Inquiry: The University of Arkansas Undergraduate Research Journal

Volume 9

Article 1

Fall 2008

Inquiry: The University of Arkansas Undergraduate Research Journal - Volume 9 - Fall 2008

Inquiry Editors

Follow this and additional works at: http://scholarworks.uark.edu/inquiry

Recommended Citation

Available at: http://scholarworks.uark.edu/inquiry/vol9/iss1/1

This Entire Issue is brought to you for free and open access by ScholarWorks@UARK. It has been accepted for inclusion in Inquiry: The University of Arkansas Undergraduate Research Journal by an authorized editor of ScholarWorks@UARK. For more information, please contact scholar@uark.edu, ccmiddle@uark.edu.
INQUIRY: The Undergraduate Research Journal of the University of Arkansas

Volume 9 – Fall 2008

2 Editor’s Foreword, Publication Board

Winners of the 2008 Undergraduate Research Award given by the Teaching Academy

3 ANTHROPOLOGY: Covington, A. K. Why Pentacostal? A Look at the Phenomenon of Rapid Pentecostal Growth in Latin America. (for full text, go to http://inquiry.uark.edu/)


16 MUSIC: Sears, D. Processing Surprise Tension in Tonal Melodies.

Remaining Table of Contents in Alphabetical Order by Last Name

23 MECHANICAL ENGINEERING: Cantrell, A. Factors Affecting the Spread of a Bioterrorist Agent Throughout a Building.

27 FINANCE: Cox, C. Flights-to-Quality: The Effects of Market Volatility on Short Term U.S. Treasury Yields.


47 ECONOMICS: Lippert, T. The Impact of Severance Taxes on the Arkansas Natural Gas Market.

57 CLASSICAL STUDIES: Merced, D. J. Roman Isis and the Pendulum of Religious Tolerance in the Empire. (for full text, go to http://inquiry.uark.edu/)

58 TRANSPORTATION AND LOGISTICS: Otto, A. Economic Literacy at the University of Arkansas.

66 HISTORY: Peyton, C. Opportunity Knocks: An Examination of the Knoxville Transient Bureau and Transient Population. (for full text, go to http://inquiry.uark.edu/)


78 COMPUTER SCIENCE AND COMPUTER ENGINEERING: Rosequist, A. Relative Searching Using an Ordered Token List.


Cover: Roma Termini, Rome, Italy. Image used with permission of Tricia R. Quinn.
The Inquiry journal is a project of the Teaching Academy of the University of Arkansas. The journal is supported financially and conceptually by the offices of the Provost and the Vice Chancellor for Research. Through print and on-line publication, Inquiry provides a forum for sharing the research and creative endeavors of undergraduate students at the U of A.

This print issue of Inquiry records the scholarly contributions of 16 U of A student/faculty mentor pairs during the 2007/2008 academic year. The full text is included for 13 manuscripts. Abstracts are provided for an additional three papers which are posted on the Inquiry website (http://inquiry.uark.edu/). These longer works were submitted originally for consideration for the Undergraduate Teaching Award and received high ratings from reviewers. Student authors of longer papers were encouraged to submit condensed versions of their work for print publication. In some instances, reducing document length significantly diminished the overall quality of the work, and these specific authors were given an opportunity to publish on-line.

As has been true in the past, the number of articles accepted for publication is controlled in part by the size of the journal, and we received many more high quality manuscripts than we could publish. The articles included in this issue are drawn from disciplines in four of the six undergraduate colleges and schools—the J. William Fulbright College of Arts and Sciences, the Sam M. Walton College of Business Administration, the Dale Bumpers College of Agricultural, Food & Life Sciences, and the College of Engineering—and are representative of the quality of research done by students in the various disciplines represented on campus. The breadth of subject matter included here is testimony to the commitment made throughout the university to research at the undergraduate as well as the graduate level. Articles are selected for publication after review by faculty members. The range of disciplines at an institution like ours makes it impossible for a few individuals to evaluate critically all the papers submitted without the assistance of experts in specific academic areas. The Editor and Publication Board of Inquiry are grateful for the assistance of those campus colleagues who have served as reviewers. As Editor, I must also thank the members of the Publication Board who gave so generously of their time, particularly at the end of the spring semester when no one has any free time.

While the papers chosen for publication vary in subject, method, length, writing style, and manuscript format, they are uniformly excellent in content. As much as possible, we have endeavored to maintain discipline-specific styles to provide students with a publication experience comparable to one they might find in their individual disciplines. The intent of the journal is to record the depth and breadth of the scholarly activities of some of the university’s best undergraduate students. I hope you enjoy the print and on-line offerings of the 2008 issue of Inquiry.

Barbara B. Shadden, Editor

INQUIRY PUBLICATION BOARD, 2007-2008 ACADEMIC YEAR

Ro Di Brezzo, University Professor of Exercise Science; Norman D. Dennis, Jr., Professor of Civil Engineering; David E. Gay, Professor of Economics; Ethel Goodstein, Professor of Architecture; Amy Jo Herzberg, Professor of Drama; John M. Norwood, Associate Professor of Accounting; Curt R. Rom, Professor of Horticulture; Charles F. Rosenkrans, Jr., Professor of Animal Science; Mary Jo Schneider, Professor of Anthropology; Murray J. Smart, Jr., Emeritus Professor of Architecture.
WHY PENTECOSTAL?
A LOOK AT THE PHENOMENON OF RAPID PENTECOSTAL GROWTH IN LATIN AMERICA

By Allison Kidd Covington
Department of Anthropology

Faculty Mentor: Stephen M. Striffler
Department of Anthropology

Abstract
I first got the idea for my thesis studying Spanish in Costa Rica in the fall of 2006. Not long after I arrived at my host family's home, my host mother asked me whether or not I was “evangélica”. I was somewhat confused by this question because it went against my previous assumption that the majority of Latin Americans—or at least Latin American Christians—were Catholic. Knowing a minimal amount of Spanish and very little about the culture, I answered yes, essentially translating “evangélica” as “Protestant”. I would soon learn, however, that the term “evangélica” had much deeper meaning and held quite different connotations to my host mother and to many other Latin Americans than my definition of Protestant did to me.

I quickly began to grasp the difference between the Latin American understanding of “evangélica” and my own understanding of Protestant as I regularly attended church with my family. On my third Sunday, a revival preacher visited the church. He stood before the lively crowd, yelling in people’s faces, “más, más, más!” as they commenced falling to the ground unconscious, where fellow worshippers quickly covered them in blankets. These experiences, combined with dozens of similar ones, caught my attention and inspired my research on the Pentecostal Church’s immense growth throughout Latin America. I soon found out that when she said “evangélica”, my host mother was not referring to being Baptist, Methodist, Presbyterian or most other mainstream Protestant denominations. She was referring instead to the relatively new wave of charismatic and Pentecostal churches that have begun to overtake Latin America at an astounding rate.

Ultimately, this paper seeks to explain why the Pentecostal Church experienced such a huge and unprecedented growth spurt in the mid to late twentieth century. Why was Pentecostalism so attractive to Latin Americans at this specific point in history? In order to answer this question, I first looked at general characteristics of the Pentecostal Church. What are the major beliefs, practices, and theological stances of Pentecostalism throughout history and in present-day Latin America? Although Pentecostal churches vary greatly from one another, I was able to deduce several overarching qualities that can be applied to the vast majority of Pentecostal congregations: an emphasis on the power of the Holy Spirit in each believer’s life; speaking in tongues; healing; and the importance of missions and community outreach. This is by no means a comprehensive list, and there are certainly Pentecostal churches that cannot be described by the above characteristics. For the purpose of this paper, however, it is important to define what exactly is meant by Pentecostal, and the previously mentioned aspects appear to be the most inclusive.

Having defined the term “Pentecostal”, I explored various social and political stimuli for the movement’s growth, keeping in mind the political turmoil that existed in the latter half of the twentieth century. Interestingly, some of the countries that experienced the most political upheaval at this time (e.g. Chile, Guatemala, Brazil) simultaneously experienced substantial Pentecostal growth. Two important social stimuli for Pentecostal growth have been the rights of both women and minority groups. Pentecostal congregations generally believe in the priesthood of all believers, or the idea that each believer, regardless of gender or social status, is equipped by the Holy Spirit to perform a vital role within the church.

The Pentecostal Church has served both as a haven to people who have been given a low position in society, as well as a place of empowerment. In general, Pentecostalism has given the masses a socially-acceptable mode of empowering themselves to live in their own society, while at the same time maintaining a degree of separation from it. In societies where the governments and social structures have been feeble, unpredictable, and in many cases threatening, Pentecostalism has given marginalized peoples a culturally-appropriate place of refuge and a vehicle for change. It has similarly served to forge new identities for individuals who have felt a sense of “social anomie” (a sense that one does not belong, or a lack of clear identity). As Latin Americans have been forced to relocate to urban areas or even the United States in search of work and/or safety, they have struggled to establish new identities. Pentecostalism has been the solution for many Latin Americans in these types of situations, giving them a community in which to be involved.

One of the last, and perhaps most important, questions that I address involves the Catholic Church. If it has held such a strong religious monopoly since the arrival of the Spanish in the sixteenth century, then why is its growth suddenly paling in comparison to that of the Pentecostal Church? The general conclusion to which my research led me was that the Catholic Church simply did not change to meet the needs of a changing society. While Pentecostal congregations consistently adhered to the priesthood of all believers, Catholic churches continued to be hindered by their hierarchical structures, which provided
the marginalized peoples no refuge from their daily struggles. The Pentecostal Church has placed considerably more power in the hands of the lay people and less in the hands of one authority figure. It is also known for its worship styles, being more oral and spontaneous in nature than the liturgies and formulaic prayers of many Catholic churches. Ultimately, Pentecostalism has molded itself to fit society, while at the same time offering a shelter from society's ills.

(For full text, go to http://inquiry.uark.edu/)

Mentor Comments
Stephen Striffler’s are brief and to the point. What more needs to be said when one’s student has received a Teaching Academy Undergraduate Research Award?

Allison Covington’s thesis is exceptional. Few anthropology majors are able to engage in original research in other countries. Allison’s project is rooted in her own research and experiences with Pentecostalism in Costa Rica. This not only provides for an engaging narrative, but a point of departure for a review of the scholarly literature on religion in Latin America that is thorough, analytically sharp, and wonderfully informed by her own research. The result is a remarkably nuanced and sophisticated understanding of the rapid rise of Pentecostalism in Latin America.
OUTCOMES OF THE TRANSITION PROCESS IN CENTRAL AND EASTERN EUROPE: THE ROLES OF CULTURE AND SOCIETY IN ADOPTING DEMOCRATIC CAPITALISM

By Brian Lee
Department of Finance

Faculty Mentor: David E. R. Gay
Department of Finance

Abstract

In this paper, the author explores the reasons why some states have achieved higher levels of progress in transitioning from a communist system to a system rooted in democratic capitalism. Unlike the majority of scholars, though, he does not fault any one government’s policies or the reform path chosen for a country’s success or supposed failure along the way. Instead, the author concludes that the outcome of the transition process is dependent upon the interaction between the new formal institutions being adopted and the prevailing informal societal institutions and rules found throughout the region. If the formal institutions are in harmony with the informal institutions, then transition should occur relatively smoothly. Transaction costs will be lower, and individuals will be encouraged to engage in economic activity, producing wealth.

Using a panel growth regression and a data sample drawn from 20 transition countries in Central and Eastern Europe over the period 1990 – 2005, the author empirically tests the validity of this hypothesis. In order to measure whether the interaction between the new formal institutions and the informal institutions has a significant impact on economic growth in each sample country, he incorporates each state’s score for overall economic freedom from the Heritage Foundation’s 2006 Index of Economic Freedom, as well as its score for political rights and civil liberties from Freedom House. The more in tune a country’s informal rules are with the formal rules being adopted — those of a market economy and pluralistic system of government — the better the scores it is awarded from each organization.

The results provide evidence in support of the author’s hypothesis. Even though the impact of political rights and civil liberties is inconclusive, the regression estimation does show that the level of economic freedom enjoyed by the citizens of a country does significantly affect levels of GDP, meaning that greater levels of success along the path towards democracy and a market-based economy in Central and Eastern Europe can be attributed to each country’s informal institutions. The more these are in tune with the formal rules being adopted, the quicker citizens in a country are able to adapt to the new institutions, allowing for a democratic, capitalist system to take root and flourish.

I. Introduction

The end of Communism in Central and Eastern Europe was a watershed event of the late 20th century. The authoritarian regimes of the region had succumbed to domestic pressures and relinquished their monopoly on power, and nearly overnight approximately 300 million Europeans were presented with the opportunity to choose their political and economic systems anew. Where should they begin? What was important? Where should they look for guidance and assistance? The citizens of each country had to answer these simple, yet crucial questions. In response, academics of all disciplines and convictions, governments, and international organizations attempted to provide guidance and advice on the steps of the reform process.

II. Success in the Transition Process – The Interaction Thesis

A small group of economists comprised mainly of Enrico Colombatto, Jan Svejnar, Jan Winnicki, and Svetozar Pejovich, attributes the success of the transition process to whether or not the new market system, with its new and different rules and values, is adaptable to the culture and history of the country in question. As Svetozar Pejovich (2003) states:

Transition means institutional restructuring. Since formal rules are not a policy variable, transition has to mean the enactment of new formal rules: that is, constitutions, statutes, common law precedents, and/or governmental regulations. The results of transition then depend on the interaction of new formal and prevailing informal rules. Of course, the rules do not interact. Individuals do. New formal rules create new incentives and opportunities for human interactions. How individuals react to these new opportunities for exchange depends on how they perceive them. And how individuals perceive new opportunities depend on the prevailing culture.

In other words, the issue on which the success of the transition process actually hinges is not the technical aspects of reform or the numerous policies, but rather the transactions costs created by these reforms, costs which are defined as the cost incurred by an individual when engaging in economic activity. Transaction costs are indirectly related to the compatibility of these transition policies and the new market system with the history of the numerous populations in Central and Eastern Europe and the customs, traditions, and culture that grew out of this history. If the new formal institutions are in tune with the informal institutions, then the transition
should occur relatively smoothly because the transaction costs of conducting business are low and individuals will find it worthwhile to engage in economic activity and produce wealth. If the new formal institutions are not in tune with the informal institutions, then individuals will find that transactions costs are higher than before, which discourages them from conducting business and producing wealth. Therefore the question needing to be answered is whether the formal institutions being developed by the transition process and the values and behavior necessary for the proper functioning of these formal institutions agree with the informal institutions already in place in these countries (Colombatto and Macey 1999, Pejovich 1997, 2001, 2003 and Winiecki 2000, 2004).

A. The Interaction Thesis

Svetozar Pejovich (1997) terms this phenomenon the interaction thesis, which states that, if the formal institutions are compatible with the informal institutions, then the incentives they create will tend to reinforce one another. As Pejovich (2003) later summarizes:

When changes in formal rules are in harmony with the prevailing informal rules, the incentives they create will tend to reduce transaction costs and free some resources for the production of wealth. When new formal rules conflict with the prevailing informal rules, the incentive they create will raise transaction costs and reduce the production of wealth in the community.

The interaction thesis describes the scenario that had been unfolding in Central and Eastern Europe for years. Following the events of 1989 and 1990, the new governments in this region undertook the process of building capitalism, with the guidance and strong support from the West. Almost overnight the formal institutions that had been in place during the communist era had been deemed obsolete. In their place a new set of rules and institutions based on those found in Western societies was installed, forming the new framework by which the people in the region would conduct business and interact amongst themselves and with the rest of the world.

B. Outcomes of the Transition Process if the Formal & Informal Institutions are Incompatible

As Colombatto (2001, 284) points out, “...in the modern world, freedom is a secondary need. Security comes first.” In the case of Central and Eastern Europe, such circumstances created the opportunity for the new system to be hijacked and run aground by those who wanted to use the political machinery for personal gain. This allowed members of the former nomenklatura and other rent-seekers to gain control of the new economic system and led to the development of a large informal economy or black market in some countries. Even though the emergence of practices such as the oligarchs in Russia or a similar business class in the Ukraine did little to promote the public welfare or the success of the transition process, the majority of society did not protest against such behavior because it provided a sense of security, which arose from being able to engage in transactions much in the same manner as they had for the past decades and even centuries. These transactions were not based on the type of trust that permitted the creation of an extended order in society, which was needed if the new formal rules were to produce the desired outcome. Instead, these economic transactions were based on one’s reputation and the personal relationships found in the close-knit communities of Central and Eastern Europe (Colombatto 2001 and Pejovich 1997, 2003).

The conflict between the new formal institutions and the region’s informal institutions also led to the election of leaders who were able to undermine the reform process. Such leaders played on the fears of the populace, promising to provide security at a time when no one was certain as to the course economic and political reform would take. They were subsequently elected, but instead of promoting the aims of the transition process, these leaders sabotaged the entire process, effectively securing a monopoly on power for themselves and in doing so taking the newly-won political and civil freedoms away from the people. Vladimir Putin in Russia and Aleksandr Lukashenka are two examples of such leaders. This process was able to occur because only in a stable society, where people feel secure, can the exercise of freedom or personal economic interests become relevant. Until such a point has been reached, though, the desire for order and comfort takes precedence over any desire for freedom (Colombatto 2001 and Woodruff 2005).

C. The Interaction Thesis and Central and Eastern Europe – A Summary

The success of the transition process is not dependent on the technical steps taken during the reform process, rather on the interaction between the new formal institutions that are being transplanted and the informal institutions already present in the countries of Central and Eastern Europe. While certain steps must be taken if the transition process is to result in a market-based economy and a state that is democratic in nature and upholds the rule of law, these steps have only a marginal impact on the eventual outcome of the reform process. They are merely steps in the reform process, necessary in that they are required to complete the transition from a socialist economy to a capitalistic one. However, these steps can take varying forms and the end result will still be the same. For example, one can look at the process of transforming an economy as similar to remodeling a house. Someone can desire to remodel the house and transform it from its original state for various reasons, and through this process, the house can gain an additional pair of new rooms or the house can be altered so that it receives a different outward appearance. Such changes are usually carried out for aesthetic, not practical reasons. Either the owner had grown tired of the outward appearance of the house or he believed that an addition was needed in order to allow for more occupants, etc. These changes are aesthetic,
in that the foundation of the house has not been altered; only its outward appearance has evolved. It is much like the process of a newly-elected government coming to power in Western Europe or the United States. The new government does not typically alter the political or economic institutions of the country; these institutions remain in place, but this does not necessarily mean that the government retains all of the policies of the previous government. It is allowed to pass laws and implement policy as it sees fit, according to its party doctrine and the promises supposedly made over the course of the campaign. The process of changing power within a system of government possesses many similarities with the process of remodeling a house after a new owner moves in. In both instances, the foundation of both the house and the system remains the same, only the appearance changes, in order to incorporate the preferences or policies of the new government or in this example, the preferences of the new owner. Such a process occurs regularly in Western societies. Power is exchanged peacefully within a system of government, and with this exchange of power, policies that have varying aims are implemented, but this does not mean that the foundation which the entire system is built upon has been altered. These governments must work within the institutional foundation of each country, a foundation that in the case of a Western society means the procurement of civil liberties as well as the property rights system and legal system necessary for a market economy. For each country that adheres to such a system, the content and aims of its policies may differ due to the preferences of its populace, but the foundation on which the entire system rests remains unaltered. There are instances, though, that require one to rebuild the foundation of the house, not for aesthetic purposes, but in order to ensure that the structure will continue to remain in existence. This could be due to various reasons: erosion, a crack in the foundation, etc. In these circumstances, one would be required to either repair the foundation of the house to a large extent or rebuild the foundation in its entirety. Following the events of 1989 and 1990, the countries of Central and Eastern Europe found themselves in such a situation. No longer would it prove adequate to simply change the exterior of its house or in this case, the policies and regulations of the state. The people demanded a change in the foundation of each country. No longer could the leaders of each country attempt to simply alter the communist system and simply improve the political environment and increase economic growth. Instead, the citizens of each country pushed for the foundation on which the state stood to be rebuilt. They insisted upon the right to political participation, a guarantee of civil liberties, and the right to make economic decisions based on their preferences, not those dictated to them from the Communist party and its apparatus. They demanded to live under a political and economic system similar to the system found in the Western world, but in order for such a system to be successfully transplanted the new formal rules would have to be in tune with the informal rules that were already present throughout the region.

Such a transition demanded more than simple adherence to a certain transition model or reform plan; it required the people of each country to be in a position to adapt to the new formal rules that were being exogenously “forced” upon them. The technical elements of the reform process did not determine the ultimate outcome; rather they served simply as a means to an end. Ultimately, the success of the transition process in each country depended upon whether the populace was able to adapt to the new formal institutions. If these new formal rules were successfully adopted then the foundation of the new market system would be able to stand firm, allowing for stable economic growth. If not successfully adopted, then the results would be different – economic growth would be volatile and the size of the informal economy would tend to increase as citizens began conducting business based on reputation and personal relationships, forgoing the voluntary exchanges with strangers that allow a capitalistic economy to operate optimally. In the case of Central and Eastern Europe, such results would derail the transition process as the people in each country reverted to methods of conducting economic and political affairs that might provide comfort and security, but were not compatible with the methods of a society supportive of a thriving market economy.

III. Support for the Interaction Thesis

The successful transition from a communist system to a democratic government and an economy based on the market was dependent upon the interaction between the new formal rules being adopted and the informal ones already in place in each country throughout Central and Eastern Europe (Colombatto and Macey 1999; Pejovich 1997, 2001, 2003; Winiecki 2000, 2004; and Zweynert and Goldschmidt 2006). Hayek (1988) maintains that exposure to the West and its notion of an extended order encourages the development of the individual and promotes the self-serving decision-making process that underlies a capitalist system. According to the interaction argument, countries with this exposure were able to put in place a system of credible and secure property rights with its supporting institutions of a constitution, an independent judiciary, and freedom of contracts. This historical connection allowed the new market economy to take root quickly and begin to thrive within a relatively short time period as the population of these countries felt comfortable to engage in voluntary exchanges in the marketplace, fueling economic growth. On the other hand, countries lacking prior contact with the West and its traditions and practices had a more difficult time adopting the formal institutions necessary for a properly functioning market economy. In these cases, the transition to a capitalistic economy was derailed at an early stage, allowing for either strong-armed rulers or rent-seekers to stall both the economic and political transition process in these countries.

A. Econometric Specifications

The empirical methodology used in this thesis is based on the estimation of growth regressions using panel data. The data sample covers the period 1990 - 2005 and includes the
following group of 20 transition countries: Albania, Belarus, Bosnia and Herzegovina, Bulgaria, Croatia, the Czech Republic, Estonia, Georgia, Hungary, Latvia, Lithuania, Macedonia, Moldova, Poland, Romania, Russia, Serbia and Montenegro, the Slovak Republic, Slovenia, and the Ukraine. Each of these countries is geographically located in Central, Eastern, or Southeastern Europe, and began the transition from a socialist economy to a market-based system in the early 1990s.

This sample utilized panel data, instead of time-series data. Unlike time-series data, panel data expand the size of the data sample by including more than one country, thus increasing the explanatory power of the regression. In doing so, the data sample increases substantially and the estimation results are therefore better able to explain the causes of the economic growth experienced by each country in the region. In working with the data sample, the author also used fixed effects techniques instead of random effects, allowing for the control of certain effects that remain constant over the chosen time period. By using random effects, it would have meant that there exist events or circumstances which cannot be controlled for and could therefore have a substantial impact on the economic growth of each country. The possibility that such events or circumstances could occur would alter the coefficient estimations and significantly decrease the explanatory power of the growth regression, and in turn damage the credibility of the empirical study. However, the author maintains that this is not the case for this data set, and thus utilizes fixed effects.

For this study, five-year running averages were used, instead of simple five year blocks. By incorporating five-year running averages into the estimations, the explanatory power of the growth regression is enhanced in two ways. First, five-year running averages, which cover time periods 1990-1994, 1991-1995, 1992-1996, etc., enable an increase in the explanatory power of the regression estimation by maximizing the number of observations. With five-year running averages, it is possible to include 15 five-year blocks, instead of three, which would be the case if the chosen time period were divided into three blocks. Second, averaging the data over five-year increments helps to eliminate the effects of short-run business cycle dynamics on growth. Failure to eliminate the effects of such short-run dynamics tends to lead to highly correlated time series and to gross overestimation of coefficients (Iradian 2007).

Following the example of Iradian (2007), Fidrmuc (2003), and Berg et al. (1999), the author uses the following growth regression model, hereafter called equation 1, to construct the econometric specifications:

$$g_{it} = \chi + \beta \hat{C}_{it} + \lambda X_{it} + \epsilon_{it}$$  \hspace{1cm} (1)

where $g_{it}$, the dependent variable, is the GDP growth rate in country $i$ during the time period $t$. $\hat{C}$ is composed of a set of control variables that are often used in growth literature, including the level of development as proxied by initial GDP per capita, inflation, the annual population growth rate, the total amount of government expenditure, school enrollment numbers, total investment, and a measurement of trade openness. $X$ is the vector of explanatory variables the author wishes to test. These are the variables (including level of economic freedom, level of political rights, and level of civil liberty) the author uses to test the validity of the interaction thesis, and $\epsilon_{it}$ is the specification error term.

**B. Explanation of Variables and Data Sources**

The empirical analysis uses data from a sample of 20 countries during the period 1990 – 2005. Values of annual population growth rate, real GDP per capita, annual GDP growth rate, secondary school enrollment rates, levels of investment, inflation, trade openness, and the share of government expenditures in GDP are taken from the World Bank’s World Development Indicators (2007). The values for overall economic freedom in each country are extracted from the Heritage Foundation’s 2006 Index of Economic Freedom, and the scores regarding the level of political rights and civil liberty for each country are from the Freedom House (2008).

The first group of variables is a basic set of explanatory variables commonly used in growth regression analysis. They are used to control for some core components that usually drive economic growth in a country. Real GDP per capita is used to control for prior levels of development. For example, in this analysis, each five-year running average includes the beginning year’s real GDP per capita level. If the five-year average corresponds to the period 1990 – 1994, then the real GDP per capita level from 1990 is used in order to see how the level of economic development in 1990 affected GDP growth over that five year period. Annual population growth rates take into account the size of the labor pool in each country. The secondary school enrollment rate measures the number of people in each country receiving a basic level of education and thus is a rough estimate regarding the skill level of the country’s workforce, while the variable measuring the total amount of government expenditure represents the extent to which a state’s fiscal policy aids or harms its economy.

By incorporating inflation into the regression, the author is controlling for overall levels of macroeconomic stability in each of the countries being tested. The level of investment is measured by the amount of capital available in each country and is used to show the amount of capital or funds earmarked for economic investment. Also included in this group of control variables is a measure of trade openness, since the level of openness in an economy can greatly affect the levels of economic growth experienced by that country (Azor and Vamvakidis 2004, and Kali et al. 2007). Trade openness is calculated as the total value of imports plus exports over total GDP (I+E/GDP). The final variable included in each five-year running average is the initial level of GDP growth. Even though GDP growth serves as the dependent variable, preceding levels of GDP growth are believed to directly affect later levels of economic growth. Thus, for each five-year
running average, the average GDP growth for the previous five-year running average was included.

The second group of variables, including economic freedom, political rights, and civil liberties, comprises the vector of explanatory variables the author wishes to test. These three variables have been chosen in order to test the validity of the interaction thesis. By incorporating these three variables into the regression specification, the author is able to observe how the process of democratization influences levels of GDP in the region. More importantly, he can also see whether the compatibility or incompatibility of the new formal institutions and the informal institutions in each country significantly affects and leads to economic growth. The first explanatory variable, economic freedom, is included in order to observe how the interaction between formal and informal institutions in each country affects economic growth over the chosen period. As mentioned earlier, the values for this variable are taken from the Heritage Foundation’s 2006 Index of Economic Freedom and are equal to the overall score for economic freedom of each country over the period 1990 – 2005. For each country considered, a score is given on a scale of 1 to 5, with lower scores translating into higher levels of economic freedom.

Due to its ability to measure the level of government participation in the economy, the author chose to incorporate the Heritage Foundation’s Index of Economic Freedom into the growth regression as a measure of how well a state has progressed along the path towards a market-based economy. The notion behind this decision is that the lower the score a state receives, the better that country has performed in its transition to a market-based economy. The new formal institutions and rules of a capitalist economy are in tune with the informal institutions already present in that country, thus, the new system has been successfully adopted by the citizens of that country. Now the question to be answered is whether the adoption of a market-based system leads to economic growth, and if so then to what extent. By including this measure of economic freedom as a variable in the growth regression, the author desires to answer this question, and see if there exists empirical support for the interaction thesis.

The final two variables included are political rights and civil liberties, and they are meant to measure the effect of democratization on economic growth in the region. Each of the values for these variables is extracted from the Freedom House’s annual publication that measures the level of political freedom, which it defines as “opportunity to act spontaneously in a variety of fields outside the control of government and other centers of potential domination” (Freedom House 2008), enjoyed by the citizens of each state throughout the world. The organization states that it “…does not maintain a culture-bound view of freedom. The methodology of the survey is grounded in basic standards of political rights and civil liberties, derived in large measure from relevant portions of the Universal Declaration of Human Rights” (Freedom House 2008).

These two measures of political freedom were included in order to determine whether the democratization process occurring in conjunction with the economic transformation throughout the region impacts economic growth. Did the transition to a democratic system of government affect the standard of living in Central and Eastern Europe? Or did the process of political liberalization have little effect on the lives of individuals in these countries, except that citizens are now empowered to choose their own leaders and are better able to hold those elected accountable for their actions once in power, and to do so without fear of punishment if one’s views digressed from that of the state apparatus? Inclusion of the two measures of political freedom in the growth regression provided a mechanism for answering these types of questions.

C. Empirical Estimation Results

As mentioned earlier, the author utilizes five-year running averages when conducting fixed effects estimations of equation. In order to check for robustness, three different specifications are run. Each includes the basic set of control variables represented by \( C \) in the equation, the difference among the three specifications being the inclusion of one of the three explanatory variables from vector \( X \). The first specification includes only the variable measuring civil liberties; the second specification incorporates only the value that quantifies the level of political rights; and the final specification includes only the variable representing economic freedom. A specification incorporating all three of the explanatory variables the author wished to test is not possible due to the high levels of correlation existing between the variables. Thus, the author is only able to examine each of these variables individually. Additionally, the results presented here have been corrected for heteroskedasticity using White’s correction method.

Table 1 presents the results of estimating equation (1) for each specification. Column 1 presents the estimation results for the first specification, which measures the impact of civil liberties on economic growth. The results for the second specification, which includes the political rights variable, are shown in column 2, and column 3 shows the results for the third and final specification, which calculates the effect of economic freedom on economic growth for the 20 transition countries included in study. In addition, located at the bottom of each column is the number of observations and the adjusted R-squared for the corresponding specification.

The results of the three specifications provide an interesting picture regarding the effects of democratization and economic liberalization on economic growth in the transition countries of Central and Eastern Europe. As one can see, the first two specifications generate similar values for the coefficients of each control variable. For each of the control variables, the impact on the GDP growth rate in each country is as expected. The initial levels of GDP growth significantly affect later levels of economic growth in a positive manner, meaning that the higher the level of economic growth at
the beginning of the five-year average, the higher the levels of growth a country experiences over that five year period. Trade openness leads to increases in GDP as does the level of investment in each country. As one can see, macroeconomic

Table 1. Dependent Variable: GDP Growth Rate for each Country

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial GDP Growth</td>
<td>0.511482*</td>
<td>0.568297*</td>
<td>0.568297*</td>
</tr>
<tr>
<td>GDP per capita</td>
<td>-2.307338*</td>
<td>-2.471234*</td>
<td>-2.304492*</td>
</tr>
<tr>
<td>Trade Openness</td>
<td>1.253509**</td>
<td>1.267780**</td>
<td>2.011660**</td>
</tr>
<tr>
<td>Government Expenditure</td>
<td>0.063596</td>
<td>0.028269</td>
<td>-0.143586**</td>
</tr>
<tr>
<td>Investment</td>
<td>7.916759*</td>
<td>7.756691*</td>
<td>8.253119**</td>
</tr>
<tr>
<td>Inflation</td>
<td>-0.002446*</td>
<td>-0.002305*</td>
<td>-0.006991*</td>
</tr>
<tr>
<td>Population/Growth</td>
<td>0.020333</td>
<td>0.213664*</td>
<td>-0.079532*</td>
</tr>
<tr>
<td>Liberal Rights</td>
<td>-0.578404 (1.488496)</td>
<td>-0.291137 (0.857902)</td>
<td></td>
</tr>
<tr>
<td>Economic Freedom</td>
<td></td>
<td></td>
<td>-2.099091**</td>
</tr>
<tr>
<td>No. Observations</td>
<td>152</td>
<td>152</td>
<td>104</td>
</tr>
<tr>
<td>Adj. R-squared</td>
<td>0.956754</td>
<td>0.937094</td>
<td>0.946245</td>
</tr>
</tbody>
</table>

Notes: The symbols * and ** mean that the estimated coefficients are significantly different than zero at the 1 and 5 percent confidence level, respectively; t-statistics are in parentheses.

stability, which is measured by inflation, is necessary in achieving economic growth. GDP per capita, as expected, possesses a negative effect on economic growth, meaning that higher levels of initial GDP or economic development tend to slow economic expansion.

Population growth proves to be insignificant in determining levels of output, but this phenomenon could arguably be attributed to the fact that the population levels in each of these countries did not significantly increase in the period 1990 – 2005. In the first two specifications, overall economic growth is not significantly affected by the total amount of government expenditure in each country, and the effects of school enrollment levels on GDP growth cannot be included in either of the specifications due to the lack of observations available.

In both of these specifications, the level of civil liberties or political rights enjoyed by citizens of a country in the region proves to be insignificant in determining economic growth, although the coefficients for these two variables possess different values. According to results for the first specification, higher levels of civil liberties lead to greater economic growth, while greater political rights adversely affect economic output. However, despite this seemingly puzzling outcome, any further analysis of these two variables is irrelevant for two reasons. First, the estimated coefficients for both of these variables do not prove to be significant at either the 1 or 5 percent level, and for the purposes of this thesis the question as to whether the process of democratization significantly impacts economic growth has been answered. Neither the level of civil liberties nor the level of political rights in each country appears to have influenced economic growth during the period 1990 – 2005. In the third specification, the estimated coefficients of the control variables do not alter significantly and the Adjusted R-squared increases only slightly from 0.936754 and 0.937094 to 0.964658, respectively. The only noticeable difference among the group of control variables is the finding that the total amount of government expenditures becomes significant at the 5 percent level. However, the estimated coefficient for economic freedom is significant at the 5 percent level. This means that, in contrast to the level of civil liberties or political rights, the amount of economic freedom enjoyed by the citizens of a country does impact economic growth, and as estimated by the regression specification it possesses an overall positive effect on GDP. In other words, if a country increases its overall score for economic freedom by one point, for example from 3.2 to 2.2, then that country should see its overall economic output increase by more than 2 percent. Such a result shows not only that greater levels of economic freedom lead to higher rates of GDP growth, but in doing so it also provides evidence in support of the interaction thesis, since the higher the score a country receives from the Heritage Foundation for overall economic freedom, the greater that country's economy operates according to market principles. Therefore, one can conclude that the culture and history, as well as the societal makeup of a country, do matter in determining the rate at which a system of democratic capitalism is adopted. The more in tune a people's culture and their corresponding informal rules are with the new formal institutions, the greater the level of economic growth a country can expect to experience.

D. Implications of the Growth Regression Results for the Transition Process in Central and Eastern Europe

Based on the results of the panel growth regression described in the previous sections, two points warrant further discussion. First, it was estimated that two factors, civil liberties and political rights, did not significantly increase economic growth for the region of Central and Eastern Europe from 1990 through 2005. Proponents of the Chinese model or the pluralistic market socialist transition model could utilize these results to support their argument that the former Communist countries in Europe should not have introduced reforms to transform the state politically, especially as these countries were transitioning from an economy run by the Communist party to one where the market was permitted to dictate the amount and types of goods to be produced. According to them, it would have been more prudent to follow a path of transition similar to that of China and other Asian countries or to implement an economic reform plan similar to the “reform socialism” pushed by the Communist governments in Europe at various times following World War II. The most notable examples being Gorbachev's policies of perestroika and glasnost. In both of these scenarios, a single political party or entity maintains a tight control over judicial affairs while supposedly liberalizing the economy in phases over an extended period of time.

However, in light of this critique regarding the transition process occurring in Central and Eastern Europe, the author...
reiterates that such arguments are irrelevant and even erroneous for two distinct reasons. In both of the cases mentioned in the preceding paragraph, complete economic liberalization is usually not the main goal of the government in power. It is only lip-service to the outside world. The commanding heights of the economy remain in the hands of those holding political power, and the individual is still only allowed a limited role in determining economic outcomes in a market that remains handicapped by the self-interests and overbearing presence of the state or those in power. In other words, the reforms implemented only go so far. The entire economic system will never be transformed completely, because such a transition would undermine the interests of the political party or individuals who possess complete control over the political machinery within the country.

In addition, proponents of this and similar views fail to take into account the circumstances and sentiments surrounding the events of 1989 and 1990 in the Eastern bloc. According to the regression results, one can state that civil liberties and political rights did not significantly impact economic growth during the first 15 years of transition, yet from the same results one can also infer that these two variables did not adversely affect economic growth either (Fidrmuc 2003). The people living in the Eastern bloc desired both a standard of living similar to that of their neighbors in the West as well as a system of government in which they could elect their leaders and openly hold views that differed from those of the state. If the Communist party was to relinquish its control over the commanding heights of the economy, then it would also have to give up the reins of political power since the two systems were tightly intermeshed. The term “centrally-planned economy” did not simply describe the method by which the economy was managed, because the economy was an inseparable part of an authoritarian state where the party, governmental, and economic hierarchies were tightly intermeshed, with the party and its apparatus possessing the dominant voice among the three entities. Reforms were therefore inevitably blocked at the point where any further change in the system of economic management might undermine the interests of the party (Kozul-Wright and Rayment 1997). Hence, economic reform could not occur without political reform and vice-versa; the two processes of reform had to occur simultaneously if either one of them was to prove successful.

The second point worth additional commentary is the role a people and their informal institutions and rules played and continue to play in determining the outcome of the transition process in Central and Eastern Europe. The growth regression results show that a country’s level of economic freedom has a significant impact on economic growth. Despite this result, though, many economists would agree with Åstrund (2002, 75) when he states that a country’s culture, history, and religion had only minimal impact on economic growth in Central and Eastern Europe from 1990-2005, especially in comparison to a country’s economic policy, geography, and levels of investment and technology. However, this author strongly disagrees. It is true that current levels of development, investment and education, as well as previous levels of GDP, have a large effect on future rates of growth. However, in order to maintain satisfactory levels of investment and economic development, a country has to be able to create an environment that encourages individuals – foreign or domestic – to invest in its economy and engage in activity within its borders that creates wealth and leads to economic growth. If such an environment requires certain rules and standards that most politicians and economists agree are necessary for investment and economic growth, then why do some countries still lag behind in the transition process? Why are some transition countries unable to effectively implement these rules and standards, while other states have been able to do so and are now experiencing impressive levels of economic expansion?

As discussed by Friedrich von Hayek (1988) and others, certain cultures and societal orders are more conducive to a market-based economy and a political order rooted in the notions of democracy and individual liberties. These societies promote self-interest, self-responsibility, and self-determination; values that reinforce the formal institutions – secure property rights, a constitution, an independent judiciary, and freedom of contracts – underlying a capitalistic system (Colombatto and Macey 1997, Hayek 1988, and Pejovich 1995, 1997, 2003). If a people have had previous contact with a societal order that nurtures those or similar values, then they will be more inclined to successfully adopt a system of government and economics that are based in these formal institutions and rules. As a result, transaction costs will be minimal, and individuals will be encouraged to engage in economic activity, which in turn creates economic output and wealth. If not, transaction costs are high, leading to situations in which people do not trust the new system and revert to what is considered “comfortable” or secure. Special interest groups gain control of the political and economic system and encourage rent-seeking actions and other behavior that undermines the transition process. As a result, the new system does not take root and a political and economic doldrums emerges. Such is the case in Central and Eastern Europe. Countries having prior contact and interaction with the West and its values have arguably fared better in the transition process, while countries lacking such contact have not achieved the same levels of progress as some of their neighbors.

V. Conclusions

The events of 1989 and 1990 opened a new chapter for Central and Eastern Europe. Almost overnight, Communism and its ideals lost favor throughout the region, causing communist parties to relinquish their control over society and a power vacuum to emerge. Elections were held shortly thereafter, and although there were varying opinions as to the new direction each state should take, the majority of citizens throughout the region desired a government modeled on the
democratic values of the West and an economic system based on the markets, allowing individuals to make consumption decisions based on their preferences and for goods to be allocated according to the law of supply and demand. These preferences represented a clear break from the past. No longer would the individual be forced to look to the state for guidance in every aspect of one’s life. Instead, the people would be able to control the political and economic destiny of their countries.

The past two decades have witnessed the new governments in Central and Eastern Europe attempt to steer their countries down the long road to political and economic transition, reaching varying levels of success along the way. Some states have been successful in adopting a democratic, pluralistic system of government and a market-based economy, while other countries have been unable to achieve similar results. Why is this? Why have some countries succeeded and others lagged behind?

Numerous scholars have attempted to answer this question over the past 18 years, but they have yet to offer an explanation that is applicable to every country undergoing political and economic transformation in the region. Among other issues, many of these scholars have concentrated on whether or not each state should have followed a certain transition model when implementing reform measures, or whether enough time was spent on institution building at the beginning of the reform process. The question remains, why do none of these explanations provide convincing evidence as to why certain countries have failed and others have succeeded to varying degrees?

This author argues that a core issue is the lack of attention given to the interaction of the new formal institutions these countries were attempting to adopt and the informal institutions already present. Contrary to the presumptions of these other scholars, history and culture did play a pivotal role in determining whether a country succeeded in adopting the new formal rules of democratic capitalism. The West is associated with the notion of an extended order society, one that encourages the development of the individual and the decision-making process that underlies a capitalist system. Countries that had prior contact with the West and its notion of an extended order society were able to adopt the new formal institutions within a relatively short period of time. As a result, transaction costs decreased in these states, encouraging individuals to engage in the type of voluntary exchanges in the marketplace that fuel economic growth.

On the other hand, countries lacking such a cultural exchange with the West have tended to lag behind their neighbors in the transition process, unable to adopt certain aspects of either a democratic system or a market economy or both. Progress in these states has occurred more slowly, as individuals have taken longer to adjust to the reforms implemented since 1989 and 1990, but such delayed progress should not be interpreted as a sign that these countries will never succeed in their quest to shed the legacy of the communist era. In every transition process, time plays a role.

As evidenced by the improved scores for overall levels of economic freedom awarded by the Heritage Foundation, it is possibly only a matter of time before certain states catch up with their neighbors in adopting the institutions and rules necessary for a properly functioning market economy and a political order rooted in the principles of democracy and pluralism.

Acknowledgments

I thank the efforts of many faculty members at UA for their support and feedback. These include Donna Daniels, David Gay, Hoyt Purvis, Javier Reyes, and Murray Smart. Additionally, this process was helped by the paper presentation in the Economic Scholars Program at the Federal Reserve Bank of Dallas and Austin College, in March.

Bibliography


Wolf, Martin. “Avoiding the Trap of Transition: The former Communist Countries that have done best have ensured that goods can be bought more freely than Judges.” Financial Times, October 11, 2000. London edition (p. 25).


**Mentor Comments**

David Gay clarifies the scope of the research project from which this article was drawn. He leaves no doubt about the excellence of the work and the reason that the thesis received an Undergraduate Research Award from the Teaching Academy.

This paper is an edited and condensed version of Brian Lee’s honors thesis which is an exceptionally well-developed volume of nearly 150 pages. The condensed version of the thesis was presented at the Economics Scholars Program at the Federal Reserve Bank of Dallas in the spring of 2008, after a competitive selection process involving submissions from students at Texas A&M University, Michigan State University, Southern Methodist University, the University of Texas at Austin, and others.

Brian and I met weekly for more than two semesters as he crafted the outline and details of his senior honor’s thesis. The topic and approach were primarily worked out by Brian Lee. Knowing how well he nuanced answers on exams from our three courses together, it was clear that his literature review and focus using his majors in economics, German, international relations, and European studies would be an epic undertaking. We agreed that a crucial element to take his thesis beyond an exemplary review of the works of others was to employ empirical testing of the interaction hypothesis and other hypotheses. He was aided by Donna Daniels in the University Library, along with Professor Javier Reyes, who teaches econometrics.
Comparing the usual measures of economic freedom in the foreground of economic development during the 1990-2005 period yielded results that put economic freedom in the foreground of economic development. The Freedom House index did not yield the explanatory power of the Heritage Foundation index of economic freedom. In addition to this publication in Inquiry, I expect that a trimmed down version of this paper will be published in a high quality economics journal under his sole authorship.

Brian's thesis was the best that I've seen in 35 years. I have seen many great ones, but this has spoiled me (most likely, forever). The same theme was echoed by Professor Hoyt Purvis in international relations and University Professor Murray Smart ("better than many doctoral theses"). The topic was compatible with my own interests, but Brian Lee did the yeoman's dogged work to organize, to compare and contrast positions and hypotheses, and to dig out the data to push his honor's thesis beyond the realm of being a great thesis into being an extraordinary one. I do not expect to see the likes of this quality again during what I hope are my next ten years at the university. It has been and likely will remain a highlight of a long academic career.

Brian Lee has received a Fulbright to study in Vienna, Austria. He will enroll in the Master of Advanced International Studies program at the Vienna School of Diplomacy. Under different circumstances, he would most likely have qualified for a Marshall Fellowship but his interests and experiences (both in Austria and as an intern with the German Bundestag) led him elsewhere.
PROCESSING SURPRISE TENSION IN TONAL MELODIES

By David Sears
Department of Music

Faculty Mentor: Elizabeth Margulis
Department of Music

Abstract

Expectation plays a vital role in understanding how we experience music. Processing music requires listeners to form expectations about upcoming events as they unfold. The formation and violation of these expectations has been empirically tested in harmonic priming paradigms, in harmonic priming during melodies, and in more general divided attention models. Recent research has also suggested that passive exposure to music leads to the development of implicit knowledge about musical expectations for untrained as well as trained listeners. Expectation has even been cited by the theorist Leonard Meyer as “the prime generator of musical affect” (qtd. in Margulis, 2005). Expectation therefore not only plays a role in how listeners process melodies, but might also provide evidence as to why musical events produce affective response in listeners. Unfortunately, much of the psychological research concerning expectation in music has appropriated a commonplace definition of the term “expectation,” and in fact fails to address the difficulties in rendering the term more accessible to empirical scrutiny. Such a decision most likely reflects a lack of awareness of the dynamic nature of listener expectations. Is it possible to consider the subtle and more nuanced variations of expectation, such as events that invoke a sense of surprise, yearning, or foreboding?

The present inquiry considers the empirical study of surprise during the experience of music, an approach informed particularly by Elizabeth Margulis’ model of melodic expectation. Margulis’ model represents an ideal theory by which to predict and quantify moments of surprise as they are defined in the model, and the priming paradigm provides a method by which to empirically test the model’s predictions. Until now, few studies outside of brain research have yet successfully produced a method to permit the study of musical events experienced in “real-time,” and at a tempo adequately fast enough to represent real music. The primary goal of this study is to provide an empirical method to study just those sorts of violations of expectation—automatic, implicit expectations that build up in a musical piece at the local rather than global level. The results of this study have suggested that the psychological categories of melodic perception provided by the Margulis model—stability, proximity, and direction—may play an important role in determining how listeners process tonal melodies, as well as provide an explanation as to why certain moments elicit particular affective responses like surprise in listeners.

Introduction

Expectation plays a vital role in understanding how we experience music. Processing music requires listeners to form expectations about upcoming events as they unfold. The formation and violation of these expectations has been empirically tested in harmonic priming paradigms (Bharucha & Stoeckig, 1986; Krumhansl, 1990), in harmonic priming during melodies (Bigand & Madurell, 2005; Loui & Wessel, 2007; Margulis & Levine, 2006; Schmuckler, 1997), and in more general divided attention models (Berent & Perfetti, 1993; Bigand & McAdams, 2000). Recent research has also suggested that passive exposure to music leads to the development of implicit knowledge about musical expectations for untrained as well as trained listeners (Loui & Wessel, 2007; Margulis, 2006). Expectation has even been cited by the theorist Leonard Meyer as “the prime generator of musical affect” (qtd. in Margulis, 2005). Therefore, it not only plays a role in how listeners process melodies, but might also provide evidence as to which events produce affective response in listeners.

To study the psychological effect of expectations in music, researchers have primarily relied on the priming paradigm. The priming paradigm theorizes that the processing of incoming events is affected by the context in which they appear. Related events are primed, thus facilitating processing. The reaction time (RT) method provides a method by which to test these explanations. It is founded on the principle that, during the perception of highly unexpected events, listeners have to reevaluate and revise their mental representations of past events. This operation is assumed to demand more attentional resources than successfully embedding a highly expected event within the current representation. As a result, listeners are slower to react to competing secondary stimuli that occur simultaneously. Response time studies have repeatedly shown that participants are slower to respond to tonally unrelated chords and highly unexpected melodic events (Loui & Wessel, 2007; Schmuckler, 1997).

Early research employing the priming paradigm privileged harmony as the primary indicator of expectancy in music (Bharucha, 1986; Krumhansl, 1990). However, recent studies have begun to investigate other potential indicators of expectancy. A divided attention model permits theorists to conceive of music as an exercise in divided attention among several different musical attributes where harmony, melody, rhythm, and timbre are allocated cognitive resources in the formation of a unified musical percept (Loui & Wessel, 2007). Newer reaction time studies have therefore approached...
expectations from the standpoint of melody (Sehmuckler, 1997), timbre (Margulis & Levine, 2006), and voice leading and polyphony (Bigand, 2000 & 2005).

Unfortunately, much of the psychological research concerning expectancy in music has appropriated a commonplace definition of the term “expectation,” and in fact fails to address the difficulties in rendering the term more accessible to empirical scrutiny. Such a decision most likely reflects a lack of awareness of the dynamic nature of listener expectations. Is it possible to consider the subtle and more nuanced variations of expectation, such as events that invoke a sense of surprise, yearning, or foreboding?

In “Surprise and Listening Ahead: Analytic Engagements with Musical Tendencies,” Elizabeth Margulis (2007) considers both musical analyses and empirical research in order to draw attention to the various uses of the term “expectation,” as well as the dangers of being unaware of how analysts and psychologists might be appropriating (or misappropriating) the term in their own research. Margulis proposed a model of melodic expectation derived from Gestalt psychology and the work of music theorists Fred Lerdahl and Eugene Narmour. The Margulis model attempts to separate the experiences stemming from expectations into three types in order to “make a preliminary step toward a richer taxonomy of the multiple dimensions of musical experience” (2005, p. 697).

The present inquiry considers the empirical study of music from a more nuanced view of listener expectations, an approach informed particularly by music theory and analysis. However, an approach that grounds the psychology of expectation in the theoretical study of music is not entirely unique to the body of research concerned with expectation, as attempts have already been made in studies concerned with the perceptual principles governing voice leading (Huron, 2001; Poulin-Charronnat & Bigand, 2005) and consonance and dissonance (Bigand & McAdams, 2000; Loui & Wessel, 2007).

Until now, few studies outside of brain research have successfully produced a method to permit the study of musical events experienced in “real-time” and at a tempo adequately fast enough to represent real music. In 2007 psychologists Loui and Wessel tested listener expectations by having subjects respond with contour information about the melody at the note-to-note level (whether the note event was higher or lower than the event before it). However, at an interval of 1600 msec between note events, the tempo was far too slow to permit subjects to store more than a note or two of the melody in short term memory, nor did that tempo adequately represent “real” music.

The flaw in employing a contour task is that subjects simply cannot accurately respond to the contour of melodies at faster tempos, especially considering that note-to-note events within melodies primarily move by step, which makes a contour decision particularly difficult. In 1986 psychologists Bharucha and Stoeckig employed an intonation task to provide data about the expectedness of harmonic events; subjects had to detect whether the melodic event was in tune or out of tune. Though this task was generally much easier both for trained and untrained listeners, the study was concerned specifically with the relatedness of a prime harmonic event to a target harmonic event. Listeners therefore only heard two chords contiguously. Margulis’ model of melodic expectation theorizes, however, that listener expectations build up over several melodic events at the local level, and that sudden violations of expectation produce a surprise response in listeners that is both automatic and largely unconscious. The primary goal of this study is to provide an empirical method to study just those sorts of violations of expectation—automatic, implicit expectations that build up in a musical piece at the local rather than global level. Secondly, this study will attempt to provide empirical support for the Margulis model of melodic expectation.

A RT method employing an intonation task should provide access to the symptoms that characterize surprise in tonal melodies as well as empirical evidence supporting predictions about violations of expectation. Margulis’ model of melodic expectation will then offer access to the potential cause of those symptoms, as it represents an ideal theory by which to predict and quantify moments of surprise in tonal music. Her model provides quantifiable predictions about the note-to-note expectancies of tonal melodies. It weighs expectancies for surprise tension through a series of musical characteristics derived from Gestalt perceptual psychology. The model presents three factors as the primary determinants of the expectancy of a melodic event: stability, proximity and direction.

It has been theorized that listeners expect relatively more stable pitches than less stable pitches (Margulis, 2006). The model further theorizes that four categories of stability exist within a key. From maximally stable to minimally stable they are: the tonic; the third and fifth; other diatonic pitches; the remaining chromatic notes (which are deemed to be maximally unstable). Stability ratings can also change according to the tonal context of the events. For example, before a melody modulates from C to G, a G will receive a rating of 5 as the fifth of the tonic chord, while after the modulation, a G will receive a rating of 6 as the root of the new tonic chord.

Proximity indicates that listeners should expect pitches that are closer more than pitches that are farther away. Finally, direction indicates that listeners should expect that melodies will reverse direction after large intervals, while melodies will continue in the same direction if the intervals between each note are smaller. These three factors make up the basis by which the model determines note-to-note expectancies for melodies (Margulis, 2006).

This study specifically considers the notion of surprise in music. It is predicted that during periods of surprise in the chorale melodies, as indicated by the Margulis model of melodic expectation, subjects will be slower to react to a competing secondary stimulus that occurs simultaneously.
Method

Subjects
Twenty students from the University of Arkansas at Fayetteville participated in the study. All of the subjects volunteered in return for credit in undergraduate psychology courses. The mean age of the subjects was 20.4 years. The subjects averaged 2.6 years of musical training, and they also reported listening to an average of 16.8 hours of music per week.

Apparatus and Test Stimuli
The chorales and melodies were recorded by the researcher with an M-Audio Keystation keyboard that used a recorded Grand Piano sample from Propellerhead Reason software. Subjects heard the recordings on stereophonic headphones at a computer terminal. DirectRT software (Jarvis, 2002) was used to record subject responses.

Each of the 64 melodies was composed in a random key (32 in major, 32 in minor), and a piano noise track was inserted between each melody to clear the subjects’ minds of the tonality of the previous melody. There were five noise tracks, each consisting of 5 seconds of random piano notes played within a 3 octave range of C2 to C5. They were chosen at random. Each melody consisted of a 16 to 20 note chorale presented in 2:1 melodic events to harmonic events, at a rate of 800 ms per note. The length of the interonset interval (IOI) was chosen so as to provide evidence of real-time expectations at a tempo adequately fast enough to represent real music. During the test melodies, the computer monitor flashed green with the word “respond” at the onset of the target event, and subjects had 1600 ms to respond. Out of tune events were pitched a ¼ tone higher than the target note.

The attack velocity of the melodies at the note-to-note level was adjusted so as to mimic human performance. However, note velocities were kept within a range of 64 and 88 on a scale of 1 to 127 in standard MIDI, and expressivity due to note velocities and rhythmic duration was carefully controlled during the “surprise” events of the melody so as to prevent it from affecting response times. The melody was also presented at a 40% louder volume than the remaining three-voice accompaniment in order to promote listener attention to the melody.

For each of the eight test groups, a visual cue was inserted after the onset of variously expected melodic events as they are indicated by the Margulis model. Each melody group was composed so as to create an A melody consisting of highly expected events that were in tune, and a B melody consisting of highly unexpected events that were in tune. The cued events in the C and D melodies were analogous to the first two melodies, but they were out of tune. The harmonic events of every version of each group were identical until the onset of the cued event. The melodic events in the first bar of all four permutations of each chorale group were adjusted so as to prevent the subject from becoming too acquainted with any one melody. However, the melodic events leading up to the target event were preserved so as to also preserve the expectations theorized by the Margulis model.
Figure 3 illustrates the expectancy ratings for the cued events in both the A and B melodies. According to the model, some of the melodies should provide better results concerning response times than others. This graph demonstrates the variability of melodic expectancy as predicted by the model for the expected (A) melodies.

Procedure
Subjects were presented with eight groups of 4 four-part chorales alongside 32 fillers. They were told to attend to the melody of these chorales. At specific moments in the music, the computer monitor flashed green, at which point they were told to respond to whether the note they just heard was in tune or out of tune. They responded by pressing one of two corresponding buttons. Subjects were provided with ten practice melodies before they began the experiment in order to acquaint them with the task. The practice generally took no longer than 5 minutes of a 45 minute session. Accuracy and RT data were collected.

Results
The graphs in Figure 4 provide accuracy and RT information for the cued events across all 4 conditions. As predicted, subjects were much less accurate and much slower in their responses to the unexpected/in tune condition, condition B, than to the expected/in tune condition, condition A. The main effects for both accuracy and response time were significant: F(3, 57) = 59.22, p < .00001; F(3, 57) = 28.843, p < .0001. Further, paired comparisons between the A, B, C, and D conditions both for accuracy and response time indicated that paired differences between the A-B, A-C, B-C, and B-D conditions were all significant. A paired samples test for the A-B, B-C, and B-D comparisons provided the most robust effect, with all three registering an effect of p < .0001.

There was a strong bias to judge targets to be in tune when related and out of tune when unrelated. All 20 subjects showed both an accuracy and a reaction time advantage for the in tune expected condition (A), while the accuracy data for the in tune unexpected condition (B) indicates that subject responses were nearly no better than chance. A paired comparison of the C and D conditions both for accuracy and response time did not approach significance.

The results from the intonation method used in this experiment were also broadly consistent with the findings of Bharucha and Stoeckig (1986). They first used the intonation method to test harmonic priming between a prime chord and a target chord. Their accuracy data is provided in Figure 5. However, the Bharucha and Stoeckig study found a robust staircase effect between the B, C, and D conditions, and a paired comparison between the C and D conditions was significant. This discrepancy between the two studies is likely caused by the fact that out of tune events were pitched an 1/8th tone higher in the Bharucha and Stoeckig study, while this study pitched events a ¼ tone higher. Events in this study were therefore so out of tune that the out of tune expected condition (C) was much easier to detect than it was in the Bharucha and Stoeckig study.

Of the 20 subjects, 7 indicated having at least 5 years of formal training, 5 of which further reported being music majors. A paired comparison of accuracy data in condition B for trained vs. untrained subjects provided significant results, F (1, 18) = 4.974, p < .04, though response time data did not also indicate this effect. Figure 6 indicates that untrained subject responses were slightly worse than chance, while
trained subjects were generally much better at identifying the unexpected condition as in tune. The high variability in the trained responses, as indicated in the standard error bar, is an effect of having data from only 7 subjects.

Discussion

The results of this experiment suggest that several components of the priming paradigm provide researchers with an accurate method to detect musical expectations in listeners. First, the intonation task offers researchers an innovative response method for testing subject responses to music for trained as well as untrained listeners, since it permits listeners to respond to single events within the melodies at an adequate tempo to represent music. Second, this experiment successfully isolated stability within the Margulis model to provide empirical evidence about its role in listener expectations.

However, in order to isolate the category of stability, both proximity and direction had to be removed from the chorale melodies. In both the A and B conditions, the interval between the preceding event and the target event was preserved so as to ensure the proximity of the event did not affect listener expectations. Direction only has a pronounced effect on listener expectations in the Margulis model if there is a large interval between the preceding event and the target event, which always implies a directional reversal. However, within the test melodies, no intervals larger than a major 2nd approached the target event. Thus, two of the three potential factors in determining surprise in the expectation of melodies were effectively removed from the experiment in order to specifically consider the role of stability in listener expectations of tonal melodies.

The results of this experiment suggest that stability plays an important role in how listeners process melodies in real time. Margulis explains that “tonal stability is a central concept in both music theory and music cognition. Stability captures the intuition that, in general, listeners expect relatively stable melodic events” (2008, 675). Margulis’ model adopts Fred Lerdahl’s event governance rules, which select an operative chord and key in which stability may be evaluated. By weighing the stability of events in terms of an operative chord or key, Margulis appears to have accurately predicted that stability is a central factor in listener expectations. However, there are a few inherent weaknesses both in the priming methodology and in the current study. First, the data indicated a ceiling effect for accuracy, as is indicated in Figure 7. While Margulis’ model is highly variable in predicting the expectedness of events in the A condition, the accuracy data for the A condition provided by the experiment was nearly always 100%. This finding indicates that the response time method does not accurately detect expectations after a certain threshold of expectedness.

Second, the experiment did not attempt to isolate specific musical events. Instead, it simplistically grouped events as either expected or unexpected, rather than producing a larger number of groups to determine how accurate the Margulis model predicts listener expectations. However, as Figure 8 indicates, RT data provides accurate data at the level of the individual melody. Future studies could arrange individual melodies into several groups and staircase those groups according to their expectedness ratings, as RT data should provide a reliable indication of whether the predicted staircase effect is consistent with listener expectations.

However, the primary goal of the present study was to...
develop an on-line method that could permit the study of real-
time processing decisions according to listener expectations,
using an approach that does not terminate musical processing.
This study suggests that listeners without formal musical
training can detect events with varying degrees of expectedness
therefore possess a tacit knowledge of the musical idiom, in
this case common practice western harmony. That knowledge
generates automatic expectations during the listening
experience, which ultimately leads to the kind of physiological,
affective response in listeners we call surprise.

References
Bharucha, J. J., & Stoeckig, K. “Reaction time and musical
expectancy: Priming of Chords.” *Journal of Experimental
Bigand, E. & McAdams, S. “Divided attention in music.”
*International Journal of Psychology*, 35, No. 6, 2000,
270-278.
Charronnat, B. & Bigand, E. “The Influence of Voice Leading
on Harmonic Priming.” *Music Perception*, 22, No. 4,
2005, 613-627.
Hogg, Nanette. “Measuring Cognitive Load,” *Handbook of
research on electronic surveys and measurements, 2007,
188-194.
Jarvis, B. “DirectRT” [Computer Software]. New York, NY:
Empirisoft.
Krumhansl, C.L. *Cognitive Foundations of Musical Pitch.
Lou, Psyche, & Wessel, David. “Harmonic Expectation and
affect in Western music: Effects of attention and training.”
*Perception and Psychophysics*. Vol. 69, No. 7, 2007,
1084-1092.
Margulis, Elizabeth Hellmuth. “A Model of Melodic
663-714.
Margulis, Elizabeth Hellmuth. “Surprise and Listening Ahead:
Analytic Engagements with Musical Tendencies,” *Music
Margulis, E. & Levine, W. “Timbre Priming Effects and
Expectation in Melody.” *Journal of New Music Research*,
Rikers, Remy. “A Critical Reflection on Emerging Topics in
Cognitive Load Research.” *Applied Cognitive Psychology*,
Schmuckler, Mark. “Expectancy Effects in Memory for
Melodies,” *Canadian Journal of Experimental Psychology*,

Mentor Comments
Professor Elizabeth Margulis speaks glowingly about the
quality of David Sears’ work, describing it as equivalent to
doctoral level research and emphasizing its uniqueness in
drawing from three diverse disciplines.

In five classes across five years, David Sears has proven
himself to be an exceptional thinker — perhaps the most
creative, probing intellect I’ve encountered in nearly a
decade of teaching. I first noticed David’s rare abilities in
a basic music theory course, when he regularly scribbled
questions and comments into the margins of homework
assignments. These questions examined the intellectual
foundation of the discipline, anticipated future topics, and
resonated with current concerns in the field. Throughout
his undergraduate years, he has pursued a fascinating
agenda of original research. I am unsurprised that he
graduates this semester with a degree in Music on top
of his degree in English, the author of a publishable
piece of research that has won him fellowships to McGill
University in Montreal and University of Pennsylvania in
Philadelphia, as well as an invitation to present the paper
at a conference at the University of Calgary.

This project began two years ago, when David and I
started meeting regularly to discuss his interest in music
 cognition. It started to take clearer shape as a semester
project for my Fall 2007 Honors Colloquium on Music
and Mind. David zeroed in on expectation as a mechanism
shared by music and language processing. He sought
to investigate empirically the way that tonal surprises
can create tension in melodies. This pursuit is original
and important; no researcher has yet made systematic
connections between expectation and affect in a laboratory
setting.

David conducted an enormous literature review for this
study. He probably knows more about expectation in music
than all but perhaps a handful of scholars. I guided him
in certain directions, on occasion, but the initiative and
the review was all his. Although David made use of my
work in his article, he conceived of his research question,
designed his study, composed the stimuli (a huge task on
its own, consisting of the writing and programming of 32 musical excerpts fulfilling stringent requirements), ran all the subjects, analyzed the data, and wrote the paper. I offered suggestions along the way, but the work was entirely David's.

David deserves this award because he has produced Ph.D. thesis-level work as an undergraduate, teaching me about what students are capable of, and encouraging me to bring more students into my lab. I have been inspired by David to get more undergraduates involved in research. But I know that it will be a while before I see another David. He represents the absolute best in undergraduate research. He made connections between the English, Music, and Psychology Departments to complete an interdisciplinary project that offers a substantial contribution to the literature. The amount of creativity, dedication, initiative, and expertise in multiple fields required to accomplish this study defies description. I feel fortunate to have had the opportunity to work with such a gifted student.
FACTORS AFFECTING THE SPREAD OF A BIOTERRORIST AGENT THROUGHOUT A BUILDING

By William A. Cantrell
Department of Mechanical Engineering

Faculty Mentor: Dr. Darin Nutter
Department of Mechanical Engineering

Abstract

Bioterrorism has become a greater concern for Americans since the 2001 anthrax letters. Recent studies have explored the possibilities of biological attacks, and most deal with possible large-scale attacks. However, there is reason to believe that small-scale attacks are more likely. Even though there have been investigations of the postal delivery system and the spread of bioagents through mail, few if any studies have looked at an attack on a single building and the resultant spread from room to room. One particular method of attacking a building would be a single-event release of an aerosol bioagent in the building. This paper describes the development of a method for studying the spread of an aerosol throughout a building in order to determine what factors most affect the time between release and the lethal exposure for an occupant in various locations. A multi-zone airflow model, CONTAM, was used to simulate and compare the effects of the air handling system operation, door position, building level, predominant wind direction, and other factors. It was found that the air handling system, building floor level, and door position changed the time to lethal exposure. For the scenarios investigated, lethal exposure times ranged from 5 seconds to nearly 15 minutes, and the air handling system was found to have the greatest effect on a contaminant’s spread through a building.

Introduction

Throughout history, biological weapons have been used to wage war. One of the earliest and possibly deadliest examples occurred in the mid-1300s in Kaffa as bubonic plague victims of the Tartar army were catapulted over the city walls. Some believe that this is what led to the epidemic throughout medieval Europe that killed 25 million. The twentieth century saw the rise of research into biowarfare among nations across the world. This led to the signing of the 1972 Biological Weapons Convention, which forbids research with offensive biological agents and stockpiling bioweapons for military purposes [1].

Bioterrorism has become a concern for everyday Americans following the September 11, 2001 attacks on the World Trade Center. The first incidents involving anthrax occurred on September 25, 2001, when an assistant to Tom Brokaw (NBC anchorman) began to develop cutaneous anthrax after handling a letter containing anthrax powder. By November 2, 2001, the Center for Disease Control (CDC) had reported 21 cases of anthrax—16 confirmed and 5 suspected [2]. Anthrax is classified as a Category A bioterrorism agent. There are three categories of possible bioterrorist diseases or agents. Category A Diseases/Agents are the highest priority risks. These agents are the worst because they can be transmitted easily, result in high mortality rates, have potential for major public health impacts, and require special action for public health preparedness. Currently there are six listed by the CDC in Category A: Anthrax, Botulism, Plague, Smallpox, Tularemia, and Viral Hemorrhagic Fevers. Category B Diseases/Agents are moderately easy to disseminate, have low mortality rates, and require enhanced disease surveillance by the CDC. Category C Diseases/Agents are those considered to be available for mass dissemination, are easily produced, and have potential for high mortality rates. [3]

A large-scale release into the atmosphere or over a large city is greatly feared. For example, the release of 100 kg of anthrax over a large city could kill millions [4]. Large-scale attacks have been attempted by terrorists in the past, but all have failed. For instance, the Japanese doomsday cult Aum Shinrikyo failed on ten separate occasions at an open-air urban attack of anthrax or botulism, despite having considerable wealth and scientific capabilities. In March of 1995, the cult eventually killed 12 people through the release of sarin nerve gas in a Tokyo subway. Experts believe that in the near term, it is considered more likely that terrorist attacks will be small-scale attempts or merely hoaxes. [5]

Problem Statement

The misuse of a building’s ventilation system to spread a biological agent throughout a building is a real possibility [6]. The purpose of this study was to simulate various small-scale attack scenarios on a typical “office” building. The time between bioagent release and the time at which an occupant is exposed to a lethal dose were compared for various scenarios. In addition, the importance of building related factors such as air handling system (AHS) operation, building floor level, door position (open or closed), and predominant wind direction were analyzed.

Modeling Method

The use of airflow model techniques was determined to be the best approach for this undergraduate research project.
The National Institute of Science and Technology developed and maintains a model that was originally designed to analyze building ventilation and indoor air quality issues. This computer simulation model is known as CONTAM and is a multizone airflow and contaminant transport model capable of determining zonal airflow, contaminant concentrations, and personal exposures rates [7]. Using CONTAM to evaluate potential building terrorism is a logical extension of its application. It has more recently been considered an appropriate tool for such building simulations [8-11]. Other published works [12-14], only somewhat related, are recent studies on exposure to passengers, airflow and pathogen transport within aircraft cabins.

For this study, a simple building was sketched to model several different scenarios. The building is two stories tall with both floors having a large open space in the center that is meant to represent a cubicle area. On each floor and along the two opposite sides, are smaller rooms representing individual offices. In Figure 1, the CONTAM sketches for each floor are shown. CONTAM inputs included sizes for walls, ducts, and airflow paths (windows, doors, wall leakages, cracks, etc.). Mechanical systems such as ducts, fans, and zone sizes are also inputs. In addition, information on the tracked contaminants (i.e., biological agents) was input along with the location and method of entering the building.

Depending on the scenario, a burst contaminant source was placed in either a 1st floor office (location A) or a 1st floor maintenance room (location B). The source considered was an aerosol burst of 0.4 kg contaminant into the model at 10:00 AM. Simulations were run with doors in the building either all open or all closed. Exposure results were calculated in 1st and 2nd floor offices (locations 1-2) and in 1st and 2nd floor cubicle areas (locations 3-4), resulting in a total of 16 simulated scenarios. The burst source is representative of an aerosol release of a bioagent.

The effect of the AHS can be seen by comparing exposure times between scenarios with the contaminant originating in the maintenance room (with no ventilation) and the office (having ventilation). For an exposure in the same room, it would take 5 to 14 times as long for a lethal dose to be reached when the contaminant burst originated in the maintenance room as compared to the office. For contaminants originating in the office, the longest time to lethal exposure was 1 minute 25 seconds, occurring in the second floor cubicle area with all doors open. No matter the scenario, a release in the office had some of the agent immediately drawn into the AHS and quickly distributed throughout the building. The worst maintenance room release case occurred with all doors open. The time to lethal exposure was 45 seconds in the adjoining cubicle area. For most contaminants originating in the office, the longest time to lethal exposure was 1 minute 25 seconds, occurring in the second floor cubicle area with all doors open. No matter the scenario, a release in the office had some of the agent immediately drawn into the AHS and quickly distributed throughout the building. The worst maintenance room release case occurred with all doors open. The time to lethal exposure was 45 seconds in the adjoining cubicle area. For most
maintenance release cases, however, several minutes passed before lethal exposure time was reached. This was because the contaminant would have to first exit the maintenance room before it could be spread through the building within the AHS.

The AHS was the most dominant factor; however, the level (or floor) an occupant is on (relative to the contaminant release location) was also found to be important. The longest time until lethal exposure occurred when burst contaminant originated in the maintenance room. Of course, the longer the time to lethal exposure, the more opportunity to evacuate. If the biological release originated in an office, a person would become lethally exposed by simply traveling through the first floor cubicle area. When the release originated in the maintenance room, however, the contaminant level in the cubicles might be low enough for a period of time for occupants to leave the building. Further, it was expected that, with a contaminant originating on the first floor, the exposure times on a different floor would be nearly equal in each room. The second floor cubicle area consistently had longer exposure times than the second floor office room. This is suspected to be caused by unequal air circulation between the rooms. If one room has a higher air exchange rate, it would follow that the AHS would deliver a contaminant to that particular room at a higher rate as well.

Having open or closed doors affected the outcome the least. In general, longer times to lethal exposure occurred with doors closed when the exposure concern was in the adjoining room to the release point. The same was true (all doors closed) in nearly every case where the contaminant originated in the maintenance room. When the contaminant originated in the office room, having the doors open was better in every exposure room excluding the adjoining cubicle area. This was due to the fact that some of the contaminant escaped through open office door (into the cubical area), leaving less to enter the AHS which supplies the rest of the building.

Conclusions and Recommendations

Simulations for common elements of a two-story office building computed times for occupant lethal exposure that ranged from a few seconds to just over 14 minutes. Shorter times corresponded to open rooms on the same floor and near the release. Longer times to lethal exposure corresponded to rooms on floors different from the release point and when the contaminant was released from rooms without ventilation. The study shows the critical nature and importance of protecting against small-scale bioterror attacks in buildings. This means that an increased ability to detect bioagents is needed. Sensing technologies must be developed to detect quickly various agents at low concentrations. In the event of a biological release in a building, it is clear that one operating strategy is to shut off AHS as soon as a threat is detected.

Many effects that were not considered in this study could be further studied with CONTAM. These include the effects of outside windows, shutting off an AHS after release of an agent, filters and filter efficiencies, multiple AHS within a building, etc. To be fully prepared for the type of bioterrorist attack examined in this study, a model of a specific building should be made, and multiple scenarios should be run for that particular building to determine what procedures will minimize the occupant’s exposures.

References


Mentor Comments:

Dr. Darrin Nutter highlights Andrew Cantrell’s independence and the breadth of knowledge required to complete this innovative research project.

As an undergraduate student in mechanical engineering at the University of Arkansas, Mr. Andrew Cantrell was awarded an Honors College Undergraduate Research Grant for the research described in this paper. Andrew, currently in the United States Navy, was a hard working and well disciplined student. The research required learning the fundamentals of aerosol dispersion, existing literature, building air-conditioning systems, and the modeling software (CONTAM) that is made available by the U.S. National Institute of Standards and Technology. Andrew’s research was completely independent and took about a year and a half to complete. As the advisor, I provided guidance, structure, and expertise for Mr. Cantrell to complete his research. The research topic is unique and not addressed in the available literature. I encouraged the publication of Andrew’s work. Finally, it should be noted that even though the presented work could be perceived as sensitive, it is important to publish. Understandings of the research findings should be used in a proactive way and to emphasize the importance of developing the necessary sensing technologies related to prevention and aerosol bioagent spread minimization within a building.
FLIGHTS-TO-QUALITY: THE EFFECTS OF MARKET VOLATILITY ON SHORT TERM U.S. TREASURY YIELDS

By Craig Michael Cox
Department of Finance

Faculty Mentor: Craig G. Rennie
Department of Finance

Abstract

Flights-to-quality are the sudden, and sometimes irrational, rebalancing of investment portfolios to include more liquid and safer investments during times of uncertainty, high market volatility, or other unusual stock market environments. While previous research has explained flights-to-quality in terms of liquidity needs and credit risk premiums, this paper examines the significant statistical relationship between the VIX Index of implied market volatility and yields on U.S. Treasury bills. I found that the VIX Index explains a significant portion of U.S. Treasury yield variability and that the models become more significant and accurate as the maturity of the Treasuries increases. In terms of mispricing, the one month U.S. Treasury exhibited the largest deviation from the theoretical yield—more than 30%. Also, as the Treasuries’ maturities increased, the degree of mispricing decreased; this parallels the tendency to see a steepening yield curve during times of higher implied volatility.

1. Introduction and Overview

Flights-to-quality are the sudden, sometimes irrational, rebalancing of investment portfolios to include more liquid and safe investments (U.S. Treasury Bills) during times of uncertainty, high market volatility, or other unusual stock market environments. As investors demand safer securities, the prices of these investments increase and the corresponding yields (the annual rate of return on the investments) decrease significantly. In the recent subprime lending market fallout, a distinct flight-to-quality was observed in the one month U.S. Treasury bill between August 8, 2007 and August 27, 2007, where the yield on a one month U.S. Treasury bill fell from 5% to nearly 2% in less than a week.

While many factors influence the yields on U.S. government securities, a significant increase in implied market volatility was most likely the underlying trigger of this particular flight-to-quality. The VIX Index of implied volatility was at a four year high during this drop in yields. It is common for investors to seek safe, fixed-income investments when the market is abnormally volatile (VIX Index value of more than 30), but how much variance in short term U.S. government Treasury yields can actually be explained by implied market volatility? A drop in yields from 5% to 2% is hard to justify given only an increase in one index of volatility. Also, if there is a significant effect, which durations experience the greatest change in yields (one month, three month, one year Treasury bills)? How is the entire yield curve affected by flights-to-quality?

The topic of market volatility and flights-to-quality is an interesting and important aspect of empirical financial research. While flights-to-quality have been examined previously, this paper sheds more light on the overall effect of increased market volatility on U.S. Treasury bills and on the dynamics of the U.S. Treasury yield curve during these periods of market uncertainty. The outcomes of the data analysis conducted here provide additional information about the market’s often irrational reaction to increased volatility and the underlying causes of investors’ flights-to-quality.

The statistical analyses conducted in this study included multivariate regressions which were used to determine the amount of variability in short term U.S. Treasury yields that can be explained by the VIX Index of market volatility and the Federal Funds target rate, a control variable. The initial regression analysis showed that the VIX Index and the Federal Funds rate variables explained approximately 98% of the variability in the yield on the one month U.S. Treasury since October 31, 2001. According to the regression equation, during the flight-to-quality in August, 2007, the one month Treasury was mispriced by approximately 53%. The data for the three month and six month Treasuries showed similar but different degrees of yield change during the August flight-to-quality, and significantly different levels of Treasury mispricing.

The statistical strategies used in this study have applications ranging from new trading strategies to economic policy making, adding to existing finance research on the topic. The following section of this article summarizes relevant research and theoretical premises. The remainder of the article provides a more detailed discussion of the research methods findings, and conclusions.

1 The ticker symbol for the Chicago Board Options Exchange (CBOE) Volatility Index, which shows the market’s expectation of 30-day volatility. It is constructed using the implied volatilities of a wide range of S&P 500 index options. This volatility is meant to be forward looking and is calculated from both calls and puts. The VIX is a widely used measure of market risk and is often referred to as the “investor fear gauge.”
2 A graph that plots the interest rates, at a set point in time, of fixed-income instruments having equal credit quality, but differing maturity dates.
3 August 20, 2007—two days after the VIX Index reached a year-to-date high of 30.83. The yield on the one month U.S. Treasury fell from 4.60% to 2.11% in only five trading days (see Appendix 1).
2. Review of literature

Relevant literature can be classified into two areas that are useful for this paper: evidence of the flight-to-quality phenomenon and causes of the flight-to-quality. Each of these is discussed briefly below.

2.1. Evidence of the flight-to-quality

In a paper on stock-bond correlation, Baur and Lucey (2006) find that “there is strong evidence for flight-to-quality” and “there is simultaneous flight-to-quality across countries [...] in several crisis periods.” Their paper examined data from eight developed countries which suggest that the flight-to-quality is a pervasive phenomenon in global stock markets.

Underlying reasons for these flights-to-quality have been investigated by Beber, et al. (2006) who find that “in times of market stress, investors chase liquidity, not credit quality” (liquidity is a key determinant of demand for short-term money market securities). The idea of a flight-to-liquidity is further supported by Longstaff (1998) who finds that there is “a large liquidity premium in Treasury bonds, which can be more than 15% of the value of some [...] bonds.”

2.2. Causes of the flight-to-quality

As for the effects of volatility on flights-to-quality and flights to liquidity, Vayanos (2004) finds that “during volatile times, assets’ liquidity premia increase, investors become more risk averse, assets become more negatively correlated with volatility, assets’ pairwise correlations can increase, and illiquid assets’ market betas increase.” Longstaff also suggests that this premium is caused by consumer confidence, the amount of debt available to investors, and movements to equity and money market mutual funds. This implies that Treasury bonds’ popularity has a direct effect on their valuation. In terms of sheer market dynamics, flights-to-quality can be explained by volatility and demand for liquidity. It is still difficult, however, to quantitatively explain such a sudden and irrational short-term demand for investments earning only 2%.

In addition to the liquidity premium suggested in Longstaff (2004), there is also evidence of a credit risk premium inherent in U.S. Treasury yields. Huang (2003) notes that up to 30% of a U.S. Treasury-corporate bond credit spread is attributable to a credit risk premium.

Other than the liquidity and credit premiums suggested in the papers above, little light has been shed on the actual behavioral causes of flights-to-quality. Avery and Zemsly (1998) find that “herd behavior can lead to a significant, short run mispricing” in certain securities. Do investors only exhibit herd behavior with U.S. government securities? What about other fixed-income investments? Why bid for government securities when you can walk to a local bank and open a CD account with a similar risk profile that earns three times as much?

In summary, previous research has concentrated on evidence and causes of flights-to-quality. The study described in this paper adds to the literature by examining more thoroughly the effects of volatility on short term U.S. Treasury yields and the changes in the yield curve that result. In addition, significant and possibly lucrative trading strategies could be developed given the results of the statistical analysis explained below.

3. Sample selection and research method

As previous research has explained flights-to-quality in terms of liquidity needs and credit risk premiums, this paper examined the significant statistical relationship between the VIX Index of implied market volatility and yields on U.S. Treasury bills. I also investigated the effects of flights-to-quality on the positioning of the yield curve and the degree of theoretical mispricing in the Treasuries during flights-to-quality.

To explore these topics, the following data were compiled: historical daily yields for three relevant U.S. Treasury bills (one month, three month, and six month maturities); the corresponding VIX volatility index values, and the target Federal Funds rate. Regression analyses were then performed to determine the level of variation in the Treasury data explained by the corresponding VIX Index values, or how much of the flight-to-quality phenomenon can be statistically justified by the underlying market volatility. The format of the regression equation is as follows, where $Yield_t$ is the yield for the short term Treasury, $VIX_t$ is the corresponding VIX Index value, $c_1$ is the regression coefficient for the VIX Index variable, $FFRate_t$ is the corresponding Federal Funds target rate, $c_2$ is the coefficient for the Federal Funds variable, and $\varepsilon$ is an error term.

\begin{equation}
Yield_t = c_1 VIX_t + c_2 FFRate_t + \varepsilon
\end{equation}

Two additional regression analyses were performed using data with the VIX Index values lagged by one and two days, respectively. In the corresponding regression equations, $VIX_{t-1}$ is the one-day lagged VIX Index value and $VIX_{t-2}$ is the two-day lagged VIX Index value.

\begin{equation}
Yield_t = c_1 VIX_{t-1} + c_2 FFRate_{t} + \varepsilon
\end{equation}

\begin{equation}
Yield_t = c_1 VIX_{t-2} + c_2 FFRate_{t} + \varepsilon
\end{equation}

Market data were compiled from Bloomberg Professional (an integrated financial platform that streams together price and trading data, financials, news, and economic statistics). Given the six year range of the data, “generic” Treasury yields were used. Generic yields are derived from the synthetic yield history that is created by piecing together observed closing yields for benchmark bills of a given maturity. Such derived yields provide a more useful and seamless data set.

After determining the correlation between the VIX Index and Treasury yields, I used the regression equations to determine theoretical yields for each U.S. Treasury (based on the given Federal Funds target rate and the VIX Index value) and compared these to the actual yields during the August 2007 flight-to-quality. This gave an approximate estimate for the mispricing of the Treasury bills during the period shortly after the spike in market volatility.
4. Findings

4.1. Relevant statistical findings

This study used multivariate regression to determine the amount of variability in short term U.S. Treasury yields that could be explained by the VIX Index of market volatility and the Federal Funds target rate. The results of univariate, bivariate, and multivariate analyses are summarized below and presented in Table 1.

4.2. Strong correlation between Federal Funds target rate and U.S. Treasury yields affects study

The correlation between the Federal Funds target rate and each of the U.S. Treasury yields is strong—approximately 99% for each maturity (one month, three month, and six month). To account for this collinearity, regressions were recalculated using the VIX Index as the only independent variable. Further, a specific data set in which the Federal Funds rate remained unchanged at 5.25% was used to account for the variability in each of the variables, I used historical trading data from an August flight-to-quality to estimate approximately the theoretical yield on the U.S. Treasuries, given that day’s VIX Index value and Federal Funds rate. I then compared this value with the actual close, to determine roughly the amount of mispricing due to the underlying flight-to-quality. Put another way, I determined what the Treasury should have closed at and compared it to its actual closing yield. This analysis showed strong mispricing in the one and three month U.S. Treasury yields—53% and 29% overpriced, respectively (see Tables 2 and 3). The six month Treasury was actually underpriced (true yield was higher than the theoretical) by only 1% (see Table 4).

This mispricing result can be explained by the direct relationship between implied volatility and the slope of the yield curve. As volatility increases, there is downward pressure on the yields of Treasuries with shorter maturities. Since the yield curve steepens in response to this increase in volatility, short maturity bills are mispriced more considerably more than longer term notes and bills. The small underpricing in the six month yield is most likely insignificant and does not represent a considerable deviation from the theoretical yield (as in the one and three month bills).

4.3. Significant mispricing in one and three month U.S. Treasuries during flight-to-quality

After determining a regression equations using both the VIX Index and Federal Funds target rate as independent variables, I used historical trading data from an August flight-to-quality to estimate approximately the theoretical yield on the U.S. Treasuries, given that day’s VIX Index value and Federal Funds rate. I then compared this value with the actual close, to determine roughly the amount of mispricing due to the underlying flight-to-quality. Put another way, I determined what the Treasury should have closed at and compared it to its actual closing yield. This analysis showed strong mispricing in the one and three month U.S. Treasury yields—53% and 29% overpriced, respectively (see Tables 2 and 3). The six month Treasury was actually underpriced (true yield was higher than the theoretical) by only 1% (see Table 4).

This mispricing result can be explained by the direct relationship between implied volatility and the slope of the yield curve. As volatility increases, there is downward pressure on the yields of Treasuries with shorter maturities. Since the yield curve steepens in response to this increase in volatility, short maturity bills are mispriced more considerably more than longer term notes and bills. The small underpricing in the six month yield is most likely insignificant and does not represent a considerable deviation from the theoretical yield (as in the one and three month bills).

4.4. Regression significance increases as U.S. Treasury maturity increases

For all three U.S. Treasuries, both regression analyses (with and without Federal Funds target rate) exhibited significance. However, as the maturity increased, there were

Table 1: Descriptive statistics and correlations

<table>
<thead>
<tr>
<th>Panel A: Descriptive statistics</th>
<th>VIX Index</th>
<th>Fed Funds Target Rate</th>
<th>1 Month Treasury</th>
<th>3 Month Treasury</th>
<th>6 Month Treasury</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>18.0200</td>
<td>2.8226</td>
<td>2.6068</td>
<td>2.7112</td>
<td>2.8573</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.1760</td>
<td>0.0428</td>
<td>0.0398</td>
<td>0.0494</td>
<td>0.0410</td>
</tr>
<tr>
<td>Median</td>
<td>16.6200</td>
<td>2.0000</td>
<td>1.8870</td>
<td>2.0120</td>
<td>2.1720</td>
</tr>
<tr>
<td>Mode</td>
<td>11.9800</td>
<td>5.2500</td>
<td>0.8650</td>
<td>0.9270</td>
<td>1.0320</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>6.8271</td>
<td>1.6211</td>
<td>1.5423</td>
<td>1.5691</td>
<td>1.5906</td>
</tr>
<tr>
<td>Sample Variance</td>
<td>46.6089</td>
<td>2.7627</td>
<td>2.3788</td>
<td>2.4620</td>
<td>2.5300</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>1.2122</td>
<td>-1.3054</td>
<td>-1.3902</td>
<td>-1.5023</td>
<td>1.5632</td>
</tr>
<tr>
<td>Skewness</td>
<td>2.2912</td>
<td>0.4114</td>
<td>0.4339</td>
<td>0.3589</td>
<td>0.2639</td>
</tr>
<tr>
<td>Range</td>
<td>35.1900</td>
<td>4.2500</td>
<td>4.5350</td>
<td>4.7500</td>
<td>4.5100</td>
</tr>
<tr>
<td>Minimum</td>
<td>9.8900</td>
<td>1.0000</td>
<td>0.7120</td>
<td>0.8500</td>
<td>0.8060</td>
</tr>
<tr>
<td>Maximum</td>
<td>45.0800</td>
<td>5.2500</td>
<td>5.2470</td>
<td>5.1750</td>
<td>5.3160</td>
</tr>
<tr>
<td>Sum</td>
<td>27120.1100</td>
<td>4248.0000</td>
<td>3923.2980</td>
<td>4080.4100</td>
<td>4500.2400</td>
</tr>
<tr>
<td>Count</td>
<td>1505</td>
<td>1505</td>
<td>1505</td>
<td>1505</td>
<td>1505</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Correlation matrix</th>
<th>VIX Index</th>
<th>Fed Funds Target Rate</th>
<th>1 Month Treasury</th>
<th>3 Month Treasury</th>
<th>6 Month Treasury</th>
</tr>
</thead>
<tbody>
<tr>
<td>VIX Index</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fed Funds Target Rate</td>
<td>-0.4591</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Month Treasury</td>
<td>-0.4803</td>
<td>0.9870</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Month Treasury</td>
<td>-0.5083</td>
<td>0.9910</td>
<td>0.9957</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6 Month Treasury</td>
<td>-0.5381</td>
<td>0.9869</td>
<td>0.9885</td>
<td>0.9974</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: Controlled correlation matrix</th>
<th>VIX Index</th>
<th>1 Month Treasury</th>
<th>3 Month Treasury</th>
<th>6 Month Treasury</th>
</tr>
</thead>
<tbody>
<tr>
<td>VIX Index</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Month Treasury</td>
<td>-0.6140</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 Month Treasury</td>
<td>-0.7368</td>
<td>0.9985</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6 Month Treasury</td>
<td>-0.7492</td>
<td>0.7951</td>
<td>0.9556</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 1: Descriptive statistics and correlations for the five variables used in the study. Panels A and B cover a period from October 31st 2001 to October 31st 2007 while Panel C includes only data from June 20th to September 17th, this controls for the Federal Funds rate as it remained at 5.25% over the period.

4 The Federal Funds rate was unchanged at 5.25% from June 29, 2006 to September 17, 2007.

http://scholarworks.uark.edu/inquiry/vol9/iss1/1
proportional increases in significance (F values closer to zero) as well as increases in the adjusted R² values. Accordingly, the six month Treasury regression showed the most significance while also explaining the most variability in the Treasury yields that could be attributed to the corresponding VIX Index values (see Table 4). Given the six month Treasury’s relatively low degree of mispricing, the result of a more accurate and significant model is not unexpected.

Table 2: One month U.S. Treasury regression analysis

<table>
<thead>
<tr>
<th></th>
<th>FFRate and VIX Index</th>
<th>VIX Index</th>
<th>VIX Indexₓ₁</th>
<th>VIX Indexₓ₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.2033*</td>
<td>5.6372*</td>
<td>5.691*</td>
<td>5.7123*</td>
</tr>
<tr>
<td></td>
<td>(7.4947)</td>
<td>(90.1104)</td>
<td>(93.4950)</td>
<td>(94.4152)</td>
</tr>
<tr>
<td>cᵣ (VIX)</td>
<td>-0.0078*</td>
<td>-0.0563*</td>
<td>-0.0601*</td>
<td>-0.0617*</td>
</tr>
<tr>
<td></td>
<td>(-7.5204)</td>
<td>(-13.5406)</td>
<td>(-14.8296)</td>
<td>(-15.2781)</td>
</tr>
<tr>
<td>cᵣ (FFRate)</td>
<td>0.9012*</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(212.0522)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Significance F</td>
<td>0</td>
<td>5.4520 x 10⁻³³</td>
<td>8.7707 x 10⁻⁵⁶</td>
<td>1.8127 x 10⁻³⁹</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.9751</td>
<td>0.3749</td>
<td>0.4186</td>
<td>0.4333</td>
</tr>
<tr>
<td>N</td>
<td>1505</td>
<td>305</td>
<td>305</td>
<td>305</td>
</tr>
</tbody>
</table>

*Indicates statistical significance at the 0.05 level

Table 3: Three month U.S. Treasury regression analysis

<table>
<thead>
<tr>
<th></th>
<th>FFRate and VIX Index</th>
<th>VIX Index</th>
<th>VIX Indexₓ₁</th>
<th>VIX Indexₓ₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.4329*</td>
<td>5.6139*</td>
<td>5.6453*</td>
<td>5.6625*</td>
</tr>
<tr>
<td></td>
<td>(20.6639)</td>
<td>(144.2953)</td>
<td>(150.2182)</td>
<td>(153.1091)</td>
</tr>
<tr>
<td>cᵣ (VIX)</td>
<td>-0.0155*</td>
<td>-0.0490*</td>
<td>-0.0513*</td>
<td>-0.0526*</td>
</tr>
<tr>
<td></td>
<td>(-19.4262)</td>
<td>(-18.9693)</td>
<td>(-20.4968)</td>
<td>(-21.3006)</td>
</tr>
<tr>
<td>cᵣ (FFRate)</td>
<td>0.9062*</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(276.1704)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Significance F</td>
<td>0</td>
<td>1.9406 x 10⁻³³</td>
<td>3.5163 x 10⁻⁵⁹</td>
<td>3.5583 x 10⁻⁶²</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.9857</td>
<td>0.5414</td>
<td>0.5796</td>
<td>0.5983</td>
</tr>
<tr>
<td>N</td>
<td>1505</td>
<td>305</td>
<td>305</td>
<td>305</td>
</tr>
</tbody>
</table>

*Indicates statistical significance at the 0.05 level

Table 4 shows the results of the regression analyses performed using one month U.S. Treasury data. Column (1) summarizes the regression analysis calculated using both the VIX Index values and Federal Funds target rate as independent variables over the period of October 31, 2001 to October 31, 2007. Column (2) summarizes the regression analysis calculated using the VIX Index values as the sole independent variable. This regression only includes data from June 29, 2006 to September 17, 2007 to control for the Federal Funds target rate as it remained at 5.25% over the period. Columns (3) and (4) summarize the regressions calculated using only lagged VIX Index values from June 29, 2006 to September 17, 2007 (to control for the Federal Funds target rate) as the independent variable. The VIX Index values for these regressions were lagged one and two days, respectively. t-statistics are in parentheses.

4.5. Regression significance increases when VIX Index values are lagged

For all three U.S. Treasuries, regression analyses using lagged VIX Index values showed higher adjusted R² values and lower significance F values, indicating more accurate models. The one month Treasury regression showed the highest increase in adjusted R² from lagging the VIX Index by two days—a nearly six percentage point increase. This result suggests that U.S. Treasury yields require at least one day to price in the effects of higher VIX Index values (this of
Table 4: Six month U.S. Treasury regression analysis

<table>
<thead>
<tr>
<th></th>
<th>FFRate and VIX Index</th>
<th>VIX Index</th>
<th>VIX Index,t-1</th>
<th>VIX Index,t-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.7772*</td>
<td>5.5903*</td>
<td>5.6047*</td>
<td>5.6160*</td>
</tr>
<tr>
<td></td>
<td>(33.8087)</td>
<td>(182.3036)</td>
<td>(184.7627)</td>
<td>(186.9325)</td>
</tr>
<tr>
<td>c_t (VIX)</td>
<td>-0.0251*</td>
<td>-0.0401*</td>
<td>-0.0412*</td>
<td>-0.0421*</td>
</tr>
<tr>
<td></td>
<td>(-28.6255)</td>
<td>(-19.6869)</td>
<td>(-20.3801)</td>
<td>(-20.9577)</td>
</tr>
<tr>
<td>c_t (FFRate)</td>
<td>0.8971*</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(249.1773)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Significance F</td>
<td>0</td>
<td>3.8305 x 10^{-5}</td>
<td>9.6044 x 10^{-5}</td>
<td>6.7001 x 10^{-6}</td>
</tr>
<tr>
<td>Adjusted R^2</td>
<td>0.9832</td>
<td>0.5598</td>
<td>0.5768</td>
<td>0.5904</td>
</tr>
<tr>
<td>N</td>
<td>1505</td>
<td>305</td>
<td>305</td>
<td>305</td>
</tr>
</tbody>
</table>

*Indicates statistical significance at the 0.05 level.

Table 4 shows the results of the regression analyses performed using six month U.S. Treasury data. Column (1) summarizes the regression analysis calculated using both the VIX Index values and Federal Funds target rate as independent variables over the period of October 31, 2001 to October 31, 2007. Column (2) summarizes the regression analysis calculated using the VIX Index values as the sole independent variable. This regression only includes data from June 29, 2006 to September 17, 2007 to control for the Federal Funds target rate as it remained at 5.25% over the period. Columns (3) and (4) summarize the regressions calculated using only lagged VIX Index values from June 29, 2006 to September 17, 2007 (to control for the Federal Funds target rate) as the independent variable. The VIX Index values for these regressions were lagged one and two days, respectively. t-statistics are in parentheses.

5. Conclusion

This study examined the amount of variability in short term U.S. Treasury yields that can be explained by the VIX Index of implied market volatility. Twelve regression analyses were performed using the one, three, and six month generic U.S. Treasury yields as dependent variables and the corresponding VIX Index values and Federal Funds target rate as independent variables. To improve the accuracy of the study, models were created both including and excluding the Federal Funds target rate. This process addressed the collinearity between the Treasuries and Federal Funds rate. Models with lagged VIX Index values were also created.

The models showed a significant level of mispricing in the one and three month yields during the August flight-to-quality (see Figure 1). Due to the proportional relationship between the slope of the yield curve and implied volatility, this result is not unexpected. The sheer degree of mispricing in the one month Treasury (more than 50%), however, is both an unexpected and interesting result (see Table 5).

The tendency for model accuracy to improve as the Treasury maturity increases is another interesting result of the analysis. Examining the flight-to-quality in August, though, shows a less volatile market for the six month Treasury (Figure 1). This smoother data set is most likely the reason for the more significant model. Regression models using data lagged by one and two days respectively explained a higher percentage of variability in the Treasury yields that could be explained by the VIX Index values. Models with lagged data also exhibited more significance (lower F values).

The analyses used in this study examined the dynamics of short term U.S. Treasury yields during flights-to-quality.
Table 5: Determination of fight-to-quality mispricing using regression equations

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Regression equation</th>
<th>Theoretical yield</th>
<th>Actual yield</th>
<th>Percent mispriced</th>
</tr>
</thead>
<tbody>
<tr>
<td>One month U.S. Treasury</td>
<td>$Y_{\text{Treasury}} = -0.0078X_\text{v} + 0.9012X_\text{f}$</td>
<td>4.53%</td>
<td>2.11%</td>
<td>-53.45%</td>
</tr>
<tr>
<td>Three month U.S. Treasury</td>
<td>$Y_{\text{Treasury}} = -0.0155X_\text{v} + 0.9062X_\text{f}$</td>
<td>4.35%</td>
<td>3.09%</td>
<td>-28.88%</td>
</tr>
<tr>
<td>Six month U.S. Treasury</td>
<td>$Y_{\text{Treasury}} = -0.0251X_\text{v} + 0.8971X_\text{f}$</td>
<td>4.05%</td>
<td>4.09%</td>
<td>0.88%</td>
</tr>
</tbody>
</table>

Table 5 shows the mispricing of U.S. Treasuries on August 20th 2007 due to a flight-to-quality. Using the regression equations that included both the Federal Funds and VIX Index variables, theoretical yields were calculated and compared to the actual yield.

The statistical results show interesting levels of mispricing and various degrees of quality and significance among the three different maturities. The models created in this analysis can be adapted to complement trading strategies and will add to the existing (and limited) research on the statistical effects of volatility on flights-to-quality.

References

Mentor Comments:
Dr. Craig Rennie draws attention to the independence of Craig Cox's work and its importance in contributing to our understanding of stock market volatility.

Craig Cox's undergraduate honors thesis titled “Flights to Quality: The Effects of Market Volatility on Short Term U.S. Treasury Yields” addresses a research topic of great interest not previously addressed in Finance. Market volatility and associated flights to quality via short-term U.S. Treasury securities is of great interest to Finance academics and practitioners. Artificially low yields and correspondingly high prices appear inconsistent with market efficiency. However, a growing body of behavioral Finance literature suggests investors often overreact to news and market volatility through excessive buying or selling of securities. The effects of stock market volatility, reflected in the relatively new VIX index, on U.S. Treasury yields and the U.S. bond market, have not previously been studied. Craig shows that stock market volatility, as measured by the VIX index, helps explain volatility in U.S. Treasury yields and thus prices. In fact, he documents mispricings of as much as 50% during periods of high stock market volatility. These results are some of the first to explain changes in U.S. Treasury yield changes associated with flights to quality following increases in stock market volatility. They represent a material contribution to the body of knowledge in the field of Finance.

Craig's research is original, and he selected his topic completely independently. His thesis “Flights to Quality: The Effects of Market Volatility on Short Term U.S. Treasury Yields” is one of the very best undergraduate research papers that I have seen.
THE SAINTE-CHAPELLE IVORY VIRGIN & CHILD: RAYONNANT STYLE AND PRIVATE DEVOTION

By Caroline Harrington
Department of Art

Faculty Mentor: Lynn Jacobs
Department of Art

Abstract
This paper examines a major shift in French Gothic sculpture of the second half of the thirteenth century, as exemplified by the Sainte-Chapelle Virgin and Child. During this period there was a new emphasis on elegance in art works, giving rise to a new style called Rayonnant, a style paralleled by a new emphasis on the humanity of Christ and the Virgin Mary. The scale and patronage of the Sainte-Chapelle Virgin and Child demonstrate the changing purpose of sculpture from a ceremonial role in church life to a private devotional object for the French elite, in particular king of France Louis IX, for whom the work was made. The changing style and function of this sculpture is explored in light of earlier treatments of the Virgin and Child in Byzantine and Romanesque art and in the context of the rise of the cult of the Virgin and of a new emphasis on internal (as opposed to external) religious experiences. In this way, the paper provides a case study of the intersection of style and iconography with patronage and function.

The Sainte-Chapelle Ivory Virgin & Child

In the second half of the thirteenth century, French Gothic sculpture underwent a major shift in style. At this time there was a new emphasis on elegance in art, paralleled by a new emphasis on the humanity of Christ and the Virgin Mary. The ivory Virgin and Child statue from Sainte-Chapelle displays this novel approach of late thirteenth century France. Rayonnant (from the French word radiating) is the name given to the style of this time, a style which is characterized by soft features, swaying postures, decorative drapery, and precious materials. The Sainte-Chapelle Virgin and Child’s small scale exemplifies the changing purpose of sculpture from serving ritual functions in the church to functioning in the realm of solitary private devotion. In addition, the work’s patronage reflects the growing importance of private devotion within the French elite of the later thirteenth century.

Today, the Virgin and Child of Sainte-Chapelle, which is made of ivory with traces of gold and polychrome, resides in Paris’ Louvre. It came from the Sainte-Chapelle treasury and dates from around 1260 to 1270. Standing at about 16 inches (41 centimeters), this charming statue was made for King Louis IX for use in his private chapel, the Sainte-Chapelle.

The Sainte-Chapelle Virgin and Child displays the elegant qualities of the Rayonnant style. The Virgin’s face is composed with grace and beauty; she has almond eyes, which convey an emotion and a passion that empowers her warm, curved smile. Her pointed chin and dainty, heart-shaped face are sheltered by delicate curls. All of these gentle features produce a quality of a pleasant refinement in the face of Mary (see Figure 1).

Figure 1. Virgin and child. From the treasury of the Sainte-Chapelle, Paris. French. 1250-60. Ivory H. 41 cm. Location: Louvre, Paris, France. Photo Credit: Scala/Art Resource, NY Image reference: ART123372

The pose, anatomy, and scale of the ivory Virgin and Child are also very sophisticated. Mary is standing – an innovation first developed within Virgin and Child statues of the Rayonnant period. The swaying “S” curve of her body...
gives the Virgin an element of poise. She is also holding Jesus, who appears to have weight rather than being suspended over her lap. Mary holds him on her hip and leans back slightly, compensating for his heaviness. Although Jesus’ head is slightly too small for his body, he is still recognizable as a baby, rather than as an oddly sized man. Mary appears to be proportional, especially to her baby, eradicating the age old use of hierarchy of scale to distinguish the most important figure in a work. Paul Williamson described the statue as having a “gentle contrapposto... [that would] permeate many other Virgin and Child works.”2 This stance became the standard for Rayonnant works. Although standing in this relaxed, contrapposto pose, Mary’s posture demands respect through its dignity.

The flowing, elegant drapery of the Virgin and Child is another trait of the French Rayonnant. Although their bodies have weight, Mary and Jesus’ clothing is less grounded to the earth and more ethereally floating. The deep, long folds work both to give the work a realistic quality and also to give the fabric an otherworldly richness. The gathered material exhibits a graceful and dainty quality. The sculptor worked with the natural curve in the ivory to create a curving, moving sensation throughout the work, especially in the drapery. Rather than looking like stiff, solid ivory, the material is transformed into soft, pliable cloth. Also notable are the actual garments of Mary and Jesus. They are clad not rather in peasants’ garbs, but capes and gowns of royalty. Their clothing could easily be made of silk or another precious material from the “large ‘beak’ folds” that convey “power and harmony”3 – as described in the Grove Dictionary of Art. These qualities appear in another, more monumental, Rayonnant sculpture, the Vierge Dorée of Amiens dating from the mid to late thirteenth century. Although made of stone rather than ivory, the drapery of the Vierge Dorée (see Figure 2) has a similar composition and similar drapery to the Sainte-Chapelle Virgin and Child. The large pleats of stone trick the viewer into believing they are a gentle fabric flowing down the bodies of the majestic duo. The Vierge Dorée and the ivory Virgin and Child are fundamental examples of the sophisticated, polished Rayonnant design.

Even the base of the Sainte-Chapelle statue displays Rayonnant refinement. Mary stands on a polygonally shaped base surrounded by diamonds and elaborate foliage. To have a statue of such elegance and sophistication on a plain base would be insulting and inappropriate, especially for work produced for a French king. Diamonds were and still are a symbol that references royalty, and Louis Grodecki explains the natural elements were used to “glorify God.”4 So, the detailed base harmonizes with the richness of the statuette, while also sending a message of piety about the patron.

Also notable is the gilded ivory composing the statue. Ivory was a rare material not native to Europe; it experienced an increase in use around the late thirteenth century. Indeed, as Peter Barnett notes, the rising desire for ivory in Western Europe at this time resulted in the establishment of trade routes with the Mozambique in Southern Africa.5 However, only the nobility and others with great wealth could afford such a luxurious material. The Virgin and Child statue is adorned with gold in addition to the ivory, thus making it all the more extravagant. Evidently, the French nobility were willing to spare no expense for their lavish taste in precious, religious devotional objects. Their willingness to pay for such art works is a sign not only of their love of luxury, but also of the importance of private devotion in the life of the elite at this time.

The Sainte-Chapelle Virgin and Child statuette is a key example within a new genre of small-scale sculptures that marked a shift in religious focus among the French aristocracy away from public worship in churches, and more toward personal devotion within private chambers.6 This statue, as Williamson noted, was intended specifically for the private worship of the king of France, Louis IX. This shift toward private worship arose out of the later medieval concern for more internal religious experiences. To foster such experiences, small statues like this developed between the eleventh and thirteenth century. The smaller proportions emphasized the humanity of the religious figures, therefore fostering personal engagement with them – The Sainte-Chapelle Virgin and Child
exemplifies this new category of image designed to facilitate the experience of private devotion by portraying a "tender relationship between the mother and her child."

The work is not only representational of a correlation between a new stylistic movement in sculpture and a different approach to religious worship, but it also displays the personal taste of Louis IX that would influence the French elite and even the common people of this time. King Louis’ religious decisions – displayed through his aesthetic choices - to participate in private devotion took hold with most people who sought out personal and private religious relationships. He was pious, and considered a saint by his people before he was ever canonized. Upon his death, “nearly 400 witnesses gathered...to testify his sanctity.” Under his forty-four year reign, Louis IX was renowned not only for making Paris a reputable “artistic and intellectual center” but also for the many churches and religious houses he established. He built the Sainte-Chapelle, his personal chapel, to enshrine the relics he had collected from Christ’s Passion. He was truly a virtuous king, full of religious fervor.8 For Louis IX, the Sainte-Chapelle with its Virgin and Child was his own safe haven for reflection and prayer, which ultimately permeated the late thirteenth century society.

The humanity of the Virgin and Child was the focal point for most private devotion. Christ’s image changed from the third century “Good Shepherd,” to a “God King,” and finally in the 1200s to a “God born of a human.” As the focus changed from his divinity to his mortality, a cult dedicated to the Virgin developed.9 Willibald Sauerländer revealed how the increasing veneration for the Virgin grew into a cult following and, as a result, Marian themes even occasionally replaced the Last Judgment on cathedral portals. In such portals Mary was a living vessel of all virtues and therefore a role model to all young women.10 It was not until the Rayonnant period, however, that Mary was shown standing and holding the Christ Child. Penny Schine Gold argued for the necessity of the transition from the seated Virgin, as the throne of wisdom, to the standing Virgin. The standing Virgin suggests a gentle, caring mother unlike the hieratic, metaphysical majesty of the throne of wisdom. When Mary stands, she no longer serves as a seat for Jesus. She is more dominant than passive. She is less of a “God-bearer” and more a human mother.11 The Sainte-Chapelle Virgin and Child presents a particularly effective image of the humanity of Christ and the Virgin.

In this statue, Mary and Jesus display naturalistic behavior through their interactions and appear to be full of life and emotion. They are playing with an apple, perhaps an allusion to the original sin Adam and Eve brought into the world – a sin for which Jesus, brought into this world by Mary, would ultimately die. Another ivory statue of a similar date, the Virgin and Child of Saint-Denis, also highlights the humanity of Christ through naturalistic behavior between him and Mary. The Virgin is depicted actively engaging with her son who is reaching for a rose in her hand (http://www.taftmuseum.org/collectionT.php?pieceid=250). While the rose here may be a Marian symbol, the key issue nonetheless is the humanity of Christ. Thus, both of these small, ivory, Rayonnant statues represent the human qualities - compassionate, forgiving, loving - attributed to Mary and Jesus.

The naturalism and human character of the Sainte-Chapelle Virgin and Child stands in stark contrast to the Byzantine images of the Virgin and Child from the sixth through the fifteenth century. Throughout Byzantine art, Jesus is usually shown with distinction and occasionally as royalty, even when on the cross. Christ was considered pure divinity and Mary was depicted as merely the human vessel used to transport Jesus to Earth. When shown as a child, there was little interaction between the two. In some instances, as in Virgin and Child Enthroned (Figure 3), Mary even looks solemn as if contemplating the fate of her first born. Additionally, the Christ child is usually depicted more like a small god sitting on Mary’s lap rather than a happy, playful baby ignorant of the future. In the Rayonnant period, this focus on Christ’s divinity shifts. Mary’s role becomes more than merely a vessel or a lamenting mother, but rather more of a real person with real human experiences.

The Sainte-Chapelle Virgin and Child is also dramatically different from the Romanesque (tenth to twelfth century) images of the Virgin and Child. Throughout this time, Mary
gained esteem as the ‘Throne of Wisdom’ (seen in the wooden sculpture in Figure 4) on which Christ sat. Daniel H. Weiss interprets the ‘Throne of Wisdom’ imagery as communicating the “[i]ncarnation, God’s living appearance on earth,”12 in other words, communicating that Mary was just a physical representation of Jesus’ life on earth. Although Christ is smaller in scale to his mother, Mary’s frontal pose highlights Jesus as the focal point of the statue, reducing the sentimental affection found in the later Virgin and Child works.

The ivory Sainte-Chapelle Virgin and Child is thus more than just an example of Rayonnant style. It is a work that demonstrates the intersection of style and iconography with patronage and function. The elegance and refinement of the style displays the aesthetics of its aristocratic owner, Louis IX. The work’s iconography, (i.e., its fresh view of Christ as human) and its depiction of the warmth and love between Christ and his mother served as a stimulus for empathic response from Louis when he contemplated the work as part of his private devotions. This exquisite work thus gives the viewer a glimpse into an extremely private world, the personal religious experience of the king of France himself.

References Cited

1Marie-Cécile Bardoz. Decorative Arts: Middle Ages
4Louis Grodecki, Sainte-Chapelle, (Paris: Caisse nationale des monuments historiques et des sites, 1979), 34.
6Ibid., 5-7
11Ibid., pp. 65-70

Figure 4. Virgin and Child in Majesty. 1159-1200. Oak, polychromy, gesso, linen
Overall: 31 5/16 x 12 1/2 x 11 1/2 in. (79.5 x 31.7 x 29.2 cm).
Location: The Metropolitan Museum of Art, New York, NY, U.S.A.
Photo Credit: Image copyright © The Metropolitan Museum of Art/Art Resource/ NY Image reference: ART334224

During this Romanesque period, art served more as part of public, rather than private devotion, often within liturgical ritual. With the shift from Romanesque to Gothic, and the shift in sculpture production from Southern to Northern France, the stiff, wooden, seated Virgin was replaced by the triumphant Virgin. Mary seated next to Christ in heaven in scenes of the Coronation of the Virgin. The popularity of the Coronation of the Virgin then led to the widespread trend of the standing Virgin.13 Throughout this evolution, Mary developed her own standing and developed a more interactive relationship with her infant son.
The paper also makes an important contribution to the field of art history in its consideration of how the function of the work – its use to stimulate the private prayers of its owner, Louis IX of France – relates both to its elegant style (preferred by the French elite) and its iconography, in which the Virgin and Child are depicted in very human ways, designed to stimulate empathic response within the experience of personal devotion. As such, the paper embodies current methodologies within art history, which have moved away from independent assessment of style and iconography, to a more nuanced sense of how both these features of an art work are embedded within issues of patronage and function. Harrington’s paper also engages with other contemporary issues relating to this statue, notably issues of gender associated with the rise of the cult of the Virgin, and issues in the relation between art and its audience.

Ms. Harrington’s study of the Sainte-Chapelle Virgin and Child thus forms an example of a paper that blends careful observation with thorough research, and object-oriented analysis with an understanding of the full cultural context in which that object is embedded. It not only gives us a better understanding of the work under consideration, but also a better understanding of the societal role played by the object. These accomplishments represent art history at its best.
RUNOFF USING A COMBINED GEOGRAPHIC INFORMATION SYSTEM AND CURVE NUMBER APPROACH

By Keshia M. Koehn
Department of Crop, Soil and Environmental Science

Faculty Mentor: Kristofer R. Brye
Department of Crop, Soil and Environmental Science

Abstract

Stormwater runoff can transport nutrients, sediments, chemicals, and pathogens to surface water bodies. Managing runoff is crucial to preserving water quality in rapidly developing urban watersheds like Northwest Arkansas. A watershed containing much of the University of Arkansas campus was identified as the target area for this study because stormwater from this location drains into the West Fork of the White River, designated as an impaired water body due to siltation. The project objective was to develop a methodology to test existing stormwater drainage infrastructure, identify potential areas of improvement, and estimate potentially contaminated runoff by combining two widely used prediction models. The U.S. Department of Agriculture’s Natural Resource Conservation Service’s curve number (CN) method was used to estimate runoff depths and volumes, while a flow-direction model integrated topography, land use, and stormwater drainage infrastructure in a geographic information system. This study combined the CN and flow-direction models in a single geodatabase to develop flow direction/quantity models. Models were developed for 5-, 10-, 25-, 50-, and 100-year floods and varied by the antecedent moisture content. These models predicted flow directions within existing drainage infrastructure, runoff volumes for each flood, and a hypothetical flood analysis model. Results showed that between 24,000 m³ (5-year flood) and 60,000 m³ (100-year flood) of runoff would be transported to the West Fork of the White River. The methodology developed and results generated will help stormwater planners visualize localized runoff, and potentially adapt existing drainage networks to accommodate runoff, prevent flooding and erosion, and improve the quality of runoff entering nearby surface water bodies.

Introduction

After a precipitation event, stormwater runoff transports nutrients, chemicals, sediments, and pathogens to surface water bodies. Managing this stormwater is crucial in preserving water quality, especially in urban watersheds under heavy development typified by Benton and Washington Counties in Northwest Arkansas. Northwest Arkansas is the home of several large-scale corporations such as Wal-Mart, Tyson Foods, and J.B. Hunt Transportation. According to the 2000 Arkansas Census, Benton County’s population experienced a 57% increase and Washington County experienced a 39% increase from the 1990 Arkansas Census, adding more than 500,000 people in both counties (US Census, 2000). Washington County is also home to the University of Arkansas, the land-grant institution in the state. Since 2003, more than nine buildings have been restored and more than twenty buildings have been erected at the University of Arkansas (Facilities Management Planning Group, FMPG, 2007). Of these newly constructed buildings, there were three dormitories built to house an increasing student population. Like the population growth in Northwest Arkansas, the student enrollment at the University of Arkansas has increased by 21%, from 15,396 in 2000 to 18,647 in 2007 (Voorhies, 2007).

All stormwater runoff within the watershed surrounding the University of Arkansas in Fayetteville flows into the West Fork of the White River, a tributary of Beaver Lake, the source for much of Northwest Arkansas’ drinking water. However, the West Fork of the White River has been identified as an impaired stream by the Arkansas Department of Environmental Quality due to “high turbidity levels and excessive silt loads” that were creating an aquatic environment incapable of supporting adequate life (ADP&E, 2007).

After the West Fork of the White River was placed on the Arkansas 303(d) list for impaired water bodies in 1998, it was determined that sediment loads were originating from a variety of sources including stream bank erosion, local construction and development, pastures, forests, and urban areas (Formica et al., 2004). According to a best management practice (BMP) study by the Environmental Protection Agency (EPA), average sediment contributions to surface water bodies ranged from 213 million mg ha⁻¹ yr⁻¹ (190 lbs ac⁻¹ yr⁻¹) in medium-density residential areas to 1.21 billion mg ha⁻¹ yr⁻¹ (1000 lbs ac⁻¹ y⁻¹) in commercial areas (USEPA, 1999). With the current high rate of expansion and development in Northwest Arkansas, strategic modeling and planning of stormwater runoff plays a critical role in preserving the quality of surface water.

One of the most common stormwater runoff prediction models is the U.S. Department of Agriculture’s Natural Resource Conservation Service’s (NRCS) curve number (CN) method (Thompson et al., 2003). Formerly known as the Soil Conservation Service (SCS) Method, the CN method calculates a net runoff depth for a specific amount of precipitation. This method is based on estimations of net runoff after initial losses of accumulated rainfall due to soil storage, interception, and infiltrated runoff (SCS, 1972).

Several parameters are used to determine the CN for an area. The land use of an area, or amount/type of surficial cover, can be used to determine the amount of runoff that can be intercepted and/or infiltrated. For example, a paved area would...
have greater runoff than a grassy area, which would result in greater infiltration. Land use can also include land treatment in agricultural settings where crop rotations, contouring/terracing, and the amount of grazing and burning affect the quantity of stormwater runoff (Anonymous, 2007).

Soil properties can also have an effect on the depth of stormwater runoff. The hydrologic soils group (HSG) is a classification of soil moisture based on the quantity of water that is able to infiltrate the soil, which in turn is influenced by the condition of the soil surface and the soil profile horizonation. Also included would be values for slope, texture, and hydraulic conductivity (USDA, 2007). Table 1 provides the HSG classification definitions by the NRCS.

Antecedent moisture content (AMC) is another soil property that has a significant effect on the quantity of stormwater runoff. AMC is defined as the level of soil moisture before a precipitation event and is divided into three classes: AMC I, AMC II, and AMC III (Anonymous, 2007). A soil with AMC I conditions is described as considerably dry, but not to the plant. A soil having AMC II conditions is described as having an average soil moisture condition, and AMC III conditions correspond to a soil that is nearly saturated (Novotny, 1995).

<table>
<thead>
<tr>
<th>HSG</th>
<th>Description of Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Soils having a high infiltration rate (low runoff potential) when thoroughly wet. These consist mainly of deep, well drained to excessively drained sands or gravelly sands. These soils have a high rate of water transmission.</td>
</tr>
<tr>
<td>B</td>
<td>Soils having a moderate infiltration rate when thoroughly wet. These consist chiefly of moderately deep or deep, moderately well drained or well drained soils that have moderately fine texture to moderately coarse texture. These soils have a moderate rate of water transmission.</td>
</tr>
<tr>
<td>C</td>
<td>Soils having a slow infiltration rate when thoroughly wet. These consist chiefly of soils having a layer that impedes the downward movement of water or soils of moderately fine texture or fine texture. These soils have a slow rate of water transmission.</td>
</tr>
<tr>
<td>D</td>
<td>Soils having a very slow infiltration rate (high runoff potential) when thoroughly wet. These consist chiefly of clays that have high shrink-swell potential, soils that have a permanent high water table, soils that have a claypan or clay layer at or near the surface, and soils that are shallow over nearly impervious material. These soils have a very slow rate of water transmission.</td>
</tr>
</tbody>
</table>

Table 1. Natural Resource Conservation Service hydrologic soils group (HSG) classification descriptions (USDA, 2007)

The amount of precipitation, land use, HSG, and AMC are used in the CN method to calculate the net runoff for an area. This method is based upon the assumption that each soil-landcover combination produces a separate curve number that can be used on catchment areas up to 1000 km² (Williams, 1976).

This project was designed to develop a stormwater runoff prediction model in an effort to simulate non-point source contamination of local rivers, like the West Fork of the White River. This study sought to develop a methodology to test existing stormwater drainage infrastructure, identify potential areas of improvement, and estimate potentially contaminated runoff volumes by combining two widely used prediction models [i.e., the NRCS CN method integrated with a geographic information system (GIS) modeling approach].

With the methodology in place, the objective of this study was to evaluate the effect of AMC on stormwater runoff for 5-, 10-, 25-, 50-, and 100-year flood events. The study area was defined as the portion of the University of Arkansas main campus in Fayetteville that contributes stormwater runoff and potential pollutants to the West Fork of the White River. It was hypothesized that the effects of AMC would increase as the flood return period increased.

Methodology

Study Area

The study area was located in the City of Fayetteville, Washington County, Northwest Arkansas, and is a delineated sub-watershed of the NRCS 12-digit Hydrologic Unit Code (HUC) Town Branch – West Fork – White River Watershed (11010001004). This sub-watershed has an area of 320 ha (~800 acres) and contains most of the University of Arkansas’ main campus.

This site was chosen because of the availability of necessary data, the diversity of land uses within the sub-watershed, and the fact that this area is a reasonably representative model of small-scale, rapid development. The region encompassing the actual study area is situated in the Ozark Highlands, where geologic ages of the underlying stratigraphic layers range from the Late-Mississippian to the Middle Pennsylvanian sandstone with underlying sequences of shale, siltstone, and limestone (USGS, 2007). Soil data for the study area, obtained from the NRCS Soil Survey Geographic (SSURGO) Database (Soil Survey Staff, 2006), indicate the most common soil surface textural class present in the study area is fine sandy loam.

The area of study was divided into four quadrants: the Northern, Western, Eastern, and Southern regions. Stormwater runoff from these quadrants flows into College Branch Creek and is transported to the West Fork of the White River (Figure 1).

The Northern region of the study area is characterized by many construction zones, residential areas, parking lots, and sports facilities, including Reynolds Razorback Stadium. This region’s stormwater runoff may contain sediments from disturbed topsoil in construction zones, chemicals from paints and fertilizers used in the stadium, pathogens from animal litter in parks and residential areas, and trash, oils, and heavy metals from parking lot runoff.

The Western region contains residential areas in addition to large wooded areas. Stormwater runoff from residential areas can transport pathogens from animal litter, chemicals from drained pools, and heavy metals from leaking vehicles, trash, and yard waste to surface water bodies. Wooded areas without development can transport leaf litter in runoff water that can clog stormwater drains. Clogged stormwater drains become a problem when runoff is forced to flow over the land surface, contributing to surficial erosion and...
Figure 1. Land use quadrants within the study area and characterized by potential sources of runoff contamination. These sources include parking lots, construction zones, residential areas, and athletic fields.

These five land uses are defined as follows: impervious surfaces include paved parking, building infrastructure, and major roadways; woodlands include light woods, wood-grass combination, and tree farms; grasses include parks, golf courses, greenways, and grazed pastures; bare soil includes gravel parking, quarries, and land under development; and water corresponds to any open surface water body (Halley, 2007).

Positions of stormwater features and infrastructure for the University campus and City of Fayetteville were identified and differentially corrected (post-processed code) using a Trimble GeoExplorer XT GPS unit. Collected stormwater features included intakes like storm grates, linear grates, culverts, area drains, floor drains, roof drains, and curb inlets. Outflow features included outflow pipes and culverts. Other stormwater features which did not play an active role in the flow direction analysis of runoff included manholes and cleanout features. Locations and attributes of pipelines were provided by the University of Arkansas Facilities Management and the City of Fayetteville GIS Laboratory.

The study area was delineated from three NRCS 12-digit HUC watershed boundaries using the LIDAR DEM. These three boundaries were selected based on their spatial proximity to the central campus: the Hamestring Creek watershed (111101030203), the Town Branch – West Fork watershed (110100010404), and the Mud Creek – Clear Creek watershed (111101030202).

Flow Direction Analysis

The first stage of synthesizing the model was to establish the flow direction of stormwater runoff within the existing drainage infrastructure. A flow direction model provides a way for planners to assess areas needing improvement and to aid in tracing potential contamination pathways.

Spatial and physical connectivity between stormwater pipelines and features was established after creating a geometric network using ArcGIS version 9.2 [Environmental Systems Research Institute (ESRI), Redlands, CA]. Geometric network development enabled the complex linear edges and point features of the existing stormwater drainage network to operate as a complete system. Weights were added to the geometric network such as pipe lengths, diameters, and elevations. In addition, the material used to construct the infrastructure was recorded. Using ArcHydro, and extension of ArcGIS, these weights were used to design a set of algorithms that were able to establish flow direction within the desired stormwater drainage network. After determining the flow direction of stormwater runoff within the drainage network, the Utility Network Analyst toolset of ArcGIS was used to determine sample contamination pathways in addition to the lengths of hypothetical contamination pathways.

Runoff Depth and Volumetric Quantification

It was necessary to generate a set of spatially distributed CN for the entire study area in order to estimate the potential amount of stormwater runoff for a specific precipitation event.
Curve numbers were generated using the NRCS CN method, an impervious surface map, surface topography data from the DEM, soils information for the area, and local precipitation data.

Volumes of runoff were calculated in 50 sub-watersheds using the net runoff depths from the CN method. ArcHydro [Environmental Systems Research Institute (ESRI), Redlands, CA] was used to delineate sub-watersheds based on the stored flow direction of the stormwater pipelines and key outflow storm features. Because the stormwater drainage network was used instead of a surface water network, sub-watersheds were developed based on flow directions within the pipelines.

The precipitation amount, initial abstraction (i.e., amount of runoff lost to infiltration, interception, and possible evaporation), and potential maximum subsurface storage were inserted as additional sub-watershed attributes. These parameters were used to estimate runoff using the CN method. Precipitation data were derived for Northwest Arkansas for five flooding recurrence intervals, 5 (142 mm), 10 (159 mm), 25 (184 mm), 50 (203 mm), and 100 year (225 mm), provided by the U.S. Weather Bureau (1948). These precipitation data were used in the CN method to provide a base level for net runoff.

**Potential Maximum Flood Model Analysis**

The potential maximum flood model was designed to be a hypothetical scenario, created to visualize runoff movement through the watershed. To assess the potential maximum flooding depth, it was necessary to generate a flood environment confined by the study area boundaries. The initial model assumed no addition of runoff water from surrounding watersheds and no loss of runoff water from the study area to surrounding watersheds. The depth of ponding in this model was used to visualize localized runoff and to determine areas that had volumetrically high runoff.

In order to calculate the total volume, or volumetric capacity, of individual sub-watersheds, ArcHydro was used to determine the volume of an empty sub-watershed based on the summation of a series of volumetric slices calculated from the original DEM at 1-m contour intervals. The calculated empty-basin volumes were added to the sub-watershed attribute table. In addition to the empty basin volumes, the maximum and minimum elevations were also appended to each sub-watershed’s attribute table.

After calculating the volumetric capacity of the study area and flow direction of runoff, the percent volume filled by runoff water in the sub-watersheds was calculated using the volumes of each empty sub-watershed basin. These volumes were determined from the DEM volumetric slice analysis and each corresponding volume of runoff calculated by the CN method. By calculating the percent of each subwatershed filled by runoff water, the maximum elevation attained by ponded water in the “bowl” environment could be calculated.

Using the maximum fill elevations for the study area as the reference, ArcScene (ESRI) was used to create a three-dimensional, seamless model capable of representing the advance and retreat of runoff water in the study area, assuming it was a confined environment. The depressionless DEM, aerial photography, and the sub-watershed and study area vector boundaries were imported into ArcScene (NAD State Plane 1987 FIPS 301 Feet, Coordinate System). The base heights for the photography and vector boundaries were set equal to the heights of the DEM in order to achieve a seamless, three-dimensional model.

The base heights for the “flood” layer were not set to that of the DEM, but instead to the minimum elevation of the study area. This boundary created a moveable, planar layer that was able to simulate the flooding capacity of runoff within the study area. This height of the “water” layer was set to the starting position of “No Flooding.” For each corresponding flood interval, this layer’s base height increased relative to the maximum height of ponded water. ArcScene was also used to model the maximum depth of ponded water for 5-, 10-, 25-, 50-, and 100-year flood recurrence intervals.

**Results and Discussion**

**Flow Direction Tracing**

The flow direction prediction model that was developed for this study is important for planning officials at the University of Arkansas and City of Fayetteville on a small scale. However, even though the study area only covered approximately 320 ha (800 acres), the methodology used to develop this flow direction prediction model can be repeated for larger scales. Since transportation routes of potentially contaminated runoff can be visualized, the model can be used to trace accidental spills, re-route runoff to treatment facilities, and to identify locations near outflow features that may be particularly sensitive to contamination.

**Curve Number Analysis**

A composite CN map was generated for each level of AMC (AMC I, II, and III) using the impervious surface and soils maps. Each land-use/soil-group polygon was assigned a CN to be used to calculate a weighted CN average for each sub-watershed. The weighted CNs calculated for each of the 50 sub-watersheds were used to determine the impact of land use and soils on the amount of direct runoff. The CN maps generated for each AMC condition illustrate not only the abundances of low-permeable land uses in the study area, but also the effects these land uses have on the volume of runoff water being transported into the West Fork of the White River.

Upon observation of the weighted, sub-watershed CN map with reference to the impervious surface map, areas that have the same land-use category, but different HSG, have different local CN. This is shown in the Western quadrant of the study area within the wooded land-use area. This situation can be compared to a sandy soil and a clayey soil under tree cover, because each soil texture has different infiltration and water-holding capacities that affect the amount of runoff. The HSG of an area has a direct effect on the CN-runoff
relationship—sandy soils are capable of being more permeable than clayey soils, thus sandy soils have a greater capacity to filter runoff water. These areas are particularly important in filtering contaminated stormwater runoff and are comparable to the drain field of a septic system. Increasing areas that have the capability of runoff infiltration and decontamination can decrease contaminant loading to nearby surface water bodies, such as the West Fork of the White River.

The mean weighted CN for the sub-watersheds increased with each increase in AMC level. Three maps were developed depicting the weighted CN for each sub-watershed varied by the level of AMC (Figure 2). The three maps show an increase in the average weighted CN for each increase in AMC level supported by the areal extents of higher CN. There is a clear relationship between the weighted CN and the AMC level—as the soil water content (i.e., AMC) increases, the weighted CN for a watershed also increases, meaning more runoff will occur because the soil has a decreasing capacity to store more infiltrating water as the water content approaches saturation (i.e., AMC III).

The minimum calculated curve numbers for AMC I, II, and III conditions were 52, 71, and 85, respectively. The maximum calculated curve numbers for AMC I, II, and III conditions were 93, 98, and 99, respectively. Average curve numbers for AMC I, II, and III conditions were 79, 89, and 94, respectively.

Figure 2. Sub-watershed weighted curve numbers for varied AMC levels. For each map, left to right, the AMC was increased, increasing the average sub-watershed CN and runoff depths.

Runoff Depth and Volume Analysis

The CN method was used to determine the amount of net runoff for each sub-watershed for a specific single-storm event. Net runoff increased for each increase in flooding-recurrence interval. As expected, for each increase in precipitation amount, there was a corresponding increase in runoff because of the decreasing soil storage capacity. Modeling depths of runoff from precipitation data is important in visualizing the quantity of stormwater being transported by existing infrastructure given the soil moisture condition (i.e., AMC I, II, or III) at the time of the event.

For each increase in AMC level, as shown on the weighted sub-watershed CN map, there was a corresponding increase in the net runoff for the same precipitation event. Table 2 summarizes the effects of increasing precipitation on the stormwater runoff in each sub-watershed per AMC level.

Net runoff depths calculated from precipitation data were used to calculate the volumes of water associated with a specific precipitation event. Runoff data showed that each increase in precipitation was associated with an increase in the volume of runoff water. In addition, the increasing trend in the volume of stormwater runoff was directly related to the net depth of runoff and to the soil moisture condition (i.e., AMC level). Calculated stormwater runoff volumes were combined with the flow direction model to visualize maximum flood water retention within the study area.
conditions, respectively (Figure 3). This is directly related to the decreasing soil storage capacity as the AMC increases.

For each increase in AMC level, there was an increase in the area extent covered by ponded water. For lower-magnitude flood-recurrence intervals (i.e., 5- and 10-year), there was a smaller range of area covered by runoff water. For greater-magnitude storm events (i.e., 25-, 50-, and 100-year floods), there was a greater range of areas covered by ponded water (Table 3).

### Table 3. Cumulative runoff depths and volumes for study area calculated by combining GIS calculations with the curve number method

<table>
<thead>
<tr>
<th>Variable</th>
<th>5-yr Flood (m)</th>
<th>10-yr Flood (m)</th>
<th>25-yr Flood (m)</th>
<th>50-yr Flood (m)</th>
<th>100-yr Flood (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depth</td>
<td>4313</td>
<td>5031</td>
<td>6159</td>
<td>7021</td>
<td>8011</td>
</tr>
<tr>
<td>Volume</td>
<td>24758</td>
<td>28975</td>
<td>35625</td>
<td>40716</td>
<td>46571</td>
</tr>
</tbody>
</table>

The average height of rise of flood water in the hypothetical closed “bowl” study area was 0.95 meters (3.11 feet). This is a relatively shallow depth of water, but its magnitude increases greatly when distributed over a low-relief region of the study area. As previously described, this model was conceived in a hypothetically closed system and was able to represent the height of rise and area extent covered by ponded runoff water. In reality, there would also be simultaneous additions and losses of runoff water in the study area that likely would keep maximum flooding depths lower than that predicted by the model.

### Significance of Research

This project’s objective was to develop a methodology to test existing stormwater drainage infrastructure, identify potential areas of improvement, and estimate the volume of potentially contaminated runoff by combining two widely used prediction models. Using these developed models as a guide, planners at the University of Arkansas and City of Fayetteville can work to improve the quality of runoff water being transported to the West Fork of the White River. Using the flow direction model and estimated runoff volumes, stormwater drainage infrastructure can be improved in sub-watersheds that have been shown to contribute the largest volumes of runoff from the study area and areas experiencing or predicted to experience localized flooding and soil erosion on land and along stream banks.

Currently, stormwater drainage infrastructure within the study area releases untreated runoff water directly into College Branch Creek, which is a tributary of the West Fork of the White River. During storm events, nutrients, chemicals, sediments, and pathogens are ultimately carried through the

---

**Figure 3.** Maximum flooding extent for varied AMC levels showing the locations of potential flooding on aerial photography of the southern quadrant of the study area.
The runoff flow direction model of the existing stormwater drainage network is particularly useful for tracing possible point-source contamination. In the case of accidental spills or leaks, planners can accurately trace the contamination pathway through the pipeline infrastructure. In addition, drainage infrastructure surrounding potentially hazardous areas can be modified to transport contaminated water away from surface water bodies in the likelihood of a spill or leak. Finally, the runoff flow direction model can be used to decrease localized flooding by re-routing runoff water away from low-relief areas that have greater potential to accumulate runoff from a precipitation event.

Because of the current rate of expansion at the University of Arkansas, construction sites are prevalent throughout much of central campus, contributing to the sediment loading of the West Fork after precipitation. The amount of sediment in stormwater runoff could be decreased by installing various BMPs that increase the percentage of permeable land, such as permeable pavement and green roofs and a series of detention ponds or grassy swales to slow the water velocity and allow sediment to drop out of suspension before entering College Branch Creek. By constructing monitoring stations along key points in the stormwater drainage network, areas with high sediment or contaminant loads could be re-routed using the flow-direction model of the existing drainage infrastructure.

Conclusions

In rapidly developing urban watersheds, improperly managed stormwater runoff can degrade surrounding surface water bodies. In Northwest Arkansas, sediment-laden stormwater runoff is transported to the West Fork of the White River, a surface water body impaired by siltation. This project established a repeatable protocol that resulted in a stormwater prediction model that was varied by potential soil moisture conditions in an effort to simulate non-point source contamination of local rivers, such as the West Fork, from urban stormwater drainage networks. The AMC of the soil had a direct effect on the amount of stormwater runoff from the study area because the soil had a decreasing capacity to store additional infiltrated water as the soil water content approached saturation.

The methodology developed by this research can be used to test existing stormwater drainage infrastructure and identify potential areas of improvement and to estimate the volume of potentially contaminated runoff. The runoff flow-direction model will be particularly useful in tracing point-source contamination within the stormwater drainage network. Volumes of runoff water from specific storm events, calculated using the CN method, can be used to gauge the effects of semi-permeable land uses on the quantity and quality of runoff transported to surface water bodies.

Designing a stormwater runoff prediction model that includes both the water flow direction and quantity of water transported is essential for not only urban stormwater management planners, but also city utility officials and urban...
developed. In Northwest Arkansas, larger-scale replicates of these prediction models could play a crucial role in preserving the quality of surface water bodies like the West Fork of the White River. With the advent of cost-effective monitoring programs, BMP construction, and education to improve water quality, the West Fork of the White River could eventually be removed from the impaired water bodies list and have its biological productivity return to normal.

Acknowledgements

I would like to sincerely thank my honors thesis committee, Ms. Cristina Scarlat, Mr. Vaughn Skinner, and Dr. Duane Wolf for their help with this project. I would additionally like to thank my mentor, Dr. Kristofor Brye, for his tireless efforts to improve the quality of my research. I also could not have finished this project without the patience and the understanding offered by my family and friends. This project was funded by the State Undergraduate Research Fellowship (SURF) and by the Dale Bumper’s Undergraduate Research Data Grant. Data contributions were made by the University of Arkansas Facilities Management and the City of Fayetteville GIS Laboratory.

LITERATURE CITED


http://scholarworks.uark.edu/inquiry/vol9/iss1/1


Mentor Comments

Dr. Kristofor Brye underscores the interdisciplinary nature of Keshia Koehn’s work and draws attention to the originality of her approach and practical impact of her findings.

Ms. Koehn undertook a rigorous project for an undergraduate that was only possible because of Ms. Koehn’s unique combined background in soils, environmental science, hydrology, and GIS. Very few undergraduates possess the abilities Ms. Koehn has acquired while still being an undergraduate. Ms. Koehn is a highly independent student. As such she worked independently on her research project getting only minor assistance from a few colleagues she works with in the Center for Advanced Spatial Technology (CAST). Ms. Koehn conducted a significant amount of field work to validate electronic data she used in her study and protocol development.

The nature of Ms. Koehn’s research project was significant in that it demonstrated the usefulness of combining various types of information in a GIS to analyze several very real, potential harmful, environmental scenarios. The end product of Ms. Koehn’s research project was a working model for land-use planners to use and an adaptable
template that can be applied to similar, but yet somewhat different, scenarios and conditions for preparing for and managing storm water runoff in an urban setting. The model she developed will be implemented and used by Facilities Management to improve existing storm water infrastructure and minimize potential contaminant transport from entering the West Fork of the White River that may be originating from the University of Arkansas's main campus area.

Ms. Koehn's accomplishments with this research project are highly unique and original in that this approach had not been attempted previously due to the lack of time and resources. Ms. Koehn's work on this project will have saved Facilities Management a substantial amount of time and money and will likely make a significant impact on how storm water and potential contaminants from the U of A campus are managed. Ms. Koehn developed her research topic on her own with little guidance from me. The area of research that she pursued was tangential to my area of expertise and interest, but we decided early on in the process that we would go through this process together, as a team, learning along the way and hopefully making the right decisions. Ms. Koehn's was true to this from the beginning as she asked many questions and thought about every decision that needed to be made along the way.
THE IMPACT OF SEVERANCE TAXES ON THE ARKANSAS NATURAL GAS MARKET

By Tammy Lippert
Department of Economics

Faculty Mentor: Katherine Deck
Center for Business and Economic Research

Abstract

The recent development of the Fayetteville Shale Play, an unconventional natural gas reservoir in Central and Eastern Arkansas, has created considerable opportunities for the state and its citizens as the industry has made substantial investment in the region. These developments have resulted in thousands of new jobs for Arkansans, billions of dollars in direct and indirect output, and millions in state and local tax revenues. One of the most visible issues in recent state news has been the controversy surrounding the severance tax levied by the state government on the extraction of natural gas. The question at hand has been whether or not to increase Arkansas’s rate. The state has had the lowest severance tax incidence in the nation causing many to speak out for a raise in the tax rate to something comparable to surrounding states in the region. These demands caught the attention of Arkansas Governor Beebe who worked with natural gas companies to find a reasonable severance tax package including some discounts for shale play wells. The legislature shortly thereafter approved this increase to be enacted January 1, 2009.

The purpose of this study was to evaluate the state’s reaction to these recent events and offer any additional recommendations that may enhance this set of decisions. A comparison of Arkansas, Kansas, Louisiana, Oklahoma, and Texas was conducted on multiple levels of economic conditions to evaluate the overall tax structure within each state. Outside research was also considered in forming these conclusions. After completing an extensive cross-state comparison and incorporating econometric research, it was determined that the Arkansas state legislature was justified in increasing the severance tax rate. However, the rationale for tax increases – specifically the reasoning that other states have higher severance tax rates – is somewhat flawed based on consideration of economic conditions, natural gas production numbers, and overall tax structures. A stronger rationale lies in additional research that suggests that an increase in liabilities for severance taxes yields minor changes in investment and drilling activity and potentially positive economic rewards.

Introduction

Costs of energy continue to soar and fears of exhausting natural resources endure as a legitimate concern. The search for alternative energy sources and less costly extraction is being pursued across the entire nation, as states and companies attempt to gain an advantage in the volatile energy market. The state of Arkansas currently has a unique opportunity to develop its previously meager natural gas market, since an unconventional gas reservoir called the Fayetteville Shale has recently been determined to be economical for gas extraction. Although production of natural gas is fairly new to Arkansas, the state does have a modest history in natural gas extraction. According to the Arkansas Geological Society (2008), natural gas was first discovered in Fort Smith in 1887, and subsequently commercially developed in 1902 near Mansfield. As natural gas exploration continued, the 1923 state legislature passed Act 118 levying a tax for severing the natural resource from the state (hence, the name severance tax) at 2.5% of cash market value, according to Ernest Dumas of the Arkansas Times (2008). Since then, there have been fluctuations in the amount and manner of tax levies on natural gas, in part due to the involvement of prominent state families in the gas industry.

Shale deposits were first recognized as a legitimate source of natural gas as early as the 1980’s, yet the concern had been the difficulty and expense of withdrawing the resource. As a result, shale gas had yet to be utilized as a supply of natural gas in the Arkansas market. However, as commodity prices continued to rise and the advancement of technology persisted, these concerns gradually began to shrink. Finally in 2004, Southwestern Energy Company announced successful drilling and production of gas from the shale. This demonstrated that the once inefficient reserve could perhaps be further developed to accommodate the growing demand for energy resources. Recently, other oil and gas companies have followed suit by seeking stake in the development of this emerging market and newly available supply, and the state has experienced significant investment as a direct consequence. The result has been the identification of what is called the Fayetteville Shale gas reservoir. This reservoir lies on the Arkansas side of the Arkoma Basin and ranges in thickness from 30 to 325 feet and ranges in depth from 1,500 to 6,500 feet. It runs across central and eastern Arkansas under multiple counties including Cleburne, Conway, Faulkner, Independence, Johnson, St. Francis, Prairie, Van Buren, White, and Woodruff.

With the development of the Fayetteville Shale reservoir, natural gas production has become a driving force in the Arkansas economy and the tax consequences of this production have been brought to the attention of the Arkansas legislature. Severeance taxes are perhaps the most relevant and are generally levied upon non-renewable resources that are removed from the earth. In Arkansas, either the producer or the purchaser of natural resources is assessed the fee at the time of severance. Arkansas’s severance tax rate has been among the lowest in...
the country at three-tenths of one cent per thousand cubic feet generating only an approximate $600,000 annually in general state revenue, according to the Arkansas Department of Finance and Administration (2008). Questions now arose regarding changes in the state's severance tax and details of how much to increase the rate, whether tax breaks or incentives would be offered, and where the revenue should be allocated after collection.

Sheffield Nelson, former gas company executive and former state chairman of the Republican Party, was among the first to recognize this opportunity and propose a plan to increase the state severance tax. Nelson prepared a ballot initiative for the November 2008 general election that would increase the severance tax rate to 7% of the market value of natural gas at the time of its extraction, as reported by John Brummett of the Arkansas News Bureau (2007). The revenue generated would be applied to higher education, highways, and local aid. Arkansas Attorney General, Dustin McDaniel, approved the initiative in February 2008, and Nelson quickly began to gather the approximately 62,000 required signatures by the July 7, 2008 ballot-qualifying deadline.

In addition to Nelson's proposal, Governor Mike Beebe diligently worked towards reaching an agreement with gas companies, according to Mark Hengel of ArkansasBusiness.com (2008). The initial pitch to gas companies and legislators was a smaller rate increase to be approved through a special session that Beebe would call only if he garnered a commitment of the necessary three-quarters majority to approve his plan. Revenues generated would be dedicated strictly to roads as an estimated $19 billion will be necessary to repair state highways and bridges over the next 20 years. After a seeming standstill in early March, negotiations with the industry halted, and Beebe began the process to submit another ballot initiative to compete with Nelson's.

Within a few short weeks, however, the Governor announced that he had finally reached an agreement with the natural gas companies to increase the severance tax rate to 5% of market value with exemptions for new discovery, high-cost, and marginal gas wells. These exemptions lower the rate for "high-cost" wells, which currently account for about 38% of Arkansas wells, to 1.5% for the first three years and "marginal" wells, about 56% of Arkansas wells, to 1.25%. This leaves only about 5% of the state's wells to be taxed at the 5% base rate, although projections have been released that estimate approximately 12% would qualify under this rate by January 1, 2009, when the increase will go into effect. On March 31, 2008, a special session of the legislature was convened and the increase was approved three days later by the legislature and signed into law by the governor. New severance tax revenue will be allocated 95% to road improvements – of that, 70% is distributed to state highways, 15% each to cities and counties, and 5% to replace the current tax that goes into the general revenue fund. Nelson had said that he would pull his initiative from the ballot should the legislature pass the Governor's proposal, as some of his intention behind the measure was to encourage the gas industry to negotiate with Beebe. He withdrew his bill immediately thereafter approval of Beebe's bill.

With this development of a new energy market in Arkansas and associated increases in severance tax revenues, there is a need for evaluation of the current severance tax structure in terms of generation of state revenue, encouragement of development, preservation of natural resources, and other indirect impacts. Comprehensive analysis of these issues is required if reliable recommendations are to be made concerning what is best for the emerging natural gas market and the state of Arkansas as a whole.

Some research has already been completed at the Center for Business and Economic Research (CBER) of the Sam M. Walton College of Business at the University of Arkansas. In a 2006 report addressing the economic impact of development from the Fayetteville Shale Play, projected economic outcomes included: state-wide economic activity of approximately $5.5 billion; the creation of nearly 10,000 jobs; and the generation of approximately $350 million in state and local tax revenues for the period from 2005-2008. In March of 2008, the CBER issued an update reporting that previous projections considerably underestimated the economic impact of natural gas exploitation and revising earlier projections. This update also reported data derived from a survey of natural gas companies. The survey instrument probed the impact of severance tax increases on economic development in many arenas.

Mitch Kunce (2003) and several other researchers from the University of Wyoming also completed an extensive econometric study evaluating how effective tax incentives are in encouraging drilling activity. In their study, reduced tax rates led to a substantial decrease in generation of state tax revenue, with mild changes in drilling and production. Kunce (2003) provides several reasons for these outcomes. First, he explains that cuts in severance tax rates offer no 'direct' incentive to increase drilling as these tax cuts are 'downstream' incentives offered at the end of the process. Thus, the benefits to this type of tax cut are only realized if the companies drill and are successful. He also makes the point that 'upstream' incentives may stimulate increased involvement more effectively as they are given at the beginning of the process. Second, because severance taxes at the state level are deductible from federal corporate income taxes, the actual impact here is only a semi-shift from state revenue to federal rather than a full decrease in total liabilities. It was concluded that increasing severance taxes is likely to generate revenue without significantly negatively impacting drilling and production activity.

The purpose of the study reported in this paper is to evaluate the state's reaction to changes in tax severance rates in order to develop additional recommendations that may improve state decision-making. A comparison of Arkansas, Kansas, Louisiana, Oklahoma, and Texas was conducted on multiple levels of economic conditions to evaluate the overall tax structure within each state. Outside research was also considered in forming these conclusions.
Methods

In order to conduct a theoretical analysis of the impact of the severance tax on the natural gas market in Arkansas, this study evaluated the conditions of four other states chosen for their similarities with Arkansas. Similarities included claims of state legislators regarding states’ severance taxes, proximity to the state of Arkansas, natural gas production levels, and overall economic profile. The chosen states were Kansas, Louisiana, Oklahoma, and Texas, with most data collected by relevant federal or state agencies.

It is first important to examine the general economic status of each state to add context to any conclusions to be drawn after evaluation of tax structure. In this case, state populations, per capita real gross domestic product in dollars, per capita personal income in dollars, a cost of living index, and unemployment rates as a percentage are provided in Table 1 for the most recent periods for which data could be found. State population provides context for comparison and allows evaluation of the number of people directly impacted by each state government. Per capita gross domestic product (GDP) demonstrates the total market value of goods and services produced in the state per population. This allows GDP to be compared among states more fairly based on the number of people contributing to the state’s economy. Per capita personal income represents the average income of state residents. Combining personal income with the cost of living is important in order to determine how much income is actually worth in that particular state. The cost of living index evaluates the cost of groceries, housing, utilities, healthcare, transportation, and a basket of miscellaneous goods as compared to the national average. Lower scores, considered superior, are then ranked nationally with a low score suggesting lowest cost to citizens of the state. Finally, unemployment data are provided to factor in the percentage of the adult population seeking employment and unable to find work. This group of economic indicators provides a fairly comprehensive representation of a state’s overall economic well-being.

As another context for results, natural gas production data are useful in evaluating the responsiveness of the industry with respect to severance tax rates and incentives. In Table 2, 2006 Energy Information Administration data are provided for the number of producing gas wells, gross withdrawals in million cubic feet, marketed production in million cubic feet, and wellhead prices in dollars per thousand cubic feet. The number of wells drilled, gross withdrawals, and marketed production provide context for the amount of drilling and production activity within a state. Wellhead price represents the value of natural gas as it is withdrawn from the ground or the price obtained by the producer for sale at the well with a higher price being most advantageous to producers. These figures offer perspectives on the natural gas industry in each state evaluated.

Next, the overall tax structure of each state was evaluated by looking at the severance tax, corporate income tax, sales tax, property tax on gas wells and/or surface equipment, and overall state tax climate. These data are presented in Table 3. The severance tax is the rate at which the state government taxes the extraction of natural gas from the ground within state borders. Corporate income tax represents the rate at which business profit is taxed. Sales tax takes into consideration the rate consumers pay when applicable goods and services are purchased within the state’s borders. Additionally, some states impose a property tax at the state level in addition to the many local and county jurisdictions that do so. It is important to recognize in the present study, as gas wells and/or surface equipment often can be taxed under this category. Finally, this data set considers the overall state tax index for business in 2008 as calculated by the Tax Foundation (a lower ranking is superior). This index evaluates the total tax burden on companies operating within the state. Each of these tax structures contribute to the climate within which the natural gas industry operates in each state.

The remaining set of data gathered for this study focuses on exemptions to the severance tax and specific incentives offered to companies involved in drilling and production activities. Table 4 provides information for royalty deductibility, whether or not exemptions for the base rate are offered for new discovery wells, high cost wells, marginal or inactive wells, horizontal wells, deep wells, and other conditions impacting profitability. In many cases, companies must lease property to drill natural gas, and royalties are then paid to landowners. Some states offer deductions for these payments which can lead to a substantial decrease in tax liabilities. Additionally, each state has different definitions and qualifications for new discovery, high cost, marginal/inactive, and deep wells and varying breaks or incentives, which all have an impact on the cost to companies conducting drilling activities.

After consideration of all of the above data, a scorecard for the state of Arkansas was created employing the balanced scorecard method. As described by the Balanced Scorecard Institute (2008), this procedure has been used since the early 1900’s to consider non-financial measures of business and government in an appraisal of their performance (seen Table 5). This method has become a recognizable research tool since Dr. Robert Kaplan and Dr. David Norton expanded the specific application for business efficiency in the 1990’s. The advantages to using this method include its derivability and its flexibility in categories considered, allowing the researcher to tailor the scorecard to the particular subject being assessed. Due to the complexity and variation in types of data considered within this study, the balanced scorecard method is appropriate to build the theoretical conclusions drawn within this study.

The scorecard developed for this project was created to determine Arkansas’ overall tax structure as compared to the four other evaluated states, including equally weighted categories for corporate income tax, sales tax, property tax, severance tax, and exemptions. In each category, Arkansas was evaluated with respect to the other states as having a higher burden (−), a similar or equivalent burden (0), or a lesser burden (+) on the natural gas industry considering both the...
previous and new severance tax structure. Then the individual categories were compiled into a total burden index with respect to each state.

**Results**

**General Economic Data**

The first set of data provided in Table 1 describe the general economic situation in each state used for comparison. The first statistic evaluated is population. Arkansas has fewer residents than any other state except Kansas which has approximately 50,000 fewer people. Oklahoma is the next closest in population size, followed by Louisiana, and then Texas far above the others at the top. In terms of US ranking, the four closest states all fall within seven places of one another, providing adequate similarity for a comparison of this sort.

Next, with respect to per capita real gross domestic product, Arkansas is at the bottom of the group comparison, almost $10,000 below the national average. Oklahoma is within $200 of Arkansas' GDP and is ranked nationally only one state above Arkansas. There is a similar spread among the remaining states, with Texas deviating less than might be expected based on population size.

All five states are below the national average for per capita personal income. Arkansas is the lowest, a little more than $8000 below the national average. This figure is perhaps the strongest indicator of how the citizens of Arkansas fare in the national economy. Being among the bottom in the nation reflects poorly on the state's economy, although it is necessary to consider personal income in conjunction with the cost of living in the state to determine the significance of this difference.

When examining the cost of living index, there is a change in the trend of finding Arkansas towards the bottom in the nation and even the group for comparison here. Instead, Arkansas leads the nation in the top ten for getting the greatest value for the dollar, but in the middle for the five specific states being considered. This does account for some of the reasoning that Arkansas pulls up the rear in the nation for personal income, since residents can get more for their dollar. However, with respect to this comparison, the state still appears to be at the bottom for the sample group in income when considering this average price index.

The final economic indicator is state unemployment. Arkansas is the only state in this assessment that has a rate greater than the national unemployment rate and more than one percentage point higher than any of the other states in this sample. In general, Arkansas appears to suffer from worse economic conditions than any of the other states considered here - Kansas, Louisiana, Oklahoma, and Texas.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Arkansas</td>
<td>2,810,872</td>
<td>27,875</td>
<td>28,444</td>
<td>90.4</td>
<td>5.4</td>
</tr>
<tr>
<td>Kansas</td>
<td>2,764,075</td>
<td>33,298</td>
<td>34,744</td>
<td>92.0</td>
<td>4.1</td>
</tr>
<tr>
<td></td>
<td>US Rank = 33</td>
<td>US Rank = 34</td>
<td>US Rank = 22</td>
<td>US Rank = 8</td>
<td></td>
</tr>
<tr>
<td>Louisiana</td>
<td>4,287,768</td>
<td>30,798</td>
<td>31,369</td>
<td>95.0</td>
<td>3.8</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>3,579,212</td>
<td>27,963</td>
<td>32,398</td>
<td>89.8</td>
<td>4.3</td>
</tr>
<tr>
<td></td>
<td>US Rank = 28</td>
<td>US Rank = 46</td>
<td>US Rank = 33</td>
<td>US Rank = 4</td>
<td></td>
</tr>
<tr>
<td>Texas</td>
<td>23,507,783</td>
<td>36,277</td>
<td>35,058</td>
<td>89.3</td>
<td>4.3</td>
</tr>
<tr>
<td>United States</td>
<td>301,139,947</td>
<td>36,842</td>
<td>36,629</td>
<td>100.0</td>
<td>4.6</td>
</tr>
</tbody>
</table>

¹US Census Bureau: State and County QuickFacts for 2006
²Bureau of Economic Analysis: Regional Economic Accounts for 2005
⁴Missouri Economic Research and Information Center: Cost of Living Index for 2007
**Table 2 – Natural Gas Production Data (2006)**

<table>
<thead>
<tr>
<th>State</th>
<th># of Producing Gas Wells*</th>
<th>Gross Withdrawals (in mcmf)*</th>
<th>Marketed Production (in mcmf)*</th>
<th>Wellhead Prices ($ per mcmf)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arkansas</td>
<td>3,811</td>
<td>193,942</td>
<td>193,258</td>
<td>6.43</td>
</tr>
<tr>
<td>Kansas</td>
<td>19,713</td>
<td>372,029</td>
<td>371,044</td>
<td>5.61</td>
</tr>
<tr>
<td>Louisiana</td>
<td>17,459</td>
<td>1,378,238</td>
<td>1,361,119</td>
<td>6.93</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>38,060</td>
<td>1,688,985</td>
<td>1,688,985</td>
<td>6.32</td>
</tr>
<tr>
<td>Texas</td>
<td>83,218</td>
<td>6,292,150</td>
<td>5,513,739</td>
<td>6.60</td>
</tr>
</tbody>
</table>


**Natural Gas Production Data**

Table 2 provides the current status of the natural gas industry in each state. Arkansas falls behind each of the comparison states, presumably reflecting the fact that it is a relative newcomer in the production of this resource. It is interesting to note, however, the marked differences between Kansas and Louisiana in the number of producing wells and withdrawals. Although Kansas leads Louisiana in gas wells by more than 2,000 individual wells, Louisiana overtakes production by nearly 100,000 million cubic feet.

The next relevant piece of information is the wellhead price of natural gas by state. Arkansas is situated in the middle of the comparison states, although each state is within 60 cents of one another except for Kansas, which falls at the bottom by nearly a dollar difference.

**Tax Structure Data**

With the understanding of each of the state’s economic and natural gas production environment, the next appropriate area of research is overall tax structure as it relates to companies in the natural gas industry. Table 3 outlines the details of each state’s severance, corporate income, sales, and property taxes. Each of the five states tax the extraction of natural gas at levels varying from 3/10 of one cent per thousand cubic feet previously levied in Arkansas to 7.5% of market value in Texas. The five states also vary in the basis for taxation from production volume to market and gross value. Kansas actually taxes natural gas extraction at a rate of 8% but offsets this high rate with a 3.67% property tax credit. Arkansas indeed held the lowest rate prior to the March 2008 special legislative session which authorized an increase that will go into effect in January of 2009. This increase will raise the rate to 5% of market value and will move the severance tax rate in the middle of those imposed by the five states sampled.

However, to accompany the range in the severance tax, corporate income tax rates also significantly vary from no tax at all in Texas to 8% in Louisiana’s highest tax bracket. Arkansas appears to fall in the middle of the five state group. Also, the basis for levying the tax varies from a flat tax to two, five, or six income brackets with increasing rates. Louisiana differs slightly as it taxes a flat 4% with a 3.5% surtax for incomes greater than $50,000.

The sales tax rate in each of the five states demonstrates less variability ranging only from 4% in Louisiana to 6.25% in Texas. Arkansas nears the top of this set at 6%, just below Texas. Kansas imposes a 5.3% sales tax and Oklahoma a 4.5% rate.

Finally, Arkansas, Kansas, and Louisiana each require state level property taxes which apply to gas wells and/or surface equipment for drilling. Oklahoma and Texas do not; however, local or counties may levy property rates in their respective jurisdictions in all five states.

An interesting index which does appear consistent with these findings is the Tax Foundation’s State Business Tax Climate Index. Included in Table 3 for reference, Arkansas, Louisiana, and Kansas are all similarly ranked at 35, 33, and 32 respectively. Texas offers the most business-friendly tax climate at number 8 nationally, with Oklahoma ranked next at 19.

**Exemptions and Incentives**

Tax breaks for certain types of natural gas wells can result in a substantial decrease in tax liabilities for drilling and production companies in the industry and are thus important to consider when evaluating the tax structure and the impacts it has on natural gas activities within each state. The exemptions and breaks are presented in Table 4.

One break offered by Texas and Oklahoma only is a deduction for royalty payments to landowners. Each of the
Table 3 – Tax Structure Data

<table>
<thead>
<tr>
<th></th>
<th>Severance Tax</th>
<th>Corporate Income Tax</th>
<th>Sales Tax</th>
<th>Property Tax (on Gas Wells and/or Surface Equipment)</th>
<th>State Business Tax Climate Index Ranking (2008)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arkansas¹</td>
<td>Old</td>
<td>$.03/mcf (approx. 0.375% of market value)</td>
<td>1.0% &gt; $0</td>
<td>6.0%</td>
<td>Yes (State-level)</td>
</tr>
<tr>
<td></td>
<td>New</td>
<td>5% of market value</td>
<td>2.0 &gt; 3K</td>
<td></td>
<td>35</td>
</tr>
<tr>
<td>Kansas²</td>
<td></td>
<td>4.33% (8% of market value less .367% property tax credit)</td>
<td>4% &gt; $0</td>
<td>5.3%</td>
<td>Yes (State-level)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7.35 &gt; 50K (3.35% surtax over $50,000)</td>
<td></td>
<td>33</td>
</tr>
<tr>
<td>Louisiana³</td>
<td></td>
<td>$.269/mcf (approx. 3.3% of market value)</td>
<td>4.0% &gt; $0</td>
<td>4.0%</td>
<td>Yes (State-level)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5.0 &gt; 25K</td>
<td></td>
<td>32</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>6.0 &gt; 50K</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>7.0 &gt; 100K</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>8.0 &gt; 200K</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oklahoma⁴</td>
<td>7% gross production &amp; .95% excise tax</td>
<td>6.0% Flat</td>
<td>4.5%</td>
<td>No (Locally Only)</td>
<td></td>
</tr>
<tr>
<td>Texas⁵</td>
<td>7.5% of market value</td>
<td>None</td>
<td>6.25%</td>
<td>No (Locally Only)</td>
<td></td>
</tr>
</tbody>
</table>

¹Arkansas Department of Finance and Administration: Tax Rates
²Kansas Department of Revenue: Tax Rates
³Louisiana Department of Revenue: Tax Rates
⁴OklahomaTax Commission: Tax Rates
⁵Texas Comptroller of Public Accounts: Tax Rates
⁶The Tax Foundation: Tax Data for 2008

The five states also offer different rate discounts with respect to different well characteristics that increase costs to the company. For example, all states except for Texas cut the severance tax for new discovery wells. Also, Arkansas and Texas both offer breaks on high-cost wells which includes shale play wells. All five states discount marginal production or partially inactive wells. Louisiana and Oklahoma also decrease rates on horizontal and deep wells which are often in shale play areas. Texas exclusively offers an incentive to market previously flared or vented casinghead gas wells which increase the risk of repeat instances of decreased productivity in those wells. The decisions to offer discounts are similar in that each state decides to do so, but they differ in the instances in which the discounts are offered. Under Arkansas' previous severance tax structure, no exemptions, breaks, or incentives were offered but they were implemented as the rate was increased.

Arkansas Scorecard

After compilation of all of the data, a scorecard was created in Table 5 to compare the overall tax burden of the natural gas industry in each state as compared to Arkansas with respect to both the old and new severance tax packages. Under the old severance tax rate of 3/10 of one cent per thousand cubic feet with no exemptions or breaks, Arkansas had a similar tax burden to Kansas and Louisiana yet a greater burden than Oklahoma and Texas. Under the new severance tax structure, Arkansas maintained a similar tax burden to Kansas but surpassed Louisiana and remains higher than Oklahoma and Texas. The effects of the rate increase appear to be somewhat offset by the newly offered exemptions, but these still fail to compensate for other tax burdens on the industry.

Discussion

The Severance Tax Debate

In the Arkansas severance tax debate, there are two main camps - those who wish to increase the rate and those who do not. Each side has a strong faction with multiple arguments in support of their position. Both also seem to hold the best interest of the state as the foundation for their convictions. The cohorts clash, however, with respect to how to uphold the state's well-being.
### Table 4 – Exemptions and Incentives

<table>
<thead>
<tr>
<th></th>
<th>Royalty Deductibility</th>
<th>New Discover y Wells</th>
<th>High-Cost Wells</th>
<th>Marginal/ Inactive Wells</th>
<th>Horizontal Wells</th>
<th>Deep Wells</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR</td>
<td>Old</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>New</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>--</td>
</tr>
<tr>
<td>Kansas</td>
<td>New</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>--</td>
</tr>
<tr>
<td>Louisiana</td>
<td>New</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>--</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>--</td>
</tr>
<tr>
<td>Texas</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Incentive to Market Previously Flared or Vented Casinghead Gas</td>
</tr>
</tbody>
</table>

1Arkansas Department of Finance and Administration: Tax Rates  
2Kansas Department of Revenue: Tax Rates  
3Louisiana Department of Revenue: Tax Rates  
4Oklahoma Tax Commission: Tax Rates  
5Texas Comptroller of Public Accounts: Tax Rates

### Table 5 – Arkansas Scorecard

<table>
<thead>
<tr>
<th></th>
<th>Kansas</th>
<th>Louisiana</th>
<th>Oklahoma</th>
<th>Texas</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corporate Income Tax</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>Sales Tax</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0</td>
</tr>
<tr>
<td>Property Tax</td>
<td>0</td>
<td>0</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Severance Tax</td>
<td>Old</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>New</td>
<td>0</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Exemptions</td>
<td>Old</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>New</td>
<td>0</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td>TOTAL BURDEN</td>
<td>Old</td>
<td>0</td>
<td>0</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>New</td>
<td>0</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>
The first group to consider is those advocating an increase in the rate. Proponents of this change include Governor Mike Beebe, Former Republican Party State Chairman Sheffield Nelson, other state leaders, and most state legislators. The support behind the policy is grounded in prioritizing increased fairness to state citizens. The higher rates of surrounding states are commonly cited as a reason to increase Arkansas’s severance tax in order to charge a more regionally comparable price for the extraction of the state’s natural gas. Additionally, the destruction of the state in terms of local roads and the environment are considered justification within this camp as well. The obvious advantage gained by increasing the severance tax is generating greater revenues for the state to remedy these concerns as well as fund other state initiatives.

On the other hand, critics of the new severance tax policy include the natural gas industry and select conservative state leaders and legislators. These opponents support the maintenance of the severance tax rate at the current level based on the increased cost to businesses and the corresponding negative impacts. One major concern is the decreased competitiveness of the state for investment when higher taxes are levied on natural gas producers. These increased costs will possibly increase rates to Arkansas consumers or reduce investment in the state. Consequences might also include decreased employment and collection from other state taxes.

One main group that tends to fall somewhere between the two sides of the issue are the royalty owners. While it appears to be obvious that this group would oppose a rate increase for fear of lost income, many recognize the need for road improvements in their areas. Thus these citizens generally advocate a compromise.

Analysis of Potential Consequences

Each side of the argument to increase the severance tax appears to have valid concerns. In order to determine the appropriate policy for the growing natural gas market, it is necessary to analyze all issues.

Based on the data gathered in this study and evaluation via the balanced scorecard, Arkansas appears to impose a similar tax burden as Kansas and a greater tax burden than Louisiana, Oklahoma, and Texas on the natural gas industry. Thus, increasing the severance tax may be risky policy with respect to the state’s economy when simply comparing this small sample of states, particularly considering the fact that development of the natural gas industry is in its infancy in Arkansas. For example, Texas experienced growth similar to Arkansas with the Barnett Shale Play just a few years earlier. Due to the uniqueness and expense of shale play drilling, Texas offered a ten year exemption of the 7.5% of market value severance tax to natural gas wells drilled in the play area to encourage investment and activity. Thus, the general argument in support of increasing the severance tax to charge an equivalent price for the extraction of natural gas from the state as neighboring states seems to come up short.

Additionally, with recent announcements of fertile reservoirs in Louisiana and continued discoveries across the region, companies are faced with expanded choices for where to invest. The overall tax burden of a state will certainly be considered when making these decisions, and companies will opt for the cheapest lands to further drilling activities. However, given finite resources, companies are restricted to areas with natural gas to drill regardless of increased discovery. It is obviously important to keep it profitable for companies to invest in exploration within the state. Many companies have already invested in the area; it is difficult to conclude that an increase in the severance tax will drive them completely out of the state, although decreasing investment is still possible.

Some of the argument against increasing the severance tax has stemmed from the fear that these new costs to companies would be passed on to the Arkansas consumer, who is already suffering when compared to the national average. However, there is little evidence to support this concern, in that the price of natural gas paid by consumers is determined by the global market based strictly on supply and demand and is not likely to be influenced by local factors such as tax rates. Even under this misconception, most of the natural gas produced in Arkansas is exported out of the state so any tax increase that would yield higher prices would, at worst, be passed on to consumers in other states. Arkansas, on the other hand, currently imports most of its supply. Therefore, under this argument, consumers already pay other states’ severance taxes in their gas bills.

Based on the general economic well-being of Arkansans, the risk of damaging the state economy is amplified when one considers the possibility of inverse impact on prosperity brought into the area by natural gas companies’ investment. Of particular concern is potential response to the implementation of a 5% severance tax rate by respondents in the CBER survey who claimed an average decrease in investment of 13%. This decrease in investment creates a corresponding decline of economic output by $2.3 billion for the 2008-2012 period. It is important to note that this decline is simply an estimation based on reporting of natural gas companies who do have a clear incentive to keep tax rates low.

When applying Kunce’s (2003) research to this scenario, empirical evidence shows vast increases in tax revenue and miniscule drops in drilling activities when raising severance tax rates, even in the states used for comparison purposes. Based on these data, state legislatures should be discouraged from offering breaks or discounts. In fact, one of Kunce’s (2003) strongest points is a demonstrated shift in federal tax liabilities to state revenue. Because state severance taxes are deductible on federal corporate income taxes, companies pay these taxes within a similar dollar range regardless of the state rate. This is compelling evidence to increase the tax despite comparability to the states in this study. Given these findings, it is unlikely that Arkansans’s economy will be damaged and/or mass investment will be discouraged.

Additional Economic Context

Another important aspect of the severance tax policy is allocation of the revenue generated. Although not directly related to the rate of the severance tax or exemptions and
Incentives, the application of these tax dollars demonstrates the benefit to the state from revenues as compared to general economic output in Arkansas. It was originally proposed that the revenue be dedicated to education, and specifically to higher education or scholarships for in-state, low-income students. Arkansas has one of the lowest college attainment rates in the nation as personal income is meager on the national scale and the costs of a college degree continue to grow. However, Governor Beebe has pointed out that funding for education was increased during the last session and that it would be legislatively irresponsible to continue to ignore the growing need for highway funding. Additionally, bursary for education currently comes from general revenue. It was noted that, should the revenue from the severance tax not be allocated to highways, the $19 billion anticipated to repair state roads and bridges would have to be pulled from general revenue. Essentially, it would just be a shift in budgeting, and possibly a risky one for education as it would be based solely on market prices.

It is also necessary to consider the concern of many Arkansans regarding the preservation of the Arkansas environment, especially being known as the ‘natural state.’ The damaging of local roads has been considered and remedied within the new structure. However, environmental issues, which include land based, surface, water, air, and noise contamination, are likely or even inevitable as a direct result of drilling activity. One of the greatest risks is pollution of the water supply in light of the millions of gallons of process water, drilling fluids, and return water that are injected into deep disposal wells below groundwater in saltwater formations. Additional study is necessary to determine the effects of these disposal practices, but the potential for catastrophic impact exists. On a less disastrous note, each well site requires a minimum of a 5 acre gravel pad that will remain for 25-40 years typically resulting in deforestation or clearing of prairie. Several states allocate a percentage of severance tax revenue to accommodate these concerns. The Arkansas legislature needs to make accommodations for these environmental impacts as well, and severance tax revenue would be an appropriate source from which to set aside funding to protect the state. This is particularly relevant as such taxes are collected from the same activities that have the potential to do harm.

Conclusions

Taking each of these pieces of evidence into consideration, the Arkansas state legislature was seemingly justified in increasing the severance tax in general. The plan passed in March of 2008 raises the rate to 5% and offers discounts for almost all wells in the state with marginal production or high cost to 1.25 or 1.5% respectively for up to three years. Projections anticipate generation of state revenues in the amount $57 million in the first year and up to $100 million annually by 2013.

If the legislature were to apply the recommendation of Kunce (2003), these discounts would not be offered. However, conservative revenue projections under Sheffield Nelson’s proposal of a 7% tax were similar to those of the governor’s plan at $60 million the first year, although they also ran up to $100 million. It is possible that different production figures were used to calculate these two forecasts, especially considering Kunce’s (2003) research suggesting that breaks signify decrease tax revenues brought in by the state and the gap is large (between 1.25 and 7%). Additionally, these projections are also based on stable natural gas prices as the basis for the tax that has switched from volume to market value. It was wise, based on Kunce’s (2003) study, for the legislature to limit the breaks to three years rather than the 10 offered by Texas. This allows the state to gauge investment, drilling, and production as influenced by the change in severance tax policy. After consideration of new figures, it may be appropriate to reevaluate the tax package.

The Arkansas state legislature seems to have indeed made a reasonable adjustment to the natural gas severance tax structure by raising the rate, however, not necessarily for the right reason – i.e., to ‘match’ surrounding states’ higher severance taxes. Instead, this policy alteration is appropriate in that econometric models based on empirical evidence demonstrate that increased rates yield minor changes in drilling activities and significant revenue generation. Although this same research warns of little benefit to offering exemptions and tax breaks, it is not unreasonable for the state to ‘test’ the rate increase with short term discounts. The downside is that it will be difficult to gather the necessary two-thirds support in the legislature to raise the tax again should Arkansas follow Kunce’s (2003) model. Finally, perhaps the greatest shortcoming of the recent change is in its oversight in allocating some of the revenue generated from the severance tax to create a reserve fund to counteract any negative consequences to the environment. Overall, however, the state responded suitably to the Arkansas natural gas boom in the Fayetteville Shale Play region.

References


http://scholarworks.uark.edu/inquiry/vol9/iss1/1


Mentor Comments

Katherine Deck points to the immediate relevance of Tammy Lippert’s work on severance taxes in Arkansas.

This letter is in reference to the work “The Impact of Severance Taxes on the Arkansas Natural Gas Market,” an undergraduate research project by Ms. Tammy Lippert. I acted as her advisor on the project. Ms. Lippert came to me last fall with the desire to engage in practical, policy-oriented economic research for her Honor’s Thesis. I suggested looking at the severance tax, which was very much in discussion at the time. The work compares the magnitude of the taxes on natural gas in Arkansas with those in surrounding states. The report acts as a summary of publicly available information that policy makers could use as a reference. The Arkansas legislature and governor’s office used similar kinds of information when deciding to raise the severance tax last month.
ROMAN ISIS AND THE PENDULUM OF TOLERANCE IN THE EMPIRE

By D. Jasmine Merced
Classical Studies Program

Faculty Mentor: Dave C. Fredrick
Classical Studies Program

Abstract

This paper examines the evolution of Isis, ostensibly the "sacred mother," as a political tool in Egypt and (especially) in Rome. Through an analysis of primary and secondary source materials, it is established that Isis' treatment by Roman politicians represented a running discourse on the contemporary political relationship between Rome and Egypt, and, at times, on Rome's complex negotiation of foreign influences on its own society. Following the deaths of the first two Roman emperors, Isis was gradually elevated from the status of pariah to an acclaimed goddess within the Roman pantheon who was deemed worthy of beneficence and protection from the imperial government.

This investigation of the religion of Isis in Rome from the Late Republic through the Early Empire encapsulates the inseparability of religious tolerance from fluctuating political climates. By analyzing the political rationale behind the oscillations between the persecution and tolerance, or even promotion of this decidedly un-Roman goddess, it becomes apparent that hostility resulted from governmental disdain for Isis' home country and also from the need to sway public opinion, while tolerance arose from harmony with Egypt, smooth trade relations, and the need to mollify Rome's diverse and dynamic lower-to-middle class population, many of whom were brought or emigrated from Egypt, Libya, and the Levant.

(For full text, go to http://inquiry.uark.edu/)

Mentor Comments

David Fredrick targets the substantial quality of the research conducted by Jasmine Merced in this paper.

In scope, argumentation, and finish, this is an outstanding thesis, which certainly could serve as the basis for a paper at a professional conference. The core of Jasmine's analysis is generated around the idea that Isis was not simply a goddess for foreigners and women that the Roman government first suppressed, then tolerated, and finally promoted. Rather, Isis had a sustained ideological function in Roman politics of the late Republic and early Empire, serving as the Oriental "Other" (in Edward Said's terms), through which the Romans could both express their reaction to immediate political events and think through larger issues of cultural identity. She makes a strong argument that periods of suppression of the Isis cult in the late Republic are tied to Rome's troubled political relationship to Egypt, and this allows us to anticipate the initially perplexing status of Isis in the Augustan period. Essentially, Augustus was compelled to take a necessarily ambiguous approach to Isis. With one eye on the senatorial elite, he suppressed her public worship as corrupt and effeminate. With his other eye on lower class Romans, whose interests he largely supported, he diffused Egyptian imagery through public and private art, and in fact his suppression of the cult was never as thorough as it might have been. The thesis is remarkable for its theoretical approach – Said's work has not been widely used in classics – and for its attention to the political complexities of the late republic and the corresponding nuances of Augustus' treatment of Isis. The thesis references an extensive bibliography of primary and secondary works, including works in Italian. Jasmine's thesis research was supported by a SURF Undergraduate Research Award, and included a semester at the University of Arkansas Rome Center for Architecture and the Humanities, where she was able to visit and photograph many of the remains of ancient Egyptian culture and Isis worship in Italy.
ECONOMIC LITERACY AT THE UNIVERSITY OF ARKANSAS

By Amanda Otto
Department of Marketing and Logistics

Faculty Mentor: Cynthia Riemenschneider
Department of Marketing and Logistics

Abstract

Students in Free Enterprise (SIFE) is an international student service organization that teaches the principles of free enterprise to people in the community. SIFE teams are encouraged to develop programs to educate different groups in the community, including children, senior citizens, and students on their own campus. The University of Arkansas SIFE, however, has no long-running, successful program targeted at college students, and so this survey research was designed to determine the topic, format and promotional strategy for a new college program.

In this study, the researchers developed a survey asking students what topics were most desired, what the ideal format of the program would be, and what would motivate students to participate. Each of these questions allowed participants to rank options within topics, formats, and motivators. The survey also collected demographic information concerning age, gender, university classification, and specific college so that the researchers could analyze the needs of different groups on campus. A total of 397 students participated.

To analyze the data, the rankings were summed to create an aggregate ranking for each question. The demographic information was used in chi-square tests of contingency tables to determine whether the rankings were in fact tied to the demographic group of the participant. The results showed that students want programs addressing success skills (such as interviewing, computer use, and oral presentation) and financial literacy topics (such as personal budgeting and investing). The program should be lead primarily by business representatives and should use an open discussion format. Extra credit was named as the primary motivator. Finally, the chi-square tests showed significant dependencies between gender and specific rankings. These results were presented to the University of Arkansas SIFE Leadership Team; they decided to develop two new programs within success skills and financial literacy to address the expressed needs.

I. Introduction

Students in Free Enterprise (SIFE) is an international student service organization dedicated to providing economic education in the community to help others develop skills needed for success in current business and financial systems. There are five main educational criteria that SIFE programs target. The first is market economics, which addresses how free markets work in the global economy; programs can cover supply and demand, globalization, and business functions like marketing and logistics. The second is success skills, focusing on helping others acquire the education and skills training needed to succeed in a dynamic, competitive global economy. Programs in this area can include technological education, resume writing, or interview skills. The third is entrepreneurship, encompassing how entrepreneurs succeed by identifying a market need and then profitably producing and marketing a product or service to fill that need. These projects tend to focus either on helping someone start a new business or helping to improve an existing one.

The fourth educational criterion addressed in SIFE programs is financial literacy, which involves teaching others the use of financial tools, such as techniques for developing a budget, the responsible use of credit, and saving/investing principles. The fifth and final criterion is business ethics, which focuses on illustrating that the long-term success of individuals, businesses, and the economy depends on the practice of good business ethics; ethics projects often involve sharing the principles of ethics as well as their real world ramifications. Each SIFE team is responsible for finding those needs within or outside of their communities that they will address, determining how they will address targeted needs, and evaluating the impact of what they have accomplished. Teams are encouraged to have a variety of programs reaching out to multiple demographic groups, including children, senior citizens, and even their own fellow students.

In recent years, University of Arkansas SIFE team has had projects targeting college students within each criterion. For example, within market economics, SIFE held a global market economics forum focusing on logistics and featuring international logistics professionals. Addressing success skills, SIFE invited a business professional to discuss different ways to promote personal image as an employee and as a student within the interview process. For entrepreneurship, SIFE has hosted local entrepreneurs, having them talk about their experiences. In the context of financial literacy, SIFE has educated students on identity theft, credit cards, and how to evaluate job offers. Finally, within business ethics, SIFE hosted a white-collar criminal who openly discussed the corporate mentality and how it can breed unethical choices among employees.

II. Purpose of Research

The University of Arkansas (UA) SIFE chapter, while a strong contender on the national scene, does not have a longstanding, successful program targeted towards college students. All of the above mentioned projects have been
discontinued, either because the audience did not respond or we could not find successful measurement tools to evaluate the impact of whatever the students had learned. UA SIFE has been trying to fill this void with a few smaller projects. However, in order to develop a project that addresses the real needs of students and thus motivates them to participate, we determined that we must first conduct market research to survey students directly concerning their needs and what they want to see in the next SIFE program on campus.

Arguably, university students today need training within each of the SIFE criteria. For example, the Association of America Colleges and Universities has conducted research that showed that more than 60 percent of employers felt that recent graduates lacked the skills to succeed in a global economy (Fischer, 2007). This finding implies that students need more training within the area of market economics to understand the changing competitive landscape.

Within success skills, successful job interviews are important to finding employment once university studies are over. Recruiters openly admit that little slip-ups can make the difference in creating a solid first impression and those errors can shift the tone of the interview (Damast, 2007). Resume writing, interview coaching, and etiquette training can minimize the occurrence of such errors, making University of Arkansas students more successful in their job searches.

Entrepreneurship continues to grow on college campuses. While still in school, many students are launching successful businesses such as Google, Yahoo!, YouTube and Facebook. Author and entrepreneur Randal Pinkett (2007) has discussed the many advantages that students have in starting a business, and University of Arkansas students have access to those same resources. If there was a program that taught students on the University of Arkansas campus how to exploit such resources, it is possible that SIFE could promote long term success in students’ chosen fields.

Financial literacy continues to be an important area for economic education at the college level. More than 100,000 students across the country leave university with thousands of dollars worth of high interest credit card debt annually. “According to Nellie Mae, the nation’s largest student loan lender, the average undergraduate has about $2,700 in credit card debt and graduate students have about $5,800” (Jennings, 2007, 46). Such trends of student spending, combined with the need to prepare for financial independence after college, demonstrate the value of sound personal financial training at the university level.

Finally, in an age of increased corporate scandals, the business ethics curriculum has become increasingly important on college campuses. These scandals and associated unethical activity often result from the combined cooperation of multiple employees, with new employees often socialized to continue these activities (Anand, Ashforth, & Joshi, 2005). Some economists point to the promulgation of “agency theory” as another cause for unethical activity, as employees consider themselves as agents for meeting the profit goals of shareholders and therefore free themselves of moral responsibility for their actions (Economist, 2007). These authors agree that training our students can prevent them from continuing such activity.

Clearly, training in all five areas is needed at the University of Arkansas. The immediate questions are: which specific criterion will provide the most impact, and which specific topic within that criterion will entice students to participate? We began this research with the idea that one reason former UA SIFE programs had limited success is that SIFE members had not taken enough time to assess the true needs and motivations of University of Arkansas students. We felt that, if UA SIFE asked students what types of training are most needed, it would be possible to create a program that met perceived needs, and such a program could then be more successful than one based on SIFE’s estimates of what is needed. Further, UA SIFE could ask what program format elements are most appealing and what would motivate students to participate in whatever program was developed.

These were the questions that motivated this research. With such a wide range of student education needs, which is the most prevalent? And once UA SIFE develops a program to meet those needs, what format should the program use and what would motivate students to participate in that program?

III. Methodology

Developing the Survey

During the first stage of preparation, we determined five primary sections for the survey:

1) The first section allowed participants to rank subtopics within each of the five SIFE criteria. This section helped us identify needed topics more specifically, as well as giving non-SIFE members a better concept of the meaning of each criterion and the types of products that fell in those categories.

2) In the second section, the participants ranked the five criteria, thereby allowing them to highlight the groups of topics from the first section that were believed to be most needed.

3) Third, questions relating to the format of the program were developed to determine what setting and educational structure would appeal most to college students (e.g., lectures, interactive discussions, information fairs or some other medium).

4) Fourth, a series of questions probed factors motivating program participation (e.g., monetary, school-related, or other), so that incentives can be developed to encourage students to participate in our program.

5) Finally, there was a section relating to personal information, specifically age, gender, university classification, and specific college affiliation. This section allowed us to analyze the data sorted by demographic variables and to determine what the specific needs are of different groups on campus.
Each of the first four sections was evaluated by asking students to rank which topics they would most like to hear about, which program formats would most interest them, and which factors would motivate them to participate. To ensure unconstrained responses, an “other” option was always provided so alternative answers could be provided in the space available. The final section, personal information, was a combination of multiple choice and free response, based on the question being asked.

The second stage of preparation was to convert the survey into an online format using the Qualtrics survey software licensed by the Sam M. Walton College of Business. We decided that an online format would help us distribute the survey around campus, particularly to colleges outside of the Walton College, since the Internet is accessible to all students more readily than paper copies. The Qualtrics survey method carried some limitations. For example, with the “other” option in the ranking, we could not include the free response space within the boundaries of the original question but rather had to create a separate question following the original to incorporate the open ended question.

After the survey was transformed into the online format, we conducted a pilot study with SIFE students in order to ensure that the survey’s flow and meaning were clear. We made the needed alterations in order to make the survey ready for general distribution. Specifically, the original survey had options ranked from one to five with “5” being the highest; however since participants were dragging their choices into a top to bottom order, we felt the natural instinct would be to rank the options with “1” being the highest. We also added more specific instructions on how to rank the items using the computer’s mouse because we found that the drag and drop format was not intuitively clear. Finally, we made some formatting changes, such as inserting breaks in the lines of text to make sure that the ranking number did not block the text.

Informed Consent Policy and Data Confidentiality

Because we were conducting an online survey, we determined that implied consent would be most appropriate. A screen detailing informed consent information would precede the online survey, and participants would imply their consent by clicking the forward arrow and beginning the survey. Regarding the confidentiality of the data, we determined that since the information was neither sensitive nor personal, we would not separate the names from the responses during analysis. However, when the data were presented to SIFE, all names would be excluded. If professors provided extra credit to participants, they would only be given an aggregate list of participant names.

Promoting and Distributing the Survey

The survey was officially launched on February 27, 2008. The survey was left active for two weeks, with an end date of March 13. To promote the survey, we completed an initial round of publicizing and then a week into the survey, a second round to encourage students to participate before the deadline.

Two primary methods of promoting the survey were used. The first was to reach the students through professors. Within the Walton College, we asked professors to publicize the survey in class and to distribute the survey via e-mail. Some went further by offering extra credit to participating students. Outside of the Walton College, we sent the survey link to professors and administrators around the university, particularly known through other research projects and activities.

The second method was to target students directly. We sent the link directly to students in our own network, and we asked that other SIFE students do the same. SIFE students made announcements in class to encourage students to take the time to complete the survey. We also printed the link on business cards so that we could distribute them while SIFE students were advertising the organization around campus. To further motivate students, SIFE provided a monetary incentive by offering gift cards to five random participants.

Analysis Methods

Because students ranked topics from 1 to 5 with 1 being the top answer for each question, we summed the ranking for each option. The option receiving the lowest total was the top response for the group. We then compared the sums for each of the five options to judge the relative importance of each option. To further analyze the data by demographic, we used chi-square tests of contingency tables with rankings treated as nominal data which are juxtaposed against personal demographics. Chi-square statistical techniques allow comparison of ranking responses on each question for different populations (Keller & Warrack, 2004 538). The demographic variables were gender, age, University college, and classification.

Data Clean Up

Before analyzing the results, data were examined for irregularities. Several surveys had no responses listed and were eliminated from the data set. An additional problem with the Qualtrics software was identified. Specifically, when students accept a given ranking and do not make changes, the “drag and drop” style of answering returns blank answers in the data set, as though the student did not answer at all. In those cases where such blanks occurred followed by preceded by responses, we felt we could safely assume the student had read through the question but had decided to leave the ranking as it was presented on the survey, and so we entered the corresponding responses.

IV. Results

Market Economics

The first question dealt with which of the following five topics would be the most interesting subject for the new college-level program: 1) supply and demand, 2) globalization,
3) marketing and consumer demand, 4) sourcing, 5) logistics and transportation, or 6) other (any additional topic suggested by the respondent). As illustrated in Figure 1, the top response for this question was globalization, followed by supply and demand, and then marketing and consumer demand. (Recall that the lowest aggregate number indicates the number one topic overall, as the number on the vertical axis represents the sum of the rankings given for that option and “1” was the highest ranking for each question.)

![Market Economics](image)

Figure 1. Potential market economic topics of interest.

There were also several responses given for other topics that could be taught. These included: sustainability, sales, information systems, advertising, the history and philosophy of economics, accounting, legal economics, management, the job market, cultural change and diversity, and elasticity of supply and demand.

**Success Skills**

The success skills topics that students were asked to rank included: 1) computer skills, 2) presentation skills, 3) résumé writing, 4) interview/self-presentation skills, 5) career planning, and 6) other (another topic specified by the respondent). Interview and self-presentation skills were identified as most needed, followed closely by computer skills, presentation skills, and career planning (see Figure 2). All four of these responses had similar point totals, suggesting that the perceived need for each is comparable.

Using the “Other” response option, students also suggested the following topics as potential subjects for a SIFE program on campus: business writing and proper communication and expression through writing, work ethic and self-management, team work skills, communication skills, finding a job for your major, researching post-graduate academic opportunities, and quantitative skills.

**Entrepreneurship**

Participants ranked the following potential entrepreneurship topics based on their need at the University of Arkansas: 1) funding a business, 2) business plan development, 3) day-to-day operations, 4) information systems in a small business, 5) consulting/help with existing business, and 6) other (open topic to be specified by participants). In Figure 3, the highest ranked topic in this area was developing a business plan, followed by funding a small business, and examining the day-to-day operations of a small business. Other topics suggested in the area of entrepreneurship included marketing for a small business, networking training, market understanding to find the need for new small businesses and taxation of a small business.

![Entrepreneurship](image)

Figure 3. Potential entrepreneurship topics of interest.

**Financial Literacy**

Within financial literacy, students ranked the following potential program topics: 1) personal budgeting, 2) personal investment, 3) credit cards, 4) identity theft, 5) differentiating job offers, and 6) other (the free response option). The top ranked topic for this criterion was personal budgeting, followed by personal investment and credit cards (see Figure 4). “Other” topics included understanding the difference between assets and liabilities, building good credit, and reading and comprehending financial statements.
Financial Literacy

Business Ethics

With respect to ethics, students ranked the following potential topics: 1) globalization issues, 2) dealing with unethical situations in the workplace, 3) corporate social responsibility, 4) legislation, 5) environmental initiatives, and 6) other (the free response option). The top ranked topic was globalization, followed by dealing with unethical situations in the workplace and corporate social responsibility (see Figure 5). “Other” topics included discretionary practices in the workplace, creating a non-discriminatory work environment, minimizing waste of resources, and ethical issues in global companies and other countries.

Educational Criteria

Having seen possible topics within each of SIFE’s core criteria, students then ranked the criteria themselves to determine which general set of topics is perceived as most needed at the University of Arkansas. As is readily apparent in Figure 6, the top ranked criterion was success skills, followed by financial literacy and market economics. The relative prominence of success skills suggests that interview and self-presentation skills were perceived as the most needed topics, as well as skills involving computers, oral presentation, and career planning. Secondary needs to address would be the top ranked topics within financial literacy, such as personal budgeting and investing, and within market economics, such as globalization and supply and demand (see Figure 6).

Program Format

There were two questions addressing the ideal program format. The first question asked students who they would want to lead a program: 1) business representatives, 2) U of A professors, 3) college students, 4) representatives from centers on campus, 5) no one (independent learning), or 6) other (some other leader to be specified by the respondent). Survey responses summarized in Figure 7 showed that students preferred business representatives to lead the session, followed by U of A professors and college students. Other potential leaders suggested included business attorneys and independent on-line modules.

Educational Criteria

![Figure 4. Potential financial literacy topics of interest.](image)

![Figure 5. Potential business ethics topics of interest.](image)

![Figure 6. Relative ranking of five SIFE educational criteria.](image)

![Figure 7. Preferred leadership for learning experiences.](image)
informational fairs, 2) structured lessons with PowerPoints, 3) open discussions/Q&A, 4) competitions, 5) reading material, and 6) other (any other possible format to be identified by the respondent). The top ranked format was open discussions/Q&A, followed by structured lessons and informational fairs as shown in Figure 8. One respondent commented that a mixture of open discussions and structured lessons often addresses topics best. Other formats suggested were receptions, hands-on activities and one-on-one counseling.

**Motivational Factors**

To determine how SIFE can best motivate students to participate, the survey asked respondents to rank the following based on which would most effectively motivate students to participate: 1) extra credit, 2) cash prizes or incentives, 3) learning, 4) résumé building, 5) networking opportunities, and 6) other (any other motivation factor the respondent identified). The top ranked motivation factor was extra credit, followed by cash prizes and networking opportunities (Figure 8). Other potential motivating factors included the possibility of improving the surrounding community and communication with other students.

**Demographic Analyses**

Four sets of contingency tables were created, one each for age, gender, university classification, and specific college. Each set of tables contained nine contingency tables, for each of the nine ranking questions in the survey (the five criteria, educational criteria rankings, program format, program leadership, and motivation). Nominal demographic categories were placed on one axis with potential question answers on the other. For example, in the age versus market economics table shown in Table 1, different age groups were on the top axis, and the six potential topics for market economics were on the other axis. In the body of the table, we calculated the number of times that a member of that demographic group ranked that option either first or second on their list (yielding a frequency count based on identifying which response options are most important or desired).

**Table 1. Sample Contingency Table Relating Market Economics Response to Age**

<table>
<thead>
<tr>
<th></th>
<th>18 and younger</th>
<th>19</th>
<th>20</th>
<th>21</th>
<th>22</th>
<th>23 and older</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choice 1</td>
<td>6</td>
<td>25</td>
<td>44</td>
<td>49</td>
<td>50</td>
<td>34</td>
</tr>
<tr>
<td>Choice 2</td>
<td>6</td>
<td>29</td>
<td>54</td>
<td>80</td>
<td>50</td>
<td>36</td>
</tr>
<tr>
<td>Choice 3</td>
<td>7</td>
<td>17</td>
<td>38</td>
<td>45</td>
<td>28</td>
<td>20</td>
</tr>
<tr>
<td>Choice 4</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>27</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Choice 5</td>
<td>1</td>
<td>4</td>
<td>13</td>
<td>26</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>Choice 6</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

Chi-square tests were completed for each contingency table in order to determine whether response distributions differed depending on the demographic variable. Significance levels were set at the .05 probability level. For chi-square tests with p values greater than .05, there was insufficient evidence to reject the null hypothesis that the two variables were independent. At this significance level there was no significant difference found with respect to age groups, University classification (freshman, sophomore, junior, senior), or University College.

Gender was also independent of rankings in market economics, entrepreneurship, financial literacy, business ethics, educational criteria, program leadership and motivation. However, the success skills educational criterion did show dependence between gender and ranking ($X^2 = 14.542, df=5$, $p=0.0125$). As a percent of responses, men ranked options 1 and 2 higher than women (computer skills and presentation skills respectively), whereas women ranked options 4 and 5 higher than men (interview/self presentation skills and career planning respectively). Regarding program format, chi-square tests addressing program format showed dependence between format rankings and gender ($X^2 = 17.516, df=4, p=0.0015$). Men ranked options 4 and 5 higher than women (competitions and reading material respectively) whereas women ranked options 1, 2, and 3 higher than men (information sessions, structured lessons with PowerPoints and open discussions/Q&A respectively.)
Topics and Learning Methods Shared with SIFE Team

On April 12, 2008, the results of this study were presented to the SIFE Leadership Team. We recommended the following three potential topics, each of which would be implemented in accordance with survey recommendations concerning program format and motivational factors (such as requesting that professors give extra credit for participation).

- Success Skills Marathon: Because the topic of success skills was the top rated educational criterion, we felt SIFE needed some project to address this criterion. As described in the above analysis, there were four different responses within success skills that were close in ranking: interview/self presentation skills, computer skills, presentation skills and career planning. Because these topics were similarly ranked, we felt the best way to address success skills would be to have a series of smaller events, addressing each topic. There would be two principal benefits with this format: first, SIFE could address all the topics students requested, and second, with multiple units, we could better sustain the project because each unit would serve as a recruiting and promotional tool for the next. In addition, we would be better able to sustain the impact on participants with continuous interaction.

- Financial Literacy Program: Financial literacy was the second ranked criterion overall. This program should stress the basics of personal budgeting, such as creating a budget and then following that budget to ensure that the participant lives within their means. The second ranked topic within financial literacy was personal investing. To meet this need, we felt the SIFE team could adapt a preexisting program, Investing for the Future, for college students. This program taught the basics of investing, focusing on stocks, bonds, and savings accounts. This program could be adapted for college students by including more practical information, such as real web sites to participate in investing activity. This second program could serve as the next level to the personal budgeting program, reinforcing those skills while adding new skills to the students’ arsenals.

- Globalization from Two Perspectives: As globalization was the highest ranked topic within both market economics and business ethics, and market economics was the third highest ranked criterion overall, we also recommended a project on globalization. Business professionals would first address globalization from a market economics perspective, managing both supply from a global logistics standpoint and demand from international consumers. They would then also address ethical encounters that they and other companies encounter in the real world, thus better preparing students for those same situations. Students would be able to engage them in discussion to clarify key ideas and gain additional insight.

New Project Development

The SIFE Leadership Team elected to adopt the first two potential projects for the coming school year. The team was enthusiastic about the results, and they concluded that these two projects were the most appropriate for two reasons. First, these projects met demonstrated needs. Second, success skills and financial literacy criteria have traditionally been two of UA SIFE’s weaker elements, needing better programs with greater impact.

Because the projects will be executed in the coming school year, they have been entrusted to the appropriate criterion managers of next year’s Leadership Team. Criterion managers are SIFE members who are responsible for overseeing all of the projects within one of the five criteria of SIFE. They ensure that the project is executed, the impact is measured, and the results are documented.

In order to fully comply with these research findings, SIFE will be selecting business representatives and teachers to facilitate the program, and they will combine structured lessons and open discussion. Because the data gathered covered five different colleges on campus, SIFE will promote the project accordingly to reach all the students that participated.

V. Further research

Now that the subject of the first projects has been addressed, SIFE must develop evaluation methods to measure the impact of these projects. Examples include, pre- and post-tests, surveys, and testimonials from participants. Using these tools, SIFE can determine whether the needs and desires expressed by the survey responses are indeed met.

Second, while the responses given in the “Other” questions of the survey were small in number, further research (whether through additional surveys or a short questionnaire included in the evaluation for the current programs) could help SIFE understand the demand for those topics. SIFE can also use further market research to benefit more programs. We have begun to use market research to evaluate needs with our CEO from Head to Toe program, which focuses on teaching entrepreneurs in Northwest Arkansas how to expand their businesses (specifically realtors for the 2008-2009 school year). Project leader Meagan Stellpflug worked to develop a survey discussing which realtors would be interested in participating and what topics would benefit them.

VI. Conclusion

The data presented in this article will benefit SIFE for the long term by allowing participants to launch programs based on real needs at the University of Arkansas. We will be able to use the data set in the future as we look for new topics for programs and as we determine the format and promotional strategy for those new programs. The SIFE projects generated from survey data will set our team apart from others because they will be based on real, quantified need, rather than the SIFE team’s projection of student need.
In summary, we had 397 students participating in this survey to generate a better understanding of students’ perceived needs with respect to SIFE projects. UA SIFE is now able to see which topics are most needed and desired by university students. We can also see what format most appeals to students and what motivation factors would appeal to students. By using post assessments when we execute these programs, the programs can evolve to remain current with student needs, thereby eliminating the need to do more market research in the near future.

Bibliography


Mentor Comments

Cindy Riemenschneider describes how she recognized Amanda Otto’s potential as a researcher early on, and how Amanda moved forward independently to complete an empirical study with outcomes of practical value to both SIFE at the University of Arkansas and to SIFE teams across the country.

During Amanda’s sophomore year she and I began our discussions of what she would do for her honors thesis. Because of her strong analytical and quantitative skills, I encouraged her to consider undertaking an empirical research study for her thesis. Her analytical skills far exceed any undergraduate student that I have previously taught and I wanted the opportunity to work with her.

Over her time at the U of A, Amanda has held multiple leadership positions in SIFE and she determined that she would like to write her thesis on a topic that would assist SIFE. She selected the thesis topic independently; I do not work with SIFE in any capacity, nor do I conduct research regarding SIFE. However, my primary contribution was to lead Amanda through the steps of conducting appropriate empirical research. The only outside assistance that Amanda had was the distribution of SIFE business cards (with the web survey information Amanda developed) during the weeks of data collection. A representation of current SIFE students helped man booths in a variety of locations on campus to encourage students of all majors to get involved in SIFE; they also distributed the business cards at these booths.

The research project that she conducted provides a unique and significant benefit to the students at the University of Arkansas. The research also can benefit SIFE teams across the United States that are creating programs to benefit their own college campus. The research is also important to the business discipline by revealing areas where students need additional training. Through her market research, Amanda was able to identify the greatest economic literacy need of the students as well as the best method for delivery. After the economic literacy project is implemented by SIFE, she will be able to measure the effectiveness of the project and make changes for improvement as needed. Amanda will be serving on the business advisory board for SIFE beginning in the fall of 2008. The economic literacy project will provide both an immediate benefit to the University of Arkansas students who participate as well as a long term benefit for future students.

In her research project, Amanda was able to implement both qualitative and quantitative analysis techniques. Even though undergraduate students have limited exposure to research design and methods, Amanda did an excellent job of completing an empirical study.
OPPORTUNITY KNOCKS: AN EXAMINATION OF THE KNOXVILLE TRANSIENT BUREAU AND TRANSIENT POPULATION

By Caroline Peyton
Department of History

Faculty Mentor: Jeannie Whayne
Department of History

Abstract

The narrative of America's economic depression in the 1920s and 1930s is often accompanied by poignant images of male transients riding the railroads. Behind the classic narrative, however, the story of transients is far more complex and varied. Although Franklin Roosevelt's answer to the Depression -- the New Deal -- is well documented and debated, most historians have ignored the Federal Transient Service (FTS), one of many New Deal programs. Although the FTS lasted a brief two years, it served the needs of 211,056 transients at the height of its operations. With approximately 270 transient bureaus, the FTS sought to provide food and shelter, jobs, education, medical care, and much more. By closely analyzing a particular transient bureau, the Knoxville (TN) Transient Bureau, the story of Depression-era transients reveals itself more fully.

As a city of crossroads, Knoxville acted as an urban gateway for the surrounding rural areas. The economic depression struck Knoxville much like other cities, with the exception of the Tennessee Valley Authority. Thus, the hope of finding work with the TVA or other industries attracted hoards of people to Knoxville. The incoming tide of transients prompted the creation of the Knoxville Transient Bureau (KTB), and established a new form of aid, unlike any other. Although the KTB was one of many transient bureaus, the importance of the KTB as a singular institution should not be dismissed. Rather, the examination of one Transient Bureau provides historians with an in-depth assessment of the problems both transients and agencies faced.

The Knoxville Transient Bureau and its director, Frances Strong, faced numerous challenges and complications during the program's existence. Chapter Three deals with one of the primary problems for the KTB, finding appropriate facilities. Most communities viewed transients with disdain. Such hostility forced the KTB to carefully determine the location of each building, as well as the 'type' of transient or function for the building. Families were often provided private shelter, while single males stayed in communal housing. The fears of pollution, moral corruption, and mixing of gender/age groups is an important indicator of the widespread anxiety of the period. More importantly, the anxieties transient populations provoked are still relevant today.

The KTB's records reveal the unique complexities of transient youth and families, discussed in Chapters Four and Five.

These chapters also help to unravel the mythic image of the depression: the lone transient male. Although adult male transients were the largest percentage of the KTB's population, these men often had families waiting for them. The fictional notion of males with wanderlust is quickly dispelled by the stark realities of the period. Transients cannot be embodied solely by one particular type; the economic depression affected people of all ages and cut across the lines of gender and race. The problems unique to individual transients deserve examination, and the KTB's records provide historians with a rare glimpse into many transient's lives. The KTB's transients were individuals with aspirations and interests, not merely silent sufferers fallen to the wayside.

Chapter Six concerns the KTB's medical care program and the unwavering sympathy the bureau had towards not only physically ill transients, but also mentally ill transients. This stands in stark contrast to common attitudes of the period. This chapter further explores the notion that transients potentially "polluted" society by their presence. The KTB's records concerning medical care are surprising in their detail, providing an excellent opportunity to analyze the beginnings of Federal medical support. The lengthy reach of the KTB extended beyond medical care, providing jobs and education, as discussed in Chapter Seven. Although other transient bureaus also had these programs, few had the success of the KTB.

As discussed in Chapter Eight, the success of the KTB is strongly tied to the network of female leaders in the Tennessee Transient Bureau that promoted and sustained the state's bureaus. Frances Strong, director of the KTB, aligned herself with her fellow female directors on the state and national level. This network truly allowed the KTB to be a successful and positive program, however brief its existence. Despite the KTB's achievements and leadership, the Federal Transient Service fell prey to the second phase of the New Deal, discussed in the conclusion. For the FTS, the bureau's success contributed to its demise. With many transients aided and assimilated, the problem of the disenfranchised was once again relegated to private charities.

Nonetheless, the legacy of the KTB lies in the wealth of information in their records. For a brief moment in time, transients were no longer elusive or mythic, but well documented and supported. The relevance of the KTB in modern times is disturbingly clear. In the wake of disasters such as Hurricane Katrina, the questions of responsibility and appropriate aid by the Federal government are still being...
debated. By examining the impact of a transient bureau on a local level, one can see the tremendous impact of the KTB on transient families, youth, and adults. But the FTS was never intended as a long-term solution to transients, rather a brief measure to restore confidence. In a capitalist society, the ability to work is valued far more than the ability to aid. Those unable to sustain themselves are often considered without value, however ill advised that opinion may be. Thus, the stories of those aided by the KTB, as well as the staff, are not only an enduring example of the Depression, but also a lesson for modern society.

For full text, go to http://inquiry.uark.edu/.

Mentor Comments
Dr. Jeannie Whayne describes the combination of analytic skills, sharp intellect, and deep empathy that makes Caroline Peyton’s thesis worthy of inclusion in its entirety on-line.

Caroline Peyton is one of those rare students who combines natural inquisitiveness with creativity and intellectual ability. She was inspired to explore transient culture in Knoxville after reading Cormac McCarthy’s Suttree, a novel featuring a character living on a dilapidated houseboat near Knoxville in the 1930s. Peyton, a double major in History and English, brought to the project an especially creative perspective and drew on both literary and historical sources. Given that this was a thesis written for the History Department, however, she privileged historiography and primary documents in her exploration of the topic. Hitting a blank wall early in the documentary research on her thesis, she patiently and persistently stayed the course. Her close reading of the scant secondary literature – and careful attention to footnotes -- revealed rarely used government records located in the National Archives. There she opened folios that likely had not been touched since deposited more the sixty years ago. The records of the New Deal’s Transient Bureau included the untapped files on the Knoxville branch which form the foundation sources for this excellent thesis.

Peyton’s use of the sources she uncovered is matched by the sharp analysis she brought to the writing phase of the project. Yet Peyton managed to convey a professional historian’s objectivity at the same time that she revealed a deep compassion for the homeless men, women, and children served by the Knoxville Transient Bureau (KTB). Her sympathetic but realistic assessment is demonstrated in one particularly perceptive passage involving Clarence Maples, an indigent man with a number of problems. Peyton carefully perused the sometimes conflicting accounts about Clarence and concluded that his family relations -- fractured by his psychological problems and complicated by his physical infirmities -- made it particularly difficult for the KTB to serve his needs. Here we see Peyton combining both a sharp critical analysis and a genuine sympathy for the individual.

The thesis consists of an introduction, eight chapters, and a conclusion. The first chapter examines the common perceptions of transients as revealed in the popular media and in literature. It also reviews the formation of New Deal policy as the number of homeless individuals reached unprecedented levels. Chapter Two portrays Knoxville as not only a “City of Crossroads” but also a city at a crossroads of its own. With two railroads and three U.S. highways intersecting there, its indigent population increased dramatically as the Great Depression intensified. It was a natural place to locate a Transient Bureau, but it was a city in crisis and this chapter foreshadows the tensions that we see emerge in sharp relief in a later chapter. Chapter Three examines the “special demands” of the KTB and highlights the race, class, and gender concerns that arose. Chapter Four furthers this analysis by focusing on the issue of adolescent male transients, and Chapter Five deals with the treatment of homeless families. Chapters Six focuses on medical care and Chapter Seven examines work, education, and recreation and highlights an experiment in social change. Chapter Eight analyzes the experience of the female administrators of the KTB, particularly the indomitable Frances Strong. The conclusion examines the liquidation of the bureau, analyzes its contribution to the problem of caring for the homeless then, and makes some keen observations about its relevance in twenty-first century America.
TRAIN STATIONS: ICONOGRAPHY, WAYFINDING, AND THE EVOLUTION OF A TYPE

By Tricia Reed Quinn
Interior Design

Faculty Mentor: Catherine Wallack
Interior Design

Abstract

The first train stations were built more than 150 years ago. Their floor plans both developed from and were informed by circulation requirements arising from the need to move passengers and trains safely and efficiently. While the prominent route of movement through stations was largely determined by their layout, certain architectural elements appeared in many stations regardless of their plans. Many of these features first appeared in response to functional needs but later acquired symbolic significance, transforming them into railway iconography. The resulting iconography not only helped distinguish train stations from other types of architecture but enhanced the legibility of the spaces within the stations, making wayfinding easier for its users.

Introduction

The particular activities supported by train stations—the coordination of arriving and departing individuals and track-dependent vehicles, as well as the services required to support both people and trains—have created unique design challenges, challenges that are further complicated by the urban context in which many stations have been erected. This study seeks to explore how the train station as a building type has evolved to allow for the efficient navigation of its occupants in increasingly complex buildings.

For their users, train stations are generally temporary spaces, frequented by arriving or departing visitors who often have no prior knowledge of the premises. Despite a lack of familiarity with a station, it seems that people generally find their way, a fact which raises an interesting question: how do people negotiate a station they have never before visited?

When the first stations were built in 1830, there was neither an architectural precedent for their construction nor the behavioral sciences to describe navigating in them. Despite a lack of precedents or a codified understanding of how to find one’s way in them, train stations evolved into very complex spaces that generally work very efficiently in enabling passengers to circulate safely and freely. One goal of this study, therefore, was to examine the development of the train station as an architectural type during its 178-year history.

An examination of the development of the station as an architectural type, however, does not adequately answer the question: how do people find their way in a station? Psychology, particularly studies pertaining to cognition and perception, helps answer the question. However, architects of the early stations did not have the discipline of psychology as a reference point. Absent scientific-based discourse on how people best navigate in complex spaces, the architects and engineers of the early stations provided meaningful cues such as clear axes and striking landmarks within an architectural context that proved helpful to wayfinding. These architectural cues were repeated with regularity from station to station and to such a large extent that they themselves became icons of stations, symbols of the buildings in which they appeared. While such features contributed to making stations unique in comparison with other types of architecture, subsequent research in the field of wayfinding supports a hypothesis that such features also made navigating through them more efficient.

1. Principles of Wayfinding

Wayfinding refers to the cognitive process of solving navigational problems in an environment. According to Passini, wayfinding consists of four stages: identification of the task, perception of environmental information, formation of a decision plan, and execution of the plan (1984, p.46). Passengers whose train delivers them to a station for the first time typically have no knowledge of the building design so retracing their steps is not an option. How do such travelers know how to exit a station they have never before visited and for which they have no previous experiences stored in memory?

Cognitive maps involve mental compilations relative to one’s perception of the environment. They allow people to determine their positions within a setting because they make it possible for individuals to organize perceived parts of the environment into a meaningful entity, permitting them to orient themselves spatially within the context of the cognitive map and, consequently, within the setting itself (Passini, 1984, p. 35). Unlike physical maps, cognitive maps change as new information about the surroundings is acquired. Shemyakin found that cognitive maps are organized in one of two ways—linear or sequential and spatial or survey (Passini, 1984, p. 37). When information is encoded via a linear cognitive map, routes are ordered sequentially and temporally relative to a person’s movement through space in time; the cognitive map resembles a AAA TripTik where the route alone is revealed sequentially, page to page. On the other hand, when information is encoded through the use of a spatial or survey cognitive map, the environment is perceived as a spatial entity, much like a floor plan.
Cognitive maps can be viewed as both product (a map) and process (mapping). Thus these maps influence not only the type of environmental data perceived but also the preferred manner of acquiring such data, creating a link between perception and behavior. Those who favor a linear approach structure the environment in terms of routes where directional signs are most helpful (Passini, 1984, p. 76). Others, who favor a spatial style, rely on topographical relationships and understand the setting as a spatial ensemble (Arthur & Passini, 1992, p. 38).

Schemata help people form cognitive maps. A schema is an abstraction of an object or of a situation but it is not specific to any particular object from within a particular classification of objects; a schema of a train station encompasses that which is generally characteristic of train stations. Earlier experiences of stations, whether first- or second-hand, are stored in memory and can be retrieved at a later time. Schemata facilitate the perception of the environment and provide a procedural basis for reacting to environmental stimuli; they are cognitive building blocks that play an integral role in the formation of cognitive maps (Passini, 1984, p. 55).

In *The Image of the City*, Lynch seeks a synthesis among the identity, the structure, and the meaning of an environmental image within an urban context. He uses the city as his model and identifies five elements common to all cities: paths, landmarks, nodes, edges, and districts. Paths are the routes along which one moves. Landmarks are “points of reference” (Lynch, 1960, p. 48). Nodes are the intersections along circulation routes. Edges are linear boundaries that separate areas. Districts have a two-dimensional extent and “are recognizable as having some common, identifying character” (Lynch, 1960, p. 47). Lynch believes that people can “structure the whole” by identifying the parts (1960, p. 13). Passini adapts these elements to an architectural model—the building. Paths are corridors, nodes are intersections between hallways, landmarks are significant spaces or objects within a building or visible from within, edges are the perimeters of the building, and districts are major zones in the building usually related by a common function.

A critical stage between orientation in space and successful navigation through it is the planning process. This stage links the cognitive map to the behavior once the appropriate script has been retrieved from memory (Passini, 1984, p. 46). During this phase of wayfinding, decision-making skills allow people to structure a plan for locating their destinations. The task to enter or to leave a station is broken down into a series of subtasks. The decision plan is sequential: there is a logical progression both in the perception of the environment and in the reaction to it in terms of formulating a decision plan and then executing it.

Passini identifies three types of environmental information: sensory, memory, and inferred (1984, p. 60). Sensory information is what people perceive through their senses. Memory information is the ability to retrieve information that involves similar events, places, decision plans, and schemata. Inferred environmental information requires the use of working memory and a manipulation of either sensory or memory information (Passini, 1984, p. 60). The strategies for obtaining environmental information correspond to the types of information (Passini, 1984, p. 70). The direct access tactic employs sensory information. The indirect access tactic relies on experience and memory. The inferred strategy is based on both sensory and memory tactics. When people arrive by train at a station the first time, an inference strategy enables them to integrate sensory and memory information to locate the exit. Upon disembarking from the train, they see the platform. Memory of similar experiences in other train stations tells them that the platform should be connected to the concourse. They have no first-hand sensory or memory information of this concourse but they can deduce this information using the inference tactic.

In order for people to reach their destination, the environment must be perceived, a decision plan must be formed, and then that decision plan must be executed. How, then, did the designers of early stations address these issues architecturally?

II. The Birth of a New Type (1830 - 1844)

The first railroad stations were erected in 1830, one in Liverpool, England, and the other in Baltimore, Maryland. Four conditions made the erection of such buildings possible: specialized tracks, the conveyance of freight, the conveyance of passengers, and mechanical traction (Meeks, 1956, p. 26). The Industrial Revolution set the stage for a new form of transportation, the railroad. There was, however, no direct precedent with respect to either the function or the design of railway stations. “Every solution had to be invented” (Meeks, 1956, iii).

In 1846, Daly, editor of the *Revue Generale de l’Architecture*, published the first classification of train station types and, in so doing, established important criteria for identifying the railroad station as a new type of building. He recognized four types, each based on the station’s parti or “basic scheme or concept for an architectural design, represented by a diagram” (Ching, 1996, p. 381). The four major types are the one-sided, the two-sided, the head type, and the L type (Figure 1). Daly identified each type on the basis of the major circulation routes of arriving and departing passengers, linking function to form and building parti to tracks. Without realizing it, Daly had also linked train stations to wayfinding, not only by identifying the major paths but by using those paths as the primary distinguishing feature among station types.

**Figure 1.** Left to right: Sketches of one-sided, two-sided, head house, and two-sided L stations. From Meeks, C.L.V. (1956). *The Railroad Station*. New Haven CT: Yale University Press. Image in public domain
One-sided stations are the simplest in terms of their floor plan. Passenger and baggage facilities appear on one side of the active train tracks. The Fayetteville (AR) Depot on Dickson Street is an example of this style of station (Figure 2). In the early years of the railway, this was the most common type and continues to be a very functional design for commuter stations.

However, as rail travel grew, both ridership and the number of tracks increased. Two-sided stations enabled departing passengers to make their way toward the tracks from one building, while arriving passengers disembarked through the another building. The first Euston Station in London, England, (Hardwick, 1839) is an example of a two-sided station.

The head house station proved to be the most versatile of stations “because it provided a pragmatic solution to the volumes of both train and passenger traffic” (Sheppard, 1996, p. 33). Functionally, the head house was important because it allowed for arriving and departing passengers to mix. The Gare de l’Est in Paris, France, (Duquesney, 1852) was the first head house station to be completed. The train platforms are connected to each other and to the ticketing and waiting areas by a cross-platform.

The architects of the early stations responded to the challenges created by the new technology using an established language. The vernacular architecture of ordinary houses and cottages proved to be a good marketing strategy for the developing industry because the public had to be persuaded that the railway was not “an object of terror” (Barman, 1950, p. 25). From the perspective of wayfinding, it makes sense that the early railroad companies articulated their stations in the comfortable vernacular because “an unpredictable, unknown, and mysterious path is not a sure path” (Passini, 1984, p. 11). If the language of the space is familiar even if the function within the space is not, a meaningful cognitive map can be formed. The passengers’ familiarity with the vernacular enhanced the formation of cognitive maps for the developing rail industry and its users.

In contrast to stations articulated in the vernacular, monumental stations that resembled temples, such as the first Euston Station were also built during the period from 1830 until 1844. While the vernacular stations assuaged fears over the new mode of transportation, the monumental ones proclaimed victory: “the early railway companies saw themselves as the standard-bearers of the new epoch” (Sheppard, 1996, p. 13). The monumental arch at Euston was iconic on two levels. From the standpoint of wayfinding, it clearly marked the entrance to the station. As an icon, the arch proved to be such a clear wayfinding device that it would be used in later stations to fulfill the same function. Some have argued that the station was to the modern city what the gate had been to the ancient city (Meeks, 1956, p. 39). Symbolically, it captured the spirit of leadership that the early railroad companies regarded as essential to building their railroad empires. Just as the Arc de Triomphe commemorated French victories, the Euston Station arch celebrated the conquests of the new technology as seen in the remnants of the arch (Figure 3).

III. Stations Find a Language (1844 – 1890)

By the mid-1840s, the railroad companies, architects, and engineers had overcome many of the initial fears raised by the new technology. Business was growing, ridership was increasing, and new stations had to be built economically and quickly. In addition, companies decided that stations should have the “right look” (Meeks, 1956, p. 39) Out of this milieu developed “the Railroad Style.” The Railroad Style was often based on the rural Italian villa and typically included a campanile that housed a bell and a clock, as well as arcaded
loggias that served as platforms (Meeks, 1956, p. 44). The Connolly Station in Dublin, Ireland (Butler, 1844) typifies this style (Figure 4). Central to Connolly Station is the tower which serves as a landmark in the urban setting. The tower alerts people to the significance of the structure. People knew that they had arrived at the station because it looked like a station; it looked like a station, in part, because the schemata of train station included such towers. The legibility of key architectural elements, like towers, makes the organization of spaces clearer. The Railroad Style helped solidify the concept of the station, both in terms of product recognition and cognitive map formation.

Figure 4. Connolly Station, Dublin, Ireland. Personal photograph.

Towers frequently housed bells and clocks. Audible cues were important sources of environmental information in the earliest days of train travel because they signaled passengers to the trains. Before the advent of rail travel, nationwide timekeeping did not exist. The arrival and departure of trains was based on a timetable which required accurate timekeeping. “Station clocks became symbols, governing the comings and going of trains and people” (Sheppard, 1996, p. 30). Furthermore, the presence of timekeeping devices underscores the temporal quality of executing wayfinding decisions and provides a frame for making such decisions in an orderly and sequentially logical fashion. Train sheds have been called the single-most important design innovation of the 19th century (Brown, 2005, p. 25). Built to shelter people and trains from the elements, train sheds evolved from simple wooden lean-to sheds into magnificent spans of iron and glass—materials that were impervious to the soot and sulphurous steam that had deteriorated earlier wooden trusses. Engineers were largely responsible for the construction of the sheds; architects, for the head houses. Some have argued that, during this period of technological advancement in the design and construction of train sheds, it was the engineers who experimented with a new language and new materials, while the architects were content to dress the new type in old clothes. In the early days of standardization, there seemed to be a disconnection between shed and station, with two notable exceptions—King’s Cross (Cubitt, 1852) and Paddington (Brunel and Wyatt, 1854) Stations, both in London, England.

Before the functional but uneaesthetic portico was added, King’s Cross’s façade clearly revealed the standard campanile and loggias seen in other stations. In a bold move, Cubitt expressed the “internal volumes” of the double train sheds (one for arrivals, the other for departures) onto the face of the building in the form of a pair of windows, simultaneously employing the iconic arch in the process (Sheppard, 1996, p. 18). Function is articulated in the form. The engineered design of the shed merges with the architectural design of the terminal. At Paddington Station, Brunel and Wyatt go one step further by integrating the shed and the terminal into a unified space, using an ecclesiastical model borrowed from the Middle Ages (Barman, 1950, p. 10). With respect to wayfinding, sheds create a visual path that connects the train to the building. The shed and the platform work together to direct passengers to their destination, be it toward the train or toward the exit.

The physical act of wayfinding occurs in two contexts—urban and architectural. The building must first be located within the urban setting before navigation through it can begin. When the train shed is articulated on the face of a building in the form of a window, two distinct functions occur which are dependent on the context. In an urban context, where the lines of the train shed are inscribed on the external façade of the station to form a large arched window, the window functions as a landmark, in accordance with Lynch’s theory. The unique physical qualities render the landmark “memorable in the context” (Lynch, 1960, p. 78). Such large windows generally are unique to stations and help distinguish them from other buildings within the urban fabric. Once the station has been identified, this type of information is no longer needed. As the context changes from an urban to an architectural one, new environmental information is required in order to complete the wayfinding task. The shed that gave form to the window as a landmark on the exterior of the building now functions as a path which delineates the way from the concourse to the platform inside the building. In addition to the dual functionality of the arched window/train shed, this icon in both roles underscores the temporal and sequential quality of perceiving environmental data, formulating a decision plan, and then executing that plan; the path to the platform does not appear before the building is identified as a station.

The success of railways in the middle of the 19th century
created a serious problem for architects: how to provide for crowds. “Churches and theaters, the principal prototypes for a building serving large numbers of people at one time, were not much help to him [the architect]: in these, the worshippers and the audiences flowed inward at stipulated times and outward at others, so that the entrances could be used as exits” (Meeks, 1956, p. 92). Circulation was the key to linking the numerous platforms with the services offered by stations. The problem of circulation and the layout of the head house set the stage for the creation of the most iconic and significant feature of railway architecture, the concourse.

Vestibules were typically present in the earliest stations. Often they were combined with the booking office. Waiting rooms were separate from the vestibules and were originally arranged by class in Great Britain and Germany. Improvements in construction techniques made it possible to span larger areas where ticket services and waiting rooms could be consolidated. The first great hall of this sort was constructed at Euston Station during the 1840s. Once wider spaces could be spanned, it was possible to reconsider the problem of circulation.

The midway or cross-platform is the predecessor to the concourse. This space allows people to move from one platform to another without having to cross the tracks. The area is perpendicular to the platforms, and gates mark the transition points between the cross-platform and the platforms. At either end of the midway are exits that allow direct egress to the street. Roma Termini (Calini, Mazzoni, Montuori, Castellazzi, Fadigati, Vitellozzi, and Pintonello, 1950) in Rome, Italy, contains a distinct midway.

The development of the concourse is significant with respect to wayfinding because it is so integrally tied to circulation. Arthur and Passini identify three kinds of primary routes: 1) main circulation between entrances and major destination zones; 2) circulation between one major destination zone and another; and 3) circulation within a major destination zone (1992, p. 48). The concourse is typical of the first kind of route. In addition, the authors find that “there is a direct link between the spatial organization of a setting with its related circulation system and the most appropriate decision plan” (Arthur & Passini, 1992, p. 49). The linear and axial nature of the concourse makes it a path; Arthur and Passini find the path the “dominating and controlling spatial element” (1992, p. 92). In Railway Stations: Planning, Design, and Management, Ross notes that “concourses are places where passengers stop to consider their next action” (2000, p. 120). Good design encourages the “free flow of passengers through public areas...in a logical order” (Ross, 2000, p. 111). Moreover, open planning allows for sight lines and open places where people can stop to orient themselves.

The years between 1844 and 1890 witnessed the standardization of railroad station architecture which included the acceptance of certain architectural conventions in the articulation of those stations. These conventions—the tower, the bell, the clock, the window, and the concourse—not only were symbolic of the stations they represented but served as valuable icons that proved useful to wayfinding.

### IV. Monumentality, Mass, and Memory (1890 - 1935)

Innovations in station design had evolved to a high level of sophistication in the first 60 years of rail travel. By 1890, many of the conventions of railway architecture had been established—in particular, the plans and the iconography. Once the basic problems created by the new technology were addressed, railroad companies began to focus their resources on increased luxury, safety, speed, and size. The conceptual image of the station changed “during the 1890s from the easily recognized one of the previous period toward that of an ordinary monumental public building” (Meeks, 1956, p. 125).

What propelled station design into a “period of giganticism” (Meeks, 1956, p. 26)? Meeks attributes this development to the Columbian Exposition of 1893 and the City Beautiful Movement (1956, p. 126). Under the influences of the École des Beaux-Arts and the City Beautiful Movement, stations began to mimic the neoclassical language used to articulate other significant public buildings of the period. Melvin refers to this architecture as “monumental urbanism” (2006, p. 86). Monumental urbanism mirrored the message railroad companies wished to convey, one of power and success.

Like the adaptation of the triumphal arch, the colonnade was another feature borrowed from classical architecture for use in railroad station design. Colonnades had always been a part of the language of railway stations as exemplified in the second Gare du Nord in Paris and the Stazione Centrale in Naples. The most exuberant expression of monumental columns, however, was employed at the Pennsylvania Station in New York City (McKim, Mead, & White, 1910). Numerous massive columns lined the façades of this station. Ironically, the preponderance of columns virtually obscured the function of the building; the neoclassical language made it more difficult to read the building as a station.

The marquee and its larger version, the portico, have been important features of railroad architecture from the early days. Like the train shed, marquees and porticos afford protection against the elements. They also provide environmental information with respect to wayfinding because they draw attention to the organization and give significance to the entrances (Arthur and Passini, 1992, p. 121). Perhaps the most spectacular portico is found at Pittsburgh’s Union Station (Burnham, 1903). The station’s rotunda or “carriage concourse” resembles the interior of the Pantheon in Rome; its skylight is reminiscent of the lunette at the Gare de l’Est. Its spacious opening pulls the visitor into the entrance and, in the process, converts the urban scale into an architectural one. On a smaller scale, the marquee performs the same function as a portico. Cincinnati’s Union Terminal (Fellheimer, Wagner, Wank, & Cret, 1933) includes not only the iconic marquee, but also the arched window, the clock, and modified towers (Figure 5).

Arthur and Passini find that a building’s layout is implied by its entrance; thus, the exterior is linked to the interior. Because the entrance is vital to understanding the space within, important environmental information is gained while...
physically approaching a building. Buildings, according to them, can be approached in one of three ways—through a frontal approach, an oblique approach, or an indirect approach, and “the legibility of an entrance varies with the angle of approach” (1992, p. 117). A frontal approach is one in which the approach is perpendicular to the building. Cincinnati’s station is blessed with a generous, axial, frontal approach. Many urban stations lack such a site. Their façades are flush with the sidewalks that lead to them, and such stations must rely, instead, on marquees to signify their entrances. Arthur and Passini refer to this approach as an oblique approach. An indirect approach occurs when the entrance is obscured from view by a corner.

The increase in size during this period was directly related to the increase in function. Hotels had long been a staple of railroad architecture but now office space and terminals for other forms of transportation were included in the program. Philadelphia’s Reading Station (Kimball, 1893) was the first to include an office building. As the size of these stations grew, wayfinding through them became more difficult. The cross-section of New York’s Grand Central Station (Reed & Stem, Warren & Wetmore, 1913) hints at the complex layout of the station. (Fig. 6) The station provides access not only to standard heavy rail but to suburban and subway trains, as well.

John Ruskin, 19th century social critic, is credited with saying that buildings are society’s memories. What is it, then, that makes a building memorable? Arthur and Passini suggest that form and size, visibility and access, frequency of use, and symbolic significance are contributing factors. By the 1890s, the period of experimentation was complete. Stations and the iconography representing them were well established. The iconography had served to provide stations with a recognizable identity. Arguably, there was less need for the iconography in the 20th century than before because, by the 20th century, a schema of a train station had formed in the minds of most people. Still, the icons long associated with stations continued to be used during the early days of modernism. Why would architects continue to employ such details at a time when many of the conventions of more traditional architecture were being eliminated from the language?
V. Transparency and Volume (1935 – present)

After World War I, Modernism affected and transformed railway architecture, as well as other forms of architecture. New materials, such as reinforced concrete, structural steel, and walls of glass, enabled architects to design spaces where transparency and volume replaced enclosures and mass. Modernism changed the form of architecture because new materials made such changes possible. Modernism also changed the discussion of architecture, as well. Form and function became distinct components of the built environment. Ross views Modernism as a rejection of architectural antecedents, a view which strips Modernism of both precedents and continuity. Meeks differentiates modern architecture from earlier periods on the basis of a distinction between form and function. He finds that there was an emphasis on the exterior (form) during the industrial era and that modern theory “gives primacy to functionalism” (1956, p. 4).

However, railway stations from the earliest days had developed in direct response to the functional needs of moving both people and trains. Meeks’s distinction between modern architecture and earlier styles is, at best, a weak one with respect to train stations. Unlike Meeks, who relates functionalism to Modernism, Thorne recognizes the functionalism of the earlier stations: “Whereas stations in the past may have been more sensitive to solving their functional requirements or reflecting, through their architecture, their mission as a transport center, today there is greater sensitivity to the role of the station in its context” (2001, p. 22). Because the issue of function had been addressed by station designers from the outset, stations today can focus on their role in the urban context. Both Meeks and Thorne emphasize the differences between the modern station and its antecedent, although for different reasons. Meeks does so on the basis of function; Thorne on the basis of context.

Transparency and volume were characteristic of the great train sheds of the 19th century, but not of the stations themselves. One of the hallmarks of 20th century rail station design is that transparency and volume replace the lithic and massive language of the monumental stations. As the language of station design continued to evolve, did the change in both aesthetic and materials really strip railroad architecture of its antecedents, as Ross suggests? Or, is there an enduring quality to the icons—the towers, arched windows, clocks, marquees, and concourses?

There are numerous examples of railroad iconography in 20th century stations. The entrance to Florence, Italy’s Stazione Santa Maria Novella (Michelucci, 1936) is marked by a long cantilevered marquee. In newer stations, tensile membrane roofing materials are used for marquees, porte-cochères, and platform canopies. While the form for such roofing is a 20th century innovation, the function remains the same.

The triumphal arch continues to figure prominently in modern station design. Stations at Chur, Switzerland (Brose & Obrist, 1992), Kowloon, Hong Kong (Farrell, 1997), Lyons, France, (Calatrava, 1994) and Rotterdam, The Netherlands (Reijnders, 1993), preserve the arch. Reijnders, the architect of the Rotterdam Blaak Station (1993), incorporates both a marquee and an arch in a radical way; the marquee is a transparent disc that is suspended from a huge arch which spans not only the entrance but a bicycle path, as well.

The clock tower at Tampere Station in Finland (Flodin and Seppälä, 1938) rises in stark contrast against the horizontal planes of the head house. In Amsterdam’s Duivendrecht Station (Kilsdonk, 1993), the tower is an open, red equilateral triangle instead of a closed, four-walled structure. The expression of the tower has changed, but the function has not. In both of these instances, the tower functions as a beacon.

The concourse continues to be an integral part of stations in the modern and post-modern period. While a comparison of rail and airport terminals seems inevitable, site context generates significant differences between the two. Railway stations are generally erected in an urban setting while airport terminals exist in relative isolation from urban structures. The correlation between the legibility of a building from its exterior and the functionality of its interior is crucial to railway stations. Both the covered walkway and marquee that lead to Roma Termini provide a direct path to the station’s main entrance and to its services, as well; the connection between exterior legibility and interior function is unbroken (Figure 7).

Figure 7. Stazione Termini di Roma. Rome, Italy. Personal photograph.

As previously noted, wayfinding is based on perceiving environmental information, forming a decision plan based on that information, and then executing the plan. Towers, marquees, arches, and concourses are important landmarks. However, as the decision process becomes more refined, signs become more important. It is not enough to know that the concourse leads to the tracks; passengers must also be able to distinguish one track from another, and the most efficient way to do that is through signs. There is, however, little accessible documentation of signage strategies in the early days of rail travel.

Signage refers collectively to signs and includes both branding (promotion of a product or service) and informational
signs. The global economy has led to multi-lingual signs. Language barriers affect the readability of signs or how well they can be understood. In order for them to be understood, they must also be legible: “legibility is the ease with which information is able to be perceived” (Arthur & Passini, 1992, p. 50).

A variety of factors can affect the legibility of signs, particularly in train stations. Typically, people move through stations in groups, especially when exiting trains. This makes it more difficult not only to find signs among a busy crush of people but also to read them once they are located. Since passengers may be unfamiliar with the layout of a train station, it is very important not only that signs be placed where they can be perceived but also that they be placed within the proximity of decision points. The Waterloo Station vestibule crystallizes the importance of proximity in decision-making (Figure 8). At the time the photograph was taken, Waterloo Station in London, England, erected in 1848, served as the station for the international Eurostar, the national rail, and the London underground. In this picture, the navy blue sign, closest to the camera position, provides directional language for the various intermodal rail lines and for tickets. Also legible from this vantage point is the clock. The information contained in this sign and the clock enables visitors to begin to make their decision plans (ultimately destination: Eurostar or metro?) within a temporal context (how much time is there to make the connection?).

Figure 8. Waterloo Station. London, England. Personal photograph.

There is insufficient architectural information from this camera position in the vestibule for an inference tactic to be of any immediate use in forming a cognitive map. The inference tactic, however, draws one further into the space, since through memory one recalls that proximity improves focus and, therefore, legibility. While the four-sided informational marquee suspended from the ceiling is not yet legible, its visual presence alerts visitors to the existence of additional information, information that may (or may not) prove useful in formulating their decision plans. The third layer of information visible from the vestibule is the train schedule. The concept of train schedule boards is likely recorded as part of the schema of train station even if one has never before visited this station. Passini finds that there is a direct correlation between the hierarchy of decision-making in wayfinding and the perception of messages on signs.

Too much signage and branding can lead to stimulation overload. When this occurs, people resort to coping mechanisms which result in either ignoring some of the stimuli or regrouping individual units of information into larger but fewer chunks of information. When people encounter a setting with which they are unfamiliar, they scan the environment, relying on fragments, to search for usable information. If the information is to be retained, then it has to be organized into a memory of longer duration. Since one of the goals of railroad architecture is to provide an environment where people can move efficiently toward their destinations, functionality is enhanced when environmental information is arranged accordingly.

There are several ways to present environmental information so that its effect is maximized. When environmental information is arranged sequentially in conjunction with the decision-making process, wayfinding is enhanced. Proper contrast between the foreground and background in signs and the way in which data are organized on them, improve the perception of the information contained in signs. Pictorial representations—arrows and pictograms—also render the space more legible. If pictorial representations in signage improve one’s ability to commit the information to memory, then it is logical to infer that pictorial representations in architecture do the same. These pictorial representations in railroad architecture are the icons long associated with stations—the arched windows, the clocks, the towers, the concourse, and so forth. The iconography of rail stations has not disappeared with the advent of Modernism; rather it continues to manifest itself in the form of landmarks, paths, and edges. In a global society where stations are part of intermodal hubs, the iconography of railroad stations may be more important than ever.

In Conclusion

“There was no functional precedent for the depot, every solution had to be invented” (Meeks, 1956, p. iii). The significance of the railway station and its contribution to the built environment has been largely overlooked in an age when the heights of skyscrapers and the plasticity of form define the criteria by which buildings are judged. Nevertheless, the genesis and the evolution of the railroad station embody the vision, the development, and the expression of a unique and complex architectural type. While railroads link cities, rail stations connect interior space and exterior context, ingress and egress, time and space, language and syntax, and, most of all, form and function.

Train stations are unique in that interior spaces and site context are dependent on each other in order to move both people and trains efficiently. Interior spaces, largely in the form of circulation routes, are integrally tied to the tracks outside the station. Beyond the tracks, stations are part of an urban
context where they must be distinguishable from other types of buildings. Consequently, stations must be legible both from the inside and from the outside. Very early in the development of railway stations, it was determined that a station ought to look like a station. Out of this decision arose a corpus of iconography that contributes to the recognition of stations, both in the formation of schema and in the formulation of cognitive maps. Iconography such as towers, arched windows, concourses, train sheds, clocks, etc., contributes to the imageability of stations, features that make stations visually distinct from other types of buildings.

Ingress and egress are the primary programmatic requirements of stations. Unlike other large-scale architectural types, such as churches or auditoriums where people generally arrive together at one time and depart in the same manner, the proximity of simultaneously arriving and departing trains requires a greater and more complex accommodation of circulation needs. Stations have always adapted to this challenge as demonstrated by their layouts. Midways, concourses, train sheds, head houses, and one- and two-sided stations are evidence of their meeting programmatic requirements in a flexible manner.

Not since the medieval cathedrals marked the time for prayer has there been a need to architecturally connect time and space. Train schedules organize the arrival and departure of trains which, in turn, affect the movement of passengers through stations. The temporal quality of stations is also reflected by the way in which people navigate through stations. Decision plans are sequentially formulated and executed. Effective wayfinding in stations must be timely; anything less than that is useless.

The language and syntax of stations must be synthetic in order to provide seamless navigation through stations. The language manifests itself in two ways—in architectural features and in signage. Architectural components—landmarks, nodes, paths, districts, and edges—constitute the language of architecture. In general, buildings are made legible through these elements. In particular, it is the language that distinguishes one type of architecture from another. Regardless of the period styles in which stations have been built over the years, stations are stations because their language is derived from recurring iconography. While the origins of the iconography may have had their roots in other architectural types, the concentration and preponderance of such forms in railroad architecture render them unique to station design. The syntax of the station is the arrangement of these forms: the tower leads to the arched window that leads to the marquee that leads to the concourse that leads to the sign that points to the platform, etc. Both the language and the syntax are intentional and significant. Environmental information is the perception of the language; moreover, the ability to form cognitive maps, to devise a decision plan, and to execute such a plan are enhanced when the syntax is coherently and logically ordered. The language and syntax of rail stations work together to facilitate wayfinding.

The function of the rail station distinguishes it from other types of architecture. Throughout its history, station design has always addressed function. While there was no functional precedent, architects and engineers identified the programmatic requirements and designed stations accordingly. The function of stations has always been supported by their form. Stations function more efficiently because the architectural forms, the iconography, work on many different levels. On a cognitive and perceptive level, the form helps people read the function of the station. On a symbolic level, the form of the iconography conveys a station’s mission. On a literal level, the form is a physical anchor that helps one find his or her way through stations. Together, form and function separate rail stations from other architectural types.

Despite the lack of a clear precedent, train stations have become very efficient spaces for moving people and trains. While the genesis of the railway station predates the study of wayfinding, station design has always responded pragmatically to the wayfinding needs of its users. The particular relationships between interior space and exterior context, ingress and egress, time and space, language and syntax, and form and function not only contribute to the evolution of the railroad station as an architectural type but serve as a model through which the principles of wayfinding and design complement each other. The manifestation of this union is the iconography which unites wayfinding and design in train stations. The iconography embodies the essential qualities of train stations, both defining the architectural type from the perspective of public recognition while simultaneously informing the personal schema. From exterior to interior and from external to internal, the study of the history of the train station provides a valuable model for understanding the practical design responses to wayfinding needs.

Works Cited and Consulted

Mentor Comments

Catherine Wallack draws attention to the interdisciplinary nature of Tricia Quinn’s work and the originality of her research, characteristics also noted by another member of her thesis committee.

I was delighted to serve as Ms. Tricia Quinn’s advisor and as her mentor for her Honors Thesis. Her thesis entitled Train Stations: Iconography, Wayfinding and the Evolution of a Type is a remarkable effort. Her work is highly original and represents a synthesis of several of Ms. Quinn’s related interests: architecture, interior design and train stations.

Ms. Quinn brings tenacity and rigor to both writing and research. She worked with great independence on the creation of this undergraduate thesis. Her thesis committee merely provided feedback and general guidance on the document in process. The idea for studying train stations in a manner that was highly relevant to her own discipline was wholly Ms. Quinn’s idea. Rather than rely on secondary sources, she took the initiative to personally document a number of critical exemplars used in the thesis.

Ms. Quinn’s thesis is an impressive interdisciplinary work. While train stations have been researched from an architectural and historical perspective, Ms. Quinn’s work brings a new viewpoint to the study of this building type. By considering these buildings in terms of phenomenology, she articulates the particular roles specific architectural elements play in navigating these complex spaces. By documenting and exploring these relationships, Tricia brings forward important material that is applicable to a number of fields. Her work will be relevant to those involved in all aspects of environmental design, including urban design, architecture, interior design, and related social sciences.

This thesis brings together a thoroughness and originality rarely found in undergraduate work. This unique document well represents the high quality of work capable by students in the University of Arkansas, and is highly deserving of publication in Inquiry.
RELATIVE SEARCHING USING AN ORDERED TOKEN LIST

By Anthony Rosequist
Department of Computer Science and Computer Engineering

Faculty Mentor: Dr. Haiying (Helen) Shen
Department of Computer Science and Computer Engineering

Abstract
Many organizations have large amounts of information, such as consumer data, that need to be processed. Traditional searching algorithms only attempt to find exact matches to particular queries. This is undesirable when data are missing, outdated, or inaccurate. Therefore, a new type of search must be developed to locate records that are considered “interesting” to the user. This research paper examines past attempts to solve this problem and explores a new method involving ordered token lists to achieve this goal. The algorithm was developed, implemented, tested, and optimized.

1. Introduction
As technology improves the ability to gather information, the quantity of data significantly increases. A prominent example is consumer records, which consist of segments of information associated with individual consumers. In addition to traditional exact match searching, businesses are interested in searching methods that find similar, relevant matches. For example, a company may have the following source records:

Ann | Johnson | 16 | Female | 1248 | Elm | St
Joe | Anderson | 25 | Male | 512 | 1nd | St
Jessica | Smith | 116 | Female | 716 | Main | St
Samantha | Anderson | 28 | Female | 1248 | Oak | Dr

When, the following query is entered, all similar records are desired:

Jessica | Johnson | 16 | Female | 1 Main | St

The algorithm should return Jessica Smith’s record first, and Ann Johnson’s record second, since they are the most similar records to the query. Through this example, we can see the importance of not requiring a strict match, since differences may be caused by missing, outdated, or inaccurate data. Jessica Smith may have been married since data was last gathered, explaining a change in her last name, or it is possible that Ann Johnson’s middle name is Jessica and she recently moved to another street. However, the chance that the remaining source records are similar is extremely slim. Therefore, it is reasonable to allow a certain degree of dissimilarity between a source record and query record and still consider them relevant.

The rest of this paper is structured as follows. Section 2 reviews representative data searching approaches. Section 3 presents a search method based on an ordered token list, including its algorithms and some experiment results. Section 4 describes three strategies to improve the performance of the basic algorithm, and the performance of the strategies. It also shows the performance of the optimized algorithm in comparison with the basic algorithm. Section 5 recommends future research directions.

2. Related Work
Previous efforts have contributed to solving this problem. The simplest solution is linear searching, which involves comparing the query record to each source record using a distance function. This is extremely inefficient.

Andoni and Indyk proposed a method called Locality Sensitive Hashing (LSH) [1]. The basic premise behind LSH is to use hashing to place similar items into the same hash buckets. Then, for the query items, the correct bucket is determined and searched, so the domain of items needing to be examined is greatly reduced. LSH is accurate in locating most similar records, but requires too much time and memory to be practical in most real-world applications [2].

Other projects have improved the efficiency of LSH through various means. One method was to use the Lempel-Ziv-Welch (LZW) algorithm, which is a string compression technique [3]. Combining LSH with LZW was successful in improving the time and memory consumption of the original algorithm, but these deductions were still not significant enough for practical implementations [4].

Another method of achieving the same goal as LSH is using min-wise independent permutations to reduce the refinement stage of LSH. This is significantly faster than LSH, although the accuracy of the returned data is somewhat diminished, since the number of false positives is increased. False positives are located records but are not actually records similar to a query. This method is currently in development, and attempts are being made to increase its accuracy [5].

The algorithm proposed in this paper differs from these methods. Instead of reducing the search domain, it directly points to matches with similar tokens in a time-efficient manner using an ordered token list.

3. Ordered Token List Search (OTLS)
The algorithm requires two input parameters: a set of source records and a set of query records, denoted by \( S \) and \( Q \), respectively. \( S \) contains records \( s_1, s_2, ..., s_n \) and \( Q \) contains records \( q_1, q_2, ..., q_m \). The records in both sets consist of a series of tokens, separated by a common delimiter. Many of the records will share common tokens, so it is useful to define a third set, \( T \), which contains all of the unique tokens that exist in
the source records, denoted as \( t_1, t_2, \ldots, t_j \), in alphabetical order. The \( T \) is the ordered token list, which is the main structure used to allow efficient searching in this process.

Algorithm 1: Pseudo-code of Ordered Token List Search algorithm

(1) create an alphabetized list of tokens
(2) for each source record \( i \) do
  (3) for each token in the source record do
    (4) add \( i \) to the Linked List for the token
  (5) end for
(6) end for
(7) for each query record \( i \) do
  (8) create new array, \( \text{results} \)
  (9) for each token in the query record do
    (10) add the Linked List for the token to \( \text{results} \)
  (11) end for
(12) sort the elements of \( \text{results} \) by number of occurrences
(13) output \( \text{results} \)
(14) end for
end

4. Performance Evaluation and Improvements

Ordered Token List Search will return every record that is relevant, sharing any tokens in common (above a particular threshold, defined by the user). Therefore, it will return no false positives, and it will miss no true positives. In this sense, the algorithm is completely effective.

However, time requirements also needed to be considered. To test Ordered Token List Search, records reflecting the nature of real world consumer data were provided by Acxiom Corporation. The Ordered Token List Search algorithm processed 423,801 source records and 10,000 query records in 2,910.98 seconds (about 48.5 minutes) on a Windows PC. This was significantly longer than expected, making this simple implementation unusable for real-world purposes.

To address this problem, we further improved the basic Ordered Token List Search algorithm to enhance its efficiency in terms of query latency. Specifically, we integrated three strategies into the basic method: binary search, list merging, and token list serializing. Instead of using sequential search to locate all related records in the ordered token list \( T \), binary search is adopted in this step. After relevant records in \( T \) are located, the list merging method is used to speed up the process of locating actual similar records. The details of the three strategies are introduced in the following sections. Finally, rather than building an ordered token list every time for data queries, a token list serializing method is developed to build a static ordered token list to save list construction time.

4.1 Binary Search

A timing analysis was used to determine bottlenecks in the code. As Figure 3 shows, a great majority of the time is spent filling the token lists with the right source indexes.

Upon further investigation, one particular section of the code can be seen as the primary performance bottleneck. Filling the token lists requires two operations: finding the list associated with the current token and appending the source
record index to the end of it. Since appending to the end of a list can be done in constant time, the main waste in the program is the time taken to find the list associated with the current token. Fortunately, this can be greatly improved.

The simple implementation of the algorithm uses linear searching: It steps through each token until it finds the correct one. However, the token list being used is in alphabetical order, so the time to find the correct list can be reduced. The linear search was replaced with a binary search. Instead of starting at the beginning, the search begins at the middle of the list, and is able to cut the search space in half after each iteration. This decreases the asymptotic runtime of this particular operation from \(O(n)\) to \(O(\log n)\), a substantial improvement. Augmenting OTLS with binary token search reduced total runtime to 617.65 seconds (about ten minutes), a decrease of almost 80%, significantly impacting the actual time complexity of the algorithm.

4.2 List Merging

Although 10 minutes is a respectable runtime for 10,000 queries, further refinement was possible. Performance profiling identified that the query processing subroutine accounted for almost 97% of the total runtime, as shown in Figure 4.

To understand why query processing takes such a large amount of time to complete, a deeper explanation of the operation is necessary. Figure 5 shows the steps that a query record takes during processing, using the example from earlier in this paper. For each token, the token list is found and appended to the end of the results list. Once all of the token lists have been gathered, the results list is sorted in ascending order. From this, the indexes that appear most often (such that their percentage of similarity is above the threshold) are calculated and returned.

\[
\text{results} = \begin{bmatrix} 2 & 72 & 492 & 2 & 52 & 492 & 2 & 72 \\ 2 & 2 & 2 & 52 & 72 & 72 & 492 & 492 \end{bmatrix}
\]

(100% match with #2, 66% match with #72 and #492, 33% match with #52)

Figure 5. Determining the closest matches of a query.

The largest amount of time in this subroutine was spent sorting the results list. Since any source record that shares any tokens in common will appear in the list, it is expected to be large, requiring a substantial amount of resources to sort. However, this can be improved by noting that the token lists are already sorted in ascending order. Therefore, the results list can be sorted while it is being built instead of waiting until the end. This is essentially the "merge" routine that is used during a merge sort operation, so a simple implementation was used. By changing the method of merging lists, the entire query file was processed in only 21.78 seconds.

4.3 Token list serializing

Twenty seconds is certainly a reasonable runtime for an operation of this magnitude; however, some corporations may process millions of records at a time and would benefit from an even faster execution time.

Looking at the results from Figure 6, it can be seen that building the token list is now the bottleneck routine in the algorithm. Before any search, it must read the entire source file and return a sorted array of all unique tokens. An approach for
Reducing this time is to serialize the list, saving it to the user’s computer, so it does not need to be generated every time. This is most optimal for static databases, where there will be few, if any, changes to the source files. However, even if the database is dynamic, changes may be done relatively quickly, since it is simply a matter of adding the tokens that do not already exist. The only time when this will not be practical is if there will be millions of source records added continually, a scenario unseen in most real-world applications of this method.

A separate file was created to hold the tokens, which were simply placed on separate lines in alphabetical order. Instead of generating this token list every time, the program simply loads this file into an array to make it usable. This dramatically reduced the overall runtime of the application. It takes around thirty milliseconds to load the token file, as opposed to thirteen seconds to build the token list previously. This lowered the total time to around 8.5 seconds.

4.4 Further Optimization

If needed, the program could be improved even more. The implementation that was used has overhead in object creation, method calling, and other operations that slightly added to the runtime. A final production version would need to reduce this overhead as much as possible; however, it would not reduce the asymptotic runtime and the results will vary across multiple systems, so it was not pursued further for this project. It is estimated that it would possibly reduce the runtime in this example another 0.5 seconds.

4.5 Performance Comparison

The Ordered Token List Search method of data searching has proven to be very effective. Due to the inflexible structure of the list orderings, it is impossible for the algorithm to miss a related entry. Similarly, it is impossible for it to identify false positives as matches, making it entirely accurate. The Ordered Token List Search is also very efficient. Although the original implementation took a significant amount of time to complete, some optimizations were made that substantially reduced the time consumption. As can be seen in Figure 7, the final version runs in a fraction of the time that the original did, and Figure 8 gives a closer look at the last three implementations. Being able to match 10,000 records with their closest matches in approximately eight seconds on a standard system is reasonable for most environments.

5. Conclusion and Future Directions

As technology to gather information improves and becomes more widespread, more advanced searching algorithms are necessary. Similarity data searching, instead of exact match searching, is increasingly needed by many organizations having large amounts of information. This paper proposed a data searching algorithm based on an ordered token list. This algorithm is able to locate not only exact matching data but also data similar to a query. Furthermore, it has high accuracy in that it returns all similar data without missing any false positives. To further improve the efficiency of the basic ordered token list search algorithm in terms of query latency, we integrated three strategies into the basic method: binary search, list merging, and token list serializing. Performance testing results show the superiority of the optimized searching algorithm in comparison with the basic algorithm.

Ordered Token List Search is an accurate, fast method of achieving this goal, running significantly faster than other routines. In the future, features can be added to improve the usefulness of the base algorithms in particular situations, such as adding weights to particular tokens (for example, matching a last name is more relevant than matching a person’s state of residence) or relative comparisons of tokens (someone that is one year older is a closer match than someone that is eighty years older than the query). These would require specific knowledge about the records and we intend to address this challenge in our future work. However, the low runtime of this method allows for additional features to be added without significantly reducing performance, providing another significant benefit.

Acknowledgements

I thank my mentor, Dr. Haiying Shen, and her graduate research assistant Ting Li for their help with this project. I also want to express deep gratitude to Dr. Tom Schweiger, Acxiom
Corporation, for his insights, and Acxiom Corporation for financial support. This work was previously published [6].

References


Mentor Comments

Dr. Haiying Shen describes the way in which Anthony Rosequist’s research and article represent the quality of outcomes generated by a combination of team effort and individual initiative.

Anthony Rosequist is an undergraduate student in the Computer Science and Computer Engineering Department, and one of our department’s top students. I am the research mentor for his research work titled “Relative searching using an ordered token list.” This research work has already produced two papers co-authored by Anthony, published in ACM/SIGCSE CCSC Mid-South ’08 and the International Conference on Data Mining ’08. Anthony’s research is part of a broader project “hash-based proximity clustering for neighbor search in Acxiom database.” The members of this project include Anthony Rosequist, another undergraduate student, and a Ph.D. student. This project is designed to study the effectiveness of locality sensitive hash function (LSH) on data searching. First, this project designed a method using LSH for data searching. Second, this project developed a simulator to test the performance of the LSH-based data searching method. Third, the performance of LSH on data searching has been analyzed, and new methods including the token-list method to improve the LSH-based method have been explored and developed. Experimental results demonstrate the superiority of the proposed methods compared with the LSH-based method with regards to memory and time consumption. Anthony has been working on the project. In each step of the project, he not only accomplished his assigned work independently, but also cooperatively worked with others. More importantly, he implemented the token-list methods by himself.

This research work focuses on data searching in a massive database. It thoroughly investigates the current data searching methods, and proposes the token-list method to achieve enhanced data searching in terms of efficiency and effectiveness. The token-list method has significant impact to our society defined as the “Information Society”, in which tremendous growth of information generates an increasing need for an efficient data searching method. In addition, this research work has many technical merits and significant contribution to the computer science and computer engineering area. This research provides critical insight into data searching, which is expected to have significant impact on data processing research. The outcome of this research is expected to serve the data processing community as a vehicle to conduct further research and experiments, and will advance the state of the art in data processing research area.

Based on the originality, novelty and contribution of Anthony Rosequist’s research work, I highly recommend this work for publication in Inquiry.
Computational Modeling of Oxidative Stress: An Analysis of NAD(P)H Effects on Nitric Oxide and Superoxide During Hypertension

By Aaron Strobel
Biological Engineering Program

Faculty Mentor: Mahendra Kavdia
Biological Engineering Program
Biological and Agricultural Engineering

Abstract

Nitric oxide (NO) is inactivated in the human body when exposed to superoxide (O$_2^-$). This reaction forms peroxynitrite (ONOO$^-$). Superoxide is produced in the cardiac system by several mechanisms, including NAD(P)H oxidase. Superoxide dismutase (SOD) breaks down superoxide into oxygen and hydrogen peroxide. This prevents superoxide from reacting with nitric oxide and allows normal function to take place. Superoxide and peroxynitrite are main contributors to vascular disease in the human body, in particular hypertension. Experiments have shown that there is an increase of superoxide production in spontaneously hypertensive rats (SHR) vs. age-matched Wistar Kyoto rats (WKY) that were normotensive. The increase in superoxide production intensifies in the presence of scavenger DETCA Cu$^{2+}$/Zn$^{2+}$. A mathematical model has been developed by Kavdia and Popel to calculate concentrations of NO, ONOO$^-$, and O$_2^-$ in the arterial and venule pair. Using this model we calculated the arterial and venule NO, ONOO$^-$, and O$_2^-$ concentration profiles for normotension, hypertension, and disease. The trends in superoxide production in this paper can help researchers and the medical community understand hypertension and vascular disease more thoroughly. Further, for future research, observed increases of Nox1 and Nox4 expression suggest specific regions where O$_2^-$ will be high and which need to be evaluated.

Introduction

Cardiovascular disease is the leading cause of death in the United States. According to the American Heart Association, this disease was responsible for taking over 870,000 lives in 2005. Increased vascular production of reactive oxygen species (ROS) is a common characteristic of cardiovascular disease. There are several factors that increase the risk of cardiovascular disease including hypertension, hypercholesterolemia, and diabetes mellitus. In the United States, hypertension affects approximately 58 million Americans.

Vascular and cardiac tissues are rich sources of ROS, including superoxide (O$_2^-$), hydrogen peroxide (H$_2$O$_2$), and peroxynitrite (ONOO$^-$). Reactive oxygen species are the by-products of oxygen metabolism and are normally present in low levels of concentration inside the cells. ROS are needed in aiding the signaling processes within the cells, and also in regulating vascular smooth muscle cell contraction and relaxation. Increasing amounts of O$_2^-$ is the most common trend in vascular diseases, such as hypertension, because it causes oxidative stress in the vascular tissue. One of the mechanisms for controlling oxidative stress in the vascular system is superoxide dismutase. Superoxide dismutase consumes O$_2^-$, and then converts it into less harmful compounds.

The major sources of vascular superoxide include xanthine oxidase, nitric oxide synthases, mitochondrial oxidases, or NAD(P)H oxidases. Each of these sources generates superoxide in a different manner in the human body.

Xanthine oxidase (XO) is an iron sulfur flavoprotein that is found in high concentrations in the endothelial cells and plasma, but not in smooth muscle cells. Xanthine generates the superoxide by catalyzing hypoanthine and xanthine to form uric acid. Nitric oxidase synthases (NOS), found predominantly in the endothelial cell region, play a major role in vascular diseases. Endothelial nitric oxide synthases (eNOS) require tetrahydrobiopterin (BH$_4$) for the transfer of electrons to the nitrone of the L-arginine. This reaction under normal conditions forms nitric oxide in the vessels. However, when BH$_4$ is not present, eNOS generates O$_2^-$ and H$_2$O$_2$. Mitochondrial oxidases utilize the uncoupling of oxygen during mitochondrial oxidative phosphorylation, which occurs during the production of ATP, to synthesize O$_2^-$. The superoxide produced with lesion development in the arteries comes primarily from the mitochondrial oxidase.

In particular, in the work reported here, we analyzed more closely the NAD(P)H oxidase as a source of vascular superoxide. NAD(P)H oxidases are present in endothelial cells, fibroblasts, smooth muscle cells, neutrophils, and phagocytic mononuclear cells. The study of NAD(P)H started at the biochemical reaction of the compound in the cardiovascular system. NAD(P)H is a multi-subunit enzyme that catalyzes O$_2^-$ production by reducing oxygen by one electron and using the NAD(P)H as the electron donor.
NAD(P)H + 2O2 → NAD(P)+ + H+ + 2O2− (1)

One of the major sources of ROS formation is from the NAD(P)H oxidases. Endothelial cells and fibroblasts express both NAD(P)H oxidase 2 (Nox2) and NAD(P)H oxidase 4 (Nox4). Vascular smooth muscle cells express Nox1 and Nox44.

In addition, nitric oxide released by endothelial cells is a key chemical that regulates blood flow. In oxidative stress conditions the availability of nitric oxides is reduced in vascular tissues, a process which is known as endothelial cell dysfunction10. With this dysfunction comes the natural response of the human body to overcome this production of superoxide. Superoxide is neutralized by a group of antioxidant enzymes. These enzymes include SOD, CAT, GPx, and thiol-disulfide oxidoreductases12.

The research described here dealt primarily with SOD as the main antioxidant enzyme. There are at least three distinctive isoforms of SOD identified in humans: mitochondrial manganese-containing SOD (MnSOD, SOD2), the cytosolic copper/zinc-containing SOD (CuZnSOD, SOD1) and the extracellular SOD (eSOD, SOD3)13. SOD dismutases superoxide to form hydrogen peroxide and oxygen. Loss of endothelial NO available to endothelial cells is caused by the reaction with O2. This reaction forms ONOO−, which is a key component in many cardiovascular diseases such as hypertension, diabetes, and atherosclerosis. Understanding the levels of nitric oxide, superoxide, and peroxynitrite should contribute to a better understanding of the diseases and the oxidative stress state of the vascular system.

In 2004, Kavdia and Popel developed a mathematical model to create the concentration profiles for NO, ONOO−, and O2 for an arterial and venule pair during microcirculation1. This model defines the geometry of the arterial and venule vessels parallel to each other, while using diffusion rates and chemical reaction rates to calculate the concentration profiles.

The objective of this study was to predict NO, ONOO−, and O2 concentration profiles for three different cases: 1) basal conditions compared to NAD(P)H stimulated conditions (2) basal conditions in normotensive and hypertensive rats with inactivation of SOD by DETCA (3) NAD(P)H stimulation in normotensive and hypertensive rats with inactivation of SOD by DETCA. We used data from Tamara Paravicini’s “Increased NADPH-Oxidase Activity and Nox4 Expression during Chronic Hypertension is Associated with Enhanced Cerebral Vasodilation to NADPH In Vivo”. In their work, the arteries from Wistar-Kyoto rats (WKY) were compared to spontaneously hypertensive rats (SHR). The O2 production in these arteries was measured by 5 μmol/L lucigenin-enhanced chemiluminescence under various conditions. The results from Paravicini’s experiment are displayed in Figure 1. The trends in these data were applied to the Kavdia and Popel model for the purpose of the current research.

### Methods

#### Model Geometry

A previous model by Kavdia and Popel was used to simulate arteriole/venule during microcirculation. This model contains six different regions in the vessels: red blood cell rich (CR), red blood cell free (CF), endothelium (E), interstitial space (IS), smooth muscle (SM), and a nonperfused parenchymal tissue (NPT). The parenchymal tissue (PT) is the region around the arteriole and venule pair. Figure 2 shows the arterial and venule vessel next to each other. The regions have increasing diameters for each separate layer. Nitric Oxide (NO) is produced at the luminal and abluminal surfaces of the endothelium10.

The steady-state mass transport equation (cylindrical conditions) can be used to solve for the NO mass transport because the convective transport of NO can be neglected and the NO profiles reach steady state within milliseconds10.

\[
D_j \nabla^2 C_j \pm \sum R_{jj} = 0
\] (2)

In this equation, \(j\) represents the particular model of interest; \(C_j\) is the concentration; \(D_j\) is the diffusivity; and \(R_{jj}\) stands for the production and consumption of the species due to chemical reactions.

### References

Reference 10 (Kavdia M. A computational model for free radicals transport in the microcirculation. Antioxid Redox Signal. 2006;8(7-8):1103-1111.)
**Boundary Conditions**

Specific boundary conditions needed to be set to model this geometry and diffusion rates. At the outer edge of the PT, a zero-flux boundary condition was fixed, and at the interfaces with the endothelium, the release of NO and $O_2^-$ were given by the following equations respectively:\(^\text{10}\)

$$Q_j = D_j \frac{\partial C_{j,\text{en}}}{\partial r} - D_j \frac{\partial C_{j,is}}{\partial r}$$

(3)

$$Q_j = D_j \frac{\partial C_{j,ef}}{\partial r} - D_j \frac{\partial C_{j,\text{en}}}{\partial r}$$

(4)

The NO and $O_2^-$ concentration profiles in the vascular tissue were obtained from these equations.

**Chemical Reactions**

The chemical reactions that are involved in the different layers of the arterial and venule are a mixture of first and second order reactions. Each region is discussed in further detail in Kavdia’s “Venular endothelium-derived NO can affect paired arteriole: a computational model”\(^\text{14}\). The areas that are rich in red blood cells contain high levels of hemoglobin. This hemoglobin reacts at a high rate with the NO in the region, as seen in the following equation where $k_{\text{CR}}$ is the effective NO reaction rate constant.

$$R_{\text{NO,CR}} = K_{\text{CR}} C_{\text{NO}}$$

(5)

In the CF region the chemical reactions are first order reactions because the hematocrit in this region is assumed to be zero\(^\text{10}\).

$$R_{\text{NO,CF}} = k_{\text{CF}} C_{\text{NO}}$$

(6)

In the remaining of the regions E, IS, and NPT the NO reaction is a second order reaction\(^\text{10}\).

$$R_{\text{NO,1}} = k_{\text{2,1}} C_{\text{NO}}^2 C_{\text{O2}}$$

(7)

Vascular smooth muscle sGC consumes the NO for the smooth muscle region (SM)\(^\text{10}\). Therefore the second-order reaction is:

$$R_{\text{NO}} = k_{\text{SM}} C_{\text{2NO}}$$

(8)

For the capillary-perfused PT region, the endothelial cells of the capillaries produce NO. For this region the reaction rate must take into account the amount of nitric oxide that is released by the capillary endothelial cell.

$$R_{\text{NO}} = k_{\text{cap}} C_{\text{NO}} Q_{\text{cap}}$$

(9)

Each region is different, so each individual chemical reaction had to be considered when the model was derived.

**Parameter Values**

All parameters that were defined in the previous equations used in the model can be found in Table 1. The geometry of the arterial and venule vessels has been discussed in previous reports\(^1\). The venule is assumed to be twice the size of the arteriole. The arteriole is 25 $\mu$m and the venule is 50 $\mu$m in diameter. This ratio was assumed due to the reported distance between the arteriole and venule vessels\(^\text{15}\). The diffusivity rates of NO, $O_2^-$, and peroxynitrite were assumed to be constant across the geometry and equal 3.3 x 10\(^{-5}\), 2.8 x 10\(^{-5}\), and 2.6 x 10\(^{-5}\) cm\(^2\)/s respectively according to Table 1. To determine the consumption of NO in CR region, Kavdia and Popel used a hematocrit of 0.45 in the region\(^1\). The reaction rate for the consumption of NO is 1.270 s\(^{-1}\) in the region\(^8\). The NO that is released by the capillary endothelial cell, $k_{\text{cap}}$ was determined using a hematocrit of 0.3 and a capillary volume of 0.0146 cm\(^3\) for the model\(^\text{10}\). The $k_{\text{cap}}$ calculated was 12.4 s\(^{-1}\) for this case\(^\text{16}\).

![Table 1. Model Parameters](http://scholarworks.uark.edu/inquiry/vol9/iss1/1)
Numerical Solutions

Flex PDE 3.0 software was used for modeling the arteriolar and venular endothelial NO, $O_2^-$, and ONOO$^-$ concentrations. Flex PDE 3.0 is computer software that can be used for modeling and solving numerical problems.

Simulations

Tamara M. Paravicini’s article “Increased NADPH-Oxidase Activity and Nox4 Expression during Chronic Hypertension is Associated with Enhanced Cerebral Vasodilatation to NADPH In Vivo” demonstrated experimental trends in the production of superoxide ($O_2^-$) in Wistar-Kyoto rats (WKY) and Spontaneously Hypertensive rats (SHR). This data was assumed for a model of the human arterial/venule microcirculation during normotension and hypertension. For the base case we used the whole tissue superoxide production as fraction $c = 0.2$ of NO production and used 10 $\mu$M of SOD.

The first simulation involved an increase of $O_2^-$ production by 22-fold due to the NADPH stimulation of the arteriolar and venule. For this case the new whole tissue superoxide production fraction of NO production was $c = 4.4$. The SOD remained at 10$\mu$M. The results of both the normotensive and hypertensive were compared for basal and NADPH stimulated vessels.

The second simulation compared the three different basal conditions with a difference in the level of SOD. We used a condition with SOD of 10$\mu$M and a superoxide production fraction $c = 0.2$ and compared it to the normotensive and hypertensive basal condition with an inactivation of SOD. The experiment used $Cu^{2+}$ chelating agent diethyldithiocarbamic acid trihydrate (DETCA) to inactive the $Cu^{2+}/Zn^{2+}$ SOD. We assumed that SOD was reduced and that the level changed from 10$\mu$M to 1$\mu$M. In normotensive basal condition with reduction of SOD the superoxide production fraction remained at $c = 0.2$. For hypertensive basal conditions the superoxide production fraction was $c = 2$ due to a 10-fold increase in superoxide production.

The third simulation compared the NAD(P)H stimulated hypertensive/normotensive vessels with the change in SOD. With SOD of 10 $\mu$M, the normotensive and hypertensive vessels both had a superoxide production fraction of $c = 4.4$, which are 22-fold greater than the basal conditions. With DETCA reduction of SOD to 1 $\mu$M, the normotensive vessels had a superoxide production fraction of $c = 8$, which was a 40-fold increase from the basal conditions. With DETCA reduction of SOD to 1 $\mu$M, the hypertensive vessels had a superoxide production fraction of $c = 18.4$, which was 2.3-fold greater than the normotensive NADPH stimulated conditions.

Results

Normotensive and Hypertensive Cases (Basal and NADPH excited) with SOD of 10$\mu$M

Profiles for the reactive oxidative species were generated according to the concentration values along the horizontal center axis of the geometry, as seen in Figure 2. The first vertical line on the graph at 0.975 cm represented the middle of the arterial vessel. The endothelial cell region was 5 $\mu$m, therefore the second line was hard to distinguish from the first because it was at 0.9775 cm on the graph. The third vertical line represented the smooth muscle cell region in the arterial vessel. The fourth vertical line represented the middle of the venule vessel at 0.105 cm on the graph.

For the base case, we used normal parameters as described in the methods section. The plots of NO, $O_2^-$, and ONOO$^-$ concentrations for the base case are displayed in Figures 3, 4 and 5, respectively. In Figure 3, the NO production of the Basal hypertensive and Basal normotensive with SOD of 10 $\mu$M is the same. This simulation of basal hypertensive and normotensive with SOD of 10 $\mu$M had the highest NO concentration and the lowest $O_2^-$ concentration. The NAD(P)H stimulated hypertensive and normotensive with SOD of 10 $\mu$M were equal to each other. The NO concentration in the lumen of the CR region of both the arterial and venule was zero. As the NO profile approached the EN region, the concentration reached 100 nM. From that region, the NO concentration profile decreased back to 0 nM in the lumen of the venule vessel. The NAD(P)H stimulation during this case only caused a noticeable change in the NO profile in the PT region surrounding the arterial and venule pair.

![Figure 3. NO CONCENTRATION PROFILE FOR CASE 1](image)

![Figure 4. ONOO$^-$ CONCENTRATION PROFILE FOR CASE 1](image)

The ONOO$^-$ concentration profile showed a visible change in Figure 4 throughout the arterial and venule pair due to
the NAD(P)H stimulation. This stimulation caused a 20-fold increase in the ONOO\(^-\) concentration. In the SM region, the ONOO\(^-\) concentration increased from 1 nM to 2.5 nM due to the NAD(P)H stimulation. The peroxynitrite concentration was 0 nM in the lumen of the venule vessel.

The O\(_2^\cdot\) concentration profile is shown in Figure 5. This concentration profile shows that there was an increase in superoxide production in both the arterial and venule EN region. The O\(_2^\cdot\) concentration had a peak at 0.433 nM and 2.29 nM in the arterial EN and venule EN, respectively.

**Figure 5. O\(_2^\cdot\) CONCENTRATION PROFILE FOR CASE 1**

Normotensive and Hypertensive Cases (Basal Only) with SOD being inactivated from 10 \(\mu\)M to 1 \(\mu\)M

We modeled the concentration profile of NO, ONOO\(^-\), and O\(_2^\cdot\) for the basal case only of normotensive and hypertensive cases, but the SOD was reduced from 10 \(\mu\)M to 1 \(\mu\)M with DETCA Cu\(^{2+}\)/Zn\(^{2+}\). The nitric oxide production in the reduced SOD was lower than that of the 10 \(\mu\)M SOD cases. Also the NO level for the hypertensive case with SOD of 1 \(\mu\)M was much lower than the normotensive case with SOD of 1 \(\mu\)M under basal conditions. Figure 6 shows that the highest concentrations of the NO were 102.3 nM, 96.1 nM, and 89.05 nM for Normotensive/Hypertensive basal case with SOD to 10 \(\mu\)M, Normotensive basal case SOD to 1 \(\mu\)M, and Hypertensive basal case with SOD to 1 \(\mu\)M, respectively in the EN arterial regions. The hypertensive basal case with SOD to 1 \(\mu\)M shows a NO concentration that was 20 nM less in the PT surrounding region of the arterial/venule pair compared to normotensive vessels in the same case.

The peroxynitrite concentration profile is shown in Figure 7. The hypertensive vessel with SOD reduction to 1 \(\mu\)M had the highest levels of ONOO\(^-\). This corresponds to the same case having the lowest levels of NO. The peak in the arterial vessel was in the EN and SM region. The concentration levels were 6.56 nM, 4.25 nM, and 0.647 nM in hypertensive basal case with SOD to 1 \(\mu\)M, normotensive basal case with SOD to 1 \(\mu\)M, and normotensive/hypertensive basal case with SOD to 10 \(\mu\)M, respectively. The peaks for the venule EN and SM region were 17.8 nM, 15.6 nM, and 2.36 nM, respectively.

**Figure 7. ONOO\(^-\) CONCENTRATION PROFILE FOR CASE 2**

The superoxide concentration profiles for these specific cases are in Figure 8. This profile demonstrates that SOD reduction caused an increase of superoxide in the vessels by

---

**Figure 6. NO CONCENTRATION PROFILE FOR CASE 2**

**Figure 8. O\(_2^\cdot\) CONCENTRATION PROFILE FOR CASE 2**
10-fold in hypertensive vessels compared to the cases with SOD of 10 μM. In the hypertensive basal case with SOD reduced, the superoxide concentration was 1.81 nM and 13.3 nM in arterial and venule EN region, respectively. In comparison the arterial and venule EN region of normotensive basal with an SOD level of 1 μM had peaks of 1.54 nM and 12.83 nM, respectively.

**Normotensive and Hypertensive Cases (NADPH excited) with SOD being inactivated from 10 μM to 1 μM**

Finally the profile for the NO, ONOO•, and O2•− for normotensive and hypertensive vessels under NAD(P)H excited conditions with SOD reduced from 10 μM to 1μM. The hypertensive NAD(P)H excited vessels with an SOD of 1 μM had the lowest concentration of nitric oxide. These results are shown in Figure 9. The endothelial region of the arterial vessel had the highest concentration of NO during microcirculation. During normotensive/hypertensive vessels with NAD(P)H stimulation and with SOD of 10 μM, the concentration was 99.6 nM. However, when SOD is reduced to 1 μM, the NO concentration in normotensive was 75.6 nM and in hypertensive was 65.2 nM. The lumen of both the arterial/venule pairs was reduced to 0 nM in all three cases.

**Figure 9. NO CONCENTRATION PROFILE FOR CASE 3**

The ONOO• concentration profile for the NAD(P)H cases with SOD reduction are presented in Figure 10. The reduced SOD, hypertensive vessels had the highest ONOO• concentrations. The lowest concentrations of peroxynitrite were in the lumen of the venule. The highest concentrations of the ONOO• were in the EN and SM region of the venule vessel. The concentrations were 29.4 nM, 23.3 nM, and 3.45 nM for the reduced SOD hypertensive, reduced SOD normotensive, and the 10 μM SOD normotensive/hypertensive, respectively.

The superoxide concentration profile for the NAD(P)H cases is shown in Figure 11. This profile shows that, when the SOD was reduced in the hypertensive vessels, the O2•− concentration peaks in the EN and SM of both the arterial and venule vessels. The arterial EN concentration level was lower than that of the venule because in the arterial the superoxide levels dropped dramatically before these regions. In hypertensive vessels with SOD reduction, the arterial EN superoxide concentration was 4.4 nM and in the venule EN superoxide concentration was 17.5 nM. In normotensive vessels with SOD reduction, the arterial EN superoxide concentration was 2.9 nM and in the venule EN superoxide concentration was 14.9 nM.

**Figure 10. ONOO• CONCENTRATION PROFILE FOR CASE 3**

**Figure 11. O2•− CONCENTRATION PROFILE FOR CASE 3**

**Discussion**

**Effects of SOD Inactivation**

SOD inactivation was shown to increase the superoxide production by 10-fold in Basal Hypertensive cases. This resulted in an increase in peroxynitrite but a decrease in the nitric oxide. SOD reduction with NAD(P)H stimulation caused an increase of 2.1-fold and 4.1-fold in superoxide production in normotensive and hypertensive vessels, respectively when compared to vessels with SOD of 10 μM. When SOD
was activated, the superoxide production was the same in hypertensive and normotensive vessels. This shows that SOD is very important in controlling $O_2^-$ production and also in vascular disease. Without SOD, the level of superoxide may increase dramatically, potentially inducing vascular diseases. When SOD is 10 $\mu$M, the vessels have a higher than normal level of $H_2O_2$ due to the dismutase of superoxide. This explains why the levels of $H_2O_2$ in plasma are higher than in normotensive patients. This level of $H_2O_2$, which is a powerful vasodilator, may have important consequences in the vascular system. Current therapies for vascular disease, such as $\beta$-blockers, angiotensin antagonists, and angiotensin-converting enzyme inhibitors, act like antioxidants in some way. New antioxidant therapies have the potential to be discovered to treat hypertension.

Effects of Normotension vs. Hypertension

Spontaneous hypertensive vessels can cause a 10-fold increase from normotensive vessels in basal conditions and a 2.3-fold increase from normotensive vessels in NAD(P)H stimulated vessels. Our data show that superoxide production is increased dramatically in hypertensive vessels. In hypertensive basilar vessels the Nox4 was 4.1-fold higher. This shows that there is a direct correlation between the increase of NAD(P)H oxidase in regions and the increase of production with superoxide and vasodilation. In the area directly before the EN region, the superoxide level decreases dramatically. However, in the endothelial region, there is a large increase in superoxide production. This can be explained because this is where Nox4 is found to be most prevalent. Nox4 mRNA is seen to be 125-fold higher in endothelial cells than in smooth muscle cells (see Table 2).

<table>
<thead>
<tr>
<th>(x10^10 copies 18S)</th>
<th>Nox4</th>
<th>Nox1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monoocyte</td>
<td>Undetectable</td>
<td>0.10</td>
</tr>
<tr>
<td>Endothelial Cells</td>
<td>270.0</td>
<td>0.87</td>
</tr>
<tr>
<td>Smooth Muscle Cells</td>
<td>2.15</td>
<td>0.22</td>
</tr>
<tr>
<td>Fibroblasts</td>
<td>6.25</td>
<td>0.45</td>
</tr>
</tbody>
</table>

Table 2. EXPRESSION OF NOX ISOFORMS

Also Nox4 and Nox1 mRNA levels are 2.5-fold and 10-fold greater in spontaneously hypertensive rats than in age-matched Wistar Kyoto rats, respectively. Published data and the results from this model, showed a direct correlation between superoxide production and Nox expression.

Effects of NAD(P)H stimulation

When NAD(P)H was used to stimulate the vessels, it increased superoxide in the cells because the NAD(P)H reacted with the oxygen in the vessels. During this reaction, superoxide was produced, which significantly increased the levels of superoxide in the vessels. NAD(P)H stimulation increased the superoxide produce 22-fold compared to normal basal conditions when SOD was 10 $\mu$M. During NAD(P)H stimulation when SOD was reduced to 1 $\mu$M, the superoxide production was 40-fold higher and 92-fold higher for normotensive and hypertensive, respectively when compared to normal basal conditions. In human coronary arteries, NAD(P)H stimulation is the major source of superoxide production, and in human coronary artery disease the NAD(P)H activity is significantly increased.

Importance and Conclusion

The purpose of this model was to provide additional information about oxidative stress in hypertension. During hypertension, there was an increase in superoxide production when SOD was decreased. The trends from the NO, ONOO-, and $O_2^-$ concentration profile show us which regions were affected by the disease and the physiological affects on the walls of the arteriole and venule. In addition, the $H_2O_2$ generated during superoxide dismutase may have a significant correlation to vasodilation during hypertension. The relevance of this research is underlined in the major role of oxidative stress in vascular diseases. The future of cardiovascular therapy is based on the balance between NO, ONOO-, and $O_2^-$ within the body.

Future Work

The data from this study could be used to further the understanding of vascular diseases. Nox1 and Nox4 are highest in the endothelial and smooth muscle cell region. The NAD(P)H oxidases are also expressed more in hypertension than during normotension. As previously stated, the expression of Nox1 is 10-fold greater and Nox4 is 2.5-fold greater. In recent years, there have been a number of publications about treating the arterioles and veins with different medications. Very similar tests have been done to measure superoxide production and Nox expression.

In one particular experiment, the vessels were treated with different hypertension medications: 1) high dose candesartan 2) low dose of candesartan 3) a dose of a combined hyralazine and hydrochlorothiazide. All three treatments reduced the expression of Nox1 and Nox4. From the results of this modeling study, we can hypothesize that a reduction of Nox1 and Nox4 will also reduce the superoxide levels produced in the arterial and venule flow. The data show that the strong correlation might suggest a stronger link between NAD(P)H stimulation and the development of vascular diseases. In future research, we plan to model the concentration profiles of medically treated vessels to evaluate the effects of the medicine on the NO, ONOO-, and $O_2^-$ profiles.

References

2. Sorescu D, Weiss D, Lassegue B, Clempus RE, Szocs K, Sorescu GP, Valpau L, Quinn MT, Lambeth JD, Vega...


Mentor Letter

Mahendra Kavdia speaks highly of Aaron Strobel’s academic excellence and of the importance of the research described in this article.

Aaron started his research in my Laboratory in April, 2007. He obtained his undergraduate degree in Biological Engineering with a concentration in Biomedical Engineering/PreMed May, 08. He is one of the best students in our Department. Though the long-term plan of Aaron is to pursue a MD degree, he had an internship offer at St. Jude, San Francisco, CA but decided to work in bioinformatics area at Little Rock, AR to better prepare him for his long-term goals of pursuing MD. Aaron was motivated to perform research. He wrote an honors proposal and got funded in Spring, 08. He participated in several research internships. The proposed research was to investigate the role of NADPH oxidase derived oxidative stress conditions on nitric oxide metabolism. This research is extremely important as oxidative stress is involved in numerous pathophysiological conditions such as cardiovascular diseases, diabetes, and aging. We had developed a computational model to understand the interaction of NADPH oxidase related oxidative stress on nitric oxide levels in vascular tissue. Aaron showed great initiative and also started communication with an experimental researcher in Australia to support his modeling research. His participation in my laboratory was very valuable for his basic knowledge of Biomedical Engineering and hands-on experience. Publication of this article in Inquiry certainly is a worthwhile recognition for Aaron’s work and he deserves this.
GIVE ME THAT OLD TIME RELIGION: NOSTALGIA, MEMORY AND THE RHETORIC OF LOSS IN BEDE'S HISTORIAL ECCLESIASTICA GENTUS ANGLORUM

By John T. R. Terry
History Department

Faculty Mentor: Lynda L. Coon

Abstract

Bede's Ecclesiastical History of the English People was composed by a monk of northeastern England around AD 731. It is a tome of nearly unparalleled prominence to this day in English church history. Many Bedan scholars have hitherto been concerned with common themes in Bede's works: biblical typology and exegesis, influences, sources, politics and even the nature of Bede's own mysterious life. This paper, however, seeks to add a definite human component to Bede and the times in which he lived where most studies have not, simply by using modern studies of nostalgia in a universal sense. Nostalgia is perhaps the deepest of human emotions, employed most often in tumultuous times. When combined with studies of collective or cultural memory, nostalgia in history can be defined as a literary or rhetorical construct of a time which may or may not have existed at all. Nearly all polemic on nostalgia as a modern literature or any form of nostalgia in a universal sense. Nostalgia is perhaps the deepest of human emotions, employed most often in tumultuous times. When combined with studies of collective or cultural memory, nostalgia in history can be defined as a literary or rhetorical construct of a time which may or may not have existed at all. Nearly all polemic on nostalgia as a human condition has focused on the "modern" period. For example, the pioneering models of nostalgia developed by Svetlana Boym have been used primarily to characterize the nineteenth and twentieth centuries, but they seem universal and have yet to be applied to the early medieval period. Virtually no medievalists are concerned with the history of emotions save Barbara Rosenwein. Most medievalists have rejected anthropologic and sociologic studies, while the latter have generally ignored the early medieval period. In his History, Bede is at times susceptible to nostalgia for an earlier idealized age and constructs England's past accordingly. Nostalgia functions most frequently in the History as a rhetoric of loss, a painful rendition of the good and orthodox past for which Bede at a dynamic and often tumultuous time in England pines. Affected by the tumult and uncertainty of his own times, Bede constructs in the History what he imagines to be the ideal past in order to define the present in somewhat imagined terms. This paper examines nostalgia in Bede's History, exploring the possibility that both modern and pre-modern periods had the propensity for the nostalgic enterprise in various media: poetry, literature or any form which is transmitted widely.

I. The Theoretical Background and Foreground for Nostalgia

This is how the present life [vita praesens] of man on earth, King, appears to me in comparison with that time which is unknown [incertum] to us. You are sitting feasting with you ealdormen and thegns in winter time; the fire is burning on the hearth in the middle of the hall and all inside is warm . . . and a sparrow flies swiftly through the hall [domum]. It enters in at one door and quickly flies out through the other. For the few moments it is inside, the storm and wintry tempest cannot touch it, but after the briefest moment of calm [parvissimo spatio serenitatis], it flits from your sight, out of the wintry storm and into it again. So this life of man appears but for a moment; what follows or indeed what went before, we know not at all [quid autem sequatur, quidue praecesserit, prosus ignoramus].

THE VENERABLE BEDE, c. AD 731

Man's life is like a Sparrow, mighty King! That—while at banquet with your Chiefs you sit Housed near a blazing fire—is seen to flit Safe from the wintry tempest. Fluttering, Here did it enter; there, on hasty wing, Flies out, and passes on from cold to cold; But whence it came we know not, nor behold Whither it goes. Even such, that transient Thing, The human Soul; not utterly unknown While in the Body lodged, her warm abode . . .

WILLIAM WORDSWORTH, c. AD 1821

As the tide ebbs and flows [accedente ac recedente], this place [locus] is surrounded twice daily by the waves of the sea like an island and twice, when the shore is left dry, it becomes again attached to the mainland. The king humbly and gladly [humiliter ac libenter] listened to the bishop's admonitions in all matters, diligently seeking to build up and extend [aedificare ac dilatare] the Church of Christ in his kingdom. It was indeed a beautiful sight [pulcherrimo spectaculo] when the bishop was preaching the gospel, to see the king acting as interpreter of the heavenly word [verbi caelestis] for his ealdormen and thegns . . .

THE VENERABLE BEDE, c. AD 731

The Venerable Bede, penning these lines towards the end of his life in 731 at the monastery at Jarrow in northern England, colored his Ecclesiastical History of the English People with many vernacular metaphors like these. Invoking Bede's famous analogy between the life of man and the flight
of a sparrow, Wordsworth was doing more than utilizing a sacred metaphor for his ecclesiastical sonnets. From the eighteenth century onwards, the extent to which nostalgia for Britain before or during the Roman occupation is allowed into English literature and polemic is striking, while modern nostalgia for early medieval England is both less prolific and dramatic. Modern polemic on nostalgia or emotions in general rarely applies to the medieval period and never to the early Middle Ages. If references are made to nostalgia, they serve only as a prop for modern perceptions of emotions in history. Since the definition of some sort of “pre-modern” nostalgia is needed, it might be best to interpret nostalgia in Bede’s Ecclesiastical History as a literary construct.

This third and seemingly disjointed passage is Bede’s description of Lindisfarne, the Holy Island, off the coast of kingdom of Northumbria in northern England. This physical description of a pseudo-island which is both connected and detached from the mainland holds true today, as in the summertime automobiles daily drive back and forth on the causeway during low tide. Bede in describing Lindisfarne immediately employs a beautiful spectacle, a pulcher spectaculum. He has no need to describe an episode with which he was never familiar—Lindisfarne was founded about a hundred years prior to the composition of Bede’s Historia—but he goes out of his way to idealize the Holy Island. Just after this passage, Bede introduces an evangelist’s utopia in which “people flocked together with joy to hear the Word (confluebant ad audiendum Verbum populi gaudentes).”

Here gaudentium, or joy, is just the type of superfluous detail often included by Bede to evoke some sort of emotional response. In his fantasy, lands were given for monasteries, Irish teachers taught the English children and the monks all lived under one rule (regula).

All, in Bede’s eyes, was right with the world. Of course, the necessary historical use of an imperfect past tense (confluebant) couches this ideal picture of Lindisfarne firmly in the past, and one must wonder how the ideal compares with Bede’s contemporary times. Bede’s narrative is equally full of loss and of recovery. The monks at Lindisfarne physically lose the mainland daily, until the tides recede and the so-called Pilgrim’s Causeway reveals itself again (see Figure 1). In this regard Lindisfarne is a perfect monastic site, removed from the world yet not completely, affording a community to experience profoundly everything as a community, including cultural loss, as it oscillates between attachment to and detachment from the world, reality and memory, loss and regain.

But when attributing nostalgia to Bede, the core of the problem here can be found in the arguably redundant phrase: “modern nostalgia.” The adjective is hardly needed at all, since the study of nostalgia is largely a modern phenomenon, presented in forms with which we are familiar. At first glance, the word is purely Greek, stemming from οὐσία, or return home, and αἴγιν, pain, with a suggestion of longing. This root would certainly make sense, considering the Homeric

![Figure 1. The Anglo-Saxon kingdoms c. 700. Lindisfarne is in the extreme northeastern portion of Northumbria, set apart from the mainland in the North Sea. Bede’s monastery at Jarrow is further south along the Northumbrian coast. (anglo-saxons.net)](http://scholarworks.uark.edu/inquiry/vol9/iss1/1)

emphasis placed on οὐσία, especially in The Odyssey. It is no coincidence, perhaps, that, one of the first images of Western literature is that of Odysseus weeping for his home on a beach. But “nostalgia” is in fact a strange construct in itself, a post-classical Latin word of Greek derivation. The term was coined by Swiss doctor Johannes Hofer in 1688 in his Dissertatio medica de Nostalgia, oder Heimweh. Indeed, nostalgia is so desperate a state of mind that Hofer considered it a curable disease, most effectively remedied by an actual return to a fatherland.

Nostalgia remained an epidemic throughout the nineteenth century, permeating the United States during the country’s own civil war in the 1860s. The only country which held fast to this now antiquated view of nostalgia as a curable disease is, almost appropriately, Israel. Interestingly enough, nostalgia might still be characterized as a disease, but the most intense states of nostalgia do not long for an actual home, but a home constructed by imagination and emotive response to contemporary times. Nostalgia is applied readily and thickly onto the modern European psyche and largely manifests itself

---

4 Ibid. 3.3: 220-221.


in a form easily recognized. The operation of nostalgia for Bede is carefully crafted to add to his narrative a rhetoric of loss within the Latin. His past is one constructed through multigenerational memory, one informed by his dealings with Irish monasticism, ascetic traditions and Britain’s pagan past, as he indicates a keen desire to recover that which he deems to have been lost by contemporary Anglia, or that which might never have existed at all. As with modern nostalgia, the past longed for is never actually past, and may be a construction in part or in whole. Bede’s view of the past both blurs it and contrasts it with the present as he painfully returns to a home imagined.

Nostalgia hitherto has been seen as strictly a modern notion, and thus calls for a modern definition. It is therefore important to define nostalgia in modern and pre-modern terms. Nostalgia in the Oxford English Dictionary is defined as

1. Acute longing for familiar surroundings, esp. regarded as a medical condition; homesickness. Also in extended use.

2. a. Sentimental longing for or regretful memory of a period of the past, esp. one in an individual’s own lifetime; (also) sentimental imagining or evocation of a period of the past.

b. Something which causes nostalgia for the past; freq. as a collective term for things which evoke a former (remembered) era.

As a modern model applied to the real world, Svetlana Boym in her pioneering study of modern European nostalgia is careful to define it at the outset as “a longing for a home that no longer exists or has never existed. Nostalgia is a sentiment of loss and displacement, but it is also a romance with one’s own fantasy.” Though the scope of her study rarely extends beyond seventh century Europe, Boym’s models for nostalgia seem universal. Initial studies of nostalgia as a disease positioned it especially among the soldiery. But for Boym, by the twenty-first century, “the passing ailment turned into the incurable modern condition,” an acute yearning for utopia. Nostalgia itself is not always manifested in the linear sense, that is, looking strictly backwards or forwards. For Boym’s purposes, the emotion can direct itself sideways: “The nostalgic feels stifled within the conventional confines of time and space . . . [nostalgia is] an affective yearning for a community with a collective memory, a longing for continuity in a fragmented world. Nostalgia inevitably reappears as a defense mechanism in a time of accelerated rhythms of life and historical upheavals.” Internal unrest as a reaction to external stress—whether a World War or simply a change in times perceived as too swift for comfort—triggers nostalgia. Prior to the twentieth century, with its tumultuous flux of regimes and advent of localized nostalgia, the reconstruction of the past served Romantic literature well.

A consideration of “collective memory” might be a useful model in which to situate the propensity for nostalgia throughout history, as Bede’s monastic existence was by definition collective. Though there is much to be said of medieval memory, the theoretical background for collective memory was most usefully and eloquently put forth by Maurice Halbwachs (d. 1945), French sociologist and philosopher responsible for developing the modern notion of collective memory. For Halbwachs, who often used the Gospels as basis for his hypotheses, collective memory is by nature fragmented, as narratives represent only a portion of what witnesses remembered and transmitted across generations largely by word of mouth.

Collective memory relies heavily on localities, or loci, usually centered on supernatural events. Christian groups therefore tried “to rediscover above all the places where Jesus had been tried, crucified, buried, and resurrected, and where he had appeared to his disciples . . . [I]f a truth is to be settled in the memory of a group it needs to be presented in the concrete form of an event, of a personality, or of a locality.” In an ever-expanding and widespread Christian community, the actual locations of Christ’s birth, transfiguration, crucifixion and the like became abstracted, but were crystallized in the imagination by doctrine and belief. Thereby were consecrated spaces thrust into prominence, as cults of deceased saints sprang up around physical loci, fusing symbolism and belief into “memory.”

Bede recognized his island’s loci and past as a treasure trove of the ideal. A working definition of nostalgia for the eighth century is reasonably a mental and literary construction of a past which may or may not have existed. For Bede nostalgia takes over the narrative in the form of brief anecdotes, linguistic and fictional binaries, ideal descriptions and the occasional emotional outburst. More generally speaking, nostalgia most often manifests itself in solitary or collective states of melancholia, based on an acute sense of personal or cultural loss. The study of nostalgia so far finds its subjects largely the twentieth century. But as a monk accustomed to communal living in times perceived to be tumultuous, the Venerable Bede may well have experienced cultural loss within his community. Nostalgia then is perhaps the deepest of emotions. Though for Boym the famous Astrological Clock in Prague, erected in 1490, represents a constant retrospective reminder of time, the argument can be made that even before modern timepieces, the agrarian or semi-agrarian life of a medieval peasant was a clock in itself, dependant on a series of cycles—sunrises, sunsets, seasons, sowing, reaping, etc. With the arrival of official church holidays, the cycle may even have been enhanced. But the church’s weapon of choice in the war against pagan cyclical

7 Ibid., xiii-xiv.
8 Ibid., xiv.
9 Ibid.
history was a Christian linear history, situating the human experience within creation, present time and judgment. Both modern and pre-modern systems, it should be remarked, have the capacity for nostalgia.

Though much doubt has been cast on the literalness of this particular passage, it is important in constructing Bede’s image of himself and his physical surroundings. His physical and emotional attachment to the monastery is indicated well by the Latin eiusdem above, where eius is the pronoun and dem is an emphatic suffix. Though Bede’s writings lend credence to the view that he was perpetually immersed in eristic studies and had little awareness of the world around him, even in an eighth-century sense, our discussion will rely heavily on the claim that Bede was acutely aware of the political scene of Anglia. The Letter to Egbert, discussed below, indicates his awareness of the spiritual and political workings of monastic communities in and around Northumbria, and in his History Bede spills copious amounts of ink on the geopolitical interactions of kings in a Christian context.

Neither Bede nor most of his subjects in the HE, however, are

25 May 735.12 The Ecclesiastical History is dedicated to King Ceolwulf of Northumbria who, with the exception of a deposition from 731 to 732, reigned from 729-737. Bede clearly respected this “most glorious King” (Gloriosissimus rex Ceolwulfus), and his deposition around the time of writing must have been shocking enough for Bede to be distressed by his own times.13 Bede probably never traveled very widely—a fact which perhaps made him even more susceptible to nostalgia—but the content of the Historia ecclesiastica (hereafter HE) indicates a substantial awareness of the politics of his day.14 Other than the rough dates of his other works, this is virtually all we know about the life of Bede.

Henry Mayr-Harting, reflecting a widely-held view, unreasonably asserts that for Bede, “the real world was to a considerable extent the world of books.”15 This almost romantic picture of Bede as isolated monk surrounded by books is an attractive one, an image bordering on the popular. Indeed, evidence for this perception of Bede comes directly from the monk’s pen: “I have spent all my life in this monastery [tempus vitae in eiusdem monasterii habitazione peragens].”16 But besides this self-constructing passage full of monastic tropes, there is really no solid foundation for Mayr-Harting’s claim.

Bede’s audience of course was clerical and aristocratic, and it is reasonable to think that he was also trying to reach a wider lay audience, given the amount of the HE devoted to vernacular or oral narrative. We have no way of knowing the capacity of Bede’s interaction with the world outside the cloister, though his work remained massively popular on the continent in the eighth and ninth centuries, especially in Carolingian Francia. Bede’s sources were prolific and the models upon which he based his history mammoth. He was probably most attuned to the styles of two major histories, those of Eusebius (d. 339) and Gregory of Tours (d. 594), but for most of his early Christian content Bede relied heavily on Gildas (d. 570) and oral sources. Other authors referenced in the beginning stages of the HE are Pliny the Elder (d. 79) and Orosius (d. 420), a disciple of St. Augustine of Hippo. A more local source, Abbot Albinus of Canterbury provided Bede with both written records and oral tradition regarding the 597 Gregorian mission.

Despite this seemingly mechanistic assemblage of sources, nostalgia seeps through the narrative as nostalgia for Irish monasticism, eastern ascetic traditions and Britain’s pagan past. Neither Bede nor most of his subjects in the HE, however, are
Irish monks, eastern ascetics or admirable pagans. These and more are all constructions for Bede, writing from the periphery in Northumbria on the outside looking in, and, in many ways, at the present looking back. The Venerable Bede was a figure larger than life, with a well-formed political thrust to his Ecclesiastical History of the English People, but he was also simply a monk looking both backwards and forwards, with an aim to situate himself and his “nation” within the linear events of Christian history.

II. Situating the Problem of Emotions in Anglo-Saxon England

Barbara Rosenwein in her fascinating article “Worrying about Emotions in History” sets up the problem in difficult terms: “Emotions have seemed tangential (if not fundamentally opposed) to the historical enterprise.”17 The historiography of emotional history is brief but complex and Rosenwein’s unique article is worth an extended discussion, especially in setting up theoretical terms for tracing Bedan nostalgia. She takes head-on Lucien Febvre’s argument as a theme. Febvre was the founder of the Annales School of historical thought in France, which sought to incorporate social sciences into history. In 1929 he founded the journal Annales d’Histoire Economique et Sociale for which his school is known. The Annales, truly pioneering for its time, fundamentally applied the study of the present to the past. Febvre viewed the progression of emotional history as one in which emotions were eventually tempered through state-building and civilizing. For Rosenwein, Febvre sought to construct a moral history or “public policy masquerading as history.”18

Norbert Elias, the great Jewish-German philologist whose historical view is illustrated well by the title of his 1938 two-volume set, The Civilizing Process, added a Freudian dimension to the discussion, observing that medieval people had not developed a “strict and stable super-ego” under a modern state.19 Rosenwein also takes exception to the views of others such as the modernists Carol and Peter Stearns of the 1980s who called for a formation of “emotinology” as a discipline. The commonality among these theorists is their dismissal of the pre-modern era in their studies. The medieval period is virtually ignored, save when Stearn notes that courteously love plays do not permeate popular society and are therefore useless to emotinology. Regardless, there was virtually nothing viable in pre-modern culture to qualify as emotinology. So far, emotinology, though useful in theoretical terms, can only be extended so far back in time.

These modernist conceptions of the “state” do not fit gracefully within the contours of the medieval state. For Rosenwein, a medievalist for whom the Foucaultian approach does not quite work, the treatment of emotions in history is troubling: “The history of the West is the history of increasing emotional restraint. Greece and Rome may be quickly dismissed: did not Homer sing of the sweet delights of anger? The Middle Ages had the emotional life of a child: unadulterated, violent, public, unashamed. The modern period (variously defined) brought with it self-discipline, control, and suppression.”20 The medieval period, in other words, is a prop for modernist polemic and its grand narrative. As most theory indicates, many historians, especially when dealing with emotions, mark a massive break at the “enlightenment” which situates an impassible disparity between “modern” and “pre-modern” periods. The problem is that medievalists have been all too hesitant to take on newer studies of sociologists, anthropologists and ethnographers, who are in turn seldom interested in the early medieval period. Nostalgia has been treated as a modern phenomenon, but the basic emotion must predate the word’s coining in 1688. This “civilizing process” to which most emotive historians adhere is simply of no use to the early medieval historian while nostalgia or emotions seem to have been of minimal use to most pre-modern scholars. The early Middle Ages have been mistreated by emotive historians by not being treated at all.

Bede’s nostalgia was indeed for an ideal but was founded firmly on intellectual practicality, as he was aware both of his audience and which sorts of carefully filtered nostalgia would be well-received. Bede’s was in a sense an inheritance of loss. Given the tenor of his Letter to Egbert, things were not quite as they seemed in the rose garden of the HE. Bede urges his bishop in light of “the harm that has been done to our people [nosstra gens]” to “strive zealously to recall to the right way of life [rectan vitae] any whom you see acting so wickedly.”21 The loss here is the right way of life and it was certainly felt personally by Bede. But the solution of recovery is an active one. It is the responsibility of Bishop Egbert “to recall [revocare]” the right way of life in Northumbria. As a nostalgic historian, Bede was wont to construct or reconstruct a history incorporating the best bits of Irish monasticism, ascetic models and pre-Christian British tradition. It is to these types of nostalgia we now turn.

III. Nostalgia Goes Native: Paradoxical Nostalgia for the Irish and Britain’s Pagan Past

How frugal and austere [quaetorum parsimoniae, cuius contentiae] he and his predecessors had been, the place [locus] itself over which they ruled bears witness. When they left, there were very few buildings there except for the church, in fact only those without which the life of a community [conversatio civilis] was impossible. They had no money but only cattle; if they received money from the rich they promptly gave it to the poor ... The sole concern of these teachers [tota sollicitudo doctoribus] was to serve God and not the world [Deo serviendi, non saeculo], to satisfy the soul and not the belly.22

THE VENERABLE BEDE, c. AD 731

18 Ibid., 823.
Here Bede breaks his narrative on the correct dating for Easter to praise Bishop Colmán’s predecessors at the monastery of Iona, just off the coast of the Isle of Mull and founded by St. Columba in 563. This passage directly follows the lengthy debate between the Irish and orthodox Christians regarding the correct dating of Easter. The fundamental problem with the Irish for Bede was that their “imperfect rules [minus perfecta institutione]” were underscored by their incorrect dating of Easter. Bede was actually instrumental in aligning the conflicting dating methods of Easter in Britain and the continental Carolingian world. Bede’s fondness for the Irish actually goes much deeper than this, since their “sole concern” was to serve God. References in the HE to the Irish are abundant and repeatedly betray a particular affinity for them. The Irish were the first real evangelists in the north of England, and this was certainly important to Bede.

But for Bede, it was a sense of time and place which he seems to have admired most in hagiography, a clear instance of Boym’s nostalgia as an “archeology of memory and place.” To frame this notion, in opening his chapter on the death of King Oswald, Bede, as elsewhere, is careful to include the time, place and circumstances of the terrible event. For Halbwachs, the location of an event triggered community memory, since “if a truth is to be settled in the memory of a group it needs to be presented in the concrete form of an event, of a personality, or of a locality.” At the foundation of Lindisfarne around 635, Bede embellishes that it was “indeed a beautiful sight [pulcherrimo spectaculo] when the bishop was preaching the gospel.” Bede obviously never witnessed the foundation of Lindisfarne, as it predated the composition of the HE by about a century. He is under no obligation to describe the event as a pulcher spectaculum, but since he does, he betrays deep sentiment for those early days of Lindisfarne monastery. This particular passage involves all three of Halbwachs’ communal memory model: event, personality and locality.

For Irish nostalgia, perhaps one of the most important passages in the HE—or in all Bedan literature—is found in 3.26, the opening passage of the chapter. Directly following a chapter in which Bede, indulging his orthodox sensibilities, could have engaged in a healthy diatribe against the Irish for not accepting the correct date of Easter, we find a poetic snapshot of Lindisfarne under the Irish: “The sole concern of these teachers was to serve God and not the world, to satisfy the soul and not the belly.” Not only this, but Bede is writing about Colmán’s Irish whom he accuses of being unorthodox. It is without doubt that Bede admires the Irish deeply, as he would never in reference to orthodoxy praise them with phrases like “Irish method” (mos Scottorum) in a favorable context and “beautiful sight [pulcher spectaculum].” Does Bede then value stabilitas or simplicity more? Organization or austerity? The Irish monks are no paradigm for the former but the mos Scottorum is transmitted in sentimental fashion in the HE.

At Aidan’s death, Bede remembers the Irishman with distinctive personal attachment. Though Aidan’s view of the Easter date was imperfect at best, Bede acknowledged warmly that he “had no other thought in his heart, reverenced and preached no other doctrine than we do, namely the redemption of the human race [redemptionem generis humani].” Aidan’s memory was soon honored by the erection of a church which housed the remains of “so great a bishop [santo pontifice dignam].” On top of this, posthumous miracles came to be associated with a primitive wooden buttress against which a dying Aidan leaned. To that primitive wooden buttress is moored a collective memory and an attempt at retrieval.

Bede certainly picked and chose the more favorable bits of Irish culture and tailored his language only as far as he needed to in order to negotiate the hills and troughs of Irish monasticism. These times were ones with which Bede was not readily familiar, and only could have idealized them for noster historia. Our history: this collective invocation was based on something Walter Goffart engages and from which he quickly shies away, the notion of “artificial perfection” in Bede’s work. This is precisely that which we have seen so far from Bede. The HE is by definition a construction, and artifice of mnemonic retrieval. The loss for Bede is Irish austerity while it is synthetically reconfigured in the HE by nostalgia.

IV. The Primitiva ecclesiae: Nostalgia for Ascetic Tradition

As soon as they had entered the dwelling-place [mansionem] allotted to them [in Canterbury], they began to imitate the way of life of the apostles and of the primitive church [coeperunt apostolicam primitiae ecclesiae vitam imitari]. They were constantly engaged in prayers, in vigils and fasts; they preached the words of life [verbum vitae] to as many as they could; they despised all worldly things [cuncta huius mundi] as foreign to them; they accepted only the necessaries of life from those whom they taught; in all things they practised what they preached and kept themselves prepared to endure adversities, even

---

23 Boym, The Future of Nostalgia, xviii.
24 HE 3.14: 254-257: “In his second year, that is in the year of our Lord 644, the most reverend father Paulinus [reverentissimus paer Paulinus] . . . departed [transiuit] to be with the Lord on 10 October having held the office of bishop for nineteen years, two months, and twenty-one days. He was buried in the sanctuary [secretario] of the church of the blessed apostle Andrew . . . .”
26 HE 3.3: 220-221. Colgrave’s translation of spectaculum as “sight” is a bit bland here. A “spectacle,” rather, implies more than simply the sensory and mundane “sight,” as it implies a more intentionally contrived and grandiose display.
27 HE 3.26: 311.
to the point of dying for the truths they proclaimed. To put it briefly, some, marveling at their simple and innocent way of life and the sweetness of their heavenly doctrine, believed and were baptized [Crediderunt nonnulli et baptizabantur, mirantes simplicitatem innocentis vitae ac dulcedimon dontrinae eorum caelestis].

This passage opens up the twenty-sixth chapter of Book I of the HE and introduces the Augustinian mission, commissioned by Pope Gregory the Great in 597 to convert the people of Kent in southern England. Here Bede straightforwardly evokes the dwelling places of heaven with mansio and couples it with this fascinating phrase, the primitiva ecclesia. This latter phrase is not at all unlike an invocation of the “ancient times” or antiquus (employed to set up the ideal), thus situating the primitive church comfortably in the past with overtones of originality (prima). Bede’s nostalgia for the simplicity of the Irish is readily apparent, and much of it manifests itself in praise for their asceticism.

Much of this chapter shall hinge on St. Cuthbert of Lindisfarne (d. 687) as an ascetic standard for Bede, presented in the Ecclesiastical History itself. Cuthbert lived from 634 to 687 and saw the extremes of Christian existence during his lifetime, as he was both an ascetic and reluctant bishop of present-day County Northumberland. An anonymous Life of Cuthbert was written in Lindisfarne around between 699 and 721, respectively. Many among his ecclesiastical audience would at least be attracted to bits of oral tradition which Bede was certainly familiar. St. Antony in Athanasius’ Life of Antony was the Athanasian superman or, in simple terms, a divine orthodox champion contrived in Athanasian literature through and through. Bede was acutely aware of this and seems to have established Cuthbert as a veritable Bedan standard for Bede, presented with older ascetic and hagiographic traditions, and his lay audience would at least be attracted to bits of oral tradition which made their way into the literature.

In the HE, Bede mentions four individuals whose bodies remained incorruptible upon their death: Ethilburg, Fursey, Ethelthryth and Cuthbert. This feature of ascetic devotion is almost peculiar to early ascetic tradition, namely that philosophy largely constructed by St. Athanasius (d. 373) with which Bede was certainly familiar. St. Antony in Athanasius’ Life of Antony was the Athanasian superman or, in simple terms, a divine orthodox champion contrived in Athanasian literature through and through. Bede was acutely aware of this and seems to have established Cuthbert as a veritable Bedan man in a loose Athanasian context. It is important, however, to note that we are not concerned so much by the influence of eastern asceticism on Bede, as that much is readily apparent. It is rather more useful with respect to Bede’s emotional fixation on Cuthbert to see how he fills the present void with the lost past.

As far as Lindisfarne was from the epicenter of Christendom, Bede was well attuned to Roman Christianity. In order to evoke the primitiva ecclesia, Bede constructed the ideal past in a literary nostalgia. In so doing, Bede positioned the primitiva ecclesia invariably in the past, a lost form of ecclesiastical utopia in which peasants flocked together to hear the word of God. With these models in mind, we now turn to St. Cuthbert, “Irish in training but Roman in his respect for the Rule of Benedict, and the reckoning of the date of Easter.”

Cuthbert for Bede is presented as a man of the people, though polarizing at times. As a near-contemporary comparison, Bede in the HE actually drops a subtle hint that even in the early days of Christianity in the South, British bishops and peasants were wont to consult their own hermits. Upon realizing that Augustine and his missionaries spoke the truth of the correct dating of Easter, the people of Kent “could not disown their former customs without the consent and approval of their own people.” This incident not only underscores the importance of local custom of the Britons, but indicates Bede’s recognition of it.

But before anything is settled, “they went first to a certain holy and prudent man [virum sanctum ac prudentem] who lived as a hermit [anachoreticam] among them to consult him as to whether they ought to forsake their own traditions at the bidding of Augustine. He answered, ‘If he is a man of God, follow him.’” Curiously enough, this advice given by our unknown holy man in Kent echoes almost exactly advice given in reference to St. Antony: “Whatever he says to you, go by his decision, for God speaks to you by him.” Bede clearly wants to illustrate here the importance of the holy middle man who, as in the East, filled an important void in local life, especially where law courts or ecclesiastical systems were remedial or nonexistent. Though about 130 years removed from the conversion of Kent, Bede felt a special connection with the kingdom’s spiritual development. In St. Augustine’s miracle of healing a blind man near the beginning of the mission, Bede goes out of his way to note that Augustine was emotionally “compelled by genuine necessity [iusta necessitate compulsus]” to pray for the man. The “primitive church” was endowed with such noble compulsions.

In many passages like that of HE 2.2, Bede betrays a peculiar affinity for pastoral country scenes of conversion, preaching and public enthusiasm for Christianity. His “good old days” were certainly introduced by St. Cuthbert at times. But Cuthbert was much more than a traditional holy man to Bede. He only spent parts of his life in solitude but often “went forth from the monastery to correct the errors of those who sinned in both these ways, sometimes on horseback but more often on foot.” The Bedan ideal for the holy man is made manifest in Cuthbert:

---

20 HE 1:26; 76-77.
21 HE 3:8; 3:19; 4:19; 4:30.
23 HE 2:2; 136-139. See below, chapter 6. “Conference” is used three times in this particular passage: colloquium (134); synodus (136); concilium (138).
Now it was the custom [moris] amongst the English people at that time, when a clerk or a priest came to a village, for all to gather [confluent] at his command to hear the Word, gladly [libenter] listening to what was said and still more gladly [libentius] carrying out in their lives whatever they heard and could understand. So great was Cuthbert’s eloquence [peritia]... that none of those present would presume to hide from him the secrets of their hearts...

Can we really believe this, that people were so enthusiastic and malleable upon hearing the Gospel from Cuthbert? In a sense, this question is irrelevant since Bede’s was not always a history concerned with what actually happened.

Regardless, what we see here in 4.27 is a glimpse at the characterization of the Bedan ideal. As Antony was for Athanasius, Cuthbert is Bede’s champion, endowed with peritia which was neither witnessed nor experienced by Bede. With respect to orthodoxy, Cuthbert is the superstar where Antony was for Athanasius. At Lindisfarne he upheld the Rule with dignity, as he “trained many in life under a Rule [regularem].”... but he also sought to convert the neighbouring people far and wide from a life of foolish customs [stultae consuetudinis] to a love of heavenly joys [ad caelestium gaudiorum convertere curabat amorem].” In this passage and others, the presence of the word regula almost acts as a signpost for Bedan fantasy, since it cannot be said with certainty whether or not any large monastic groups lived under one “Rule.” The contrast here of “foolish customs” and “heavenly joys” can also be seen as a binary of all times. Bede in a sense is projecting his contemporary time backwards, assuming the binary of good and evil which Cuthbert mitigated and where the leaders of Bede’s time failed. The holy duality of Cuthbert’s life—part evangelist and miracle worker and part Farnie Island anchorite—is paramount for Bede, as in Cuthbert Bede establishes a Northumbrian champion. Like the passage on Lindisfarne, the binary of present and past is one of loss and gain—“foolish customs” and “heavenly joys”—with Bede struggling to reconcile the two in an historical context.

With words attributed to Cuthbert, the modern reader is naturally inclined to be dubious, but this reaction is not so certain for Bede’s contemporary audience, whether lay or clerical. Upon leaving Lindisfarne to become a hermit, Cuthbert parts with this declaration: “If by divine grace it is permitted to me to be able to support myself in this place [in loco] by the work of my hands, I will gladly stay there; but if matters turn out otherwise, I intend, God willing, to return [revertar] to you forthwith.” Not only does Bede indicate his own attachment to his Lindisfarne in this chapter, but this passage also illustrates Cuthbert’s own attachment to the monastery. But when Cuthbert does leave for his eremitic locus, he fashioned a cell with no roof so that he could see nothing but heaven. This echoes the architectural design of the round church at the place of Christ’s ascension, described later by Bede and situated between descriptions of various holy loci in the life of Christ himself.

Such attachment to place is perhaps the ultimate nostalgia, as Bede moors the beliefs of the monastic group in the place of Lindisfarne. Cuthbert’s attachment to his monastic home is underscored further by Bede in the following chapter in which a monk named Herbert engages in an emotionally-charged farewell with Cuthbert, imploring him not to leave. Bede here certainly is willing to exploit human bonds in “spiritual friendship [spiritalis amicitiae].” Even though Cuthbert during his life separated himself from his beloved Lindisfarne, he later returned by unanimous vote to be bishop of the see and, though reluctant, was eventually buried there, in Bede’s eyes, at home in his patria.

Cuthbert’s cult was perpetuated long after his death, as Bede’s nostalgia for Cuthbert is channeled through his cult. By focusing on his incorruptible body and place of burial, Bede supports and perpetuates the cult, fostering belief through attachment to place. Furthermore, by adding a component of realism of humanity to Cuthbert’s life story (of which Bede had no first-hand knowledge), he disconnects the life of Northumbria’s most venerable saint with the present through drastc contrasts of the past ideal and present spiritual malaise, only mitigated in literature through nostalgia. This is Bede’s champion, an individual who combated demons and communicated with animals.

Bede’s nostalgia for asceticism speaks through his narrative especially when he deals with Cuthbert. At the time of Bede’s writing, cults of saints in England were certainly widespread. The cults depended heavily upon popular support and good advertising, and miracles performed by saints while alive or dead were less likely to be recorded if they were not believed. Regardless of whether Bede expected his audience actually to believe that seals ministered to Cuthbert following a conversation with the divine in the North Sea as reported in his Life of Cuthbert, he certainly expected the elements of extreme askesis and divine communication to be appreciated. These inventive and invented stories belonged to him and to the Northumbrians and would not be considered falsehood. The tales presented, as claimed, a highly admirable life and cemented in Northumbrian memory the spiritually attractive qualities of a man of the past, a man of his times, a man of all times.

Bede’s transmission of miracle stories in the HE and in his lives of Cuthbert underscores a wider problem he sees

---

35 HE 4.27; 432-433.
36 Ibid.
37 Ibid. 4.28; 434-435.
38 Ibid., 436-439.
39 Ibid., 5.16-5.17; 508-513.
40 Ibid., 4.29; 440-441. See above.
with contemporary society. Bede is reluctant to include any outstanding stories from the period roughly c. 700-730 (around the time of writing), perhaps indicating that Bede marked the cessation of miracles with a decline of morality in Northumbria’s recent times. Cuthbert, who died some 45 years prior to the composition of the HE, had already gained legendary status in Northumbria and his extreme holiness brought on miracles, even after his death in 687. Holiness had been lost but kept on life support through Cuthbert’s cult and Bede’s literature, and, though Cuthbert in his exceptionality exercised dominion over nature by calming storms and communicating with animals, creation by the time of Bede’s writing has been lost. Invoking Cuthbert as an outstanding exemplum in a long, universal history of asceticism, Bede seeks to reconstruct the past, to regain that which is lost, to resuscitate it through nostalgia.

VII. The Politics of Nostalgia and Mnemonic Mechanics

Great is the power of memory, extraordinarily great, my God, a sanctuary vast and infinite. Who has penetrated its depths?42

SAINT AUGUSTINE, AD 397

Susan Stewart in her pioneering 1984 study notes that “nostalgia is the repetition that mourns the inauthenticity of all repetitions and denies the repetition’s capacity to define identity.”43 Svetlana Boym fleshes out these ideas with a brilliant synthesis of modern nostalgia in her study meant “to unearth the fragments of nostalgia one needs a dual archeology as it is extremely useful in positioning Bede as a troubled traditionalist or reformist at the end of his life The Letter, composed in 734 to the bishop of York and the last of Bede’s writings extant before his death in the following year, reflects most accurately the monk’s cynicism of the state of affairs in England. Virtually nothing in the letter is in exuberant praise of monks or laymen, as the purpose of the letter is to admonish and, for our purposes, contrast with the past and identity Bede has constructed in his Historia. While Bede’s Letter engages with contemporary issues directly, the HE illuminates them through the illustrious past. Monastic abuses in Bede’s eyes intensified from the beginning of the eighth century until Egbert’s rise to power in York in 731, which Bede may have hoped would spur on reform in the Northumbrian Church. This in part explains why the HE is so lacking in detail and especially miracle stories from the period c. 700-730 and establishes the background against which Bede articulated his concerns to Egbert in 734.

Bede in his old age finds it “shocking to say [turpe est dicere] how many places that go by the name of monasteries have been taken under the control of men who have no knowledge of true monastic life.” In an extensive reform, these monasteries “that only in the most foolish way [stultissimo] deserve the name of monastery . . . [should be] turned by the authority of a council from luxury to chastity, from vanity to verity, from indulgence of the stomach and the gullet to continence and heartfelt piety [de luxuria ad castitatem, de vanitate ad veritatem, de intemperantia ventris et gulae ad continentiam et pietaetem cordis].”45 It is tempting to pass this off simply as biblical typology, but Bede might well be up to more.

In his Letter Bede has shed his use of good-evil binaries until this last passage in which they abound. By far the most marked contrast between the tenor of the Letter and the HE is that Bede’s indictments of the past in the latter, though fewer, are generally put forth retrospectively. The Letter bemoans the present with urgency, while the HE simply looks backward to redeem and reclaim the present. Gregory’s 597 mission, though ushering in the exaggerated “golden age” fantasy of English monasticism, had yet to come to a head on the island, a fact which frustrated Bede to no end in his Letter. For Bede, the contemporary Church had failed the English people where the Church of the past had not. Bede’s Letter indicates much deeper anger than had been conveyed in the HE, as the audience has now drastically changed and Bede sees more fit to use words like stultissimus (perhaps “most foolish” is too bland a translation). It might also, however, illustrate that England by 734 for Bede has worked backwards through this sentence, lurching gracelessly from good to evil, from “heartfelt piety” of, say, Cuthbert or the laymen eager for the Word near seventh-century Lindisfarne, to the “indulgence of the stomach” of contemporary “monks.” Bede’s disenchanted view of Northumbria and England is that of a playground for the corrupt in contrast with his HE stories of the grand Anglo-Saxon conversion and its holy men. In the Letter Bede’s acute sense of loss is confirmed and in England, the blind now lead the blind.

All of this does well to add a considerable human dimension to the Venerable Bede. It might be reasonable to see Bede’s nostalgia as part of his case for legitimacy, especially considering that his History is arguably that of the Northumbrians in an English context. If Bede is a nostalgic for the right times and places, he might have been able to speak to his audience on a more personal, familiar level. It is not as if collective memory and nostalgia here are one in the same, but they might be closely tied. Collective memory is the mechanism by which nostalgia and the rhetoric of


44 Boym, Svetlana. The Future of Nostalgia. xviii.

45 Bede, McClure, and Collins, The Ecclesiastical History, 350. See also Plummer, Venerabilis Baedae, 414.
loss attach themselves to places, memories, oral traditions and abstractions. The rhetoric of loss has a special place in collective memory and best represents its emotional, human side.

Bede attributes many of these features of loci in describing places in the HE. In the Anonymous Life of St. Cuthbert, the Lindisfarne author is obsessive when associating events with places. Two seemingly random events in the life of Cuthbert for this author suddenly make sense juxtaposed simply because they took place near Carlisle. As has been discussed, topography and geography are apparently paramount for medieval memory and also indicates collective memory, especially in monastic settings. Curiously enough, though drawing heavily from it, Bede departs in his Prose Life of Cuthbert from the Anonymous Life in that he omits many of the specific local details of the latter. Bede’s omission of geographical detail perhaps situates St. Cuthbert closer to Lindisfarne where the anonymous author was fixated on Melrose and surrounding areas.

As far as reaction for loss and a longing for past time, Bede’s nostalgia is roughly parallel to that which historians have put forth for modern history. To quote Svetlana Boym once again, nostalgia is necessarily “a longing for a home that no longer exists or has never existed. Nostalgia is a sentiment of loss and displacement, but it is also a romance with one’s own fantasy.” The question of why Bede does not fit into this model is a simple one, both to ask and to answer. He does, as a matter of fact. Though an exaggerated and thoroughly fabricated notion for Bede, the “golden age” is invariably put firmly in the past, and Bede employs collective memory and rhetoric in order to retrieve it. It is then difficult to distinguish Bede’s beliefs from his politics of legitimization, but when it comes to collective memory, the distinction seems not to matter as much. Nostalgia is situated in models of collective memory, i.e. emphases on places to which abstract beliefs are moored. To Maurice Halbwachs’ work I might add this literary model of nostalgia, that collective memory almost invariably introduces rhetoric of loss and literary effort to recover. Most historians, perhaps unintentionally, adhere to post-“enlightenment” models no longer tenable, given the progress made in the history of emotions. The literature of Bede itself disproves the grand narrative of history. The human propensity for nostalgia, especially for historians both past and present, mitigates the rough patches between compartmentalized historical “periods” perpetuated by “enlightenment” theory.

Theoretically then, Bede fits well into a model for nostalgia usually reserved for modern history. He understood that as a proponent of vera lex historiae he both enjoyed the freedom to write down in a constructive sense rather than a truthful one, and the responsibility to position his see at Lindisfarne in the wider geopolitical scene. It is a fact of the human condition that would-be reformers or social theorists who feel alienated and oppressed by an ever-changing world almost invariably look backward for salvation of their times. In his perception, Bede’s was an age of loss, but loss and recovery coalesced in the HE by constructed models of the past applied directly to the present. To this end Bede’s image of Lindisfarne proves ideal, as the path to the monastery is lost and recovered daily at the mercy of the tides of time.

Mentor Comments

Mentor Lynda Coon continues a theme seen in other mentor comments when she notes that John Terry’s work goes beyond what might typically be expected from an undergraduate student. She paints a picture of a bright future for this young scholar.

John Terry earned a summa cum laude for his thesis on historical and religious nostalgia in the writings of the Anglo-Saxon monk and biblical exegete, the Venerable Bede (d. 735). Terry came to the complex study of Bede uniquely prepared. A double major in Classical Studies and History (with a specialization in late ancient and medieval history), Terry was able to tackle Bede in the original Latin. Moreover, Terry used his Sturgis Fellowship to fund a year at Cambridge University, where he sat examinations for the History of Roman Britain, the Classical Body, Death in the Middle Ages, and the Transformation of the Roman World. Terry brought all of his superb training to bear on the study of Anglo-Saxon monasticism—its intellectual and ascetic cultures.

In the thesis, Terry broke new scholarly ground, and he did so at a level beyond that which one might expect from an undergraduate, even an Honors student. Terry grappled with the hagiographical traditions surrounding the life and cult of Cuthbert of Lindisfarne, a site he spent some time visiting while living in the United Kingdom. Terry also took on Bede’s famous Ecclesiastic History and he did so in the service of theorizing how the exegetical imagination of ascetic writers and practitioners played a role in the production of a sacred landscape and how that landscape worked as a tool of conversion. His angle was singular—Terry made the case that early medieval writers, like the Venerable Bede, approached ascetic landscapes and even the bodies of the ascetics themselves with a certain degree of nostalgia. In arguing for an early medieval category of nostalgia, Terry countered current scholarly views that nostalgia as an emotion is an invention of the Enlightenment. I found Terry’s corrective to the Enlightenment model of nostalgia to be very persuasive. Indeed, this Fulbright College Honors student covered some of the ground that the top scholars in the field of early medieval history are just now beginning to research. No one to date has applied the kinds of theories of collective memory and landscape as mnemonic tool (Halbwachs, Nora) to early medieval texts that Terry

46 Colgrave, Two Lives of Saint Cuthbert, 116-117
applied to the study of Bede. In all, the theoretical scope of Terry’s project—combined with his rigorous training—marked him off as exceptional among our very best Honors students in the Humanities.

Mr. Terry’s overall goal is to pursue a career in academia as an early medievalist. As he told me one day, there ‘simply isn’t any other option for him’. Happily, both Cambridge University and the University of Virginia offered him positions in their graduate programs (the M.Phil. at Cambridge and the Ph.D. at Virginia). After a long struggle making the decision where next to cloister himself in a major research library, Mr. Terry settled upon the University of Virginia, which has an excellent program in early medieval history, art history, and late Latin studies. Terry intends to spend more time at Cambridge University, working with top scholars there and continuing his passion for libraries (Cambridge houses an impressive collection of Anglo-Saxon manuscripts). Mr. Terry will be missed, both in the halls of Old Main and in his current cloister, the Special Collections division of Mullins Library.
MARKET CORRELATION: EFFECT OF HISTORICAL EVENTS ON THE WORLD’S LARGEST FINANCIAL CENTERS FROM 1983-2003

By Thomas Vo
Department of Economics

Faculty Mentor: Robert B. Stapp
Department of Economics

Abstract

This paper studies the effect of various historical events on the market correlation among the three largest financial centers of the world: New York City, London, and Tokyo for the time period 1983-2003. The analysis focuses on those correlations associated with four historical events: the U.S. stock market crash of 1987, the London IRA Bombing of 1990, the Asian Currency Crisis of 1997 (particularly the day the Thai Baht fell), and the September 11th, 2001 terrorist attacks of 2001. The purpose of this study was to provide additional information that will provide the investment community with insights about maintaining market stability during periods of economic crises. With this information, investors may be able to avoid large losses and hedge their systemic risk of global events. Market data associated with each historical event were analyzed using correlational and statistical procedures. Results suggest that, during times of economic crises, the S&P 500 and the FTSE 100 tend to be more correlated with each other than with the Nikkei 225. At the end of the tested time period, the September 11th terrorist attack shows a similar correlation among the markets although the degree of correlation is slightly less. However, the fact that their correlations are very similar shows that there have not been significant changes in market correlation during economic crises over time.

Overview

The Dow Jones Industrial Average was down 251 points. That was the outcome on October 24, 2007 after Merrill Lynch announced third quarter earnings for 2007 with $8 billion in write-downs associated with the subprime market. In addition, Citigroup had $6.5 billion in write-offs. With days like this, it is clear that the subprime market in the United States is currently problematic. With this part of the U.S. economy in flux, the U.S. stock market has continued its downhill slide after closing at an all-time high of 14,000 on July 19th of 2007.

While the stock market serves as a leading indicator of economic health, it is important to determine how all the markets around the world are responding to this crisis. Current economic conditions in the United States, especially with the continuous slump of the housing market and the ongoing subprime market crisis, are leading to wealth deterioration, since the primary source of individual wealth is housing. But how is it affecting the international markets, especially since there has been an increase in globalization over the past decade (Reyes, 2007)? This increased globalization makes it particularly pertinent to examine the market correlations among the largest financial centers in the world: the markets of New York, London, and Tokyo. In establishing how these markets correlate with each other, it is possible to see which market was less volatile during times of economic crises, allowing investors to be better informed about investing options.

Prior research suggests that the financial markets of developed nations are highly correlated (Sarkar, 2002). However, the financial markets are not always in perfect synchronization at every moment in time. This raises the following question: are markets highly connected during times of economic crises? Economies of every country are forced to react to historical events that occur both domestically and abroad, but they are not forced to react in the same manner. In this study, the market correlations from 1983-2003 among the financial centers were determined at the time of several historical events, specifically the U.S. stock market crash of 1987, the London IRA Bombing of 1990, the Asian Currency Crisis of 1997 (particularly the decline of the Thai Baht), and the U.S. September 11 terrorist attack of 2001. The year 1983 is the benchmark year because it is the year when the United States stock market was entering a bull market and beginning a period of growth (Madrick, 2002). This is used as the foundation point in time since the great stock market crash did not occur until 1987. Following 1983, the historical events of interest are separated by a few years to control as much as possible for contagion effect of the crises. Contagion is simply a cross-market linkage of significant increases after some economic event in another market (Forbes, 2002).

Using these data, the study reported here was designed to determine which market is the most stable during times of economic crises and whether there is a relationship between the performance of markets at these times. Stock indices from each of three financial centers were analyzed, specifically, the S&P 500 (New York City), the FTSE 100 (London), and the Nikkei 225 (Tokyo). Using the market closes for each of the indices, the percentage change in value was calculated and used to determine how the markets relate among themselves. In order to understand the rationale for raising these research questions and for exploring market response to contagious events, it is first important to explore further the concept of market correlation and of contagious events.

Correlation simply characterizes the nature and degree of relationship between two variables. Market correlation studies the relationship between two different stock markets or indices. During times of downside moves, correlations among U.S.
stocks and the aggregate U.S. market are much greater, thus making the impact of negative historical events more powerful (Ang, 2002). Markets typically close in the same direction most of the time, whether up or down (Buttonwood, 2007). Most fund managers divide their portfolios into different regions and countries to allow international diversification (Kearney, 2004). However, investors need to understand the importance of market correlation because, in order to be well diversified, the securities need to have little or negative correlation. Further, there is evidence that suggests that stock correlations rise during negative returns, compounding the notion that diversification strategies do worse during bear markets (Kearney, 2004). However, previous research has not focused on bear markets or bull markets; instead, it has focused on shorter periods of historical events and their immediate impact on market correlation among the financial centers of New York City, London, and Tokyo.

When any type of historical event occurs, it is rare that the incident and its consequences remain isolated in a domestic region. Instead, there is research that suggests that contagion effects are phenomena that affect international markets. During times of an economic crisis, the stock markets tend to have increased volatility, placing more importance on a diversified portfolio. The contagion effect directly impacts one of the events in this study. After the Asian Currency Crisis, Henk Jager and others explored whether or not there was a contagion effect on the surrounding countries (Jager, 2006). They studied contagion using regression analysis, as opposed to the standard correlation coefficient analysis (Jager, 2006). In their study, contagion was actually shown in Thailand to spread to both Indonesia and Malaysia. The contagion effect is an important phenomenon that may affect some of the major historical events that have occurred since 1983.

In theory, knowledge about which markets are stable during times of economic crisis should allow investors to hedge their risks during times of uncertainty and potentially increase the diversification of their portfolios. Investors may find it useful to look at correlations among the markets and to invest in negatively correlated markets for greater diversification. This greater diversification can increase the benefit to any portfolio and, combined with low volatility, could potentially increase a portfolio’s total return.

Hypothesis Development

This study was designed to test the theory of market correlation that compares the effect of historical events on the world’s three largest financial centers controlling for the time period of each respective event. The independent variables are defined as each of the historical events measured in this study and the control variables are the time frame measured for each of the historical events.

Based on prior research, a set of hypotheses was developed for this study. This set of hypotheses describes market correlations and is designed to determine whether or not the correlations among the markets have changed over time. The primary research hypotheses are:

- **H₀**: There is no relationship among market correlations of the S&P 500, FTSE 100, and Nikkei 225 over the tested time periods for the historical events, or r = 0
- **H₁**: There is a relationship among market correlations of the S&P 500, FTSE 100, and Nikkei 225 over the tested time periods for the historical events, or r ≠ 0

If any of the relationships among market correlations was significant, a subset of hypotheses was created to determine:

- **H₂**: The correlations among the S&P 500, FTSE 100, and Nikkei 225 have, in general, increased over time.
- **H₃**: The correlations among the S&P 500, FTSE 100, and Nikkei 225 have, in general, decreased over time.
- **H₄**: The correlations among the S&P 500, FTSE 100, and Nikkei have not changed over time.

This subset of hypotheses was investigated by examining trends in the correlation data.

Methodology

From 1983-2003, a variety of historical events have occurred that have had a profound impact on the world’s economies. While it was difficult to select specific crises, events were chosen based upon economic impact and time frame. The market analysis begins in 1983, a time the U.S. market was beginning its period of growth (Madrick, 2002). After a few years of normal growth, the first economic crisis occurred in the form of the stock market crash of 1987. This event had a significant impact on the U.S. market, so future historical events test each respective market’s resolve. The data were collected from Finance.Yahoo.com, a reliable website for daily market news. Because of the historical indices, Yahoo! Finance had daily market close data for the S&P 500, the FTSE 100, and the Nikkei 225. This source of data was extremely important because daily market closes were required to determine the correlation among the financial centers during the times of economic crises. The time frame for data collection was the month before the event until three months after the event to provide a sufficient number of observations.

Once the data were collected, the three financial centers’ market closes were analyzed. Data for New York City, London, and Tokyo were measured both on a daily and a weekly basis. Daily data were computed using the beginning date as the benchmark, set at 100, with the rest of the data expressed as a percentage of the benchmark. Graph 1 displays market data for the 1987 U.S. Stock Market Crash after adjustment to set start dates at 100. In this example, it is clear that markets reacted differently throughout the four months tested, but that the S&P 500 and the FTSE 100 followed each other closely, while the Nikkei 225 displayed some disparity in the latter parts of the graph. The value of the markets on September 1st, 1987 was used as a benchmark and each subsequent market close was taken as a percentage of that value to provide data to measure market correlation.

As a robustness test, the weekly percentage changes in market closes were also used for comparable correlation.
analyses. If the data were not adjusted for the percentage change in market close, and the magnitude of change was compared, then the S&P 500 would have a distinct disadvantage because the numerical value of its index is about one-fifth of the FTSE 100, while the Nikkei 225 is about one-hundred times larger than the S&P 500. Therefore, raw market data for the financial centers did not allow for proper comparison. The market closes had to be adjusted and measured as the percentage change in market close from one day to the next using the basic equation below.

\[
\text{% Change in Market Closes} = \frac{\text{Final market close} - \text{Initial market close}}{\text{Initial market close}}
\]

Results and Discussion

Primary Data Analyses

While it is well-known that economic crises create greater levels of uncertainty in the financial markets, the correlation among various financial markets is not clearly defined. Other studies have discussed the impact of historical events on the correlations of the stock market. However, most of the studies that conclude that market correlation increases during times of a financial crisis usually examine one historical event and compare volatility (Hameo, 1990).

Results of the current study examine market correlations across four historical events. Adjusted market close data for the four historical events examined in this study are shown in Graphs 1, 2, 3, and 4.

In general, based on visual inspection of the data, the market most closely linked to the historical event has the largest single day percentage decrease in market close. This is expected since these world markets are not completely dependent on any of each other. With the exception of the Asian Currency Crisis of 1997, the relationships between any two indices are relatively high, especially during the London IRA Bombing of 1990. A plausible explanation for the relatively low correlation between the Nikkei 225 and either the S&P 500 or FTSE 100 could be the fact that the Asian countries were hit the hardest during this economic crises.
Even though all of the markets may have been affected, the Nikkei 225 could have suffered from more of a contagion effect due to their close proximity to the Asian countries.

Pearson product moment correlations were calculated for the daily market closes of New York, London, and Tokyo for the time periods surrounding each of the historical moments selected in this study. Analyses used the adjusted market return data described above (see Table 1). Correlations measure the strength of relationships among variables. In this study, significance level was set at $\alpha = .10$ for a two tailed test. Any correlation of 0.150 or higher was considered significant. All Pearson product moment correlations in this study were associated with probability levels below 0.10, and all null hypotheses were rejected. When data trends across the four historical events were examined, the basic pattern or relationship is similar for all markets.

In this study, market correlation over this time period shows that the three markets measured have remained relatively constant. Because their correlations have remained similar, investors need to diversify their international holdings more carefully during times of economic crises. From visual inspection of the data and consideration of the size of the correlations, it is also clear that the financial markets of these developed nations are still subject to contagion, and degree of contagion will vary from market to market. During these times of an economic crisis, it is important for investors to properly diversify portfolios and have appropriate risk management applied to the portfolio. In order to properly allocate securities, the investor needs to understand the basic relationship among financial markets, namely market correlation.

For example, if investors use this knowledge, they could invest in the FTSE 100 and hedge it with investments in the Nikkei 225, because of their lower correlation. However, this is only valid during times of economic crises. Thus investors should be able to decrease the amount of systemic risk in their portfolios and hopefully attain a more optimal return. This is especially relevant during this recent subprime crisis that the United States market is facing. The data also show that the correlation between S&P 500 and the Nikkei 225 has remained consistent over time, holding a 0.8604 correlation in the 1987 U.S. Stock Market Crash and a 0.8617 correlation in the 2001 September 11 Attack.

**Alternative Explanations/Sensitivity Analysis**

This study suggests that the market correlation among the world's three largest financial centers tend to remain consistent over time with respect to historical events. However, several factors may have influenced the correlations in each of the models. As a robustness test for the correlation data, the correlations among the financial centers were analyzed using daily and weekly data. The daily data were used to determine the correlations, as they would be the most accurate. Furthermore, daily data were measured as a percentage based on a benchmark date. In order to test the accuracy of the correlations, a separate correlation matrix was developed for each of the markets, comparing the weekly percentage change in market closes, as shown in Table 2.

**Table 1. Correlation matrices examining daily relationships among New York, London, and Tokyo stock markets for four historical events.**

<table>
<thead>
<tr>
<th>Historical Event</th>
<th>S&amp;P 500</th>
<th>FTSE 100</th>
<th>Nikkei 225</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1987 U.S. Stock Market Crash</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTSE 100</td>
<td>0.936342166*</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Nikkei 225</td>
<td>0.86040085*</td>
<td>0.918773894*</td>
<td>1</td>
</tr>
<tr>
<td><strong>1990 London IRA Bombing</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTSE 100</td>
<td>0.981574725*</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Nikkei 225</td>
<td>0.98217989*</td>
<td>0.975210072*</td>
<td>1</td>
</tr>
<tr>
<td><strong>1997 Asian Currency Crisis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTSE 100</td>
<td>0.845510039*</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Nikkei 225</td>
<td>-0.512928005*</td>
<td>-0.715090227*</td>
<td>1</td>
</tr>
<tr>
<td><strong>2001 September 11th Attack</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTSE 100</td>
<td>0.949202634*</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Nikkei 225</td>
<td>0.861660848*</td>
<td>0.852065544*</td>
<td>1</td>
</tr>
</tbody>
</table>

*Statistically significant Pearson product moment correlations at .01 level.

**Table 2. Correlation matrices examining averaged weekly relationships among New York, London, and Tokyo stock markets for four historical events.**

<table>
<thead>
<tr>
<th>Historical Event</th>
<th>S&amp;P 500</th>
<th>FTSE 100</th>
<th>Nikkei 225</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1987 U.S. Stock Market Crash</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTSE 100</td>
<td>0.587782268*</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Nikkei 225</td>
<td>0.456410067*</td>
<td>0.766419154*</td>
<td>1</td>
</tr>
<tr>
<td><strong>1990 London IRA Bombing</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTSE 100</td>
<td>0.819118245*</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Nikkei 225</td>
<td>0.550768894*</td>
<td>0.708344143*</td>
<td>1</td>
</tr>
<tr>
<td><strong>1997 Asian Currency Crisis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTSE 100</td>
<td>0.560489952*</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Nikkei 225</td>
<td>0.214382891*</td>
<td>0.372009271*</td>
<td>1</td>
</tr>
<tr>
<td><strong>2001 September 11th Attack</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S&amp;P 500</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTSE 100</td>
<td>0.731072372*</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Nikkei 225</td>
<td>0.509313356*</td>
<td>0.617683684*</td>
<td>1</td>
</tr>
</tbody>
</table>

*Statistically significant Pearson product moment correlations at .01 level.
The Pearson product moment correlations are lower for the weekly data, but the trend remains the same and all correlations remain significant. Thus, there is a relationship among the S&P 500, FTSE 100, and Nikkei 225 even using the daily measures for greater sensitivity.

Regulatory procedures affect sentiments among investors within a specific market. For example, the United States has the strictest type of regulations, most notably the Sarbanes-Oxley Act of 2002 (Gumbel, 2007). In contrast, London has more lenient regulations, which attract more business (Gumbel, 2007). In this study, regulation may serve as a confounding variable, as its influence cannot be directly observed. If the three markets adapted the same international accounting standards, then this confounding variable would practically cease, albeit the efficiency and correlation of the markets may or may not change (Grimes, 2007). This could help explain why the New York market was more strongly affected than the other markets. In this study, some of the errors presumed to exist in the models are beyond the scope of analysis. Future research should include additional variables that may influence market performance and interrelationships among major world markets.

**Conclusion**

Based on this study, the market correlations among the S&P 500, FTSE 100, and the Nikkei 225 have not significantly changed over time. From the time period of 1983-2003, several historical events have occurred that have directly affected the financial markets. However, only a select few were chosen for this study. In the beginning of the time period, the correlation among any two markets was relatively high, especially between the S&P 500 and the FTSE 100 at 0.9363. As time progressed and other historical events were measured, market correlations remained relatively high and consistent, with evidence that these developed markets have continued to affect each other in the same manner. Therefore, contagion during times of economic crisis for one market may affect each of the other markets, as they may be more correlated. Even with the findings of this study, however, investors should still use caution when investing in order to properly hedge their risks.

These results have important implications. The current economic conditions in the United States are challenging. The subprime crisis from the summer of 2007 has placed a sharp strain in the U.S. economy. This most recent economic crisis will no doubt be felt in other markets around the world, both developed and emerging. As a result, investors may need to reconsider their asset allocation during times of economic crises or anticipated times of economic crises.

**References**


**Mentor Comments**

Faculty mentor Robert Stapp provides an overview of the economic conditions that led Thomas Vo to pursue this examination of aspects of market volatility. Dr. Stapp identifies this work as highly relevant to understanding the impact of recent historical events in order help investors reduce risk.

*Stock market volatility generates uncertainty and leads to economic losses. Recent cataclysmic events such as the*
Asian currency crisis and 9/11 generated capital losses to shareholders of the world that have rivaled the most renowned occurrences in history. Investors have therefore been seeking a 'safe haven' strategy to maintain their accumulated wealth. The major stock markets of the world, namely Tokyo, London and New York, along with their principle indexes were examined in this research in order to determine which market provides the greatest degree of stability during turbulent times. This is a current subject of keen interest in financial economics. The dynamic relationship of the world’s security markets are certainly in flux. The principle subject matter of Thomas Vo’s research is perhaps one of the most relevant to this entire discussion as it provides an analysis of which market performs the best during unexpected shocks to the system. His research offers investors a strategy to reduce their risk in the face of uncertainty and thereby maximize their welfare.