startup of new firms, examining whether the students could be taught to think differently about such scenarios. The results showed that the entrepreneurship course had a statistically significant impact on the decision process (specifically, the criteria used for making decisions) for the majority of students.

The effects of entrepreneurship education have been studied in other countries as well. Brockhaus (1991) reviewed and compared the entrepreneurial education environment in a multitude of countries including Japan, Egypt, South Africa, Korea, and many others, and found several common factors, such as the need for the private sector to provide more jobs and the recognition that entrepreneurship can help provide these. Autio, Keelley, Klofsten, and Ulfstedt (1997) studied education and the relationship with entrepreneurial intent across university settings.
in the United States, Finland, and Thailand. The results of the study showed a supportive university environment positively affected the conviction of respondents toward entrepreneurship, which was then positively related to the intent of the students to start their own firm within one year. Kristiansen and Indarti (2004), however, found no correlation between the educational backgrounds of university students and their intention to become an entrepreneur. Lee, Chang, and Lim (2005), in their study of American and Korean university students, found students from both countries benefited from entrepreneurship education, although Korean students benefited more than their American counterparts. While finding entrepreneurial orientation important, Parnell, Shwiff, Lei, and Langford (2003) did not find a significant difference among American and Chinese students’ entrepreneurial orientation.

What Aspects of Entrepreneurship Can Be Taught?

There are many different aspects that aspiring entrepreneurs can learn as they seek to identify, evaluate, and exploit opportunities. DeTienne and Chandler (2004) investigated approaches for teaching opportunity recognition and found that potential entrepreneurs could be taught to generate more ideas, and generate ideas that were more innovative. De Faoite, Henry, Johnston, and Sijde (2003) present that portions of entrepreneurship could be considered an “art” while other portions are a “science.” Utilizing work by Jack and Anderson (1998), De Faoite et al. (2003) suggest that the “art” portion of entrepreneurship (the creative and innovative aspects of entrepreneurship) cannot be taught, but the “science” portion (the business and management functional skills) can be taught to potential entrepreneurs. Alberti, Sciascia, and Poli (2004, p. 454) also provide a similar view, stating “we cannot make a person another (Richard) Branson, but the skills and creativity needed for being a successful entrepreneur could nevertheless be anyway enhanced by entrepreneurship education.”

Many of the studies previously discussed examine the effects of entrepreneurship education on the thought processes of potential entrepreneurs or their intentions to start a new business. However, few study the effects on the actual start-up of new firms. As noted by Honig and Davidsson (2000) in their review of how human capital affects startup of new firms, there is room for additional research on the effects of factors such as taking specific courses. Additionally, Gorman and Hanlon (1997) suggest more work on the effects of multiple courses in entrepreneurship is a worthwhile pursuit for future research. The lack of research incorporating education and other exogenous factors (such as attitude toward risk) to examine the pursuit of new ventures has also been noted in the literature (Luthje & Franke, 2003). In response to these calls, this research seeks to provide new insights on the influence of entrepreneurship education.
HYPOTHESES AND FRAMEWORK

The concept of opportunity is suggested to be the distinguishing characteristic of the field of entrepreneurship research (Shane & Venkataraman, 2000; Venkataraman, 1997). From the opportunity perspective, the entrepreneurial process associated with the creation of a new enterprise is categorized into three stages: opportunity recognition, exploration, and exploitation (Brush et al., 2003). Brush et al. (2003) draw on a range of prior research in development of these three stages. They note opportunity recognition includes the cognitive process of individuals as they scan for and identify opportunities. Exploration involves developing expectations of the viability of these opportunities. Finally, exploitation centers on perceptions of the likelihood of success based on the ability to effectively pursue these opportunities (Brush et al., 2003). Figure 1 depicts a typical model of the entrepreneurial opportunity process.

The intersection of individuals and opportunities posits the question of “why, when, and how some people and not others discover and exploit the opportunities for the creation of goods and services” (Shane & Venkataraman, 2000). With the opportunity recognized, the decision stage tries to answer the important question “why, faced with an identified opportunity, entrepreneurs will act and non-entrepreneurs will not” (Busenitz et al., 2003). This research investigates whether the perceptions of college students about their ability to recognize, evaluate, and exploit opportunities is associated with their post-graduation start of a business. It is plausible to assume that those who start businesses are likely to believe that they can recognize, evaluate, and exploit opportunities better than those who chose to not start businesses. It is also probable that these individuals can be found among those who elect to take entrepreneurship courses while in college. This leads to a second aspect of this research: to examine what effect taking college courses in entrepreneurship and small business has on the startup of new firms. Given the ongoing debate in the literature over the precise definition of entrepreneurship (Gartner, 1990), we chose to include small business and entrepreneurship courses in the analysis. Figure 2 presents the guiding framework for the study.

Opportunity Recognition Self-Confidence

Self-confidence is defined as a person’s conviction about his or her abilities to execute a given task within an identified context (Moreno, Castillo, & Masere,
According to Neill (2005), self-confidence is a combination of self-esteem and general self-efficacy and refers to belief in one's personal worth and likelihood of succeeding. In a study of Indonesian and Norwegian students, Kristiansen and Indarti (2004) found that self-efficacy, a dimension of self-confidence, had a significant positive relationship with intent to start a firm. Self-efficacy, a person's judgment of how capable he/she is to execute a course of action (Bandura, 1982), has also been found to moderate the relationship between entrepreneurial intentions and entrepreneurial action (Boyd & Vozikis, 1994). In addition, individuals with greater self-efficacy are more likely to pursue an identified opportunity (Chen, Greene, & Crick, 1998). In light of the difficulty in starting a new business, Cooper and Lucas (2006) note it is not surprising that a high level of confidence is central to starting companies. Self-confidence is important for entrepreneurial vision (Ensley, Carland, & Carland, 2000) and the belief that one can recognize an opportunity worth pursuing (Krueger & Brazeal, 1994). Many opportunities for an entrepreneurial discovery may exist. However, a potential entrepreneur must be able to believe that he/she can be successful in order to recognize an opportunity and the value present in the opportunity before he/she can exploit it. This discussion leads to the following hypothesis:

**Hypothesis 1:** There is a positive relationship between post-graduation startup of new firms and the opportunity recognition self-confidence of business college graduates.

**Willingness to Accept Risk**

Deciding whether an idea is an opportunity involves judgments made under conditions of uncertainty and complexity (Allinson, Chell, & Hayes, 2000; Das & Teng, 1997). As noted in the literature, it is commonly accepted that entrepreneurs face risks due to the more variable and less certain rewards they will receive when compared to traditional wage earners (Cramer, Hartog, Jonker, & Van Praag, 2002). Keh, Foo, and Lim (2002) note that risk, the probability that an entrepreneur will be able to successfully turn an idea into an opportunity, is closely associated with uncertainty. From a practical standpoint, perceived risk
is a significant aspect of how entrepreneurs evaluate available ideas and make a decision to act (Brouwer, 2000).

Every enterprise has to accept some level of risk in order to do business. Risk represents both an opportunity to make profit and the potential to make losses in any decision between becoming self-employed and becoming a wage earner (Keh et al., 2002). What is important is risk acceptance—the ability to balance the opportunities against the potential losses and be willing to act. Shane and Venkataraman (2000) present that not every opportunity that is discovered will be exploited and individual differences (such as risk perceptions) will have an impact on this decision. As Kliem and Ludin (1997) note, people with a risk acceptance orientation view risk as neither good nor bad, but a fact of life. They accept risk as it arises and prepare for the most likely cases. Their risk management has a balanced, even optimistic, perspective. The results of Cramer and colleagues (2002) show those who display less aversion to risk are more likely to pursue entrepreneurship, although causality cannot be determined. We expect that business school college graduates who start firms may be more willing to accept the risks than those who do not start businesses and thus present the following hypothesis:

**Hypothesis 2:** There is a positive relationship between post-graduation startup of new firms and the business college graduate's willingness to accept the risks involved.

**Willingness to Act**

In addition to the ability to recognize an opportunity and the willingness to accept risk, the actual launch of a new business requires a decision to act and exploit the opportunity (Wu & Knott, 2006). Entrepreneurial exploitation has been addressed primarily from the resource-based view of the firm (Alvarez & Busenitz, 2001; Choi & Shepherd, 2004; Newbert, 2005; Wu, 2007). Often, opportunities are not exploited and boats are missed (Mullins & Forlani, 2005), indicating it is not the accumulation of resources that results in exploitation, but rather the entrepreneur’s willingness to make the choice to start (Shaver & Scott, 1991). It is suggested here, following Bateman and Crant (1999), that a potential entrepreneur’s willingness to act will be associated with the exploitation of an opportunity. A proactive person has a willingness to act, to take the initiative to do something to create change (Bateman & Crant, 1999).

Individuals who are willing to act do not just wait for things to happen; they actively seek to be involved of their own accord. Put another way, Crant (2000, p. 436), in a review of the proactive personality literature, identify these individuals as “taking initiative in improving current circumstances or creating new ones; it involves challenging the status quo rather than passively adapting to present conditions.” A person who is willing to act is one who, in relative terms, is not
constrained by the situation and who causes environmental change (Bateman & Crant, 1993). In one of the first works to examine the willingness to act and entrepreneurship, Crant (1996) found proactivity accounted for 17% of the variance in entrepreneurial intent among the students in his study. A willingness to act represents a construct at the interface between a potential entrepreneur’s individual orientation and his/her view of the environment. The above suggests that a person’s willingness to act could be a factor involved in the decision to exploit an opportunity that has been identified and evaluated; therefore, we offer the following hypothesis:

Hypothesis 3: There is a positive relationship between post-graduation startup of new firms and the business college graduate’s willingness to act.

Effect of Taking Courses in Entrepreneurship—Is More Better?
Block and Sandner (2008) note that the role of formal education has been a significant part of the analysis of human capital and entrepreneurship. Formal education is seen as providing the necessary cognitive skills to adapt to environmental changes (Hatch & Dyer, 2004). Education is a source of knowledge, skills, discipline, motivation, and self-confidence (Cooper, Gimeno-Gascon, & Woo, 1994). Highly educated entrepreneurs may be better able to deal with complex problems. They may also leverage their knowledge and the social contacts generated through the education system to acquire resources to identify and exploit business opportunities (Arenius & De Clercq, 2005; Shane, 2003). Rosa (2003) suggests several benefits that education can provide entrepreneurs, including a foundation for life-long learning and credibility when seeking resources for their new venture.

Human capital theory maintains that a higher stock of knowledge provides individuals with a higher cognitive ability, which leads to more productive and efficient activity. As a dimension of human capital, education can be a source of both knowledge and motivation (Cooper et al., 1994), two essential ingredients Shane (2003) identifies as important for entrepreneurial opportunity recognition. Hao, Seibert, and Hills (2005) found that formal education would lead to greater entrepreneurial intentions and Delmar and Davidsson (2000) found that education has been generally accepted as having an effect on the decision to start a new business. Hence, individuals with more entrepreneurially-related knowledge or with a stock of knowledge of higher quality are better at perceiving and exploiting entrepreneurial opportunities than are entrepreneurs with less human capital (Davidsson & Honig, 2003; Shane, 2000).

While entrepreneurs and non-entrepreneurs do not differ in the level of their education (Robinson, Stimpson, Huefner, & Hunt, 1991), research suggests that education specifically related to venture creation can affect an individual’s deci-
sion to start a business. Clark et al. (1984) discovered taking an entrepreneurship class had a significant effect on the motivation to actually start a venture. In the study, taking an introductory entrepreneurship/small business course was important in the decision of 67% of those who advanced from intending to start a small business to actually opening the venture. Moreover, studies have shown aspects such as opportunity recognition and innovation can be effectively taught in a class (DeTienne & Chandler, 2004). We thus present the following hypothesis:

**Hypothesis 4a:** There is a positive relationship between the post-graduation startup of new firms and the number of entrepreneurship courses taken by business college graduates.

A recent meta-analysis by Van der Sluis, Van Praag, and Vijverbert (2005) found that the level of education influences the propensity to become self-employed and that, as suggested by Calvo and Wellisz (1980) and Lucas (1978), education enhances managerial ability, which increases the probability of entrepreneurship. Van der Sluis et al. (2005) further find that an additional year of schooling raises enterprise income for new firms in developing economies by an average of 5.5%. This result was similar to those in the United States, where the average return to schooling in entrepreneurial pursuits is 6.1%. Related to this, Bates (1995) found that post-graduate education was strongly associated with those who had chosen to be self-employed. Crant (1996) found that MBA students displayed higher entrepreneurial intentions than undergraduates. This evidence suggests that graduate students are more likely to start new ventures than their undergraduate counterparts, leading to the following hypothesis:

**Hypothesis 4b:** The positive relationship between the post-graduate startup of new firms and the number of entrepreneurship courses taken by business college graduates will be greater for graduate students than for undergraduate students.

The framework presents business school college students’ attitudes toward risk, self-confidence, perceived resources and the network available to the individual, willingness to act, and perceived innovativeness as having a direct effect on the formation of new firms. Finally, the number of entrepreneurship and small business courses taken is expected to be directly related to start of new firms, with MBA courses having a greater effect than undergraduate courses.

**METHODOLOGY**

**Sample and Procedure**

The sample population for the survey was all students in the College of Business Administration at a major public Midwestern university in the United States who took the business capstone course (required for graduation) from 1996 and 2006.
At this university, all entrepreneurship courses are taught by current or former business operators; 60–70% of the course offerings were consistent over the time that students in the sample attended the university. This provided a population of 5,145 students. From this population, the physical and email addresses of 2,921 students were available. A comparison of the students for whom addresses were available and for whom addresses were not available indicated the two groups were not significantly different in age, gender, major, graduation year, or courses taken. All of those with contact information available were contacted on four separate occasions, twice by email and twice using postcards. In both the email and postcard requests for their participation, potential participants received an invitation to participate and a link to the online survey instrument. This link outlined the purpose of the survey, presented a consent form, and provided the survey instrument itself. A total of 850 completed the survey, a response rate of 29.1%; 690 provided all of the data necessary for a useable response, a final response rate of 23.6%. Approximately 28% of all respondents started a business by the time of the survey. This suggests a response bias in favor of students who started businesses.

Table 1 presents the descriptive statistics of the sample. The mean age of the respondents was approximately 29 and the majority were male (60%). Most of the respondents (75%) had obtained a bachelor’s degree. The mean number of years since graduation for the sample was 5.36 (SD = 2.95). In addition, respondents had taken an average of 0.59 undergraduate entrepreneurship courses (SD = 0.937) and 0.13 graduate entrepreneurship courses (SD = 0.425). These numbers suggest that few students in the business college take any entrepreneurship courses. Tables 2 and 3 present the percentages of businesses started, broken down by major and type of course taken. Businesses were started most frequently by business administration majors (9.3%) and least frequently by accounting (0.7%) and international business (0.4%) majors. Students who took no entrepreneurship
Table 2  
Number of Businesses Started by Major

<table>
<thead>
<tr>
<th>Major</th>
<th>Start Count</th>
<th>Not Start</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business Administration</td>
<td>63</td>
<td>131</td>
<td>194</td>
</tr>
<tr>
<td>% within Major</td>
<td>32.5</td>
<td>67.5</td>
<td>100.0</td>
</tr>
<tr>
<td>% within Start</td>
<td>33.3</td>
<td>26.6</td>
<td>28.5</td>
</tr>
<tr>
<td>% of Total</td>
<td>9.3</td>
<td>19.2</td>
<td>28.5</td>
</tr>
<tr>
<td>Management</td>
<td>34</td>
<td>73</td>
<td>107</td>
</tr>
<tr>
<td>% within Major</td>
<td>31.8</td>
<td>68.2</td>
<td>100.0</td>
</tr>
<tr>
<td>% within Start</td>
<td>18.0</td>
<td>14.8</td>
<td>15.7</td>
</tr>
<tr>
<td>% of Total</td>
<td>5.0</td>
<td>10.7</td>
<td>15.7</td>
</tr>
<tr>
<td>Marketing</td>
<td>33</td>
<td>75</td>
<td>108</td>
</tr>
<tr>
<td>% within Major</td>
<td>30.6</td>
<td>69.4</td>
<td>100.0</td>
</tr>
<tr>
<td>% within Start</td>
<td>17.5</td>
<td>15.2</td>
<td>15.9</td>
</tr>
<tr>
<td>% of Total</td>
<td>4.8</td>
<td>11.0</td>
<td>15.9</td>
</tr>
<tr>
<td>Finance</td>
<td>17</td>
<td>91</td>
<td>108</td>
</tr>
<tr>
<td>% within Major</td>
<td>15.7</td>
<td>84.3</td>
<td>100.0</td>
</tr>
<tr>
<td>% within Start</td>
<td>9.0</td>
<td>18.5</td>
<td>15.9</td>
</tr>
<tr>
<td>% of Total</td>
<td>2.5</td>
<td>13.4</td>
<td>15.9</td>
</tr>
<tr>
<td>MIS/Comp Sci</td>
<td>13</td>
<td>14</td>
<td>27</td>
</tr>
<tr>
<td>% within Major</td>
<td>48.1</td>
<td>51.9</td>
<td>100.0</td>
</tr>
<tr>
<td>% within Start</td>
<td>6.9</td>
<td>2.8</td>
<td>4.0</td>
</tr>
<tr>
<td>% of Total</td>
<td>1.9</td>
<td>2.1</td>
<td>4.0</td>
</tr>
<tr>
<td>Accounting</td>
<td>5</td>
<td>48</td>
<td>53</td>
</tr>
<tr>
<td>% within Major</td>
<td>9.4</td>
<td>90.6</td>
<td>100.0</td>
</tr>
<tr>
<td>% within Start</td>
<td>2.6</td>
<td>9.8</td>
<td>7.8</td>
</tr>
<tr>
<td>% of Total</td>
<td>0.7</td>
<td>7.0</td>
<td>7.8</td>
</tr>
<tr>
<td>International Business</td>
<td>3</td>
<td>23</td>
<td>26</td>
</tr>
<tr>
<td>% within Major</td>
<td>11.5</td>
<td>88.5</td>
<td>100.0</td>
</tr>
<tr>
<td>% within Start</td>
<td>1.6</td>
<td>4.7</td>
<td>3.8</td>
</tr>
<tr>
<td>% of Total</td>
<td>0.4</td>
<td>3.4</td>
<td>3.8</td>
</tr>
<tr>
<td>Other</td>
<td>21</td>
<td>37</td>
<td>58</td>
</tr>
<tr>
<td>% within Major</td>
<td>36.2</td>
<td>63.8</td>
<td>100.0</td>
</tr>
<tr>
<td>% within Start</td>
<td>11.1</td>
<td>7.5</td>
<td>8.5</td>
</tr>
<tr>
<td>% of Total</td>
<td>3.1</td>
<td>5.4</td>
<td>8.5</td>
</tr>
<tr>
<td>Count</td>
<td>189</td>
<td>492</td>
<td>681</td>
</tr>
<tr>
<td>% of Total</td>
<td>27.8</td>
<td>72.2</td>
<td>100.0</td>
</tr>
<tr>
<td>Course Type</td>
<td>Start</td>
<td>Not Start</td>
<td>Total</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-------</td>
<td>-----------</td>
<td>-------</td>
</tr>
<tr>
<td>None</td>
<td>67</td>
<td>302</td>
<td>369</td>
</tr>
<tr>
<td>% within Courses</td>
<td>18.2</td>
<td>81.8</td>
<td>100.0</td>
</tr>
<tr>
<td>% within Start</td>
<td>35.3</td>
<td>60.4</td>
<td>53.5</td>
</tr>
<tr>
<td>% of Total</td>
<td>9.7</td>
<td>43.8</td>
<td>53.5</td>
</tr>
<tr>
<td>Business Plan</td>
<td>45</td>
<td>61</td>
<td>106</td>
</tr>
<tr>
<td>% within Courses</td>
<td>42.5</td>
<td>57.5</td>
<td>100.0</td>
</tr>
<tr>
<td>% within Start</td>
<td>23.7</td>
<td>12.2</td>
<td>15.4</td>
</tr>
<tr>
<td>% of Total</td>
<td>6.5</td>
<td>8.8</td>
<td>15.4</td>
</tr>
<tr>
<td>Small Business Mgmt.</td>
<td>27</td>
<td>87</td>
<td>114</td>
</tr>
<tr>
<td>% within Courses</td>
<td>23.7</td>
<td>76.3</td>
<td>100.0</td>
</tr>
<tr>
<td>% within Start</td>
<td>14.2</td>
<td>17.4</td>
<td>16.5</td>
</tr>
<tr>
<td>% of Total</td>
<td>3.9</td>
<td>12.6</td>
<td>16.5</td>
</tr>
<tr>
<td>Tools Courses</td>
<td>3</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>% within Courses</td>
<td>75.0</td>
<td>25.0</td>
<td>100.0</td>
</tr>
<tr>
<td>% within Start</td>
<td>1.6</td>
<td>0.2</td>
<td>0.6</td>
</tr>
<tr>
<td>% of Total</td>
<td>0.4</td>
<td>0.1</td>
<td>0.6</td>
</tr>
<tr>
<td>Plan &amp; Management</td>
<td>15</td>
<td>19</td>
<td>34</td>
</tr>
<tr>
<td>% within Courses</td>
<td>44.1</td>
<td>55.9</td>
<td>100.0</td>
</tr>
<tr>
<td>% within Start</td>
<td>7.9</td>
<td>3.8</td>
<td>4.9</td>
</tr>
<tr>
<td>% of Total</td>
<td>2.2</td>
<td>2.8</td>
<td>4.9</td>
</tr>
<tr>
<td>Plan &amp; Tools</td>
<td>12</td>
<td>4</td>
<td>16</td>
</tr>
<tr>
<td>% within Courses</td>
<td>75.0</td>
<td>25.0</td>
<td>100.0</td>
</tr>
<tr>
<td>% within Start</td>
<td>6.3</td>
<td>0.8</td>
<td>2.3</td>
</tr>
<tr>
<td>% of Total</td>
<td>1.7</td>
<td>0.6</td>
<td>2.3</td>
</tr>
<tr>
<td>Management &amp; Tools</td>
<td>5</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>% within Courses</td>
<td>25.0</td>
<td>75.0</td>
<td>100.0</td>
</tr>
<tr>
<td>% within Start</td>
<td>2.6</td>
<td>3.0</td>
<td>2.9</td>
</tr>
<tr>
<td>% of Total</td>
<td>0.7</td>
<td>2.2</td>
<td>2.9</td>
</tr>
<tr>
<td>Plan, Mgmt. &amp; Tools</td>
<td>16</td>
<td>11</td>
<td>27</td>
</tr>
<tr>
<td>% within Courses</td>
<td>59.3</td>
<td>40.7</td>
<td>100.0</td>
</tr>
<tr>
<td>% within Start</td>
<td>8.4</td>
<td>2.2</td>
<td>3.9</td>
</tr>
<tr>
<td>% of Total</td>
<td>2.3</td>
<td>1.6</td>
<td>3.9</td>
</tr>
<tr>
<td>Count</td>
<td>190</td>
<td>500</td>
<td>690</td>
</tr>
<tr>
<td>% of Total</td>
<td>27.5</td>
<td>72.5</td>
<td>100.0</td>
</tr>
</tbody>
</table>
courses started businesses most frequently (9.7%) while the business plan course had the highest frequency of new firm startups (6.5%) among students.

Based on a literature review of prior research examining similar constructs, questions for the survey instrument were developed or adapted from previous work. Specific sources for the questions in the survey are discussed along with the appropriate constructs in the following section. Following development, the survey was submitted to and approved by the two appropriate research boards for the university tasked with ensuring responsible research conduct and compliance within the regulations where the study was performed. Following university approval, the survey was pilot tested by six members of the entrepreneurship field. After the pilot study participants had reviewed the survey, they were contacted and their feedback regarding the survey instrument led to clarification of wording on some of the questions, as well as the elimination of some questions.

This following section briefly defines the constructs used in the study and discusses the development of measures used in the survey instrument. All items were self-report measures on a five-point scale (strongly disagree to strongly agree).

Measurement of the Dependent Variable:
Post Graduation Startup of New Firms
We define entrepreneurial action in this study as the actual startup of a new venture. We specified startups that occurred after graduation. Respondents to the survey were asked “Since graduation, have you ever started your own entrepreneurial venture?”

Measurement of the Independent Variables
A list of all graduate and undergraduate courses in entrepreneurship and small business available at the university where the study took place were provided in the survey. Respondents were asked to indicate the courses they had taken. These results were then used to tabulate the total courses taken for each respondent by course level (undergraduate or graduate). The courses defined as entrepreneurship courses for this study were all those listed by the university as part of their entrepreneurship category including “Identifying and Assessing Entrepreneurial Opportunities,” “Entrepreneurship and Venture Management,” “Small Business Management,” and “Introduction to Entrepreneurial Management.”

To measure opportunity recognition self-confidence in the survey instrument, participants were asked to respond, using a five-point scale, to five questions about their self-assessment of their ability to recognize opportunities. The five items (Chronbach’s Alpha = .816) can be found in Table 4.
<table>
<thead>
<tr>
<th>Items</th>
<th>Ready</th>
<th>Opp</th>
<th>Risk</th>
<th>Pro</th>
</tr>
</thead>
<tbody>
<tr>
<td>I can establish and achieve goals and objectives.</td>
<td>.771</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am hard-working and can do what it takes to succeed.</td>
<td>.761</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am motivated by success and the desire to do well.</td>
<td>.617</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I can formulate a plan of action to pursue an opportunity.</td>
<td>.578</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I can spot a good opportunity long before others can.</td>
<td>.756</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am constantly seeing business opportunities or ideas that have potential commercial value.</td>
<td>.738</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I excel at identifying opportunities.</td>
<td>.705</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am creative and often come up with ideas involved in starting and managing a business.</td>
<td>.0701</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am willing to take the financial risks involved in starting and managing a business.</td>
<td>.738</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have the ability to acquire financial capital.</td>
<td>.736</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am willing to take the career risk of leaving a job to start my own business.</td>
<td>.705</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I have other options to fall back on if my venture does not work out.</td>
<td>.576</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My family and friends support me and my business undertakings.</td>
<td>.551</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel driven to make a difference in my community and maybe the world.</td>
<td>.722</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>If I see someone in trouble, I help out in any way I can.</td>
<td>.709</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am constantly on the lookout for new ways to improve my life.</td>
<td>.633</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I love being a champion for my ideas, even against others’ opposition.</td>
<td>.602</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am always looking for better ways to do things.</td>
<td>.597</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nothing is more exciting than seeing my ideas turn into reality.</td>
<td>.579</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Principal Components, Varimax Rotation  | Items 4 5 5 6  
N = 690  | Chronbach’s Alpha: .722 .816 .746 .779
The items measuring risk acceptance asked if the respondents were willing to accept some of the risks inherent in the startup of a new firm (i.e., career risk). As noted in Table 4, five items were used to measure this construct (Chronbach’s Alpha = .746). Willingness to act was measured using six items from Bateman and Crant’s (1993) proactive personality scale (Chronbach’s Alpha = .779; see Table 4).

### Measurement of the Control Variables

Three variables were used as control variables. Respondents were asked the year they were born, their gender, and their assessment of their managerial readiness. Managerial readiness was measured by four items (Chronbach’s Alpha = .722; see Table 4 for the item descriptions and factor loadings).

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Mean, Standard Deviation, and Bivariate Correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>Start</td>
<td>0.280</td>
</tr>
<tr>
<td>Age</td>
<td>28.712</td>
</tr>
<tr>
<td>Gender</td>
<td>0.601</td>
</tr>
<tr>
<td>Mnr</td>
<td>0.000</td>
</tr>
<tr>
<td>Read i</td>
<td>0.000</td>
</tr>
<tr>
<td>OppRec</td>
<td>0.000</td>
</tr>
<tr>
<td>Risk</td>
<td>0.000</td>
</tr>
<tr>
<td>Accept</td>
<td>0.000</td>
</tr>
<tr>
<td>ActWill</td>
<td>0.000</td>
</tr>
<tr>
<td>TUgrad</td>
<td>0.590</td>
</tr>
<tr>
<td>Tgrad</td>
<td>0.130</td>
</tr>
</tbody>
</table>

*Correlation is significant at the 0.05 level (2-tailed)
**Correlation is significant at the 0.01 level (2-tailed)
<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 1</td>
<td>Model 2</td>
<td>Model 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted</td>
<td>Predicted</td>
<td>Predicted</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% correct</td>
<td>% correct</td>
<td>% correct</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>Observed</td>
<td>Observed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>%</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>463</td>
<td>100</td>
<td>455</td>
<td>8</td>
<td>98</td>
<td>417</td>
</tr>
<tr>
<td>Start</td>
<td>1</td>
<td>175</td>
<td>0</td>
<td>0</td>
<td>171</td>
<td>4</td>
</tr>
<tr>
<td>Overall</td>
<td>72.7</td>
<td>71.9</td>
<td>77.7</td>
<td>80.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percentage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>-2.45</td>
<td>0.56</td>
<td>0.00</td>
<td>-3.72</td>
<td>0.68</td>
<td>0.00</td>
</tr>
<tr>
<td>s.e.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sig.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>exp(b)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Block</td>
<td>Constant</td>
<td>-0.80</td>
<td>0.20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Management</td>
<td>Readiness</td>
<td>-0.09</td>
<td>0.09</td>
<td>-0.17</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>Personal</td>
<td>Diff</td>
<td>Opportunity</td>
<td>Regression</td>
<td>0.60</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>Risk Acceptance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Willingness to Act</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>Total Undergrad Courses</td>
<td>0.36</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Grad Courses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hosmer &amp; Lemeshow Test</td>
<td>Chi-square df</td>
<td>Sig.</td>
<td>Chi-square df</td>
<td>Sig.</td>
<td>Chi-square df</td>
<td>Sig.</td>
</tr>
<tr>
<td>Table 6</td>
<td>Results of Binary Regression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
RESULTS AND DISCUSSION

The bivariate correlations for the variables used in the study are presented in Table 5. The results from the binary logistic regression are presented in Table 6. Because one of the purposes of this study was to determine whether taking a number of entrepreneurship courses added value to a student’s launch of a new business after graduation, the data were analyzed using a block entry approach in logistical regression. Block 1 contains only the control variables, Block 2 the individual characteristics variables, and Block 3 the entrepreneurship course variables.

Results of Correlation Analysis

A review of the Pearson correlations in Table 2 indicates that all control and independent variables, with the exception of managerial readiness, are significantly associated with the start of new businesses by the graduates in the sample population. Perhaps reflecting some common methods, the number of undergraduate courses taken is significantly correlated with other independent variables. These results provide general confidence that the variables in the study do affect the post graduation launch of new businesses.

Results of the Logistical Regression Analysis

Model 1: Control variable block. As presented in Table 3, the sample population contained a total of 690 subjects, of which 190 started businesses after graduation. In Model 1, which tests the influence of the control variables on the dependent variable, age \( (B = 0.06, p = 0.00) \) and gender \( (B = -0.80, p = 0.00) \) are significant and managerial readiness is not \( (B = -0.09, p = 0.35) \). Of some interest in this first model is the inverse association between starting a business, gender, and readiness. In this study, this suggests female business college graduates are less likely to start businesses after graduation than males.

The classifications presented in Table 3 are all 2 x 2 tables, which tally correct and incorrect estimates for the null model with only the constant. The columns are the two predicted values of the dependent, while the rows are the two observed (actual) values of the dependent. In a perfect model, all cases will be on the diagonal and the overall percent correct will be 100%. If the logistic model has homoscedasticity, the percent correct will be approximately the same for both rows. Here it is not, with the model predicting non-start cases, but not predicting any start cases. While the overall percent correctly predicted seems moderately good at 71.9%, it must be noted that estimating the most frequent category (non-start) for all cases would yield a slightly better percent correct (72.6%) and only about 2% of the start cases are predicted.

The Hosmer–Lemeshow fit test was performed by dividing the predicted probabilities into deciles and then computing a Pearson chi-square that compares the predicted to the observed frequencies. Here the test is computed from the chi-
square distribution \((df = 8)\) and indicates whether the logistic model is a good fit. Lower values (and non-significance) indicate a good fit to the data and, therefore, good overall model fit. That is, if the Hosmer and Lemeshow Goodness-of-Fit test statistic is 0.05 or less, we reject the null hypothesis that there is no difference between the observed and predicted values of the dependent variable.

The Nagelkerke \(R^2\) is a pseudo \(R^2\)-squared. Logistic regression does not have an equivalent to the \(R^2\)-square that is found in ordinary least squares (OLS) regression. This statistic does not provide the same information as \(R^2\)-squared in OLS regression (the proportion of variance explained by the predictors). Because a large deviance can be thought of as a measure of how poorly the model fits (i.e., lack of fit between observed and predicted values), an analogy can be made by the sum of squared residual in ordinary least squares. The proportion of unaccounted variance that is reduced by adding variables to the model is the same as the proportion of variance accounted for, or \(R^2\). Here, the models indicate that the addition of each category of variables improves the predictive value of the model, albeit not dramatically in the final two steps of the model.

The Hosmer and Lemeshow Goodness-of-Fit Test has a significant score of \(p = 0.95\). Here, the data suggest we cannot reject the null hypothesis and that the model is not a good fit. The Cox-Snell \(R^2\) and Nagelkerke \(R^2\) are attempts to provide a logistic analogy to \(R^2\) in OLS regression. The Nagelkerke measure adapts the Cox–Snell measure so that it varies from 0 to 1, as does \(R^2\) in OLS. The Cox and Snell \(R^2\) is .05 and Nagelkerke pseudo-\(R^2\) is only .07. This too indicates that the control variables are not related to which students started businesses after graduation.

Model 2: Personal characteristics block. This model tested the inclusion of the three personal characteristic independent variables: opportunity recognition self-confidence, willingness to accept risk, and willingness to act. H1 stated the opportunity recognition self-confidence of a business college student would be positively associated with the startup of a new firm. This hypothesis is supported in the binary logistic regression (\(B = 0.60, p = .00\)). The effect of a college student’s acceptance of the inherent risks in the startup of new firms was the focus of the second hypothesis. H2 is also supported (\(B = 1.27, p = .00\)). Those who were more willing to accept the risks start a new firm more often. In fact, the odds ratio (ExpB) of 3.445 suggests that acceptance of risk might contribute most to the start of a new business indicating that for each 1-unit increase in their willingness to accept risk, the odds of an individual starting a new firm are expected to increase by more than three times. Willingness to act was the focus of the third hypothesis and indicated that willingness to act would be associated with the start of a new business by business college graduates. H3 is supported (\(B = 0.24, p = .03\)). College students who tend to be willing to act are more likely to start a new business than those who are not. Finally, the analysis of Model 2 indicates the only significant control variable is age.
The classification for Model 2 presented in Table 3 indicates that the overall prediction percent improves somewhat to 77.7%. This individual characteristics model, however, improves the classification of starts to more than 45%. The Hosmer and Lemeshow Goodness-of-Fit Test has a significance score of $p = .34$, indicating that the data suggest the null hypothesis can be rejected and that the model is a reasonably good fit. The Cox and Snell $R^2$ is .25 and the Nagelkerke pseudo-$R^2$ is .36, demonstrating that this model has better explanatory power. Once again, the analysis of Model 2 indicates that the control variables are not good predictors of which students started businesses after graduation.

**Model 3: Number of entrepreneurship courses block.** Block 3 adds the undergraduate and graduate course variables into the model. H4a suggests that the number of entrepreneurship courses taken by business college undergraduates and graduates would be directly associated with their post-graduation start of a new business, and thus indicate that taking more courses adds value to one's personal entrepreneurial characteristics as an antecedent to new venture creation. As presented in Table 3, both total number of undergraduate courses taken ($B = 0.36, p = .00$) and graduate courses taken ($B = 0.56, p = .03$) by business college students are significantly associated with the post-graduation start of a new business. Thus, H4a is supported. Hypothesis 4b suggested that graduate level courses would have a stronger relationship with business starts than undergraduate courses. H4b is also supported. The effect of graduate courses ($B = 0.56$) was markedly larger than for undergraduate courses ($B = 0.36$) in the regression. The difference in the odds ratios for graduate ($B = 1.747$) versus undergraduate ($B = 1.42$) courses also supports this hypothesis.

The classification for the full model, Model 3, presented in Table 3 indicates that the overall percent correctly predicted improves to 80.4%. The full model also improves the correct classification of starts to 53%. The Hosmer and Lemeshow Goodness-of-Fit Test has a significance score of $p = 0.12$, indicating that the data suggest the null hypothesis can be rejected and that the model is a very good fit. The Cox and Snell $R^2$ is .26 and the Nagelkerke pseudo-$R^2$ is .38, demonstrating that this model has somewhat better explanatory power than the individual characteristics model. Once again, the analysis of Model 3 indicates that the control variables (gender and managerial readiness) are not good predictors of which students started businesses after graduation, with age as the only significant predictor, suggesting older graduates are more likely to start new firms.

**Discussion**

The results of this study suggest several implications for entrepreneurship education, which we discuss in terms of the categories presented in our review of the literature (effects of education on entrepreneurs and aspects of entrepreneurship that can be taught). When considering the effects of entrepreneurship education on development of new ventures, these results lend support to the body of evidence that education does matter. In fact, they suggest more is better.
supports and extends the findings of other researchers, such as Clark et al. (1984) and Charney and Libecap (2000). While these results do not definitively answer the question of whether entrepreneurship can be taught, they do provide answers to the three questions the study raised: Is the increasing availability of entrepreneurship courses actually increasing the number of entrepreneurs? What is the impact of the large amount of resources being expended to offer more and more entrepreneurship courses and programs? Are students who take entrepreneurship courses more likely to start a new firm than those who do not? To each, the results of this study suggest the answer is “yes.” At least some aspects of entrepreneurship that students learn in these courses make them more likely to start a firm.

Perhaps the most important implications are related to how education can lead to additional students deciding to exploit the opportunities they have identified and start new firms following graduation. This is another study producing results that show self-confidence is important in the startup of new firms, in this instance, that one has a keener ability to recognize opportunities that can be turned into successful new ventures. Those educating potential entrepreneurs should include activities that improve the self-confidence of these individuals that they can succeed as entrepreneurs. This increase in self-confidence in recognizing opportunities may be one way in which education can impact the individual's decision to start a new firm.

A focus on those individuals willing to accept the risks of starting a new venture could be another implication. Thus, programs should be directed toward those willing to take the risk in order to be more effective. It is noted that the reverse relationship may also be at work with regard to risk. Those who have taken entrepreneurship courses may feel they are better prepared to undertake the risk, leading them to be more likely to start a new firm. This may be an additional way in which entrepreneurship education has an impact on the individual's decision to start a new firm.

The results demonstrate that willingness to act is an important predictor of the start of a business. People who are more proactive are relatively unconstrained by situational forces and willing to affect environmental change (Bateman & Crant, 1993). Students who are willing to act are efficiently motivated from their inner self and take advantage of opportunities to improve their situations. Proactive behavior involves stepping forward to either improve current situations and circumstances or to create new ones. Thus, a willingness to act helps people deal with expected or unexpected events and changes as well as enables them to influence and transform their environment. A proactive person is one who scans for opportunities, shows initiative, takes action, and perseveres until he or she reaches closure by bringing about change (Bateman & Crant, 1993). Crant (1996, p. 47) suggests one explanation for these findings is that more proactive people tend to envision creating situations that will allow them to capitalize on their
personality. In combination, these three dimensions and the entrepreneurship courses suggest that entrepreneurship education might help those leaning toward a career in entrepreneurship to understand more clearly that they often both see what others do not see and are willing to act when others will not.

Limitations
One of the limitations of this study is a concern over range restriction in the number of courses, noting the small standard deviations of the two course variables. Because the sample was taken from only one university, the generalizability of the results is limited, with the results only being applicable to those institutions with similar programs. Another limitation stems from the structure of the program at the university where the sample was obtained. In this case, the graduate and undergraduate courses are taught concurrently, providing similar instruction to graduate and undergraduate students. This may suggest that the results obtained for graduate courses is due to graduate students being more likely to start firms as opposed to graduate courses themselves having a greater impact. The higher response rate from more recent graduates provides another limitation to the study. The reclassification issues present may also be considered a limitation, although this issue is less of a concern in the full model. The study also suffered from a high degree of multicollinearity across a number of the variables. Perhaps this is due to the variables selected. We have considered that this may be due to those who already intend to start a firm self-selecting into the courses. Another possible limitation is that the responses to all items in the study were obtained at the same time (following graduation). It is possible that some respondents provided the responses they did because they have already started a new firm successfully.

Directions for Future Research
Several future directions for research are available. A longitudinal study would be a good next step, thanks in part to the ability to observe effects that occur over time (Brockhaus, 1987; Clouse, 1990). This would enable researchers to obtain data on the independent variables before students had successfully started a new firm, eliminating concerns that successful startup influenced responses to these items. As noted by several other authors who have proposed similar frameworks, we do not claim to have examined all of the possible antecedents, moderators, characteristics, and so on which may be applicable to the relationships being studied (Autio et al., 1997; Luthje & Franke, 2003; Robinson et al., 1991). Another avenue for future research would be to identify these other factors, which may lead to a deeper understanding of the relationships studied in this paper (e.g., prior family history of entrepreneurship). Future research could also examine the causal relationships between the aspects considered here (Baum & Locke, 2004; Gist, 1987; Luthje & Franke, 2003). For example, do individuals who enroll in entrepreneurship courses develop a more positive view of entrepreneurship or do they self-select into courses because they already have a positive view
of entrepreneurship? A longitudinal study would be beneficial in determining the nature of these causal relationships (Autio et al., 1997).

Adding the performance of the new firm to the picture could also provide beneficial insights on the impact of entrepreneurship education. While this study suggests how education may lead to additional firms being started, it does not examine if the most promising businesses are being pursued or if the education obtained by the founder results in improved performance for the new firm. Future research could consider how the entrepreneurship education of the founder or members of the new venture impact its performance initially and in the long run. In regard to the education aspect of this study, future research could contribute by examining the pedagogical impact on the outcomes studied here. For example, do entrepreneurship courses on innovation have a greater impact than courses on small business management. Methodology used in the courses themselves (i.e., lecture versus case studies) may also prove a fruitful aspect for future research. Another potential direction for future research would be to consider the interaction of willingness to accept risk and willingness to act to help explore what combinations of these constructs will lead to the startup of a new business.

Since the precursors of intentions are not static, the possibility exists to affect an individual’s intent to engage in entrepreneurial activity (Krueger & Carsrud, 1993). It may be possible through education to encourage more individuals to consider entrepreneurship as a viable career option. The intent of this paper is to enrich the extant literature that aims to answer the question of why some individuals choose to start a business while others do not. We have synthesized a variety of areas with the hope of providing additional insight into the many aspects involved in why individuals start new firms. We encourage others to continue the development of this often complicated area and look forward to the future insights yet to be revealed.
REFERENCES


ETHNIC DIVERSITY, TECHNICAL EFFICIENCY, DEMOCRACY MEASURES, AND ECONOMIC GROWTH IN SUB-SAHARIAN AFRICA

SEYMOUR PATTERSON
TRUMAN STATE UNIVERSITY

ABSTRACT

The paper examines the hypotheses of the effects of ethnicity, efficiency, and democracy on economic growth, and their effects on the efficiency of capital. Word Bank 2004 data source for 1982–2002 on real gross domestic product, the labor force, and investment were used in capital computation. Ethno-linguistic fragmentation—one minus the Herfindahl Index—is the probability of two persons drawn randomly from a country’s population are from different ethnic groups. Higher ethno-linguistic fragmentation values mean more ethnic diversity. The Posner (2004) politically relevant ethnic group, which measures competition among ethnic groups on policymaking decisions, was also used. Kekic (2007) EIU Index of Democracy is the data source for flawed democracies, and hybrid regimes. The paper also examines the effects of ethnic diversity and democracy on the efficiency of capital.

The methodology is based on a Cobb–Douglas equation augmented for ethnic diversity and measures of democracy. The regression results indicate capital and labor have positive effects on economic growth. Politically relevant ethnic group had a positive effect on economic growth, while ethno-linguistic fragmentation had a negative effect. Flawed democracies and democracy index had positive effects on economic growth; but hybrid regimes had no effect on economic growth. Generally, the effects of politically relevant ethnic group and ethno-linguistic fragmentation on the efficiency of capital are ambiguous—politically relevant ethnic group is positive, and ethno-linguistic fragmentation is negative. Flawed democracies and hybrid regimes are positive and negative on economic technical efficiency, respectively. The results support the notion that ethnic diversity and democracy have positive effects on economic growth. Policy makers should promote democracy to increasing economic growth in sub-Saharan Africa.
Key Words: growth, efficiency, fragmentation, ethnic, diversity, democratic, technical

INTRODUCTION
The paper sets out to address three hypotheses that are generally accepted as true—ethnic diversity reduces real economic growth, ethnic diversity impairs the efficiency of capital, and democratic regimes use resources more efficiently and have faster real growth. While sub-Saharan Africa (SSA) countries are not monolithic, poor economic performance appears to be a common intractable problem. The problem of slow economic growth might be related to ethnic diversity or fragmentation growth in SSA countries. Ethnic diversity (ED) or fragmentation can be sources of political instability (i.e., conflicts), which creates an uncertain climate for investment. ED can lead to competition over resources that are sources of income and rent. ED can also be a motivation for less than optimal policy choices, for example, efforts to keep control of the government, policies that favor members of the ethnic group in power, and practices of corruption. Among other things, the development literature seeks to find the nexus between economic growth and ED. Ultimately, the goal appears to be one of establishing whether there is indeed a connection between economic growth and ethnic diversity; that is, whether the more ethnically diverse societies suffer from slow growth performance.

While much research has analyzed the connection between economic growth and ethnicity, there appears to be scant effort to address how ethnicity impacts efficiency (of capital). This is despite the fact that there is a belief in a causal process leading from ethnicity to political instability to investment to slow economic growth. There is also the question of the role of institutional factors—such as democratization—on the efficiency of inputs and on growth. It is not clear whether growth leads to democracy, or democracy leads to economic growth and economic development. This debate will not be addressed in this study, nor is there any intent to resolve it here. However, an interesting take on this debate can be found in Quinn (2003, pp. 243–246). He notes that “Sklar has argued that democracy is developmental (i.e., leads to development).” The driving force behind it is accountability as from the ruler to the ruled (vertical accountability) through the voting process. If fragments of democracy are in place “they promote or deepen other fragments.” A different take on this is that when the political viability of a politician depends on voters, this makes “the implementation of painful, and politically unpopular reforms unlikely” (Quinn, 2003, pp. 243–246). In addition, Quinn (2003) finds that generally freer societies perform better than autocratic ones.

There are elections in SSA but they are not necessarily free and fair. At the core of the performance anomaly of SSA, some would argue, is the presence of ethno-linguistic diversity. A widely used measure of ethnic diversity, ethno-linguistic fractionalization (ELF), gives sub-Saharan Africa an index of 0.65, i.e., 65% prob-
ability that two individuals randomly chosen belong to different groups, compared with an index of 0.29 for all non-African cases (Posner, 2004, Table 4).

**LITERATURE REVIEW**

The economic literature contains studies that examined the link between ED and economic performance. For instance, Easterly and Levine (1997) found that ethnic diversity has a negative effect on economic performance. In fact, they have argued that moving from a homogeneous society to a heterogeneous society can result in more than a 2% reduction in economic growth. Others to investigate this causal process with respect to ethnic diversity on economic performance include Collier and Gunning (1999), Hall and Jones (1999), and Rodrik (1999). These studies used the traditional ELF index, but different computations of ethnic diversity can be found in Roeder (2001).

The ELF index has some shortcomings, among which are “specific problems that stem from the underlying ethnographic data from which the index is calculated, general problems that arise from attempting to summarize a country’s ethnic diversity with a single index (and that apply to all single-measure ethnic fractionalization indices, including the one introduced here), and problems of application that arise from the way the measure is used” (Posner, 2004). To deal with these shortcomings other measures of ethnic diversity have been devised—for instance, the Politically Relevant Ethnic Group (PREG) by Posner (2004) and the Ethnic Diversity and Clustering (EDC), and Ethnic Clustering (EC) constructed by Matuszeski and Schneider (2006).

The ELF index is described in the literature and is calculated based on the population of each ethnic group in the country as a whole. It is constructed using a Herfindahl concentration index of the shares of each ethnic group in the total population \( (n_i/N) \). In this context, the ELF index is equal to the probability that two citizens chosen at random from the country’s population will be from different ethnic groups. The ELF index can be computed from \( ELF = 1 - \sum_{l=1}^{L} \left( N_l/N \right) \), for \( l = 1 \) to \( L \). Because the Herfindahl index measures concentration, the 1-Herfindahl provides a measure of diversity or fragmentation—when applied to ethnic concentration. Thus, it gives us a measure of ethnic diversity. For example, Nigeria has 250 ethnic groups and its ELF index is 0.87; at the other extreme, Swaziland is homogenous (97% African and 3% European). This country has an ELF index of 0.06. In general, the higher the index the more ethnically diversified is the country.

Table 1 below provides a ranking of countries by two indexes of ethnic diversity: PREG and ELF. It shows the 10 least and the 10 most fractionalized countries in sub-Saharan Africa. Botswana is the least fractionalized country in the group among African countries according to PREG but Burundi tops it according to ELF. The most fractionalized country is Zaire (the Democratic Republic of the Congo) by the PREG but Tanzania is the most fractionalized country accord-
I have paired PREG with gross domestic product (GDP) growth rates (from 1982–2002) for both groups. The countries chosen were the Democratic Republic of the Congo (DRC), Cameroon, Zambia, Chad, Nigeria, Angola, and Mauritius for the PREG ranking. The correlation coefficient between PREG and real GDP is -0.73 in the group of the 10 most fractionalized countries. Similarly, using ELF and real GDP growth for DRC, Cameroon, South Africa, Nigeria, Cote d’Ivoire, Chad, and Kenya the correlation coefficient is -0.77 for the 10 most fractionalized countries. These two results suggest that ethnic diversity can lead to slow economic performance in SSA. The source for PREG and ELF is Posner (2004).
economic growth rates. The correlation coefficients for the 10 least fractionalized countries (Botswana, Burkina Faso, Lesotho, Madagascar, Seychelles, Swaziland, Guinea-Bissau, Mali, and Senegal) is -0.33 for PREG and real GDP growth, but and for ELF and real GDP growth for Burundi, Madagascar, Rwanda, Lesotho, Mauritania, Botswana, Zimbabwe, Mauritius, and Benin, the correlation is 0.67. In isolation, these numbers appear to challenge the hypothesis that ethnic diversity reduces economic growth. However, compared with the correlation of -0.77 for the 10 most fractionalized countries, appeal to reason is that the least fractionalized societies will grow faster than more fractionalized ones. But this is a very tenuous conclusion from small sample data sets. Expanding the data sample size through data pooling would perhaps lead to more reliable results. A study by Kekic (2007) “examined the state of democracy in 167 countries and rated those with an Economist Intelligence Unit Index of Democracy which focused on five general categories: free and fair election process, civil liberties, functioning of government, political participation, and political culture. Sweden scored a total of 9.88 on a scale of 0 to 10, which was the highest result; North Korea scored the lowest with 1.03. ‘Full democracies,’ ‘Flawed Democracies,’ and ‘Hybrid Regimes’ are considered to be democracies and ‘Authoritarian Regimes’ are considered to be dictatorial.”

One aim of this paper is to examine how ethnic diversity leads to policies that reduce economic efficiency. For instance, when ownership of a country’s resources is a source of income and rent, ethnic groups will compete for control of the government—this could have two adverse effects: (1) the desire of the incumbent party (ethnic group) to retain the levers of control of the government, and (2) a disinclination of the people in power to increase efficiency through the privatization of national enterprises. Either outcome—retention of power or nationalization of resources—could result in undermining efficiency and in reducing economic performance. According to Collier (2003, p. 150), “Diverse societies are alleged to find it more difficult to reach cooperative solutions and to be more likely to waste resources in distributional struggles.” The question: Is democracy the solution to the adverse effects of ethnic diversity on growth? According to Economic Intelligence Unit (EIU), the correlation coefficient between these two variables is 0.6. When the democracy index is regressed on income per capita the coefficient of the explanatory variable accounts for 40% of the inter-country variation in democracy. In addition, the direction of causality is not clear. One position on this is that development leads to democracy (Kekic, 2007). Further, the possession and reliance of a country on its natural resources can have a negative effect of development on democracy (Kekic, 2007) because the source of income is from extraction not production. The people in control of resources must employ violence, patronage, and corruption to maintain this control. The result is the maintenance of an undemocratic political system in which decisions are for the narrow interests of a few, property rights are nonexistent, and social promotion depends on political associations.
HYPOTHESES

The present paper empirically tests the hypothetical relationship between four variables—ethnicity, efficiency, democracy, and economic growth. Ethnicity is expected to influence economic growth negatively, but democracy and efficiency are expected to have positive effects on economic growth. The channels through which these variables exert their influence on economic growth are not the same. For instance, ethnicity can impact efficiency through corruption and conflict. If corruption is analyzed in terms of costs and benefits two things surface. Corruption’s ambiguous effect on economic performance becomes clearer. But it also suggests an equilibrium quantity of corruption, implying that it might be uneconomic to completely get rid of corruption. Klitgaard (1998) discusses three ways corruption might be beneficial. One is the economic view that corruption introduces a kind of market mechanism by allocating goods where willingness and ability to pay are greater. From a political point of view corruption in “payments, appointments, and policies may have political benefits. Politicians may use corruption to foster the political integration of various tribes, regions, elites, or parties, which may in turn lead to political harmony in the face of fragmented political authority, disunity, and hostility” (Klitgaard, 1998). Finally, in terms of management, “if bureaucratic rules are constraining, the organization may sometimes benefit by the employees’ corrupt circumvention of the rules. A limited amount of employee theft, embezzlement, misreporting of expenses, kickbacks, ‘speed money,’ and so forth may be tacitly allowed by top management, on the one hand, because controlling these illicit activities would be prohibitively expensive, and on the other hand, because such illicit sources of income may in the long run substitute for higher wages” (Klitgaard, 1998). Finally, democracy influences efficiency through the rule of law, the protection of private property, and through a process of making the rulers accountable to the ruled.

The three hypotheses to be tested are: (1) the connection between growth and ethnicity; (2) efficiency and ethnicity; and (3) democracy and efficiency. The conclusion of Easterly and Levine (1997) of an inverse relationship between ethnicity, democracy and economic growth leads us to the first hypothesis to be tested.

Hypothesis 1: A higher degree of ethnic diversity is associated with a slower rate of economic growth in SSA.

Barro (1991) found that political instability has a negative effect on investment. Political instability includes coups, revolutions, strikes, assassinations, and wars (civil wars and international conflicts). Therefore, political instability can discourage both foreign and domestic investment. Foreign factors can limit the inflow of funds. For instance, foreigners will be averse to putting their funds at risk in an uncertain environment of political instability. Foreign direct investment might run the risk of expropriation by the state with (but more likely, without) proper compensation. Domestic factors come into play within the framework of political instability. Banks will be reluctant to make loans, and local entrepre-
neurs would find it difficult to secure seed money for new ventures whose profitability is more uncertain because of the existing political instability. There could also be a flight of funds to escape the possible devaluation of the currency. Posner (2004) shows that ethnic diversity and conflicts (i.e. political instability) are positively correlated. In either case, capital formation will suffer and with it economic performance, on the assumption that capital is necessary for raising the productivity of labor and raising living standards. According to Collier (2003) diverse societies find it more difficult to cooperate and tend to waste resources in distributional struggles. Here the tensions among the ethnic groups affect the distribution of income. The countries’ productive resources are both a source of production of goods and services, and also a source of rent. Thus, ethnic diversity causes political instability. This leads to the second hypothesis to be resolved.

**Hypothesis 2:** A higher degree of ethnic diversity is associated with greater inefficiency (lower efficiency) in the use of inputs.

Scully (1988), Quinn (2003, p. 253), and others have shown that democracies have performed better than authoritarian regimes. Democracies have three virtues (justice, liberty, and peace)—see Reiter and Stam (2002)—that enable them to be effective in winning conflicts (wars) and to generate higher incomes. An implicit part of the three virtues, particularly embedded in justice is the notion of the rule of law, enabled by ideas of individual freedom, the definition and protection of property rights, and the presence of free and fair elections. Another important characteristic of democracies is they do not fight among themselves (perhaps more accurately, “they do not often fight among themselves”). Democracies enter into conflicts with the approval of citizens, and the conflicts are winnable going in and are expected to be of short duration. There is a political cost to long wars for the incumbent government. One other aspect of democracies is that soldiers have a vested interest in the outcome of the conflict—protection of liberty and equality that can be compromised with the victory of a non-democratic state. The implication of this is that democracies use resources more efficiently. Thus, the third hypothesis can be stated as follows:

**Hypothesis 3:** Democratic governments (based on free and fair elections) that enforce individual freedom and property rights—countries where more individual freedom and property rights are enforced—use inputs more efficiently.

Hypothesis 3 implies that democratic forms of government and economic growth are positively correlated. One rationale for this view is the Modernization Theory which states that “... countries with higher per capita gross domestic product (GDP), higher levels of education, more urban dwellers, more unions, and other indices of more modern society also had more demands from society for political inclusion” (Quinn, 2003, p. 234). Also, because democracies are more
prosperous, they can generate the financial resources necessary for the successful prosecution of conflicts. One would expect to find a positive correlation between indexes of democracy and economic performance. Countries with more democratic attributes can be expected to perform better than countries with fewer democratic qualities. It is worthwhile noting that democracies are essentially political entities and do not produce goods. The corresponding economic paradigm which goes hand in hand with democracies is the free market—that is, the idea and promotion of competition, freedom of choice, and freedom of enterprise.

**METHODOLOGY**

Assume the economy can be represented by a simple neoclassical production function that is continuous, twice differentiable, and homogeneous of degree one in the inputs. Such a production function can be written as a Cobb–Douglas equation,

\[ Y = f(K, L) = K^\alpha L^\beta \]

(1)

where \( Y \) is real output, \( K \) is the capital stock, and \( L \) is the labor supply. The exponents of capital \( \alpha \) and labor \( \beta \) can be thought of as capital and labor shares of real output \( Y \) or with respect to the growth in output the elasticities of capital and labor, respectively.

Differentiating the logarithm Equation (1) yields the rate of growth of real output as the weighted average of the inputs.

\[ g_Y = \alpha g_K + \beta g_L \]

(2)

Equation (2) is the benchmark equation. More efficient countries will combine these inputs more efficiently to produce higher rates of economic growth and more efficient countries will be more ethnically homogeneous—there will be fewer conflicts over resources, and more democratic countries will also be more stable and, thus, be more able to convert inputs more efficiently into output. Thus, the effect of ethnic diversity on economic growth can be found from Equation (3).

\[ g_Y = \alpha g_K + \beta g_L + \delta ED \]

(3)

(where \( ED \) denotes ethnic diversity)

The coefficient of the ethnic diversity variable \( ED \), \( \delta \), is expected to be negative, implying that countries with higher incidence of ethnic diversity will have slower growth rates. Similarly, the effect of democracy on economic growth can be computed by entering \( DR \) in Equation (2), thus

\[ g_Y = \alpha g_K + \beta g_L + \theta DR \]

(4)

(where \( DR \) represents democratic regime)
The coefficient of the democracy variable \( \theta \) is expected to be positive, implying that countries with more democratic characteristics will have higher growth rates.

In order to examine the effects of ethnic diversity on efficiency, following Scully’s approach let technical efficiency be \( Y \leq f(K,L) \). “Under the assumption that all measurement errors are negligible, the error term strictly captures technical efficiency differences and is computed from the vector of residuals” (Scully, 1988). Alternatively, one might think of efficiency as an upward shift of the production due to technical efficiency for given amounts of capital and labor. If democracy causes a rise in efficiency in the use of capital and labor, the production function will shift up. On the other hand, if ethnic diversity has a negative effect on efficiency, the production function would shift down. The democracies can be characterized as hybrid, flawed, and fully functioning. Fully functioning democracies are hypothesized to perform better than hybrid or flawed ones.

The form of the equation for the regression analysis includes labor and capital in addition to various measure of ethnic diversity and indexes of democracy. This equation can be written as

\[
g_Y = c + \alpha g_K + \beta g_L + \delta ED + \theta DR + \epsilon
\]

where \( c \) is a constant and \( \epsilon \) is the error term (mean zero and constant variance) and the usual assumption of Gauss Markov.

**DATA**

The data come from 2004 World Development Indicators (World Bank, 2004). Real GDP is reported in constant United States 1995, purchasing power parity dollars. The labor force is the economically active population. Capital is calculated using the perpetual inventory method—\( K_t = (1-\delta)K_{t-1} + I_t \), where depreciation \( \delta \) is 4% a year from O’Connell and Ndulu (2000). That is, the change in capital stock equals gross investment in the current year less the amount of depreciation of last period capital stock—\( K_t - K_{t-1} = I_t - \delta K_{t-1} \). The initial capital stock \( K_0 \) is computed using Hall and Jones’ (1999) methodology \( K_0 = (K/Y)_0 = (I/Y)/(g_i+\delta) \), (see Owyong, 2005) where \( g_i \) is the growth rate of gross investment (I) over a 20-year period (from 1982–2002). A more straightforward approach is to accept that in the initial year, gross investment equals the “depreciation” of capital—that is, \( (n+d)K_0 = I_g \), where \( n \) is the population growth rate and \( d \) denote the depreciation rate. It follows immediately from this that the initial capital stock is equal to gross investment divided by depreciation—thus, \( K_0 = I_g/(n+d) \).

For real GDP, labor force, and investment, the data cover the period from 1982 to 2002, except for three countries—the Democratic Republic of the Congo, Niger, and Nigeria. The data for the Congo cover the years 1982 to 1991, the data for
Niger cover the years 1982 to 1999, and the data for Nigeria cover the years 1982 to 1998.

This study (the relation of growth to ethnic diversity, the effect of ethnicity on technical efficiency, and the effect of democracy on technical efficiency) makes use of the availability of data on ethnicity (that focuses on political relevant groups), data on real economic growth (from World Bank, 2004), and data that measure the efficiency of input use (computed from World Bank data). Measures of democracy are taken from the 2004 Economic Intelligence Unit’s Democracy Index data.

**EMPIRICAL EVIDENCE**

When using time series data, it is necessary to determine if the variable is non-stationary. Nonstationarity has important implications for economic theory and modeling. The augmented Dickey–Fuller test was used to check for nonstationary of all the variables—real GDP, labor, and capital. For the test for stationarity, the null hypothesis is that the series has a unit root (i.e., the series is nonstationary); the alternate hypothesis is the series is stationary. The results of the augmented Dickey–Fuller (ADF) test are reported in Table 2. The critical value for each variable exceeds the results of the ADF test and show that the logarithm levels of the variables are stationary. The results of the ADF test are reported in Table 2, where the names of variables appear in column 1, the ADF test in column two, and the critical value at the 10 percent level of significance is in column 3. The critical value for each variable exceeds the results of the ADF test and show that the logarithm levels of the variables are stationary.

<table>
<thead>
<tr>
<th>Variables</th>
<th>ADF Test</th>
<th>Asy.Critical Value 10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>LGDP t-test</td>
<td>-4.6081</td>
<td>-3.13</td>
</tr>
<tr>
<td>LLF t-test</td>
<td>-4.1243</td>
<td>-3.13</td>
</tr>
<tr>
<td>LD t-test</td>
<td>-5.3749</td>
<td>-3.13</td>
</tr>
</tbody>
</table>

Note: LGDP is the log of GDP, LLF is the log of the labor force, LK is the log of capital, and ADF = augmented Dickey–Fuller test.

In panel data estimation, consistent estimation of the structural coefficients rests on the stochastic properties of the error term, which may be correlated with the independent variables and might also be serially correlated. The error term has both a country-specific and a time-specific component. If the independent variables are not correlated, ordinary least squares can be used to obtain consistent estimates of the coefficients. Preliminary estimates of Equation (2) using
ordinary least squares indicated that the coefficients of the independent variables were not consistent—the Durbin–Watson statistic = 0.1262. After correcting for autocorrelation of the error term using a Cochrane–Orcutt iterative estimate, the estimate of the coefficients of the independent variables improved—the Durbin–Watson = 1.9098.

From Equation (5), the inclusion of PREG and ELF is a way to compare the effects of both measures of ethnic diversity (fragmentation) on real GDP growth. PREG should influence growth because it is designed to capture the inclusion of politically relevant ethnic groups on economic policymaking. By contrast, ELF is primarily a measure of ethnic diversity of a country and is about characteristics of the country on economic performance where the prevailing view is that ethnic diversity harms economic growth. The regression results in Table 3 were calculated by regressing the logarithm of real GDP on time and taking the coefficient of the time variable as the average annual growth rate (AGR) was regressed on the logarithm of the labor force (LLF), the logarithm of capital, the politically relevant ethnic group (PREG), ethno-linguistic fractionalization (ELF), flawed democracy (FD) and hybrid regimes (HR).

In the first column of Table 3, the numbers in the parentheses are t values calculated under the assumption that the true population values of each regression coefficient individually is zero, and the numbers below the parentheses are p values, which are the exact level of significance, the exact probability of committing Type I error. The p values were included because they represent the lowest significance level at which the null hypothesis can be rejected. For instance, under the null hypothesis that the coefficient of PREG is zero, the exact probability (p value) of getting a t value of 7.74 or greater is 0.000. Thus, if we reject to null hypothesis, the probability of Type I error is 0. Going forward, p values (rather than t values) are used to interpret the significance coefficients in Table 3. The first column of Table 3 shows that the adj. $R^2$ coefficient indicates that the model explains 84% of the variation in economic growth from 1982–2002 in SSA. The effect of labor on growth computed as an elasticity with a coefficient of 0.489 is significant at all levels of significance ($p = .000$). This means that a 1% increase in the labor force would have increased output during the period by 0.489%. The effect of capital on output is also positive—the coefficient is 0.375, which is significant at all levels of significance ($p = .000$). Thus, a 1% rise in the capital stock would have raised output by 0.375%. Both the labor and capital coefficients are less than one, (i.e., inelastic), and their sum 0.864 is also less than unity, implying decreasing returns to scale. Had the region doubled their labor force and capital stock during the period 1982–2002, output would not have doubled.

The effect of ethnic diversity of economic performance is captured by the variables PREG and ELF. The PREG variable measures competition among ethnic groups on policymaking decisions and is expected to have a positive effect on economic growth. In Table 3, column 1, the coefficient of the PREG variable
Table 3
Determinants of Economic Growth in Sub-Saharan Africa

<table>
<thead>
<tr>
<th>Dep. Var. = AGR</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>6.90</td>
<td>6.83</td>
</tr>
<tr>
<td></td>
<td>(22.45)</td>
<td>(22.70)</td>
</tr>
<tr>
<td></td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>LLF</td>
<td>0.489</td>
<td>0.472</td>
</tr>
<tr>
<td></td>
<td>(25.46)</td>
<td>(25.82)</td>
</tr>
<tr>
<td></td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>LK</td>
<td>0.374</td>
<td>0.369</td>
</tr>
<tr>
<td></td>
<td>(24.80)</td>
<td>(24.88)</td>
</tr>
<tr>
<td></td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>PREG</td>
<td>0.847</td>
<td>0.638</td>
</tr>
<tr>
<td></td>
<td>(7.72)</td>
<td>(6.09)</td>
</tr>
<tr>
<td></td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>ELF</td>
<td>-0.363</td>
<td>-0.178</td>
</tr>
<tr>
<td></td>
<td>(-3.80)</td>
<td>(-1.98)</td>
</tr>
<tr>
<td></td>
<td>.000</td>
<td>.048</td>
</tr>
<tr>
<td>FD</td>
<td>0.339</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.68)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>HR</td>
<td>-0.069</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(-1.58)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.114</td>
<td></td>
</tr>
<tr>
<td>DEMIND</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>.091</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(7.73)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>.838</td>
<td></td>
</tr>
<tr>
<td>$F$-values</td>
<td>576.66</td>
<td></td>
</tr>
<tr>
<td>$Df$</td>
<td>668</td>
<td></td>
</tr>
<tr>
<td>$N$</td>
<td>675</td>
<td></td>
</tr>
</tbody>
</table>

Note: AGR is the average growth rate of real GDP, LLF is the log of the labor force, LK is the log of capital, PREG is political relevant ethnic groups, ELF is etho-linguistic fragmentation, FD is flawed democracy, HR is hybrid regimes, and DEMIND is democracy index. Values for $t$ are in brackets and $p$ values are immediately below the $t$ values.
is positive and significant at all levels of significance \( p = .000 \). Thus, in SSA countries where ethnic groups participated in the policymaking decisions, the effect on growth was 0.845. PREG indicates that ethnic diversity did not reduce economic growth. This is a result implied by Collier (2003, p. 165)—that is “. . . in democracies, except in circumstances of dominance, ethnic diversity does not significantly adversely affect economic performance.” When ethnic diversity is the presence of ethnic competition, it might improve performance (Collier, 2003, p. 165). The ethno-linguistic fragmentation variable, ELF, is a measure of ethnic diversity with an implied rivalry for control of the government and the county’s resources as a source of rent. The coefficient of ELF is -0.363 and it is significant at all levels of significance \( p = .000 \). During the period 1982–2002, the ELF coefficient indicates that countries in the SSA region of the study reduced growth by 0.363.

The first column of Table 3 also reports results for variables FD and HR, which are two measures of democracy. FD is a measure of flawed democracy and HR is a measure of hybrid regimes. The coefficient for FD is 0.339 and significant at all levels of significance \( p = .000 \). It includes countries such as Lesotho, Namibia, and South Africa, which although flawed democracies (rather than authoritarian regimes) still had a positive effect on growth. The HR coefficient is -0.069 and is not significantly different from zero \( p = .11 \) at the 10% level of significance. Thus, we conclude that hybrid regimes had no effect on growth. HR includes countries such as Mozambique, Madagascar, and Senegal.

The second column of Table 3 reports regression results for the effects of labor, capital, ethnic diversity, and an index of democracy, which had to be separated from the regression computation in Equation (3) because democracy index (DEMIND), FD and HR are highly correlated. Eighty-four percent (adj. \( R^2 = .84 \)) of the variation in growth is explained in the second column of Table 3. The coefficients of the labor force LLF variable and the capital stock LK variable are positive and significant at all levels of significance \( p = .000 \). The sum of these two coefficients is 0.83—less than unity. The PREG coefficient is positive and significant at all levels of significance (explanation as before). The ELF coefficient bears a negative sign and it is significant at the 5% level of significance \( p = .048 \).

DEMIND is a democracy index variable which is based on a scale from 0 to 10. It contains FD, HR, and fully-functioning democracies. A higher value of DEMIND measures more democracy indices for the country. For instance, Central African Republic has an index of 1.61 and is reported as an authoritarian regime. But Mauritius has an index of 8.4 and is reported as a fully functioning democracy. Hybrid regimes HR have index values of between 4 and 6, flawed democracies between 6 and 8, and fully functioning democracies have index numbers between 8 and 10. Table 3, column 2 reports a DEMIND of 0.091 which is statistically significant at all levels of significance \( p = .000 \). DEMIND suggests that countries that were more democratic had higher real economic growth performance.
TECHNICAL EFFICIENCY

In the first column of Table 4, technical efficiency, is reported as the residual of the regression of Equation 1. That is, per capita output regressed on the capital-labor ratio in a Cobb–Douglass production function. The model has a poor fit of the data—the adj. $R^2$ is .08. The coefficients of the independent variables PREG and ELF are positive and negative, respectively, and statistically significant at all levels of significance ($p = .000$). These results are counterintuitive. The presence of PREG increases inefficiency, while the presence of ELF reduces it. For the democracy measures, FD is positive and statistically significant at all levels of significance. It implies that flawed democracies were technically more inefficient. The coefficient of HR is negative and significant at the 5% level of significance ($p = .013$). Since HR is neither fish nor fowl, it is possible that the presence of hybrid regimes would have increased efficiency.

<table>
<thead>
<tr>
<th>Dep. Var. = AGR</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E</td>
<td>LKLF</td>
</tr>
<tr>
<td>Constant</td>
<td>0.082</td>
<td>8.31</td>
</tr>
<tr>
<td></td>
<td>(1.54)</td>
<td>(59.49)</td>
</tr>
<tr>
<td>PREG</td>
<td>0.694</td>
<td>-0.461</td>
</tr>
<tr>
<td></td>
<td>(6.21)</td>
<td>(-1.58)</td>
</tr>
<tr>
<td>ELF</td>
<td>-0.545</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>(-5.66)</td>
<td>(.11)</td>
</tr>
<tr>
<td>FD</td>
<td>0.273</td>
<td>1.040</td>
</tr>
<tr>
<td></td>
<td>(4.56)</td>
<td>(6.67)</td>
</tr>
<tr>
<td>HR</td>
<td>-.111</td>
<td>-0.465</td>
</tr>
<tr>
<td></td>
<td>(-2.48)</td>
<td>(-3.99)</td>
</tr>
<tr>
<td>adj. $R^2$</td>
<td>.077</td>
<td>.126</td>
</tr>
<tr>
<td>$F$-values</td>
<td>15.2</td>
<td>25.49</td>
</tr>
<tr>
<td>$Df$</td>
<td>670</td>
<td>670</td>
</tr>
<tr>
<td>N</td>
<td>675</td>
<td>675</td>
</tr>
</tbody>
</table>

Note: The dependent variables are E residuals and LKLF (log to the capital-labor ratio); the independent variables are PREG (political relevant ethnic groups), ELF (ethno-linguistic fragmentation), FD is (flawed democracy) and HR (hybrid regimes). Values for $t$ are in brackets and $p$ values are immediately below the $t$ values.
Table 4, column 2 gives an alternative look at efficiency in terms of the capital-labor ratio. The argument here is ethnic diversity and democracy can affect the capital labor ratio (LKLF). Table 4 also shows that the model explains 13% of the variation in the capital-labor ratio. In our sample of SSA countries for the period 1982–2002, the effect of ethnic diversity on the capital-labor ratio was zero—the coefficient of PREG -0.4612 is not significantly different from zero at the 10% level of significance. Also, the coefficient of ELF is 0- .028 is not significantly different from zero ($p = .910$).

The democracy measures FD and HR have coefficients of 1.04 and -0.465, respectively. The coefficients are significantly different from zero at all levels of significance ($p = .000$). Thus, the presence of flawed democracies in the sample had a positive effect on capital formation and hybrid regimes had a negative effect on capital-labor ratio.

The two models used to measure the effects of ethnic diversity and democracy on efficiency offered counterintuitive and contradictory answers. PREG was hypothesized to have a positive effect on growth and it did, but its effect on efficiency was contradictory. ELF was hypothesized to have a negative effect on growth and it did, but its effect on technical efficiency was counterintuitive. Flawed democracies increased inefficiency as noted in the first column of Table 4, but increased the capital-labor ratio in the second column in Table 4. Hybrid regimes increased efficiency in Table 4, column 1 but reduced the capital-labor ratio in Table 4, column 2.

**SUMMARY AND CONCLUSION**

The paper sets out to address three hypotheses that are generally accepted as true—H1: ethnic diversity reduces real economic growth, H2: ethnic diversity impairs the efficiency of capital, and H3: democratic regimes use resources more efficiently and have faster real growth.

Hypothesis 1 posits that ethnic diversity retards real economic growth. The results of the regression analysis in Table 3, column 1 show that the effect of ethnic diversity on growth depends on the measure of ethnic diversity. The PREG variable is positive and significant which indicates that when people competed for participation in decision-making policies economic growth was positive. However, ELF implies that there was a struggle among ethnic groups for control of the country’s resources to earn rent. Thus, in the first column of Table 3, ELF had a negative effect on economic growth.

Hypothesis 2 does not appear to hold up. Our models for testing the effects of ethnic diversity and democracy on efficiency did not produce robust results. First, Table 4, column 1 reports results of the residuals from Equation 1. The coefficients of PREG and ELF are significant but have counterintuitive signs, implying that PREG increases inefficiency and ELF reduces it. Employing an alternate
measure of efficiency—capital-labor ratio—in Table 4, column 2, we find that the coefficients of PREG and ELF are not statistically different from zero, implying that ethnic diversity had no effect on the capital-labor ratio in our sample of SSA countries for the period of the study.

Hypothesis 3 does appear to hold up. Measures of democracy—FD and HR did have some effects on both measures of efficiency. For example, the first column of Table 4 reports that the presence of flawed democracies FD increased inefficiency and hybrid regimes HR reduced it. In the second column of Table 4, FD increased the capital-labor ratio while HR reduced it. Table 3, column 1 shows that flawed democracies had a positive effect on economic growth for the sample of SSA countries in the study. The presence of hybrid regimes in our sample had no effect on economic growth.

Recommendation: In addition to labor and capital inputs that are positively correlated with economic growth, the evidence from this study suggests that ethnic diversity and democracy (even when flawed) exert positive effects of economic growth. Thus, policies that encourage democratic regimes in sub-Saharan African countries might be seen as paths to economic growth.
REFERENCES


Hedge Fund Returns, Risk, and Fees and Systematic Equity Factors

Rosemary Walker
Robert A. Weigand
Washburn University

ABSTRACT
We investigate the extent to which hedge funds earn alpha—the component of returns that can be attributed to manager skill—in excess of a set of systematic equity factors. We find that a 4-factor performance attribution model explains between 15% and 45% of hedge fund returns, depending on hedge fund style. We also document a strong positive relation between alpha-generation and systematic risk, implying that funds must take on more systematic risk exposure to earn larger alphas. There is no clear-cut relation between hedge fund fees and alpha, however, as the average alpha earned by high-fee funds is approximately equal to the average alpha earned by low-fee funds. Moreover, no hedge fund style consistently earns sufficient alpha to cover average management fees. The dispersion of hedge fund alpha supports the idea that manager skill exists, but in small supply, implying that access to skilled managers is critical. After accounting for other database biases identified by previous researchers, hedge funds’ net alpha generation is significantly negative. We find little evidence that, as an overall industry, hedge funds create value for investors.

Keywords: Hedge funds, management fees, alpha, systematic risk
INRODUCTION
We investigate the extent to which hedge funds earn alpha, and evaluate the evolution of ideas and methods used to separate alpha from the beta, or systematic, component of returns. Numerous authors have made the case that one of the major factors contributing to the recent financial crisis was too much complexity—particularly in security design and risk modeling (e.g., Schwartz, 2009). We assert that, like many aspects of the 21st century investing environment, estimation of hedge fund alpha may have become so complicated that the process now adds little value for trustees and board members of smaller endowment and pension funds who need a framework that allows them to understand the extent to which their hedge fund allocation may have generated excess returns, and how much they should pay for those returns. We propose a simpler, more direct method for estimating alpha, and use that method to investigate the magnitude of hedge fund alpha from 1980 to 2008.

The alternative asset/hedge fund space suffers from a gradually increasing perception of over-promising and under-delivering, similar to the reputation of long-only mutual funds (as documented by Elton, Gruber, Das, & Hlavka, 1993; Carhart, 1997; Bogle, 1998; and, more recently, Barras, Scaillet, & Wermers, 2010). Hedge fund returns are “... lower than commonly supposed” according to Malkiel and Saha (2005, p. 80), who also find that hedge funds are significantly riskier than more conventional investments. Fung, Xu, and Yau (2004) also report negative average alphas for hedge funds, and Pojarliev and Levich (2008) find negative mean risk-adjusted alphas among a sample of currency managers. Wallerstein, Tuchschmid, and Zaker (2010) conclude that many hedge fund replication products deliver performance that is competitive with hedge funds at significantly lower fee levels.

These results are at odds with the way hedge funds have positioned themselves as absolute return vehicles that can largely immunize their performance from the effect of systematic risk factors and deliver pure alpha (vs. relative return vehicles, like long-only mutual funds, which earn most of their returns from exposure to systematic factors). Numerous academic studies focus on how effective hedge funds are at separating alpha from beta. The majority of these find that a significant proportion of hedge fund returns can be explained by market-related factors, including work by Fung and Hsieh (1997, 2001, 2002, 2004a, 2004b), Agarwal and Naik (2004), Hasanhodzic and Lo (2007), and Wallerstein et al. (2010). Of course, these results contradict hedge funds’ claims regarding their ability to hedge beta and mainly earn alpha, and beg the question of whether effective alpha-beta separation is feasible. For example, Mladina and Coyle (2010) find that even the acclaimed Yale Endowment owes much of its historical performance to “heavy exposure to common equity risk factors” (p. 43). Jarrow (2010) goes a step further and questions whether alpha even exists, concluding that “... the existence and persistence of positive alphas is more a fantasy than a fact” (p. 18).
Research also indicates that hedge fund beta is more complex than equity beta because some of the strategies hedge funds employ induce additional risks beyond those inherent in the assets in which they invest. Hedge funds often use derivatives, short-selling, and leverage to generate alpha during periods of extreme returns in financial markets, which can lead to exposure to higher-moment equity risks such as skewness and kurtosis (Fung & Hsieh, 2001; Weisman, 2002; Bondarenko, 2004; Diez & Garcia, 2006). Other studies have found that hedge funds’ complex strategies cause their returns to be non-linearly related to systematic factors (Fung & Hsieh, 1997, 2001, 2004b; Mitchell & Pulvino, 2001; Amin & Kat, 2003; Agarwal & Naik, 2004; Hasan hodzic & Lo, 2007; Fung, Hsieh, Naik, & Ramadorai, 2008; Agarwal, Bakshi, & Huij, 2009).

The complexity of hedge funds’ investment strategies is matched by researchers’ attempts at modeling hedge funds’ exposure to systematic factors. For example, Fung and Hsieh (2001, 2004b) emphasize the option-like traits of hedge fund performance and recommend the inclusion of lookback straddle returns as systematic factors in their model. Mitchell and Pulvino (2001) show that the returns from risk arbitrage resemble the payoff from selling uncovered index put options. Both studies find that hedge funds’ risk-return characteristics are indeed non-linear, and stress the importance of taking these option-like features into account when analyzing hedge fund returns.

We propose that complex theoretical and empirical frameworks such as these suffer from several weaknesses. First, because hedge funds’ exposure to higher-moment risks often results from positions taken during extreme market conditions, effectively modeling exposure to these risks requires knowledge regarding hedge funds’ strategies and the timing of these strategies—information that few hedge funds report. Second, because this information would be hedge fund-specific, nonlinear higher-moment risk exposure cannot effectively be reduced to a general model that could be applied objectively to most hedge funds in a certain style category. Third, one of the key purposes of an alpha-beta performance attribution analysis is to report results to pension fund and endowment officers in order to demonstrate that performance fees are based on true alpha generation. It is unlikely that the average trustee or board member is going to understand the complex statistical techniques required to model nonlinear exposure to higher-moment sources of risk. Fourth, for many smaller pension funds and endowments, the exotic indexes used in the modeling process are not investable alternatives, because their allocation to hedge funds often represents the riskiest position in their portfolios. From the perspective of these investors, the beta exposure of their hedge fund allocation is more relevant when it can be compared to the beta exposure of their equity allocation.

In this study we take a step toward technological simplicity by modeling hedge fund returns’ relation with systematic equity factors only—a topic that has received little coverage in the academic literature, but is used more often in
industry. The benefits of such an approach include: the results are more likely to be understood by pension trustees and board members, the technique can be applied objectively to hedge funds in all the major style categories, and it allows us to address a list of research questions likely to be of interest to decision-makers at smaller pension and endowment funds, whose boards typically include many members who do not analyze or monitor their funds’ investments full-time.

It should also be noted that modeling hedge fund returns as a linear function of a limited set of factors is not without problems. As mentioned previously, hedge funds often use derivative securities that induce nonlinearities in their risk/return profiles, so linear models may only approximate the true relation. Another problem involves the effect of missing variables, which can induce bias in the estimation of key model parameters. The typical approach in the academic literature, however, is to assume that employing more technology to solve these problems always improves the analytical process. We argue that the recent history of financial modeling offers a compelling counterexample to this view, and that overly complicated analyses often induce more costs than benefits. For example, consider the conclusions of Martinelli and Ziemann (2010) in their recent study of portfolio construction:

In trying to . . . make their risk evaluation more sophisticated, many asset managers increase the number of risk parameters to be estimated, which in turn leads to less robust and less relevant results than if they had stuck with a simple measure (p. 3).

The simpler, more direct method for estimating alpha and beta we propose (presented in the next section) allows for direct comparability with the risk and return of a fund’s equity allocation, requires a relatively small number of parameters to be estimated (which leaves less room for error), and will undoubtedly be more intuitive to trustees and board members at many smaller pension and endowment funds.

Previewing our major results, we find that a Fama-French (1993)-Carhart (1997) 4-factor performance attribution model explains between 15% and 45% of hedge fund returns, depending on hedge fund style, and that hedge fund alphas estimated from a 4-factor model are considerably smaller than those from a 1-factor model. The dispersion of hedge fund alpha supports the idea that manager skill exists, albeit in small supply, which implies that access to skilled managers should be an important criterion when making a hedge fund allocation decision. We find that 4-factor alphas increase with exposure to the Mkt – Rf factor, indicating that managers must take on more systematic risk to earn higher alphas. Managers’ ability to earn alpha initially increases with assets under management (AUM), but alpha begins to diminish past a certain threshold of AUM, consistent with the framework of Berk and Green (2004). This effect has mitigated over time, however—our results show that it now affects only a small subset of
hedge funds. Our most damning finding is that hedge funds do not earn enough alpha to cover management fees, and that fees are only weakly related to alphas—generation—the average alpha of high-fee managers is approximately equal to the average alpha of low-fee managers. After taking fees and database biases into account, hedge fund alphas are negative for all style categories, suggesting that, as an overall industry, hedge funds do not create value for investors.

DATA AND METHODOLOGY

Our primary source of data is the Hedge Fund Research database (HFR, 2009), which contains monthly returns, assets under management, fees, and other information on a universe of hedge funds from 1980 to 2008. Use of the HFR database allows us to investigate alphas and betas by hedge fund style. Much of the existing literature focuses on consolidated Fund-of-Funds data, so examining fund risk and return across different styles provides a unique perspective.

As noted by Fung and Hsieh (2006), hedge fund databases suffer from a variety of biases, including selection bias (induced by both data vendors and the manner in which hedge funds self-report their performance), survivorship bias (funds that close and get deleted from databases tend to have lower returns), backfill bias (managers wait to see if a new fund is successful before deciding to report results to data vendors), and liquidation bias (hedge fund managers usually stop reporting to a database before final liquidation of a fund). Fung and Hsieh (2006) find that survivorship and incubation bias have the most significant effect on reported returns, resulting in an average upward bias of 2.5% and 1.5%, respectively. We take these upward biases into account when interpreting our empirical results in the following section.

The Fama-French (1993) and Carhart (1997) monthly systematic return factors, the excess market return, and a time series of the risk-free rate are obtained from Ken French’s data library (French, 2009). The factors are constructed to capture the effects of market forces thought to systematically affect equity returns:

1. Mkt – Rf — the excess return earned by a broadly diversified portfolio of US equities over the risk-free rate;
2. SMB, or “Small Minus Big” — the excess return earned by small stocks over large stocks, often referred to as the size premium;
3. HML, or “High Minus Low” — the excess return earned by portfolios of stocks with high book-to-market ratios, often referred to as the value premium; and
4. UMD, or “Up Minus Down” — a factor constructed to capture the excess return earned by stocks with positive price momentum.
For all the months \((t)\) for which data are available, we estimate both a 1-factor market model regression for each fund \((i)\):

\[
R_{i,t} - r_{ft} = \alpha_{i,t}^{1F} + \beta_{i,1}^{1F} \left( R_{MKT,t} - r_{ft} \right) + \epsilon_{i,t} \tag{1}
\]

and a Fama-French/Carhart 4-factor market model regression (FF4 hereafter):

\[
R_{i,t} - r_{ft} = \alpha_{i,t}^{FF4} + \beta_{i,1}^{FF4} \left( R_{MKT,t} - r_{ft} \right) + \beta_{i,2}^{FF4} \left( R_{SMB,t} \right) + \beta_{i,3}^{FF4} \left( R_{HML,t} \right) + \beta_{i,4}^{FF4} \left( R_{UMD,t} \right) + \epsilon_{i,t}. \tag{2}
\]

The alpha for firm \(i\) is computed as the monthly return of hedge fund \(i\) minus the fund’s expected return based on a 1-factor regression:

\[
\hat{\alpha}_{i,t}^{1F} = \bar{R}_{i,t} - \left[ \bar{r}_{ft} + \hat{\beta}_{i,1}^{1F} \left( \bar{R}_{MKT} - r_{ft} \right) \right] \tag{3}
\]

and a FF4 regression:

\[
\hat{\alpha}_{i,t}^{FF4} = \bar{R}_{i,t} - \left[ \bar{r}_{ft} + \hat{\beta}_{i,1}^{FF4} \left( \bar{R}_{MKT} - r_{ft} \right) + \hat{\beta}_{i,2}^{FF4} \left( \bar{R}_{SMB} \right) + \hat{\beta}_{i,3}^{FF4} \left( \bar{R}_{HML} \right) + \hat{\beta}_{i,4}^{FF4} \left( \bar{R}_{UMD} \right) \right]. \tag{4}
\]

The results reported in the exhibits that follow are calculated as the average alpha, beta, \(t\)-statistic, and \(R^2\) from these regressions. Note that first estimating the parameters on the fund level and then averaging across all funds avoids a “diversification effect” that would bias our results toward finding higher beta coefficients and lower alphas.

**EMPIRICAL RESULTS**

**Descriptive Statistics**

Examination of the HFR database reveals that half of all hedge funds managed $175 million or less in the 1980s, which declined to $148 million or less in the 2000s, with mean NAV increasing significantly over the same period. The increase in mean NAV and the slight decrease in median NAV reflect the startup of many new, smaller funds and an increase in the size of the largest, longer-lived funds. Mean and median fund age increases only slightly over the decades; the data indicate that many hedge funds liquidate after three to four years. Only a small number of hedge funds continue operating over anything resembling the long run.

We also find that the mean management fee is 1.4\% per year, and the mean incentive fee is 15.0\% of profits per year. The mean and median management and incentive fees self-reported by hedge funds to the HFR database remain constant over time and are lower than the anecdotal “two-and-twenty” frequently cited by the financial press. This contradicts the popular notion that as the number of hedge funds increased (there were over 10,000 in 2007), the fees they charge would be competed downward. The magnitude of the fees charged by hedge funds will be used later in the paper to investigate the extent to which investors receive added value from their hedge fund allocations.
Hedge Fund Betas by Style

Table 1 shows the average betas of the hedge funds in our sample, reported by fund style. In this and most of the exhibits that follow, results are reported from both a 1-factor and 4-factor model. This exhibit addresses two questions: the degree to which the returns of the various hedge fund styles are related to systematic equity risk factors, and which factors affect each style. (It should also be pointed out that significant exposure to a systematic equity factor does not imply that a hedge fund style is invested in a specific type of stock, just that the strategy employed by the fund creates an exposure to the same market forces that affect portfolios of stocks with certain characteristics, such as size, value and momentum.)

Long/Short Equity Hedge funds generally maintain a positive market beta exposure (mean Mkt – Rf beta = +0.53), a large-cap emphasis (mean SMB beta = -0.76), a growth stock tilt (mean HML beta = -0.36), and long exposure to the price momentum factor (mean UMD beta = +0.11). Over one-third (36%) of the returns of the average Equity Hedge fund can be explained by systematic equity factors.

The positive mean market beta exposure of Event-Driven funds (Mkt – Rf = +0.40) is consistent with these funds’ strategies, as a rising stock market implies greater deal flow. Event-Driven funds have small but significant exposure to the small-cap factor (SMB beta = +0.10) as takeover targets tend to have smaller market capitalizations than the average stock. The positive HML factor (+0.12) suggests that these funds hold more positions in value stocks (takeover targets are more likely to be undervalued), and the UMD factor (+0.08) indicates the stocks in which these funds invest also tend to have positive price momentum.

Funds-of-Funds have positive market beta exposure (Mkt – Rf = +0.37), a slight tilt toward the large-cap factor (SMB = -0.02), no significant exposure to the value factor, and positive exposure to the price momentum factor (UMD = +0.14). The Macro style has the lowest market beta exposure (Mkt – Rf = +0.17), large-cap exposure to the size factor (SMB = -0.08), a value tilt (HML = +0.09), and positive price momentum exposure (UMD = +0.19). The Relative Value style has the second-lowest market beta coefficient (Mkt – Rf = +0.27) and no significant exposure to the size, value, or momentum factors.

The results reported in Table 1 confirm that the returns of all hedge fund styles are related to the basic equity risk factor Mkt – Rf, although the estimated coefficients on this factor (ranging from 0.107 to 0.525) are significantly lower than that of the average diversified equity portfolio (1.00). The 4-factor model explains 15%–45% of hedge fund returns, depending on style, and the explanatory power of the 4-factor model is several times larger than that of the 1-factor model for all styles. Not surprisingly, the most diversified style of hedge fund, the Funds-of-Funds (FOF), has the highest adjusted $R^2$. Almost half (45%) of the returns of
this hedge fund style can be explained by exposure to systematic equity factors. Macro style hedge funds have the lowest 4-factor $R^2$ (14.8%), followed by Relative Value at 23.8%. Based on the results presented in Table 1, we conclude that the 4-factor performance attribution model provides a better fit for the returns of most hedge fund styles, and that hedge fund returns are significantly related to most of the equity factors, with the exception of the Relative Value style, whose returns are, on average, only related to the basic excess return factor Mkt – Rf.

Hedge Fund Alphas by Style

Table 2 shows average annual hedge fund alphas, reported by style and sorted into quartiles (Q1–Q4 progresses from highest to lowest), with average $t$-statistics reported beneath each value. In this exhibit we investigate the extent to which hedge funds earn alpha over and above the common equity risk factors, and which hedge fund styles generate the largest alphas.
Our results indicate that alphas measured with the 1-factor (low $R^2$) model are positive and significant for all strategies for the first 3 quartiles (except FOFs), and negative and significant for the fourth quartile. Funds-of-Funds have the

<table>
<thead>
<tr>
<th>Equity Hedge</th>
<th>Event-Driven</th>
<th>Fund of Funds</th>
<th>Macro</th>
<th>Relative Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>0.43</td>
<td>0.51</td>
<td>0.05</td>
<td>0.61</td>
</tr>
<tr>
<td></td>
<td>18.17***</td>
<td>11.75***</td>
<td>4.07***</td>
<td>12.60***</td>
</tr>
<tr>
<td>1st Quartile</td>
<td>1.54</td>
<td>1.63</td>
<td>0.54</td>
<td>1.92</td>
</tr>
<tr>
<td></td>
<td>26.08***</td>
<td>16.51***</td>
<td>20.10***</td>
<td>17.65***</td>
</tr>
<tr>
<td>2nd Quartile</td>
<td>0.57</td>
<td>0.53</td>
<td>0.17</td>
<td>0.73</td>
</tr>
<tr>
<td></td>
<td>130.88***</td>
<td>63.74***</td>
<td>89.23***</td>
<td>82.92***</td>
</tr>
<tr>
<td>3rd Quartile</td>
<td>0.21</td>
<td>0.27</td>
<td>-0.01</td>
<td>0.35</td>
</tr>
<tr>
<td></td>
<td>48.23***</td>
<td>41.98***</td>
<td>-3.33***</td>
<td>46.91***</td>
</tr>
<tr>
<td>4th Quartile</td>
<td>-0.60</td>
<td>-0.37</td>
<td>-0.53</td>
<td>-0.54</td>
</tr>
<tr>
<td></td>
<td>-14.52***</td>
<td>-5.76***</td>
<td>-26.65***</td>
<td>-4.99***</td>
</tr>
<tr>
<td>% &gt; Mngt. Fee</td>
<td>8.6%</td>
<td>10.2%</td>
<td>0.4%</td>
<td>12.9%</td>
</tr>
</tbody>
</table>

---

Table 2
Average Annual Hedge Fund Alphas (in percent) by Hedge Fund Style

The exhibit reports mean hedge fund alphas by style, with mean $t$-statistics below each alpha. In the table below, “% > mngt. fee” indicates the percentage of funds that generate alpha greater than the average management fee of 1.4%.

---

1-Factor Model

<table>
<thead>
<tr>
<th>Equity Hedge</th>
<th>Event-Driven</th>
<th>Fund of Funds</th>
<th>Macro</th>
<th>Relative Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>0.29</td>
<td>0.27</td>
<td>-0.09</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>4.09***</td>
<td>2.46**</td>
<td>-6.64***</td>
<td>6.27***</td>
</tr>
<tr>
<td>1st Quartile</td>
<td>0.87</td>
<td>0.67</td>
<td>0.14</td>
<td>1.14</td>
</tr>
<tr>
<td></td>
<td>3.33***</td>
<td>8.21***</td>
<td>3.75***</td>
<td>4.58***</td>
</tr>
<tr>
<td>2nd Quartile</td>
<td>0.30</td>
<td>0.47</td>
<td>0.02</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>12.98***</td>
<td>6.29***</td>
<td>1.92*</td>
<td>10.24***</td>
</tr>
<tr>
<td>3rd Quartile</td>
<td>0.24</td>
<td>-0.15</td>
<td>-0.08</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>9.51***</td>
<td>-0.39</td>
<td>-7.96***</td>
<td>11.24***</td>
</tr>
<tr>
<td>4th Quartile</td>
<td>-0.23</td>
<td>0.08</td>
<td>-0.45</td>
<td>-0.08</td>
</tr>
<tr>
<td></td>
<td>-2.14**</td>
<td>-0.50</td>
<td>-12.01***</td>
<td>-0.51</td>
</tr>
<tr>
<td>% &gt; Mngt. Fee</td>
<td>6.7%</td>
<td>9.0%</td>
<td>0.7%</td>
<td>10.6%</td>
</tr>
</tbody>
</table>

***, **, * indicates significance at the .01, .05, and .10 levels, respectively.
lowest alphas—this style’s overall mean alpha is negative when measured with the 4-factor model. We find that the 4-factor alphas are smaller than the 1-factor alphas in most cases, indicating that accounting for hedge funds’ exposure to all common equity risk factors results in lower reported alpha, and by implication, lower incentive fees. With the lowest exposure to the equity factors ($R^2 = 14.9\%$) and highest 4-factor alpha (0.48%), the Macro Strategy has the best performance record over time.

The results reported in Table 2 are not adjusted for management fees, however—which averaged 1.4%—or the database biases reported by Fung and Hsieh (2006)—which averaged 4.0%. Adjusting the Table 2 results for average management fees puts all hedge fund categories underwater. No hedge fund style generates sufficient alpha over and above the effects of the FF4 common equity factors to cover average fees. Subtracting another 4.0% for the effects of survivorship and incubation bias allows us to conclude that, as an overall industry, hedge funds do not add value proportional to the fees they charge investors.

Of course, it is not the case that all hedge funds fail to generate alpha consistently greater than average management fees. The last rows in Table 2 show the percentage of each hedge fund style with alpha greater than the average fee of 1.4%. The results are enlightening. Between 0.7% (Funds of Funds) and 10.6% (Global Macro) of hedge fund styles have consistently generated 4-factor alphas greater than average management fees. This finding is consistent with the idea that manager skill exists in the hedge fund space, albeit in surprisingly small supply. Over 90% of hedge funds fail to earn alpha greater than their annual management fee, indicating that access to the elite subset of skilled managers should be a critical consideration—probably the most important consideration—when making a hedge fund allocation decision. Investors who do not have a strong conviction that they are getting access to the most skilled managers would do better by avoiding allocating to hedge funds altogether. Moreover, if the Fung and Hsieh (2006) database biases apply equally to the reported returns of this subset of high-performing hedge funds, then our conclusion must be further amended, as we find no hedge funds in any style category that consistently generate alphas greater than management fees after adjusting returns downward by an additional 4.0%.

Hedge Fund Alphas by Management Fee
Table 3 shows average hedge fund alphas reported by quintile of management fee. If management skill exists and markets are reasonably efficient at disseminating news regarding the distribution of skill, we would expect to see alphas increasing with fees. Alphas measured via the 1-factor model are largest in the highest and lowest fee categories, however. The 4-factor alphas are also generally larger in the highest and lowest fee ranges, a result that is difficult to reconcile with the idea that more skilled managers command higher fees. (These results are also consistent with the hypothesis that management skill exists, but a sufficient
number of investors do not have information regarding the distribution of skill. Once again, we find that the 4-factor alphas are significantly smaller than the 1-factor alphas, and that Funds-of-Funds consistently deliver the lowest alphas. The high-fee Relative Value style hedge funds are the only category of fund that has delivered alphas in excess of average management fees. (Technical note: The Relative Value funds’ average alpha is larger when measured using the 4-factor

<table>
<thead>
<tr>
<th>Fee Range</th>
<th>1-Factor Model</th>
<th>4-Factor Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0 to 0.5%</td>
<td>0.78</td>
<td>0.73</td>
</tr>
<tr>
<td>3.56***</td>
<td>3.18***</td>
<td></td>
</tr>
<tr>
<td>0.5 to 1.0%</td>
<td>0.37</td>
<td>0.18</td>
</tr>
<tr>
<td>11.62***</td>
<td>5.32***</td>
<td></td>
</tr>
<tr>
<td>1.0 to 1.5%</td>
<td>0.39</td>
<td>0.23</td>
</tr>
<tr>
<td>17.16***</td>
<td>5.19***</td>
<td></td>
</tr>
<tr>
<td>1.5 to 2.0%</td>
<td>0.47</td>
<td>0.43</td>
</tr>
<tr>
<td>7.74***</td>
<td>1.88*</td>
<td></td>
</tr>
<tr>
<td>2.0 to 20%</td>
<td>1.13</td>
<td>0.69</td>
</tr>
<tr>
<td>2.26**</td>
<td>1.23</td>
<td></td>
</tr>
<tr>
<td>1.80*</td>
<td>2.45*</td>
<td></td>
</tr>
</tbody>
</table>

***, **, * indicates significance at the .01, .05, and .10 levels, respectively.
model because, on average, this style of hedge fund has no exposure to the HML, SMB or UMD factors.)

Hedge Fund Alphas by Market Beta and Time Period
Table 4 shows average alphas for each hedge fund style sorted by Mkt – Rf beta. This exhibit provides insight into the question of whether managers have to take on greater exposure to systematic risk to generate higher alphas. The 1-factor alphas show a slight tendency to increase with the Mkt – Rf beta throughout the quartiles, but the 4-factor alphas increase monotonically for each hedge fund style (except Q3 to Q4 for Equity Hedge). These findings provide strong evidence that the average hedge fund must take on increased exposure to common equity

Table 4
Average Annual Hedge Fund Alphas (in percent)
Sorted by Mkt – Rf Beta

The exhibit reports mean hedge fund alphas by Mkt – Rf Beta, with mean t-statistics below each alpha.

1-Factor Model

<table>
<thead>
<tr>
<th>Equity Hedge</th>
<th>Event-Driven</th>
<th>Fund of Funds</th>
<th>Macro</th>
<th>Relative Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Quartile</td>
<td>0.52</td>
<td>0.46</td>
<td>-0.06</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>10.73***</td>
<td>4.16***</td>
<td>-2.06**</td>
<td>3.12***</td>
</tr>
<tr>
<td>2nd Quartile</td>
<td>0.41</td>
<td>0.46</td>
<td>-0.01</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>13.45***</td>
<td>8.41***</td>
<td>-0.70</td>
<td>10.14***</td>
</tr>
<tr>
<td>3rd Quartile</td>
<td>0.45</td>
<td>0.52</td>
<td>0.09</td>
<td>0.58</td>
</tr>
<tr>
<td></td>
<td>15.68***</td>
<td>6.17***</td>
<td>5.99***</td>
<td>14.26***</td>
</tr>
<tr>
<td>4th Quartile</td>
<td>0.38</td>
<td>0.64</td>
<td>0.17</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>7.26***</td>
<td>7.21***</td>
<td>6.30***</td>
<td>9.21***</td>
</tr>
</tbody>
</table>

4-Factor Model

<table>
<thead>
<tr>
<th>Equity Hedge</th>
<th>Event-Driven</th>
<th>Fund of Funds</th>
<th>Macro</th>
<th>Relative Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Quartile</td>
<td>0.17</td>
<td>0.16</td>
<td>-0.24</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>2.24**</td>
<td>0.99</td>
<td>-7.49***</td>
<td>0.78</td>
</tr>
<tr>
<td>2nd Quartile</td>
<td>0.20</td>
<td>0.30</td>
<td>-0.17</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>6.38***</td>
<td>5.48***</td>
<td>-8.81***</td>
<td>5.76***</td>
</tr>
<tr>
<td>3rd Quartile</td>
<td>0.32</td>
<td>0.42</td>
<td>-0.04</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>9.68***</td>
<td>4.64***</td>
<td>-2.53**</td>
<td>10.22***</td>
</tr>
<tr>
<td>4th Quartile</td>
<td>0.30</td>
<td>0.59</td>
<td>0.07</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td>4.17***</td>
<td>6.60***</td>
<td>1.79*</td>
<td>6.77***</td>
</tr>
</tbody>
</table>

***, **, * indicates significance at the .01, .05, and .10 levels, respectively.
risk factors to earn larger alphas, which runs contrary to the view that hedge funds are effective at immunizing their performance from systematic risk effects. As above, Funds of Hedge Funds consistently deliver the lowest alphas.

Table 5 shows alphas sorted by the 1980s, 1990s, and 2000–2008 period. Four-factor alphas for all hedge fund styles decrease from the 1990s to the 2000s—the decade in which the hedge fund concept attracted the most AUM. Fund-of-Funds’ average 4-factor alpha in the 2000s is negative even before management fees, and this style’s 1-factor alpha is not significantly different from zero. No hedge fund style generated average alpha larger than average management fees in any decade.

### Multivariate Analysis of the Determinants of Alpha

Table 6 shows the result of a multivariate regression analysis that models the determinants of hedge fund 4-factor alphas:
\[ \alpha_{FF4} = \gamma_0 + \gamma_1 AUM + \gamma_2 AUM^2 + \gamma_3 \sum_{i=1}^{4} d_i Type + \gamma_4 \sum_{j=1}^{3} d_j Decade + \gamma_5 MngtFee \quad (5) \]

- \( AUM \) and \( AUM^2 \) = the level of assets under management and assets under management squared;
- \( Type = 0,1 \) indicator variables for type of hedge fund, with FOFs omitted;
- \( Decade = \) indicator variables for the 1980s, 1990s, and 2000s; and
- \( MngtFee = \) the percentage annual management fee.

All hedge fund style types have positive and significant regression coefficients, consistent with the results reported in Table 2, which shows that all styles earn small positive alphas on average (except FOFs) before accounting for management fees and database biases. The negative coefficient on the \( Decade \) indicator variable indicates that alphas have been declining over time (although, as described above, fees have remained constant). The small positive coefficient on the \( MngtFee \) variable suggests that, on average, funds that charge higher fees earn slightly larger alphas.

The variables \( AUM \) and \( AUM^2 \) are both significant, indicating that hedge fund alphas initially increase as a fund attracts more AUM (positive coefficient), but alpha-generation decreases past a certain threshold of AUM (negative coefficient on \( AUM^2 \)). This finding is consistent with the diminishing returns to scale argument made by Berk and Green (2004), whereby most funds find it difficult to scale up their unique strategies across a larger capital investment. Re-estimating Equation (5) by decade and omitting the \( Decade \) indicator variable allows us to calculate the point of inflection at which more AUM suppresses alpha. The point of inflection averaged $1.03 billion in the 1980s, $5.5 billion in the 1990s, and $44.0 billion from 2000–2008. This allows us to estimate that 19.8% of funds earned lower alpha due to increases in AUM in the 1980s, but this problem only affected 1.5% of funds in the 1990s, and only 0.1% of funds in the 2000s. These findings suggest that, in practical terms, investors should not be highly concerned about allocating funds to a hedge fund that has large AUM, as the diminishing returns to scale effect has mitigated over time.

**CONCLUSIONS**

We investigated the extent to which hedge funds earn alpha—the component of returns that can be attributed to manager skill—in excess of a set of systematic equity factors. Hedge funds have incentive to maximize alpha because it determines the size of the performance fees they can charge. We find that a Fama-French (1993)–Carhart (1997) 4-factor performance attribution model explains between 15%–45% of hedge fund returns, depending on hedge fund style, and that hedge fund alphas estimated from a 4-factor model are considerably smaller than those from a 1-factor model. All hedge fund styles have significant exposure...
The dispersion of hedge fund alpha supports the idea that manager skill exists, albeit in small supply, and from the perspective of investors, access to skilled managers should be an important criterion when making a hedge fund allocation decision. Average alphas increase with exposure to the Mkt – Rf factor, indicating that managers must take on more systematic equity risk to earn higher alphas, a finding that contradicts hedge funds’ claim regarding their ability to hedge beta risk and earn pure alpha. Managers’ ability to earn alpha increases with assets under management (AUM), but alpha begins to diminish past a certain threshold of AUM, consistent with the framework of Berk and Green (2004), although we find that this effect has mitigated over time—the threshold at which AUM exerts a negative influence on alpha averaged over $40 billion in the 2000s, representing only 0.1% of all hedge funds.

Table 6
Multivariate Regression Analysis of Factors Affecting 4-Factor Alphas

The exhibit shows regression coefficients, standard errors, t-statistics and the adjusted $R^2$ from estimation of the following model:

$$\alpha_{FF4} = \gamma_0 + \gamma_1 AUM + \gamma_2 AUM^2 + \gamma_3 \sum_{j=1}^4 d_j Type + \gamma_4 \sum_{j=1}^3 d_j Decade + \gamma_5 MngtFee$$

Variable Definitions:
- $AUM$ and $AUM^2$: the level of assets under management and assets under management squared;
- $Type = 0, 1$ indicator variables for type of hedge fund, with FOFs omitted;
- $Decade =$ indicator variables for the 1980s, 1990s, and 2000s; and
- $MngtFee =$ the percentage annual management fee.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Std. Coefficient</th>
<th>Error</th>
<th>t-statistic</th>
<th>Significance</th>
<th>Adj. $R^2$</th>
<th>F-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.800</td>
<td>0.510</td>
<td>15.55***</td>
<td>0.000</td>
<td>0.40%</td>
<td>177.79</td>
</tr>
<tr>
<td>AUM (bil)</td>
<td>0.510</td>
<td>0.007</td>
<td>7.65***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AUM 2</td>
<td>0.000</td>
<td>0.000</td>
<td>-6.24***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equity Hedge</td>
<td>0.305</td>
<td>0.015</td>
<td>20.39***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macro</td>
<td>0.469</td>
<td>0.021</td>
<td>22.03***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relative Value</td>
<td>0.185</td>
<td>0.021</td>
<td>8.95***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Driven</td>
<td>0.294</td>
<td>0.023</td>
<td>12.71***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80s, 90s, 00s</td>
<td>-0.311</td>
<td>0.017</td>
<td>-18.56***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mngt. Fee</td>
<td>0.104</td>
<td>0.011</td>
<td>9.45***</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

***, **, * indicates significance at the .01, .05, and .10 levels, respectively.

to the full range of systematic equity factors except for the Relative Value style, which is only related to the Mkt – Rf factor.
We further find that aggregate alphas are not sufficiently large to cover management fees, and that fees are only weakly related to alpha-generation. The average alpha of high-fee managers is approximately equal to the average alpha of low-fee managers. We also find that average hedge fund alpha has declined moderately over time. Four-factor alphas for all hedge fund styles decrease from the 1990s to the 2000s—the decade in which the hedge fund concept attracted the most AUM. After taking fees and database biases into account, hedge fund alphas are negative for all style categories, suggesting that, as an overall industry, hedge funds do not create value for investors.
REFERENCES


A Comparison of Multiple Regression and Neural Networks for Forecasting Real Estate Values

Daniel Friesen  
*University of North Texas at Dallas*

Mike Patterson  
Bob Harmel  
*Midwestern State University*

**ABSTRACT**

We compare the accuracies of multiple regression and artificial neural networks in forecasting the sale prices of single-family residential properties using data from Wichita County, Texas. Typical property descriptors were used in the forecasts, along with property location as identified by the *neighborhood group code*. Thirty neighborhoods were randomly selected for analysis. We used NeuralTools and NeuroShell Predictor as our artificial neural network software tools.

Two analyses were used. First, a blocked ANOVA was performed on the mean absolute deviations between forecasted and actual sale prices. The factor of variation is the forecasting method: artificial neural networks and multiple regression. The blocking factor is neighborhood. The performance differences between forecasting methods were statistically significant.

Second, in analyzing each neighborhood, the sale prices were forecast using artificial neural networks and multiple regression. Forecast method accuracy was assessed using the absolute deviations between forecast and actual prices. Paired *t*-tests determined if there existed significant differences between forecasting method accuracies. NeuralTools was significantly more accurate in 24 of the 30 neighborhoods compared to multiple regression, which only provided a significantly more accurate forecast in 2 of the 30 neighborhoods. Likewise, NeuroShell Predictor proved to be significantly more accurate than multiple regression in the 30 neighborhoods by a 17 to 3 advantage, with 10 neighborhoods showing...
no significant difference in forecasting accuracy. In comparing the two artificial neural network tools, NeuralTools was more accurate than NeurShell by a count of 19 to 4, with 7 neighborhoods showing no significant difference.

ANNs are shown to be effective in forecasting prices of single-family residences.

**Key Words:** neural networks, real estate valuation, multiple regression comparison, forecasting, Texas

**INTRODUCTION**

In this paper, we compare the accuracy of traditional multiple regression with that of artificial neural networks (ANNs) in forecasting the sale prices of single-family residential properties. Use of multiple regression models has been ubiquitous in this field (Thibodeau, 2003; Chica-Olmo, 2007) for many years. Hair, Black, Babin, Anderson, and Tatham (2006) describe dependence techniques as those that predict dependent variable values using a set of independent variable values. They say that “regression analysis is by far the most widely used and versatile dependence technique, applicable in every facet of business decision making.” Worzala, Lenk, and Silva (1995) report that multiple regression often has “serious problems” when used for real estate appraisal, including multicollinearity, outliers, and nonlinearities in the data. In this paper, we compare the familiar multiple regression technique with the latest implementation of user-friendly ANNs.

White (1989) defines a neural network as a “type of input-output model.” That is, ANNs are also dependence techniques, as far as they are used in this paper. According to Kohonen (1990), not all neural networks are dependence models. ANNs have captured the imaginations of many researchers and academics. Gately (1996), a researcher in the financial disciplines, calls ANNs “the most important and powerful analytic tool to arrive in the past fifteen years.” Further evidence of a general captivation with ANN techniques includes:

- In separate reports, the *European Journal of Operational Research* (volume 93, issue 2 in 1996), the *Journal of Systems Engineering* (volume 6, issue 3 in 1996), and the *International Journal of Adaptive Control and Signal Processing* (volume 13, issue 4 in 1999) devoted entire issues to these tools. Examples of journals strictly devoted to neural networks include *IEEE Transactions on Neural Networks, Neural Networks and Neurocomputing*. ANN-specific discussions at scholarly conferences are well attended.

- ANNs have been shown to be both inherently flexible (White, 1989) and able to generalize (Burke & Ignizio, 1992).
ANNs are by nature essentially non-parametric (Willis, Di Massimo, Montague, Tham, & Morris, 1991; Smith & Dagli, 1991; Tam & Kiang, 1992), which is usually an advantage for modeling techniques.

There is always interest in legitimizing, debunking, or otherwise probing the limits of new approaches (Burke & Ignizio, 1992; Vaithyanathan & Ignizio, 1992).

Easy-to-use software is now readily available for the construction of ANN models. Many software packages are designed with easy-to-use graphical interfaces as opposed to requiring knowledge of a programming language (McAnany, 1993). Indeed, one of the ANN software tools used in this analysis functions as an Excel add-in. Cox (1990) refers to such availability and ease-of-use as “the ultimate seal of approval.”

The housing market is an important component of the national economy (Sklarz, Miller, & Gersch, 1987; Guirguis, Giannikos, & Anderson, 2005). Developers, investors, prospective homeowners, real estate appraisers, and mortgage lenders—indeed, all participants—are key stakeholders in the process of accurately predicting housing prices (Clapp & Giaccotto, 2002; Knight, Hill, & Sirmans, 1992; Worzala, et al., 1995). Pace, Barry, Gilley, and Sirmans (2000) note that:

[M]ost local tax assessments in the US rely upon statistical predictions of housing values. In addition, primary and secondary mortgage lenders have begun exploring the use of statistical forecasting of housing prices as opposed to employing appraisers.

Forecasting approaches in this field are discussed in the Review of Literature, as are examples of ANN approaches in other fields.

**REVIEW OF LITERATURE**

The literature review proceeds in three parts: (1) real estate value forecasting, (2) artificial neural networks (ANNs), and (3) recent uses of ANNs in the real estate field.

**Real Estate Value Forecasting**

In discussing the importance of real estate asset value forecasting, Krystalogianni, Matysiak, and Tsolacos (2004) note the needs of fund managers who make strategic investment decisions for models that incorporate or define medium- and long-term market signals. Hendershott and Weicher (2002) decompose housing demand into three components: short-run adjustments to equilibrium, medium-run demographic changes, and long-run shifts in economic policy and structure. They state:
Conversely, the demand for housing seems much more strongly related to real income than to any other factor, although rising ownership with age and the aging of the baby boomers clearly accounts for the U.S. homeownership rate’s current all-time high.

ANN approaches are capable of discerning such signals, even when there is no explicit definition of independent variables.

Chica-Olmo (2007) argues for inclusion of information about neighboring houses when modeling housing sale prices. He successfully incorporates this information into a model that uses “kriging” and “cokriging” to predict housing sale prices. Thibodeau (2003) comes to similar conclusions about neighborhoods based on arguments citing neighborhood similarities in public service levels, driving distances, and home structural attributes. The comparison of models described in this paper proceed in a neighborhood-by-neighborhood fashion.

Artificial Neural Networks

In this section, the following ANN topics are discussed: relevant background, terminology, functional issues that must be addressed, and dependence modeling uses.

**Relevant Background.** ANN study is known by many names including connectionism, neural computing, and parallel distributed processing. Connectionism encompasses several scientific disciplines including mathematics, cognitive psychology, biology, and statistics (Hart, 1992). Connectionism enjoyed a period of high interest during the late 1950s. By the 1970s many researchers felt that the inherent limitations of prevalent connectionist methods had become fatally constraining to further research. Findings during the early 1980s removed many limitations of, and rekindled interest, in connectionism. Allman (1989) relates the history and major contributors, as do many other authors. A list of seminal literature appears in *Byte* (“Neural Networks: Theory and Practice,” 1989).

ANNs have been applied in most of the disciplines that make up the management and decision sciences (Sharda, 1994). They have been studied and used as competitors to the following statistical methods: linear regression, monotonic regression, time series analysis, time series prediction, function mapping, and discriminant analysis (Burke & Ignizio, 1992; Sharda, 1994; Wang, 1994; Murtagh, Wurtz, & Zurich, 1990).

ANNs are reported to have many advantages over their competitors. A potential benefit is the abrogation of the necessity of specifying a mathematical model of the function in question. The “learn by example” nature—or “model-free” nature—could generate significant cost and time savings for modelers. Another potential advantage is the “universal approximator” benefit espoused by Hornik.
He describes the approximating ability of ANNs as generally high, both in terms of approximating the true function and in terms of approximating the true function’s derivatives.

**Terminology.** (Please note that key terms have been italicized.) An ANN is a set of connected processing units called *nodes* or *units*. Individual nodes generally perform simple computations; however, due to the large number of interconnections, the total computing power of the ANN is large. This network model (i.e., many simple units that are highly interconnected) resembles the gross structure of the human brain, thus the term *neural* (Venugopal & Narendran, 1992). Jones and Hoskins (1987) and Gately (1996) discuss in more detail the *brain metaphor* of ANN study.

The ANN term for dependence model is *feedforward network*. This type of ANN transforms independent variable values into estimates of dependent variable values (Venugopal & Narendran, 1992). Independent variable values are called inputs or input signals and they are introduced into the network at the *input nodes*. The network produces estimates, called *outputs* or output signals that are delivered at *output nodes*. Nodes that are neither input nor output are called *hidden nodes*. Hidden nodes do not interact with the environment (Tam & Kiang, 1992). The calculations that nodes perform are configured through a process called *learning* or *training*. Figure 1 shows a general feedforward ANN with m input nodes, n nodes per hidden layer, p hidden layers, and q output nodes.

Each hidden layer need not have the same number of nodes. Figure 2 shows a specific version of the Figure 1 ANN with the following configuration: 2 input nodes (m=2), 3 nodes per hidden layer (n=3), 1 hidden layer (p=1), and 1 output node (q=1).

The ANN in Figure 2 is said to be *fully connected*; that is, each node in a layer is connected to every node in the previous layer. Nodes in the same layer are not connected. ANNs can be partially connected, that is, not fully connected.

The connections between nodes are usually called *links* but are sometimes called *edges*. A *weight* is associated with each link. A node in either the hidden layer or the output layer generates an output signal in two steps. Aggregating the input signals is the first step. The most-often used approach is to sum the products of incoming signals and weights associated with the respective input links:

\[
\text{node input} = \sum_{\text{links}} \left[ \text{weight}_i \times \text{output from previous node}_i \right]
\]

Thus, each neuron input is “a weighted sum of the outputs from neurons in the previous layer that are connected” (Willis, et al., 1991). In the second step, the aggregated node input signal is transformed into the node output signal through a *transfer function*. The most widely used transfer functions are hyperbolic tangent.
and sigmoid (or logistic). The sigmoid transfer function is shown in equation (2) (Sieger & Badiru, 1993).

\[
node\ output = \frac{1}{1 + \exp(-\text{node input})}
\]  

(2)

Neurons or nodes in the input layer generally do not transform the input signal; however, input values are often mapped to a 0 to +1 scale prior to use by the network.
An example node is shown in Figure 3. If configured with a sigmoid transfer function, the node effects the transformation in equation (3):

\[ C = \frac{1}{1 + \exp(-(Aw+Bv+\text{bias}))} \]  

(3)

where \( A \) and \( B \) are output signs from upstream nodes; \( v \) and \( w \) are weights associated with links; and \( \text{bias} \) is a constant which is computed by the network. If the node in Figure 3 is a hidden node then \( C \) multiplied by the appropriate weight would become part of the input of another node. If the node in Figure 3 is an output node, then \( C \) is a network output.

The weights and biases are computed through the process known as learning or training. The learning process can be thought of in different ways. In one view, learning is the process of adjusting the weights of the network such that the network produces the best results that it can. The quality of results will be quantitatively evaluated using some error function. In another view of learning, the ANN model “encodes” the relationship between inputs and outputs in the weights (Yoon & Peterson, 1992).

In yet another view of learning, the hidden nodes are said to capture the important features of the data set. The number of hidden nodes in the ANN architecture will play a key role in the number of features encoded by the ANN (Touretzky & Pomerleau, 1989). This view offers an interesting contrast between ANNs and traditional statistical models such as multiple regression models. In
statistical models, both the number of, and the order of interactions between independent variables are specified directly. In ANNs, what is controlled is the number of features detected, which is configured by controlling the number of hidden nodes. A feature may encompass one or more interactions of varying order (McAnany, 1993).

Learning strategy is usually considered to fall into one of two categories: supervised or unsupervised (Venugopal & Narendran, 1992). In supervised learning—the only category considered here—ordered sets of input and output are presented to the network. The learning algorithm uses the actual output to adjust the network weights so that the differences between ANN-calculated outputs and actual outputs are reduced (Velasco & Rowe, 1993). The difference between an actual and a calculated output is called the error. How well the network has learned the data set is assessed by examining a function of the errors. The error function that is used most often is the average sum of squared errors. There are others that can be used, particularly when the output values are discrete (White, 1989). The learning or training process is iterative; thus, it continues until the error function is judged to be sufficiently small. Each iteration through the entire data set is called an epoch (Gately, 1996).

One popular supervised method by which ANNs learn is backpropagation (BP) (Sharda, 1994; Burke & Ignizio, 1992). BP is a gradient search over the error surface for the optimal weight set. Unfortunately, the error surface is not well behaved in general. BP can diverge or stop at local minima (Rumelhart, Hinton, & Williams, 1986; White, 1989; Burke & Ignizio, 1992; Wessels & Barnard, 1992; Tang & Koehler, 1994). White (1989) discusses the statistical inefficiency of BP. Fortunately, local minima are seldom significantly worse than the global minimum. Apparently, it is the networks with too few hidden nodes for the task that tend to get stuck at poor local minima (Rumelhart, et al., 1986).

**Functional Issues.** The following issues must be addressed for this study: (a) the appropriateness of ANNs as a modeling technique, (b) ANN architecture specifics—especially the number of hidden nodes and connections, and (c) ANN training levels.

In most situations, the selection of a modeling technique is open to questions of appropriateness. One aspect of the appropriateness question is generalization. Ideally, the ANN will generate a surface that not only fits the data set well but also fits any other data obtained from the same population. Such a surface is capable of generalization. Wang and Archer (1994) report high generalization abilities in ANNs. White (1989) describes a proof by Hornick (1991) showing that general feedforward networks configured with one hidden layer “can approximate any measurable function to any accuracy, given sufficiently many hidden units.” Using an ANN with a single hidden layer and single output neuron is shown to
be a statistically consistent procedure under some general training conditions (Mielniczuk & Tyrcha, 1993).

ANNs are reported to be robust in the face of noisy data (Yoon & Peterson, 1992). The reason for this advantage is that ANNs can learn the process without learning the noise in the process. Conventional statistical techniques do not train iteratively and may therefore be learning or encoding noise as a result.

One aspect of appropriateness is the black-box effect, that is, the problem of relating the estimated weights—or even the ANN architecture—to the physics of the process being modeled. According to Andersen, Cook, and Ramaswamy (1990), this task is “usually impossible.” They characterize this impossibility as the price of capturing complex interactions among process parameters, that is, the price of a generic, distribution- and assumption-free model. While the hidden nodes are helpful for identifying the important features of the process being modeled, the large number of interconnections causes difficulty in determining the contribution of input variable to output value (Yoon, Swales, & Margavio, 1993). The black-box nature may not be a problem if accuracy of predictions is the prime concern. Unfortunately, there are yet no means of determining either (a) the significance of individual independent variables, or (b) the physics of the process being modeled (Tam & Kiang, 1992). The black-box aspect of ANNs has been discussed by many researchers; for examples, see Touretzky and Pomerleau (1989).

Architecture refers to aspects such as the number of hidden layers, the number of hidden nodes per layer, the way the layers are connected, and the node transfer functions. There is little agreement regarding what constitutes an optimal network architecture. ANN researchers and users seem to rely on heuristic knowledge, empirical evidence, and trial and error (Fogel, 1991; Seitz & Stickel, 1993), or software tools that self-configure.

Given that an ANN with a single hidden layer and sigmoid transfer functions can map any function to an arbitrarily close precision given sufficiently many hidden nodes (White, 1989), it seems inappropriate to specify multiple hidden layers. Gately (1996) recommends that the single hidden layer architecture be the initial network attempted because such networks have “now become a benchmark for judging the results of other architectures.” Furthermore, more hidden layers mean more hidden nodes and, thus, more weights to estimate—a situation that can sharply increase training time (Fogel, 1991). Parsimony is a concern. The “minimal network strategy” postulates that the simplest ANN is the best one, all other things being equal (Weigend, Rumelhart, & Huberman, 1991). Other researchers support “the desirability of having a small network . . .” (Patuwo, Hu, & Hung, 1993).
Too many hidden nodes allow an ANN to memorize the training data set—both data and noise—which will lead to poor generalization (Burke & Ignizio, 1992). Further, large numbers of hidden nodes can result in long training times. Many researchers have noted the problems associated with too many hidden nodes (Touretzky & Pomerleau, 1989; Patuwo, et al., 1993), which is sometimes called the “over-fitting problem” (Weigend, et al., 1991). It seems that an optimum number of hidden nodes is required: not too few, but not too many. Some researchers have used an iterative method for determining an appropriate number of hidden nodes (Fujita, 1992; Gingrich, Kuespert, & McAvoy, 1992; McAnany, 1993). Indeed, one software tool used in this analysis forces a single hidden layer that it configures with the optimum number of hidden nodes.

As previously defined, the BP learning procedure is essentially a gradient search over the error space for the values of the weights that minimize the error function. Since basic approaches to BP are offered in many ANN-related publications and BP is in widespread use in ANN software, the mathematics of the algorithm are not presented here. Clear development and discussion can be found in Rumelhart, et al. (1986); Jones and Hoskins (1987); Kung, Diamantaras, Mao, and Taur (1991); and Stern (1996).

Training time is an issue when using feedforward ANNs. Training is an iterative process; thus, the time that the network spends training must be specified. Overly long training times are problematic in the senses that (a) training time often represents a resource, and (b) excessive training time can imply overtraining which can result in poor generalization. Larger networks will require longer training times.

Some researchers avoid overtraining by examining the rate of change of the error; that is, training is stopped when the decrease in successive errors becomes small (Yoon, et al., 1993). For most researchers, however, overtraining is avoided by using a holdout data set for monitoring the performance of the ANN periodically during training. Gately (1996) recommends testing after every 50 to 200 epochs. When the ANN’s performance on the holdout set starts to degrade, training is discontinued (Ball & Jurs, 1993). Hutton (1992) reports that ANN performance on the test set is the most important measure of performance.

**Dependence Modeling Uses.** Chen, Peng, and Abraham (2006) use a neural approach to forecast three “major international currency exchange rates.” Their approach utilizes multiple hidden layers while assigning different node functions to different parts of the network. Genetic programming is utilized to optimize network weights. The rationale for selecting the ANN approach is that the determinants of exchange rates are influenced by many “highly correlated economical, political and even psychological factors” that interact in “a very complex fashion.”
In another foreign exchange application, Leung and Chen (2005) report the performances of two ANN models for the purpose of predicting currency exchange correlation. They note that “the selection of proper architectural design may contribute directly to the success in neural network forecasting.”

Peters, Schmitz, and Cullmann (2006) used ANNs to model flood routing for the Freiberger Mulde River. In describing the problem, they note that “recent extreme flood events in central Europe, e.g. the flood of August 2002 . . . led to the increased demand for fast and robust prediction tools.” Their ANN was trained using data generated by a computationally intensive hydrodynamic simulation. Essentially, the ANN generated much more timely results than the simulation but was able to replicate the simulation’s accuracy.

Murdianto, Sutattyo, and Haris (2005) use single-hidden layer, feedforward ANNs to predict the characterizing properties of oil wells using seismic data. Gan, Limsombunchai, Clemes, and Weng (2005) compared two ANN approaches with a logistic model to predict consumer choice between electronic banking and non-electronic banking. Their results revealed that the ANN models were superior to the logistic approach. They also note the following previously mentioned limitations of ANN approaches:

Firstly, the neural networks lack theoretical background concerning the explanatory capabilities. The connection weights in the networks cannot be interpreted or used to identify the relationships between dependent and independent variables. Secondly, there are no formal techniques for non-linear methods to test the relative relevance of the independent variables and to carry out the variable selection process. Lastly, the neural networks [sic] learning process can be very time consuming.

Recent Uses of ANNs in the Real Estate Field

Some real estate-focused studies have suggested that neural networks might offer better—more accurate—forecasts than multiple regression. Taffese (2007) discusses an overview of two artificial intelligence systems and concludes that ANN-based analysis is superior to traditional multiple regression forecasting. Other studies have come to different conclusions. Worzala, et al., (1995) analyzed the sale of 288 homes in Fort Collins, Colorado and concluded that neural networks did not provide significantly more accurate forecasts than more traditional tools. One of the ANN software tools used by these authors, NeuroShell Predictor, is used in this paper as well, although their version of the software is much older. When the authors restricted the analysis to a more homogeneous group of houses (i.e., in one particular zip code), the ANNs performed on par with the multiple regression model.
DATABASE

Data for this study were drawn from the offices of the Wichita County Tax Appraisal Office. Wichita County is located approximately 135 miles northwest of Dallas, Texas, approximately 15 miles from the Texas–Oklahoma border (i.e., the Red River). The county seat and largest city in the county is Wichita Falls. According to the city’s website (Wichita Falls, n.d.), the estimated population of the city was 101,313 at the end of 2009. Other towns located in the county are Burkburnett, Iowa Park, Electra, Kamay, and Pleasant Valley (Wichita County, 2006). According to the U. S. Census Bureau (2007), the estimated population of Wichita County in 2005 was 125,894, down from 131,664 in 2000.

The database provided by the Wichita Appraisal Office included variables related to the sale of single-family residential homes for the calendar years 2003 and 2004. Variables used in the analysis and by appraisers included the sale price (independent variable), appraised value, square footage, age (in years), the presence of a pool, number of bedrooms, number of bathrooms, and two categorical variables dealing with the quality of construction (scale from 1 to 10) and condition of house (scale from 0 to 4). The last two variables, construction quality and house condition, are determined, with some degree of subjectivity, by experienced appraisers employed in the tax office. One additional variable deals with the location of the property in Wichita County. Variables included were those provided by the tax appraisal office. Other potential variables such as lot size, landscaping, or whether the property has a basement were not available from the database. The appraisal office (Wichita Falls Tax Appraisal District, 2005) identifies the neighborhood group code (NGC), which is similar in concept to a subdivision. For the time period studied, there were 141 NGCs. In some NGCs, the houses were extremely homogeneous. In other words, the sales price, appraised values, size in square footage, classification and condition codes, bedrooms, bathrooms, and other features were quite similar for most houses. Houses in other NGCs displayed wide variation in the independent variables. In these neighborhoods, wide differences were noted in elements like the appraised value, sales prices, and house size.

For purposes of this study, the NGC proved to be a critical variable. The analysis utilized by the authors followed the approach by officials in the appraisal office. Only those neighborhoods which had sufficient sales activity to provide an adequate sample size were considered. Thirty of the NGCs were randomly selected for analysis. For the time period studied, the sample sizes (number of houses sold in each NGC), ranged from 36 to 151, with a mean of 80. Mean sale price for the houses in the database was $79,619. Calculating the mean sale prices for each NGC yields a low of $25,967 and a high of $143,954.
MODELING AND ANALYSIS PROCEDURES
For each NGC, a multiple regression model was developed utilizing the best subsets approach to identify the appropriate independent variables for inclusion. Likewise, ANN models were developed to predict sales price for the individual properties in each of the 30 NGCs. Independent variables included in model development are listed in Table 1.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appraised Value</td>
<td>Value of house according to appraisal district</td>
</tr>
<tr>
<td>Size</td>
<td>Square feet of house</td>
</tr>
<tr>
<td>Age</td>
<td>Age of house in years at time of sale</td>
</tr>
<tr>
<td>Pool</td>
<td>Dummy variable (0 = no, 1 = yes)</td>
</tr>
<tr>
<td>Bedrooms</td>
<td>Number of bedrooms</td>
</tr>
<tr>
<td>Baths</td>
<td>Number of bathrooms</td>
</tr>
<tr>
<td>Classification</td>
<td>Categorical for construction quality (1 = lowest, 10 = highest)</td>
</tr>
<tr>
<td>Condition</td>
<td>Categorical for condition of house (0 = lowest, 4 = highest)</td>
</tr>
</tbody>
</table>

The two neural network software tools used were NeuralTools and NeuroShell Predictor. NeuralTools was developed by Palisade Corporation (see http://www.palisade.com/neuraltools/), which is the same company that developed @Risk, a widely used software package used in risk analysis. NeuroShell Predictor was developed by Ward Systems Group (n.d.), which, according to its website, has been an industry leader in artificial intelligence since 1982. Both ANN software packages are implemented as Excel spreadsheet add-ins. Both implement a single hidden-layer network that trains through a combination of BP and genetic algorithms in order to reduce training times. These software tools iteratively optimize the number of hidden nodes. The type of node functions are not mentioned in the NeuroShell Predictor documentation (Ward Systems Group, personal communication, November 15, 2007); however, Ward Systems personnel (2007) report that the ANNs are configured using “a variant of the Cascade Correlation method.” This method is described by Lahnajärvi, Lehtokangas, and Saarinen (2004) as using “sigmoidal” node functions.

Two analysis approaches were used. First, a blocked ANOVA analysis was performed on the mean absolute deviations (MADs) between (a) predicted sales prices that result from applying a modeling technique to the data set and (b) the actual sales prices. Thus, three sets of MADs were generated: one for each modeling technique. The ANOVA factor of variation is the modeling technique; the factor levels are: multiple regression, NeuralTools, and NeuroShell Predictor. The
blocking factor is NGC. A follow up Tukey test was administered; the result led to the second analysis approach.

Second, in comparing results for a particular NGC, the two ANN product outputs and the multiple regression model output were used to forecast a sale price for every house in the NGC. Forecasting accuracy for each model on a given house in a particular NGC was calculated by finding the absolute value of the difference between the forecast sale price and the known actual sale price (i.e., the absolute deviation). The paired $t$-test was used to determine if there existed a statistically significant difference between the absolute deviations produced by the multiple regression model and those absolute deviations developed utilizing NeuralTools. Likewise the paired $t$-test was utilized to test for statistical significance in the absolute deviations of the multiple regression model and the absolute deviations obtained using NeuroShell Predictor. The ANN tools were run at their default settings. An alpha value of 0.05 was used throughout the study.

**RESULTS**

The randomized blocked ANOVA test for significant differences in MAD between the modeling techniques—with NGC used as the blocking factor—yielded an observed significance of 0. Using Tukey’s Honestly Significant Difference test for blocked designs reveals that the average MADs for each modeling technique are significantly different among all pairs of modeling technique. In this test, the MADs for NeuralTools had the smallest average while MADs for the multiple regression analysis had the largest average. The relative efficiency for the blocked design versus the one-way design was 1.32; that is, the one-way design would have required a factor of 1.32 more observations to achieve the same variance level on any contrast (Neter, Wasserman, & Kutner, 1990). The hypothesis for homogeneous variances cannot be rejected (observed significance = 0.36).

As a group, Tables 2, 3, and 4 contain the results for paired $t$-test analyses for each NGC. Table 2 is essentially a legend for Tables 3 and 4. Table 3 presents a

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Column Heading Descriptions for Tables 3 and 4</strong></td>
</tr>
<tr>
<td><strong>Description</strong></td>
</tr>
<tr>
<td>n</td>
</tr>
<tr>
<td>Mean Price</td>
</tr>
<tr>
<td>tcalc</td>
</tr>
<tr>
<td>$p$-value</td>
</tr>
<tr>
<td>Sig.Neural Network</td>
</tr>
<tr>
<td>Sig.Mult. Reg.</td>
</tr>
</tbody>
</table>

Regional Business Review
summary of the results from comparing the absolute deviations from multiple regression forecasts to the absolute deviations from forecasts obtained from NeuralTools. The data in Table 3 has been sorted by the mean home sale price for each of the NGCs.

The results displayed in Table 3 illustrate clearly that the ANN-derived forecasts are superior to the forecasts from multiple regression modeling. In only 2 of the 30 NGCs (NGC numbers 8 and 27) did the multiple regression models provide a statistically more accurate forecast than the neural network models. In 4 of 30 NGCs (numbers 3, 10, 12, and 18), the difference between the absolute deviations for the two approaches proved to be statistically insignificant. These four NGCs are shaded in Table 3. However, in 24 of the 30 NGCs, the forecast from NeuralTools was statistically more accurate than forecasted price derived from the multiple regression models. NeuralTools significantly out-performed multiple regression in forecasting the sale price for these residences.

A similar approach was utilized to compare the forecasting accuracy for multiple regression and NeuroShell Predictor. Table 4 summarizes the comparison for forecasting accuracy for multiple regression and NeuroShell Predictor. In 17 of the NGCs, NeuroShell Predictor provided a statistically more accurate forecast than the regression models. Multiple Regression forecasts were statistically more accurate than the NeuroShell forecasts in 3 of the NGCs (NGC numbers 21, 25, and 29). In 10 of the (NGC numbers 3, 6–8, 12–14, 17, 27, and 30) the forecasts using NeuroShell were not statistically significantly different from those found via multiple regression. These “ties” are shaded in Table 4.

While it was not the primary purpose of the paper to compare the two ANN software tools, it is a logical extension of the research to do so. Table 5 presents a summary of the statistical analysis of the comparison of the forecasting accuracy for NeuralTools and NeuroShell Predictor.

Neural Tools provided a statistically more accurate forecast in 19 of the 30 NGCs. NeuroShell outperformed NeuralTools in 4 of the NGCs. Seven of the NGCs had no significant difference in the forecasts of the two neural network tools: these “ties” are shaded in Table 5.

Table 6 shows the overall results of comparing prediction methods. The row heading “ANN Software Tools” refers to the result where at least one of the ANN tools is superior to multiple regression, but neither is superior to the other. Given the default settings of the NeuroShell Predictor tool, perhaps results could be improved by fine tuning the options. The NGCs that produced inconclusive results deserve a closer look for conditions offering explanatory value (e.g., variability among homes within the NGC).
### Table 3
Summary of Forecasting Accuracy: Multiple Regression and Neural Tools

<table>
<thead>
<tr>
<th>NGC</th>
<th>n</th>
<th>Mean Price</th>
<th>Tcalc</th>
<th>p-value</th>
<th>Significant Neural Tools</th>
<th>Significant Multiple Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>51</td>
<td>25,967</td>
<td>-6.92</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>51</td>
<td>28,461</td>
<td>-2.57</td>
<td>0.01</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>108</td>
<td>38,183</td>
<td>-1.74</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>82</td>
<td>40,439</td>
<td>-4.61</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>64</td>
<td>44,409</td>
<td>-2.19</td>
<td>0.03</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>111</td>
<td>47,666</td>
<td>-3.80</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>79</td>
<td>48,945</td>
<td>-2.96</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>151</td>
<td>50,554</td>
<td>3.36</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>36</td>
<td>58,138</td>
<td>-7.23</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>69</td>
<td>64,334</td>
<td>-1.07</td>
<td>0.29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>63</td>
<td>65,507</td>
<td>-8.04</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>68</td>
<td>65,654</td>
<td>-1.94</td>
<td>0.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>116</td>
<td>70,467</td>
<td>-3.74</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>76</td>
<td>72,338</td>
<td>-2.62</td>
<td>0.01</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>54</td>
<td>72,426</td>
<td>-4.16</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>63</td>
<td>76,589</td>
<td>-3.83</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>66</td>
<td>79,524</td>
<td>-4.68</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>118</td>
<td>85,287</td>
<td>-1.65</td>
<td>0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>113</td>
<td>94,675</td>
<td>-8.11</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>47</td>
<td>95,390</td>
<td>-3.83</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>137</td>
<td>95,875</td>
<td>-5.20</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>74</td>
<td>101,488</td>
<td>-4.91</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>34</td>
<td>103,625</td>
<td>-4.04</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>114</td>
<td>105,523</td>
<td>-4.93</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>151</td>
<td>113,677</td>
<td>-4.22</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>41</td>
<td>114,801</td>
<td>-5.22</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>73</td>
<td>118,313</td>
<td>3.14</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>77</td>
<td>126,503</td>
<td>-6.73</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>65</td>
<td>140,114</td>
<td>-4.05</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>59</td>
<td>143,954</td>
<td>-4.88</td>
<td>0.00</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

Overall 79,619 24 2
### Table 4
Summary of Forecasting Accuracy
Multiple Regression and NeuroShell Predictor

<table>
<thead>
<tr>
<th>#</th>
<th>n</th>
<th>Mean Price</th>
<th>Tcalc</th>
<th>p-value</th>
<th>Significant Neuro Shell</th>
<th>Significant Multiple Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>51</td>
<td>25,967</td>
<td>-3.555</td>
<td>0.001</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>51</td>
<td>28,461</td>
<td>-5.896</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>108</td>
<td>38,183</td>
<td>-0.165</td>
<td>0.870</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>82</td>
<td>40,439</td>
<td>-4.076</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>64</td>
<td>44,409</td>
<td>-2.646</td>
<td>0.010</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>111</td>
<td>47,666</td>
<td>-0.810</td>
<td>0.420</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>79</td>
<td>48,945</td>
<td>-1.293</td>
<td>0.200</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>151</td>
<td>50,554</td>
<td>-1.560</td>
<td>0.250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>36</td>
<td>58,138</td>
<td>-5.961</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>69</td>
<td>64,334</td>
<td>-3.391</td>
<td>0.001</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>63</td>
<td>65,507</td>
<td>-2.881</td>
<td>0.005</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>68</td>
<td>65,654</td>
<td>-0.981</td>
<td>0.330</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>116</td>
<td>70,467</td>
<td>-0.762</td>
<td>0.448</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>76</td>
<td>72,338</td>
<td>-0.884</td>
<td>0.380</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>54</td>
<td>72,426</td>
<td>-3.800</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>63</td>
<td>76,589</td>
<td>-3.300</td>
<td>0.002</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>66</td>
<td>79,524</td>
<td>-1.610</td>
<td>0.112</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>118</td>
<td>85,287</td>
<td>-4.646</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>113</td>
<td>94,675</td>
<td>-2.401</td>
<td>0.018</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>47</td>
<td>95,390</td>
<td>-4.115</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>137</td>
<td>95,875</td>
<td>4.304</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>74</td>
<td>101,488</td>
<td>-4.011</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>34</td>
<td>103,625</td>
<td>-2.319</td>
<td>0.027</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>114</td>
<td>105,523</td>
<td>-2.002</td>
<td>0.048</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>151</td>
<td>113,677</td>
<td>2.514</td>
<td>0.013</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>41</td>
<td>114,801</td>
<td>-2.037</td>
<td>0.048</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>73</td>
<td>118,313</td>
<td>-1.017</td>
<td>0.312</td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>77</td>
<td>126,503</td>
<td>-3.664</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>65</td>
<td>140,114</td>
<td>2.684</td>
<td>0.009</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>59</td>
<td>143,954</td>
<td>-0.959</td>
<td>0.341</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Overall</td>
<td>79,619</td>
<td></td>
<td>17</td>
<td>3</td>
</tr>
</tbody>
</table>
Table 5
Summary of Forecasting Accuracy NeuralTools and NeuroShell Predictor

<table>
<thead>
<tr>
<th>#</th>
<th>n</th>
<th>Mean Price</th>
<th>Tcalc</th>
<th>p-value</th>
<th>Significant Neural Tools</th>
<th>Significant Multiple NeuroShell</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>51</td>
<td>25,967</td>
<td>-5.952</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>51</td>
<td>28,461</td>
<td>5.477</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>108</td>
<td>38,183</td>
<td>-1.504</td>
<td>0.136</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>82</td>
<td>40,439</td>
<td>-3.563</td>
<td>0.001</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>64</td>
<td>44,409</td>
<td>0.040</td>
<td>0.968</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>111</td>
<td>47,666</td>
<td>-2.857</td>
<td>0.005</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>79</td>
<td>48,945</td>
<td>-2.426</td>
<td>0.018</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>151</td>
<td>50,554</td>
<td>3.549</td>
<td>0.001</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>36</td>
<td>58,138</td>
<td>-7.667</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>69</td>
<td>64,334</td>
<td>1.360</td>
<td>0.178</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>63</td>
<td>65,507</td>
<td>-7.827</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>68</td>
<td>65,654</td>
<td>-1.648</td>
<td>0.104</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>116</td>
<td>70,467</td>
<td>-4.498</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>76</td>
<td>72,338</td>
<td>-2.385</td>
<td>0.020</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>54</td>
<td>72,426</td>
<td>-0.277</td>
<td>0.783</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>63</td>
<td>76,589</td>
<td>0.078</td>
<td>0.938</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>66</td>
<td>79,524</td>
<td>-4.391</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>118</td>
<td>85,287</td>
<td>3.407</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>113</td>
<td>94,675</td>
<td>-3.765</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>47</td>
<td>95,390</td>
<td>0.061</td>
<td>0.951</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>137</td>
<td>95,875</td>
<td>-7.077</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>74</td>
<td>101,488</td>
<td>-2.141</td>
<td>0.036</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>34</td>
<td>103,625</td>
<td>-3.692</td>
<td>0.001</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>114</td>
<td>105,523</td>
<td>-4.068</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>151</td>
<td>113,677</td>
<td>-5.768</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>41</td>
<td>114,801</td>
<td>-5.345</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>73</td>
<td>118,313</td>
<td>3.168</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>77</td>
<td>126,503</td>
<td>-4.773</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>65</td>
<td>140,114</td>
<td>-4.931</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>59</td>
<td>143,954</td>
<td>-4.058</td>
<td>0.000</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>79,619</td>
<td>19</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ANNs are shown to be a promising tool in comparison to multiple regression for forecasting sale prices for single-family residences. The two ANN software tools used in the study, NeuralTools and NeuroShell Predictor, both provided more accurate forecasts than the traditional multiple regression models in a significant number of the 30 neighborhoods analyzed.

### Table 6
**Comparison of Predictive Model Performance**

<table>
<thead>
<tr>
<th>Significantly Better-Performing Model</th>
<th>Number of Times Method was Significantly Better</th>
<th>NGC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple Regression</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>NeuralTools</td>
<td>19</td>
<td>1, 4, 6, 7, 9, 11, 13, 14, 17, 19, 21–26, 28–30</td>
</tr>
<tr>
<td>NeuroShell Predictor</td>
<td>2</td>
<td>2, 18</td>
</tr>
<tr>
<td>ANN Software Tools</td>
<td>4</td>
<td>5, 15, 16, 20</td>
</tr>
<tr>
<td>Inconclusive</td>
<td>5</td>
<td>3, 8, 10, 12, 27</td>
</tr>
</tbody>
</table>

**Total** 30

**SUMMARY**

ANNs are shown to be a promising tool in comparison to multiple regression for forecasting sale prices for single-family residences. The two ANN software tools used in the study, NeuralTools and NeuroShell Predictor, both provided more accurate forecasts than the traditional multiple regression models in a significant number of the 30 neighborhoods analyzed.
REFERENCES


Electronic Pedigrees and Counterfeit Pharmaceuticals: The U.S. Experience with RFID

Bill Roach
Gene Wunder
Washburn University

ABSTRACT
The U.S. Food and Drug Administration has mandated the use of electronic pedigrees based on radio frequency identification technology to protect the American public from counterfeit pharmaceutical drugs and to prevent the diversion of legitimate pharmaceuticals into illegal markets. Congress passed the Pharmaceutical Drug Marketing Act in 1987. The Pharmaceutical Drug Marketing Act requires the pharmaceutical drug manufacturers to implement a “track and trace” capability. The Food and Drug Administration did not attempt to enforce the Pharmaceutical Drug Marketing Act until late 2006. By that time, radio frequency identification technology had advanced to the point where radio frequency identification-based compliance would be possible in the near future.

This article describes the evolution of radio frequency identification technology in the pharmaceutical drug manufacturing industry and describes the politics of implementing the Pharmaceutical Drug Marketing Act. While many important players are pushing radio frequency identification, implementing the Pharmaceutical Drug Marketing Act is still far from a fait accompli. Small drug wholesalers are resisting the implementation of the Act in court. The recession, difficulties with radio frequency identification technology, and information technology problems with the large data sets involved are all slowing the process.

Keywords: Big data, RFID (radio frequency identification), EPC (electric product code), counterfeit drugs, diversion, PDMA (Pharmaceutical Drug Manufacturing Act), track and trace, electronic pedigree
INTRODUCTION
This paper looks at the use of electronic pedigrees based on radio frequency identification (RFID) technology to protect the American public from counterfeit pharmaceuticals and the diversion of legitimate pharmaceuticals into illegal markets. The World Health Organization (WHO) defines a counterfeit pharmaceutical as one that is “deliberately and fraudulently mislabeled with respect to identity and/or source” (Gren, 2007). Counterfeit prescription drugs can include:

- Drugs which are deliberately and fraudulently mislabeled with respect to identity and/or source.
- Either a branded or generic drug.
- Drugs with correct ingredients, wrong ingredients, without active ingredients, with insufficient amounts of the active ingredient, or with falsified packaging.

Counterfeit pharmaceuticals pose a significant problem in the world economy because (Morris, 2002; Margolies, 2003):

1. Lost sales and depressed prices caused by counterfeit products represent an economic loss to the company whose product is being counterfeited.
2. Patients taking counterfeit prescription drugs may face significant health risks:
   A. They may not be getting the drug recommended by their physicians in the recommended strength.
   B. Counterfeit drugs could include contaminants that are harmful to patients.
   C. Counterfeit drugs may be expired or have a fraudulent expiration date.
   D. Counterfeit drugs may not have been handled, packaged, or labeled properly making them potentially ineffective or toxic.
   E. Counterfeit drugs could cause a loss of faith in the healthcare system which results in patients not seeking appropriate care.
3. The pharmaceutical industry may divert funds from research to anti-counterfeiting measures.
4. The pharmaceutical industry may not fund research to create new drugs to the same extent it has in the past.
5. Counterfeit drugs may expose the maker of the legitimate product to product liability litigation.
Some of the most recent counterfeit drugs found in the United States (and their common use) include (Runy, 2007):

- Combivir* (AIDS)
- Procrit* (kidney disease, effects of chemotherapy)
- Diflucan* (fungal infections, AIDS)
- Serostim* (AIDS wasting)
- Epivir* (AIDS)
- Sustiva* (AIDS)
- Epogen* (kidney failure, side effects of chemotherapy)
- Trizivir * (AIDS)
- Lamisil* (nail fungus)
- Viagra* (erectile dysfunction)
- Lipitor* (high cholesterol)
- Zerit * (AIDS)
- Norvasc * (calcium channel blocker)
- Neupogen* (side effects of chemotherapy)
- Zyprexa* (schizophrenia, acute bipolar syndrome)
- Propecia* (hair growth)

Outside of the United States, counterfeit Artusenate®, an anti-malaria drug, is a major problem. A 2005 article in the British medical journal, The Lancet, estimated that 40% of the Artusenate® available in developing countries is counterfeit. Experts estimate that 200,000 deaths could be avoided annually if the genuine Artusenate® were available in appropriate quantities (Stockholm Network, 2004).

The latest estimates jointly elaborated by WHO, the Organization for Economic Co-Operation and Development, and the Pharmaceutical Security Institute show (Gren, 2007):

- More than 30% of medicines in some areas of Latin America, Southeast Asia, and Sub-Saharan Africa are counterfeits.
- In emerging economies the proportion is estimated at 10%; however, in many of the former Soviet Republics it can be as high as 20%.
- In wealthy countries with strong regulatory mechanisms, counterfeits account for less than 1% of the market value, but 50% of Internet sales are counterfeit.

The U.S. Food and Drug Administration (FDA) reports that counterfeit pharmaceuticals may make up more than half of some countries’ drug supply. For example, up to 50% of anti-malarial medication in Africa is counterfeit (FDA, 2006c). The WHO reports that approximately 8% of the drugs sold in the United States are counterfeit (Gren, 2007).
Counterfeiting is a problem for commodities other than pharmaceutical drugs. U.S. Customs statistics indicate that about 17% of their seizures are media, 10% are computer hardware, and 10% are wearing apparel (U.S. Customs and Border Patrol, 2008). Statistics on pharmaceutical drugs are aggregated into an “other” category by U.S. Customs.

**PRESCRIPTION DRUG MARKETING ACT (PDMA) OF 1987 AND PRESCRIPTION DRUG AMENDMENTS OF 1992**

Congress passed PDMA in 1987 “because there were insufficient safeguards in the prescription drug distribution system to prevent the introduction and retail sale of substandard, ineffective, or counterfeit drugs and that a wholesale drug diversion submarket had developed that prevented effective control over, or even routine knowledge of, the true sources of the drug” (FDA, 2001).

A key element of PDMA is a legal requirement of drug traceability back to the manufacturer. Every pharmaceutical drug dispensed to a patient must have a pedigree; it must be traceable back through the distribution system to the manufacturer and a particular manufacturing batch. “[The] FDA has concluded that an electronic pedigree should accomplish and surpass the goals of PDMA and is potentially a more effective solution for tracing the movement of pharmaceuticals than a paper pedigree” (FDA, Department of Health and Human Services, 2004).

The FDA has identified radio frequency identification (RFID) as the most likely technology to meet the requirements of PDMA. The FDA believed that RFID technology could be implemented for pharmaceutical drugs by 2007. For pharmaceutical drugs, the track and trace requirement will involve:
Assigning serial numbers to individual packages purchased by consumers and

An information infrastructure of tags and readers.

The actual EPC RFID chips used by Walmart are now smaller than a grain of rice (Christiansen, 2007). One of the authors was informed that Walmart’s cost for RFID chips was US$ 0.18 per chip in the summer of 2005. Current tag prices average US$ 0.11 per tag when purchased in lots of one million (Odin Technologies, 2009). This cost is expected to fall as RFID manufacturing technology is improved.

The FDA required compliance with PDMA by December 2006. The actual requirement is that pharmaceutical drug manufacturing companies start implementing the electronic pedigree system beginning with the products most vulnerable to counterfeiting. The bulk of e-pedigree activity has been at the state level. Arizona, Colorado, Indiana, Oregon, and Florida had e-pedigree requirements as of 2006. Similar electronic pedigree laws in California, Nevada, and Texas required compliance by 2007. In the first few months of 2007, Idaho, Wyoming, Maryland, and North and South Dakota enacted similar requirements. California, Iowa, Kansas, Mississippi, Nebraska, New Mexico, Texas, and Vermont have all proposed legislation or planned a timeline for enacting e-pedigree (VeriSign, 2006; Harrison, 2007).

**HOW TRACK AND TRACE WORKS WITH RFID AND EPC**

The Electronic Product Code (EPC) was created as an eventual successor to the Universal Product Code (UPC) or “bar code.” The aim was to create a low-cost method of tracking goods using RFID technology. The benefits of RFID/EPC for pharmaceuticals are:

1. The EPC has larger information capacity than the UPC, making it possible to identify each unit of a product with its own EPC.
2. The product can be scanned without each individual product being scanned, generating a savings in labor costs.
3. Both the product and the RFID tag can be inside a tamper proof container on a pallet containing many stock-keeping units (SKUs), each with their own RFID tags.

Passive tags have practical read distances ranging from about 10 mm (0.39 inches) up to about 6 meters (19.7 feet) at speeds up to 160 miles per hour (260 km per hour). Because passive tags are cheaper to manufacture and have no battery, the majority of RFID tags in existence are of the passive variety. As of 2008, these tags cost from US$ 0.16 in high volumes. Universal RFID tagging of individual products will become commercially viable at very large volumes of 10 billion units per year, driving production cost to less than US $0.05 according to one manufacturer (Airvenue, 2006).
What is an EPC?
A Universal Product Code (UPC) is the sequence of numbers in a U.S.-standard barcode today. An Electronic Product Code (EPC) is a standard for data formats in RFID tags that is meant to replace the UPC barcode. For example, an EPC-96 code has four components:

1. A version number, indicating the tag type (e.g., 96-bit EPC Class 1);
2. A domain manager specifying the entity that administers the tag code, e.g., “Acme Products Co”;
3. An object class specifying the type of product the RFID tag is attached to, e.g., “Model L rocket”; and
4. A unique identifier, a number that, in combination with the other EPC components, uniquely specifies the tag (and object).

Figure 2 summarizes the EPC structure. The numbers are bit lengths for the various EPC components.

In addition, a tag contains two non-readable data elements: A 16-bit checksum (cyclic redundancy code) used to identify transmission errors, and an identification number used for such operations as “killing” (i.e., permanent disablement) the tag for privacy enforcement.

In addition to RFID/EPC tags on the package of medicine, some producers are printing EPC codes on individual pills. This means that the pedigree of individual pills can be obtained (Basta, 2006). Data on the tag can be encrypted so that it is not readable by unauthorized readers.

Track and Trace
To realize the “track and trace” potential of RFID/EPC technology, retailers, distributors, and manufacturers need access to comprehensive, real-time product
information. VeriSign Corporation, the company that maintains the Internet’s “.com” and “.net” registry, has been chosen by EPC Global to host the EPC network for the global exchange of EPC information. The cost of the EPC starter service is US$ 10,000 to sign up plus US$ 4,000 per month or a flat of US$ 50,000 for the first year (VeriSign, 2005; Gilbert, 2004). The EPC Network is composed of three key elements:

- EPC Information Services (EPC-IS)
- EPC Discovery Services
- Object Name Service (ONS)

When an RFID tag is manufactured with an EPC, the EPC is registered within the ONS. The manufacturer attaches the RFID tag to the product and the EPC becomes a part of that product as it moves through the supply chain. Particular product information is added to the manufacturer’s EPC-IS, and the knowledge that this data exists within the manufacturer’s EPC-IS is passed to the EPC Discovery Service.

As the product leaves the manufacturer’s facility, its departure is automatically registered with the EPC-IS. Similarly, when the product arrives at the next point in the supply chain (e.g., a distributor site) the EPC is automatically read and registered with the distributor’s EPC-IS and with the EPC Discovery Service.

![Figure 3: Electronic Product Code Discovery Service](image-url)

When the distributor needs product information, it asks the ONS for the location of the manufacturer’s EPC-IS. The root ONS provides the location of the manufacturer’s ONS, which in turn provides the location of the manufacturer’s EPC-IS. This query process is transparent to the supply chain member and takes only milliseconds to execute. With the manufacturer’s EPC-IS location, the distributor’s application can request specific product information.

As products progress through the supply chain, the supply chain is in constant communication with the EPC-IS. The ONS is the “glue” that holds the EPC...
Network together; this service tracks every EPC-IS and EPC Discovery Service. Authorized users have access to complete, up-to-the minute intelligence about any product. The ONS will run on the same infrastructure as the Internet’s Domain Name System (DNS), which routes domain names to specific Web servers around the world. The result is real-time full visibility of the supply chain.

**RFID Tests**

RFID technology loses some of its effectiveness with metal and liquid products. Read rates are virtually 100% at the pallet level, and 95–97% at the case level. Item level reads have been reported at a much lower level (51% to 80%). Read rates would be 100% if it were permitted to infer data from the pallet levels (Pharmaceutical Commerce, 2006b).

This has been difficult to achieve because of different definitions of what constitutes a successful read. The problem is that to achieve a high read rate requires that antenna arrays and the size and shape of the tags must be tuned for the material (e.g., glass, liquid) that is being shipped. Within the RFID supply chain, the key players are Label Converters. Label Converters put the tags onto adhesive strips and onto other mediums. If a Label Converter was using one array, one type of reader, at one distance and one read speed for their inlays, they well could get a 95% read rate, while end users might have a far lower read rate! The key is to define more specifically when a 95% read rate is achieved (what inlay combined with what antenna array past what kind of reader moving at what speed gives 95% read rates), to certify the partners according to specific definitions, and then to communicate this to potential customers (Banker, 2006).

Currently, the progress of adoption of RFID has been very uneven. Paper product manufacturers such as Kimberly Clark have been deploying RFID very rapidly because both its products and its packaging are RF inert, so Kimberly Clark can achieve virtually 100% read rates. Pharmaceutical drug manufacturers are having a more difficult time. Walmart is claiming 99.5% read rates at the case level in its pilot program (O’Connor, 2007). TAGSYS and Owens-Illinois are both guaranteeing Six Sigma read rates (Barlas, 2007).

**Diversion**

Diversion of a pharmaceutical drug is the practice of drug abusers diverting the drug from its medicinal purpose to use it to feed their addiction. Experts estimate that 30% of the drug problem in the United States is accounted for by the abuse of pharmaceutical drugs (RXDiversion, 2006). RFID will enable healthcare institutions to implement a chain-of-custody that will go all the way to the patient. All pharmaceutical drugs are tagged with a HIPA-TAG™ that identifies the NDC number, lot/date, dosage, special instructions, the intended patient’s name, the physician, and the pharmacist. During the admission process, a patient
HIPA-TAG™ is placed inside the wristband identifying the patient name, gender, SSN, birth date, blood type, allergies, admission date, contact name and phone, attending physician, the diagnosed ailments, and other relevant information. An RFID scanner reader automatically checks patient’s tag against the prescriptions at the bedside.

As the volume of RFID tags increase, the cost of the tags, readers, and the rest of the infrastructure is expected to decrease. The development of interoperable readers that can read both RFID tags and bar codes will simplify the transition from UPC bar codes to RFID codes.

For years, drug manufacturers have sold the bulk of their drugs (90%) through the three major wholesaler/distributors: AmerisourceBergen, Cardinal Health, and McKesson. These wholesalers buy large quantities of drugs in hopes that manufacturers will increase drug prices. Large purchases by the wholesalers help consumers pay less for drugs, but the practice makes it difficult for pharmaceutical companies to accurately forecast demand. Because of this, Pfizer and Merck took revenue hits upwards of $300 million to reduce their wholesaler inventories in 2003 (Miller, 2004).

A wave of mergers has meant that the drug distribution market has undergone significant consolidation over the past five years. Many pharmaceutical manufacturing companies now have the option of distributing directly to pharmacies and bypassing the wholesalers. For the wholesalers to avoid disintermediation, they must make their services more valuable. Some examples of new services and innovation might be in the handling of controlled substances, which require extra security, and in the handling of cold chain drugs, such as some vaccines that need to be stored at temperatures between two and eight degrees Celsius.

At the same time as the number of drug manufacturers has dropped dramatically, there has been an even more dramatic increase in the number of licensed drug wholesalers. The large number of licensed wholesalers makes it extremely difficult to monitor what is happening in the drug supply chain (Eban, 2005).

**FDA Action**

The FDA ordered drug manufacturers to start implementing PDMA in December 2006 (FDA, 2006d). While FDA did not offer any further delays in implementation of the PDMA, its order seemed to anticipate that immediate compliance was impossible and that the FDA would allow the drug companies to implement the PDMA first for the drugs most likely to be counterfeited.

So far RFID looks like the most feasible way, both economically and technically, to meet the requirements of PDMA. The FDA has indicated a willingness to work with the industry to ease the implementation of RFID by reducing or streamlin-
Pure and unadulterated drugs are made by the pharmaceutical manufacturers and then shipped to either drug wholesalers or directly to drug dispensers. Drugs are also exported overseas, often at a significantly lower price because of negotiated price controls.

Counterfeited drugs, including fully fake, adulterated and falsely relabeled drugs, are made by drug counterfeiters. Drug diverters help the counterfeiters, both by supplying them with legitimate drugs that can be relabeled and repackaged and by serving as the means for counterfeiters to enter the legitimate drug supply.

Diverters and their gray market distributors, can mix legitimate medicine with drugs that are counterfeit, substandard, expired or have been degraded by conditions in storage and transport.

Compromised gray market drugs move into the mainstream drug supply chain by first passing through a complex network of small distributors and secondary wholesalers.

Consumers’ purchases from online drug stores—or direct purchase of imported drugs—may also yield drugs that are unapproved by the FDA, counterfeit, or dangerous, since counterfeits are far more prevalent in foreign drug supplies (Yankus, 2006).
ing its regulations. The expectation is that RFID standards will emerge from the ongoing tests and that the pharmaceutical drug industry is following these studies and will be ready to promptly implement the standards.

Technical problems with RFID include:

- Obtaining 100% read rates.
- Assuring that the RF readers have no affect on the chemical compounds or biological substances in pharmaceutical drugs.
- Assuring that RFID does not violate the privacy rights of individuals.
- Making appropriate arrangements for the ownership, security, and access to the RFID database.
- Economically storing and processing the very large volumes of data involved in an RFID pedigree system.

One consequence of the implementation of RFID is that the participants in the pharmaceutical drug industry are likely to be unwilling to partner with firms that have not implemented RFID technology. Similar phenomena happened earlier with Electronic Data Interchange (EDI) technology. Companies refused to partner with companies that could not provide an EDI interface. In the case of EDI, it was relatively easy for a company to obtain access to the technology. RFID demands a significantly higher level of sophistication and probably a higher level of investment. Oracle, SAP, Baan, i2, and others currently have such relationships. As time goes by, the Enterprise Resource Planning (ERP) vendors will provide RFID data and software in a seamless manner. Temporarily, there are a number of firms providing middleware, a software interface between RFID and the various ERP systems.

**RFID Performance**

Implementing RFID/EPC tags involves much more than just putting RFID tags on products. Retailers buying the merchandise (e.g., Walmart) have performance requirements. Thus it is appropriate to experiment with tag encoding, tag placement, and pallet configuration. The manufacturer needs to plan for the RFID scanning configuration at distributors and retailers. Products may be exposed to humidity, heat, cold, and abrasion in transit. The manufacturer needs to plan RFID tag placement and technology in the light of the transit environment.

Implementation of RFID/EPCs for pharmaceuticals is complicated by the regulation of the pharmaceutical industry by the FDA and other government agencies in the area of product packaging, product labeling, and information privacy. A trial of RFID in the pharmaceutical industry, conducted by Accenture in 2004, led to a series of recommendations.
The project team identified and brought to the attention of EPCglobal™ and the FDA the need for standards and business practices relating to the use of RFID/EPC technologies that address the unique requirements of the pharmaceutical industry (Collins, 2005).

When the process of RFID introduction started, speculation put the size of the RFID database at 150 terabytes as compared to the 136 terabytes in the entire Library of Congress collection (University of California Berkeley, 2008). More contemporary speculation suggests that a large retailer could generate 15 terabytes per day (Michahelles & Ilic, 2009). The technology to store and process such large volumes of data economically may not be available for some time.

**RFID Standards**

Part of the problem in implementing an RFID based electronic pedigree system for prescription drugs is the absence of standards. Initially, the problem was the standard for coding the tag. There was an International Standards Organization (ISO) standard tag (ISO 18000-6) backed by vendors, but it was rejected by major users such as Walmart, Gillette, and Proctor and Gamble who went on to develop the EPC. GEN 1 of the EPC had many variations; it was essentially a proprietary standard with many variations, that is, there was no standard. Recently, EPC GEN2, a true standard, has been implemented, and it has won wide acceptance. The transition to GEN2 was still underway at the start of 2011 (Frost & Sullivan, 2011).

Still there are variations in the frequencies used by the tags and tag readers. UHF (Ultra High Frequency) seems to dominate for pallet and case tags. HF (High Frequency) dominates for SKU tags. Electronic pedigree systems implemented Pfizer for Viagra® used UHF (915 MHz) tags for cases and pallets and HF (13.56MHz) tags for SKUs (O’Connor, 2005). Before the Viagra® trial, McKesson, the drug wholesaler involved, had been experimenting with UHF (915 MHz). Purdue Pharma used UHF tags on bottles of OxyContin® shipped to Walmart. Purdue chose UHF because it was what Walmart wanted to use (Roberti, 2006). HF vendors claim that HF tags work better around metal and water and hence are better for item level tagging. Recent advances in UHF RFID tags which put a new kind of antenna on the UHF chip, so-called “near field” UHF tags, have yielded much better results.

The International Standards Organization has adopted a number of standards both for the RFID chips themselves and the RF interface (ISO, 2008). To some extent, the implementation of RFID technology is dependent on the development of standards. But this has not been a rapid process. While EPC (AUTO ID Center) has broad support from the likely supply chain customers of RFID/EPC, it is essentially a proprietary standard. The history of proprietary standards, even those backed by major players (e.g., General Motors and the Machine Automa-
tion Protocol (MAP) standard, IBM, and SAP) is not encouraging. The history of proprietary standards is a history of failure, and many players stay well on the sidelines until a real standard emerges.

On July 11, 2006, the ISO adopted the EPC Gen 2 standard as ISO 18000-6C (O’Connor, 2006). The adoption of a standard (an industry standard not a proprietary standard) should ease the way for the rapid development of an electronic pedigree system. Retailers, pharmaceutical drug manufacturers, pharmaceutical drug wholesalers, software houses, and hardware manufacturers no longer need fear that they would be gambling on the wrong technology. There is a clear path forward, and ISO has said where it lies.

**RESOLUTION OF RFID CONTROVERSY**

There are a number of 800-pound gorillas in the room:

- The World Trade Organization (WTO)
- The U.S. Department of Defense (DOD)
- Walmart and other EPC partners
- The Food and Drug Administration (FDA)
- Several states including Florida and California
- The United States Congress
- American hospitals
- Wholesale drug distributors

All of the gorillas are pushing for rapid deployment of a standard system RFID tags and readers with varying emphasis on the rapid deployment and standard part of the requirement. The small, independent, secondary wholesale drug distributors are not supportive of the change. Secondary wholesalers are wholesalers who are not authorized distributors of record (ADRs) for the drugs they distribute. In May of 2007, the U.S. Senate passed the Food and Drug Administration Revitalization Act (S1082; Juvan, 2007). The act provided a provision which requires anti-counterfeiting technology analogous to those technologies used with U.S. currency and specifically excluding anti-counterfeiting technologies such as RFID or barcodes which require an infrastructure of readers and scanners. Lobbying Congress is likely to play a more important role in determining the course of the future of RFID technology. California’s threatened implementation of its own e-pedigree requirement is likely to motivate the pharmaceutical drug manufacturers to get behind a federal (FDA) program that would pre-empt state requirements.

Some independent drug wholesalers are also resisting the movement to RFID adoption. Immediate adoption of RFID technology would probably lead to the exclusion of many independent drug wholesalers from the pharmaceutical drug supply chain and their consequent bankruptcy (FDA, 2006d).
Follow the Money!
Economic benefits associated with EPC/RFID adoption range between $500 million and $1 billion annually for pharmaceutical manufacturers, and between $200 million and $400 million annually for healthcare distributors, according to a new report released by the Healthcare Distribution Management Association (HDMA) Healthcare Foundation. The report, entitled “Adopting EPC in Healthcare: Costs and Benefits,” finds additional annual benefits of between $200 million and $400 million in avoided incidents of counterfeiting, which can damage brand value and public confidence (A.T. Kearney, 2004).

Potential benefits of this magnitude will be impossible to ignore. The A.T. Kearney (2004) study shows that virtually everyone in the system (manufacturers, wholesalers, retailers, and hospitals) will share in the gain.

While the pecuniary benefits are large, some of the non-pecuniary benefits are more compelling. Drug manufacturers will reap substantial cost savings from greatly simplified and focused recalls of potentially dangerous drugs. The entire society will benefit from the rapid resolution of drug recall situations. The hospitals can reap significant cost savings from the reduction in medication errors. The larger society will probably value this improvement more highly than pharmaceutical drug manufacturers.

The benefits of RFID are not as compelling for individual community pharmacies and small pharmacy chains. Accenture recently completed a study funded by National Community Pharmacists Association (NCPA) and National Association of Chain Drug Stores (NACDS; Bacheldor, 2008). The study estimated that the costs of implementing an e-pedigree system based on RFID would be much higher than previously estimated. The study was a response to the proposed Safeguarding Pharmaceuticals Act of 2008 (HR 5839), introduced in April of 2008. HR 5839 was an attempt by Congress to put a federal track and trace system on the fast track and avoid the delays created by the *RxUSA Wholesalers, Inc. v. HHS* (2006) suit. It seems likely that RFID does have negative implications for community pharmacies and small pharmacy chains. However, the public’s right to a secure supply of ethical pharmaceuticals seems likely to prevail.

RFID technology will be implemented for the pharmaceutical drug supply chain, and it is likely to be implemented more quickly than the EDI or UPC precedents would indicate for two reasons:

1. the large financial incentives for adopting a standard, and
2. the large organizations (e.g., the FDA, Walmart, the DOD, and the WTO) that are aggressively championing RFID.
Technical difficulties in achieving 100% read rates and in economically processing and storing very large volumes of data are slowing the process. In addition, the recovery from the United States recession is slow and has not provided much incentive for the investment in RFID infrastructure.

Implementation
On June 9, 2006, the FDA announced that it would not delay the implementation of the PDMA of 1987 beyond December of 2006 (FDA, 2006d). While the FDA did not delay compliance with the PDMA, it also realized that compliance will take place over time. In its compliance guide, the FDA told pharmaceutical drug manufacturers to prioritize (FDA, 2006d) their implementation of the PDMA. The priority criteria for track and trace to be implemented include:

1. high dollar value in the United States market,
2. history of counterfeiting,
3. for a new drug, reasonable likelihood of counterfeiting, and
4. other law violations involved.

Less than two weeks later, June 20, 2006, pharmaceutical industry associations threw their weight behind the FDA recommendations (Bacheldor, 2006). The states theoretically could set a stricter timetable than the FDA; however, at first they gave way to the FDA by delaying the implementation of its drug pedigree regulation until January 1, 2008 (Basta, 2006; National Conference of State Legislatures, 2008). It makes a great deal of sense for the states to give way to the FDA because:

1. the states do not have the expertise to enforce an electronic pedigree requirement,
2. the states would be hard pressed to finance the enforcement of such a pedigree requirement, and
3. the FDA must now shoulder the responsibility for the tragic incidents that may occur during the implementation of PDMA.

More recently, California decided to implement its own e-pedigree requirement with an effective date of January 2009 (FDA, 2008). The California law applies to “dangerous drugs” (i.e., prescription medications); however, California later delayed implementation of its requirement until 2011 (NCPA, 2008).

RxUSA Wholesalers, Inc. v. HHS and Drug Wholesaler Licensing
On December 8, 2006, a federal district court in the Eastern District of New York issued a preliminary injunction in RxUSA WHOLESALE v. HHS to prohibit FDA from implementing a
regulation that requires that certain information be included in a pedigree, which documents the custody of certain prescription drugs in the drug supply chain. The regulation, (21 CFR § 203.50(a)), which went into effect on December 1, 2006, was issued by FDA to implement the Prescription Drug Marketing Act of 1987 (PDMA), as amended by the Prescription Drug Amendments of 1992 (PDA; FDA, 2006a).

Jayne Juvan, an attorney who runs a blog on issues related to the PDMA and RFID, says that the courts are likely to support the injunction until the case is resolved, probably several years from now (Juvan, 2006). The injunction does not completely suspend the implementation of the PDMA, but it does have a substantial effect.

The supply chain for pharmaceutical drugs is quite complex. The FDA implementation of PDMA would have simplified the supply chain by eliminating many of the small wholesalers. The small wholesalers could not sell their current inventories because they did not have pedigrees for them, and they could not return them to the manufacturer. The small wholesalers had to supply pedigrees with the drugs they sold, but they could not procure pedigrees with the drugs they purchased. The court enjoined enforcement of portions of the PDMA because of these problems. Nevertheless, channel management issues are at the heart of the counterfeit drug problem.

In effect, the FDA implementation of PDMA made the large pharmaceutical drug manufacturers the ultimate licensing authority for drug wholesalers. If the drug manufacturers will not provide wholesalers the pedigree for the drugs the wholesalers handle or grant the drug wholesalers authorized distributors of record (ADR) status, then the drug wholesalers cannot do business.

Secondary wholesalers have historically played an important role in facilitating price competition in the retail market for pharmaceutical drugs. By buying large quantities of drugs and reselling them to independent pharmacies and small pharmacy chains, the secondary wholesalers have helped the small pharmacies to be more competitive in the retail pharmacy market. Elimination of the secondary drug wholesalers would have significant anticompetitive implications.

One alternative being considered by Congress is to supersede state licensing of drug wholesalers by federal licensing of drug wholesalers through the FDA. The Plasma Protein Therapeutics Association (PPTA) advocated such a plan in a letter to Senators Ted Kennedy and Michael Enzi (Birkofer, 2007). This proposal has appeared in various incarnations and is included in the current healthcare reform proposals being considered by Congress.
Repackaging

RFID, by itself, cannot secure the American pharmaceutical drug supply from counterfeiting. If drugs are repackaged at the pharmacy, the most an electronic pedigree system implemented through RFID can do is to safeguard the drug supply as it enters the pharmacy (FDA, 2006b). Drug repackaging is currently licensed in the United States. Most international, online pharmacies are not licensed repackers. These online pharmacies are the primary source of concern right now, but the outright prohibition of repackaging should be considered. Drug repackaging, although common, is not the norm everywhere. In Argentina, patients routinely purchase medicine in containers sealed by the manufacturer.

Fees for drug repackaging represent a substantial share of the income earned by pharmacies and some physicians.

According to a DispenseXpress price list, doctors can clear $65.50 every time they dispense a 90-tablet bottle of 800-milligram ibuprofen. They can buy the common painkiller from DispenseXpress for $9.19, but under current workers’ comp rules, they can bill insurers at a rate of $74.69 per bottle (Senate Democratic Caucus, State of California, 2005).

Elimination of repackaging (and repackaging fees) might happen, but so far there has been very little discussion of how this will happen and what will substitute for repackaging fees. In some countries, there is much less repackaging than there is in the United States. The new e-pedigree requirement in California makes specific provision for repackaging.

... a single pedigree shall include every change of ownership of a given dangerous drug from its initial manufacture through to its final transactions to a pharmacy or other person for furnishing, administering or dispensing the drug, regardless of repackaging or assignment of another National Drug Code (NDC) Directory number ... (California Business and Professions Code section 4034(c), 2011).

CONCLUSIONS

The implementation of track and trace through and RFID based e-pedigree system has not been as smooth and orderly as consumers might have hoped. Nevertheless, it is possible to draw certain conclusions about this process:

1. An RFID based e-pedigree system for pharmaceutical drugs is going to happen at the federal level. The people who are lined up in support of such a system are quite formidable. They will not be denied. The supporters of RFID include the U.S. Congress, the Food and Drug Administration (FDA), California and the 10 other states considering e-pedigree regulations, the Department of Defense (DOD), Walmart and other major retailers, major manufacturers of consum-
er goods, and American hospitals. The manufacturers of pharmaceutical drugs and the large drug wholesalers can probably be added to this list. If they do not get behind a federal program, they may be stuck with a patchwork of state systems.

2. Not everyone will gain from the federal e-pedigree system that is likely to be implemented. Community pharmacies and small pharmacy chains will incur costs that are a larger fraction of their sales than the corresponding costs incurred by the large pharmacy chains. It may not be economically feasible for small pharmacies to implement the other changes that result in large economic benefits for large pharmacy chains. The survival of the independent, secondary drug wholesaler needs to be a priority. Historically, the secondary drug wholesaler has played an important role in making independent pharmacies and small pharmacy chains competitive.

3. The e-pedigree regulations have probably not worked out a working arrangement for repackagers. Repackagers include physician groups, freestanding pharmacies, and hospital pharmacies. Repackagers and “unit doses” have been with the pharmaceutical drug supply chain for a long time. It seems improbable that pharmaceutical drug manufacturers can do all of the packaging, including unit doses, that needs to be done.

4. California’s entry into this process will probably have a significant impact on the emerging e-pedigree standard. The California law spoke specifically to the repackaging issue, and that will be a critical part of the system that will emerge. Years ago, California led the way in setting strict emission standards for automobiles. The California standard is now a de facto national standard. The California e-pedigree requirements will probably have a significant impact on the federal e-pedigree system. The FDA is really not free to ignore what happened in California.

5. The U.S. approach to e-pedigrees should have a significant impact on a global standard. Much of the technology being implemented already corresponds to international (ISO) standards. Developed economies have an interest in implementing similar systems and facilitating international trade in pharmaceutical drugs.

6. Some countries may not have an interest in eliminating trade in counterfeit drugs. Unless the drug manufacturers license AIDS and anti-malarial drugs on very advantageous terms, it seems unlikely that the developing countries will support international anti-counterfeiting agenda.
7. The e-pedigree system will work to eliminate counterfeits from the legitimate pharmaceutical drug distribution channels. The e-pedigree system is not likely to have a comparable effect on the Internet pharmacies.

8. The e-pedigree systems demand a large investment in market infrastructure. That investment will be difficult to justify until the recovery from the recession is complete. The forecasters are expressing some uncertainty about the speed and magnitude of the recovery. The forecast uncertainties translate into uncertainties about the implementation of an e-pedigree system for pharmaceutical drugs.
REFERENCES


U. S. Food and Drug Administration [FDA]. (2006b, June 9). *Drug supply security expert points to weakness in RFID. FDA Week, 12(23).*


Book Review
While K–12 education continues to take hits from movies like *Waiting for Superman*, higher education’s woes are not so widely publicized. Complaints about escalating costs are common but the causes are not often as clearly depicted as in the book *Higher Education? How Colleges Are Wasting Our Money and Failing Our Kids—and What We Can Do About It*. Dreifus’ and Hacker’s view of the ills of higher education are often sadly on target and well articulated. The outcry regarding faculty tenure and athletic budget sinkholes have been covered extensively in the mainstream media so this review will instead focus on the less publicized problems with higher education discussed in this book.

In an economy where we often hear “you get what you pay for,” high-profile Ivy League institutions (referred to by the authors as the “Golden Dozen”) are ridiculed for being considered places a student would receive one of “the best” educations in our nation. Harvard’s undergraduate education is depicted in an especially poor light in the chapter *Teaching: Good, Great, Abysmal*. The chapter opens with a comparison of Harvard and Oregon State that have 8% and 84% acceptance rates, respectively. Harvard spends $125,684 per student while Oregon State spends $29,990 on each of its students (p. 77). According to the authors, both schools do, however, have one powerful similarity “what they have in common—connects them with many other colleges—is the teaching their students receive. All-in-all, it’s pretty poor” (p. 77). Harvard students are quoted as complaining that faculty will openly admit they are at Harvard for the research...
opportunities and teaching responsibilities are often seen as a nuisance. Hacker and Dreifus draw the conclusion that a majority of the faculty at both institutions are apathetic about student learning and, perhaps more startling, they do not feel like they have to be concerned about it. However, later in the book Harvard was given credit for its willingness to address its teaching problem with the appointment of a task force that put together a report titled “A Compact to Enhance Teaching and Learning at Harvard.” Three basic suggestions from the report are: reward faculty who are willing to devote more time to the classroom, encourage peers to visit each other’s classrooms to provide feedback regarding teaching effectiveness, and “all junior faculty should be urged to do individual videotape sessions” (p. 93). I had to laugh out loud at the authors’ questioning of why only the junior faculty would be subject to this suggestion. Their research suggests concerns about senior faculty are much more widespread and frequent than that of the junior faculty but the 839 full professors were not asked to commit to this recommendation (p. 94).

The primary factor attributed to lackluster instruction at institutions of higher learning is that faculty are focused on research and publication to ensure they are promoted. After all, it is human nature to focus on what we are rewarded for and at most institutions of higher education faculty are rewarded for research. While this revelation is not surprising for a place like Harvard, I was disheartened to learn that Oregon State rewards the same element. According to the authors, as Oregon State is working to reach parity with its fellow state University of Oregon, its faculty are being encouraged to focus their energy on research at the expense of teaching. The authors record a response a student at another research university shared with them. The student claims the response he heard from a department chair when he complained about the quality of teaching was, “If you want that sort of thing you should have gone to a liberal arts school. That’s not what we do here” (p. 79). To back up their case that good teaching and academic research are actually inversely related (and thus counterproductive) the authors use statistics from the National Survey of Student Engagement (2008–2009) with respect to faculty ratings on “helpfulness” and “availability.” Smaller schools such as Centre College in Kentucky (81), Earlham in Indiana (78), Augsburg in Minnesota (77), and Hendrix in Arkansas (75) earned significantly higher scores than research schools such as: University of Virginia (46), University of Iowa (45), University of Michigan (40), and University of Minnesota (39; p. 90).

According to Hacker and Dreifus, two fairly simple skills are needed for good teaching. Faculty need to be able to explain their material and care enough to ensure their students are grasping it. Retention rates are often used to measure the results of these factors. The book uses statistics from three state campuses in Missouri to highlight the high non-return rates for freshmen with 35%, 36%, and 44% drop-out rates noted (p. 92).
I did find the authors’ criticism of the use of contingent faculty (adjuncts and lecturers) in the chapter entitled *Contingent Education* in stark contrast to their approval of the use of preceptors in their chapter *Visiting the Future in Florida* which promotes the use of grading preceptors who are paid quite poorly. Their coverage of the transformation of Humanities 2510 to accommodate sections of 1400 students seems to contradict one of their twelve proposals to end exploitation of adjuncts, “It is immoral and unseemly to have a person teaching exactly the same class as an ensconced faculty member, but for one-sixth the pay” (p.240). “These preceptors are not content experts, in fact, a bachelor’s degree in any field usually [suffices] as their basic qualifications.” The book provided one example of a preceptor being paid $25 per student or $2,250 for grading papers and other assignments (p. 197). The final two chapters highlight schools that make Hacker’s and Dreifus’s *Schools We Like—Our Top Ten List*. 

This book is a must-read for anyone interested in the quality of higher education. Administrators, faculty and staff, and parents of students who will attend a college or university would be well-advised to read and consider some of the suggestions provided in the book.