Contents

2  About The Journal

7  Editor’s Note

9  The Crisis and Emancipation of the Modern Corporative Executive: How the Bhagavad Gita reinforces Edward Freeman’s Stakeholder Theory
Shreyans Parekh
University of Pennsylvania

35  Drive for Show, Putt for Dough: Rates of Return to Golf Skills, Events Played, and Age on the PGA Tour
John R. Watkins, Jr.
Davidson College

61  Determinants of New York City Residential Rental Prices
Larissa Marco
Washington University in St. Louis

85  Counterfeit Goods and Their Potential Financing of International Terrorism
Zachary A. Pollinger
Harvard University
About The Journal
The Michigan Journal of Business is a bi-annually published academic journal run by undergraduate students at the Stephen M. Ross School of Business at the University of Michigan. The Journal is intended to provide undergraduate students worldwide with a platform for exceptional work in the field of business. The Michigan Journal of Business seeks to publish distinguished theses, empirical research, case studies, and theories in issues relating to areas of Accounting, Economics, Finance, Marketing, Management, Operations Management, Information Systems, Business Law, Corporate Ethics, and Public Policy.

Mission
The contemporary business environment is exceedingly complex. Analyzing this real world phenomenon through traditional applications of theories often yield a suboptimal understanding of the world. The Journal, accordingly, encourages work that takes an interdisciplinary approach to understanding a topic and emphasizes the importance of incorporating the knowledge of liberal arts into an area of interest. By providing a venue to recognize high quality work, the Michigan Journal of Business gives an incentive for students to explore their area of interest, rewarding them with the experience to share the power of knowledge with others. The Journal’s mission and philosophy parallels the mission of the University of Michigan, the premier research university in the United States. This project will involve not only students from the University of Michigan, but will seek involvement from a worldwide audience.

Review Process
The organization is entirely student-run, with an editorial staff of about fifteen of the top students at the Department of Economics and the Stephen M. Ross School of Business at the University of Michigan. Each semester, the Michigan Journal of Business calls for papers from undergraduate students around the world. Throughout the semester, the editorial board carefully reviews, selects, and edits exceptional work for publication. Faculty willing to advise the Journal is formed from each department to give minor oversight for the project. Throughout the process, a blind review process is implemented to ensure an impartial review of all submissions.

Article Submissions
The Journal only accepts works from undergraduate students or works completed during undergraduate study. Each manuscript submitted should
include a short abstract, author information, and any acknowledgements. Papers will be evaluated based upon sound analysis, originality of argument and novelty of research. For more information, please visit michiganjb.org.

**Editorial Board**
The editorial board is composed of the top undergraduate students at the Stephen M. Ross School of Business and the Department of Economics at the University of Michigan. They are selected through a competitive application process based on academic merit and aptitude in writing. This diverse group of students includes summer associates at Morgan Stanley, a future law student at the University of Michigan Law School, and Class Valedictorian.

**EDITOR-IN-CHIEF**
William Moon

**SENIOR EDITORS**
Michael Billington
Dara Leung
Alexandria Zhang

**DIRECTORS OF FINANCE**
Justin Kim
Charley Li

**DIRECTORS OF MARKETING**
Russell Shevins
Robert W. Weishar

**DIRECTOR OF WEB DEVELOPMENT**
Paul Sinkevics

**ASSOCIATE EDITORS**
Andy Balan
Emily Liu
Anthony Rubin
Franklin Shaddy
Chris Singh
Ming Tu
Charles Zvibleman
Faculty Supervision
The Journal is supervised by Professor Scott Moore, Author F. Thurnau Professor of Business Information Technology and BBA Program Director at the Ross School of Business. The following members of the Ross School of Business faculty also provide minor oversight to the editorial board during the review process:

Anocha Aribarg
Ph.D., University of Wisconsin
Assistant Professor of Marketing

George Cameron III
J.D., Ph.D., University of Michigan
Professor Emeritus of Business Law

David Hess
J.D., Iowa, Ph.D., University of Pennsylvania
Assistant Professor of Business Law and Business Ethics

Aneel G. Karnani
Ph.D., Harvard
Associate Professor of Strategy; Chair of Strategy

Scott A. Moore
Ph.D., University of Pennsylvania
Arthur F. Thurnau Professor; BBA Program Director; Associate Professor of Business Information Technology; Supervisor for Michigan Journal of Business

Andrea Morrow
M.A., University of Michigan
Director of Writing Programs; Co-Coordinator BBA Communication Program

Dana M. Muir
J.D., University of Michigan
Professor of Business Law
Anu Nagarajan  
*Ph.D., University of Michigan*  
Lecturer of Strategy

Lisa A. Pawlik  
*M.B.A., University of Michigan*  
Adjunct Lecturer of Business Communication; Co-Coordinator of BBA Communication Program

Josh Pierce  
*Ph.D., Michigan State University*  
Assistant Professor of Finance at University of South Carolina

Lloyd Sandelands  
*Ph.D., Northwestern*  
Professor of Management and Organizations & Professor of Psychology

Dennis G. Severance  
*Ph.D., University of Michigan*  
Accenture Professor of Business Information Technology

Amitabh Sinha  
*Ph.D., Carnegie Mellon University*  
Assistant Professor of Operations and Management Science

Valerie Y. Suslow  
*Ph.D., Stanford*  
Associate Dean for Degree Programs; Louis and Myrtle Moskowitz Research Professor of Business and Law; Associate Professor of Business Economics and Public Policy

David B. Wooten  
*Ph.D., University of Michigan*  
Associate Professor of Marketing

Lynn P. Wooten  
*Ph.D., University of Michigan*  
Clinical Assistant Professor of Strategy and Management & Organizations
List of Institutions

Submitted papers for our inaugural edition came from the following institutions:

University of Birmingham, U.K.
Bowdoin College
Cornell University
Davidson College
Duke University
Edinburgh University, U.K.
Harvard University
Indiana University in Bloomington
The London School of Economics
Middlebury College
University of California in Berkeley
University of Nottingham, U.K.
University of Georgia
University of Michigan
University of Pennsylvania
University of Toronto, Canada
University of Virginia
University of Washington
Villanova University
Washington University in St. Louis
Editor’s Note
I am pleased to present the inaugural edition of the Michigan Journal of Business. This page gives me an opportunity to reflect upon two important goals underlying the premise of our existence: first is to provide a much-needed platform for undergraduate students in business and economics to transition from consumer of knowledge to producer and consumer both; second is to select and distribute cutting-edge pieces that add to the intellectual capital in academic literatures. Over thirty submissions were considered for publication, and only space limited our desire to share many of the fine articles we were privileged to review.

Our inaugural issue contains four important articles that give us the pleasure to think about business within the context of the world, and as a product of human behavior. “The Crisis and Emancipation of the Modern Corporative Executive” is a management piece concerning corporate social responsibility and the stakeholder’s management theory, a topic of growing scholarly interests. The second article, “Drive for Show, Putt for Dough: Rates of Return to Golf Skills, Events Played, and Age on the PGA Tour” incorporates rigorous economic analysis to investigate the rates of return to golf skills on PGA Tours. The third article, “Determinants of New York City Residential Rental Prices” performs an empirical analysis on the physical and demographical characteristics of New York City to better understand pricing in the real estate market. Finally, the journal ends with, “Counterfeit Goods and Their Potential Financing of International Terrorism,” a thought-provoking piece that uses a number of economic control studies to draw public policy implications of counterfeit goods.

I would like to thank Professor Scott Moore for his enthusiastic support and leadership that allowed this project to take off. I would also like to thank our faculty sponsors at the Stephen M. Ross School of Business for their supervision for our editors. In particular, I am indebted to Professor Lloyd Sandelands, Professor Anocha Aribag, Professor Damian Beil, and Professor George Siedel for sharing their knowledge and expertise in running the business aspects of the project. I would also like to thank Dr. Michael Kamen, Professor Jim Walsh, Professor Penny Von Eschen, and Mr. Thomas C. Jones for their mentorship, whose wisdom allowed me to concentrate on this project throughout the semester. Mr. Nathan Rupp from the Kresge Library and Mr. Alex Edelson from the Michigan Journal of Political Science have also been of a tremendous help for this project. I would like to thank the editorial board at the Michigan Journal of Business, a distinguished group
of students at the University of Michigan, whom I admire greatly for their intellectual devotion and friendship. Finally, I would like to thank all of our readers who give the editorial board at the journal the remarkable gift to experience academic journalism.

William J. Moon
Editor-in-Chief
Abstract

Business leaders are looking towards corporate social responsibility as a critical means to satisfying their stakeholders, while facilitating their firm’s long-term sustainability. As part of this holistic trend, more managers are looking towards the Bhagavad-Gita for lessons on leadership and character development in place of the classic Sun Tzu’s Art of War. Companies that currently lack strong CSR initiatives must follow suit in developing socially responsible products and projects due to the tremendous value that such actions can unlock for stakeholders. The Gita requests leaders to be the sattvic frame of reference for ethics and morality that the corporate world has lacked for many years. Important implications of CSR are the cooperation of corporations and NGOs with like-minded social objectives, and the creation of new firms and careers dedicated towards the research and development of socially responsible projects. With the greening effect sweeping across the United States, CSR will be a crucial moving force in the corporate world, and business executives must acknowledge that social value can be an amplifying force for corporate value.
I. Introduction

It is the year 506 B.C. and Sun Tzu’s armies are descending upon the ancient Chinese city of Ying to claim victory over the warring Chu Empire. Near the same time, during the climactic Kurukshetra War in ancient India, Prince Arjuna of the Pandava clan is faced with a dilemma on the battlefield and turns to his revered mentor and godly incarnation Krishna for advice. The enemies that Arjuna faces on the battlefield are similar to the foes that Sun Tzu encounters. For Arjuna, they are his own friends and family members who happen to be part of the rival Kaurava clan. For Sun Tzu, they are his fellow compatriots who have grown up in the rebellious southern state of Chu.

Faced with parallel situations, Sun Tzu and Arjuna approach their battles with contrasting ideologies. Sun Tzu devises clever tactics to conquer the enemy at all costs and institutes “iron discipline” in order to train his troops. Arjuna feels that enlightened leaders should be compassionate and fair. He hopes to concentrate more on the purity of his thoughts and actions, rather than on solely achieving victory.

Today, it is the year 2007, and many corporate boardrooms are experiencing a transformation. Managerial styles have evolved from the era twenty years ago when Gordon Gekko argued that avarice was the only way to survive in business, in the movie *Wall Street*. Nowadays, more executives are adopting a more corporate social responsibility-minded approach when generating profits instead of the ruthless, bottom-line model. They are realizing that a firm can only persist in the long-run if it satisfies needs of all stakeholders, and not solely the shareholders.

In the modern-day corporate battleground, the “generals of the private sector” are faced with a similar dilemma as Prince Arjuna; there are many other factors involved in a corporation’s wellbeing than solely the end goal of “victory” by achieving profits. Some business leaders who once adopted their management style from Sun Tzu’s *The Art of War*, and subsequently, Gerald Michaelson’s *The Art of War for Managers*, are now turning towards the ancient Indian text *Bhagavad Gita* for strategic advice. The text that leaders such as Mahatma Gandhi and scholars such as Ralph Waldo Emerson considered to be a practical guide to living a prosperous life is now transforming the way business is conducted globally.

For Arjuna, the “corporation” consisted of not only his own Pandava

---


clan, but all of his loved ones, including those who would not be on the glorious end of battle. He could have adopted the same approach as Sun Tzu, with victory at all costs as his sole ambition, but instead wished to take all factors of his “corporation” into account when achieving his end aspiration. From today’s business managerial perspective, this dilemma surfaces when executives must decide whether the objectives of the corporation should include all stakeholders in the firm, or just its shareholders, who have traditionally been the primary focus.4

This transition from Sun Tzu’s \textit{Art of War}, which relates to the late Milton Friedman’s Shareholder Theory, to the \textit{Bhagavad Gita}, which is associated with Edward Freeman’s Stakeholder Theory of the Modern Corporation, is evidence that more business leaders are adopting a holistic approach to business – one that takes into consideration all the factors that may influence a company’s outcome. This trend must continue, however, even for firms that do not see the stakeholder approach as a priority to conducting business. “The best way to describe it is inclusive capitalism,” says C.K. Prahalad, a professor at the University of Michigan’s Ross School of Business. “It’s the idea that corporations can simultaneously create value and social justice.”5 The success that these firms have had over the past few years with their examples of corporate social responsibility (CSR) and a more universal stakeholder approach should be evidence that more firms should include CSR as part of their operations because it would reap tremendous rewards for them. Furthermore, companies must take a proactive, offensive-minded approach in implementing CSR instead of hiding behind community actions as a defensive mechanism.

With the many scandals that have sullied the corporate world in the past decade, today’s modern corporate executive faces a crisis. The ideology of the relentless pursuit of the bottom line that has propelled the business world for so long has been tainted by many forms of unethical behavior.6 More business leaders should follow the footsteps of those who have already crossed over, and should look towards the \textit{Bhagavad Gita}’s lessons as spiritual and mental guides for leadership. The greening effect that is taking

hold on many firms today is not a fad; it is a testament to the fact that many corporate executives acknowledge that long-term sustainability for a firm depends on socially responsible action today. Analogous to the philosophy of Arjuna, this mindset will lead to a sense of personal emancipation for the manager and satisfaction of all stakeholders in the organization.

II. Comparisons between Bhagavad Gita and Freeman’s Stakeholder Theory

Work for the common good without selfish interests; the rewards of selfless work will take you to a supreme state.

Krishna, The Bhagavad Gita

Today’s global companies are taking a more holistic approach to the notion of “green” than solely meeting the bottom line. Business leaders are faced with the challenge of how to achieve a competitive advantage while promoting the prosperity and wellness of their stakeholders. Corporate Social Responsibility (CSR) is the idea that organizations are responsible for considering the interests of not just shareholders, but also of customers, employees, communities, fellow businesses and the local environment that is affected by business practices. Organizations must consider the social and environmental ramifications of their business activities. By integrating CSR into core business processes and stakeholder management, companies can achieve their ultimate goal of creating both corporate and social value.

Central to both Edward Freeman’s Stakeholder Theory of the Modern Corporation and the Bhagavad Gita is the theme that managers bear a synergetic relationship to stakeholders, who are all groups who have a stake in or claim on the leader’s decision making and the fruits of the corporation. Freeman argues that this relationship is based on a fiduciary duty while Krishna asserts that this is relationship is based on karma yoga. A leader who practices karma yoga thinks less of personal advantage and more of the fulfillment of the common good. As Krishna advises in chapter 2 of the Gita, “the righteous one who accepts and shares the rewards of his actions is freed from sins, while he who works for self-interests incurs sin.”

Taken from a managerial perspective, an executive should be concerned with the best interests of the entire corporation, and not solely on pursuing profits. Opponents of this theory point out a major flaw in the Stakeholder Theory about conflicting interests. After all, a company cannot

The Crisis and Emancipation of the Modern Corporate Executive

surely please *everyone* involved with its operations: what should a firm do if stakeholder interests conflict? There have been many studies proposed regarding analysis of the Stakeholder Theory and Conflict Management. Stakeholder analysis is the answer to critics. This analysis has been developed by experts in fields of study ranging from business management to international relations, policy development, ecology, and natural resource management. It is used to differentiate and study an organization’s stakeholders on the basis of their attributes and the criteria of the analyst during the assessment of a specific situation to a range of tools for the identification and description of stakeholders of a firm on the basis of their attributes, interrelationships, and interests related to a given issue or resource.

When interests in a firm collide, as they often do, managers need to implement stakeholder analysis within their organization to resolve these conflicts. This analysis answers the following questions: the relative power and interest of each stakeholder, the importance and influence they have on the firm’s operations, the multiple “hats” they wear and the networks and coalitions to which they belong.8 Ricardo Ramirez of the University of Guelph in Ontario, Canada, divides stakeholders in an organization into four major categories: those with claims to legal protection, those with political clout, those with power to block negotiated agreements, and those with moral claims to public sympathy. The stakeholder analysis described by Ramirez considers nine major propositions that are organized under a conceptual framework. A manager must proceed through the framework step-by-step to resolve conflicts between stakeholders.

One of the propositions that managers must address in their analysis is the notion that three stakeholder attributes merit the most attention in conflict resolution: a stakeholders’ power, urgency and legitimacy. This proposition follows that stakeholders with two or more attributes are likely to be noticed and participate in the firm’s decision-making first, while those without them will be on the fringes of the decision-making process. The analysis proceeds to describe this and other propositions in intricate detail.

Certain dilemmas arise when Ramirez’s analysis is taken to its logical extremes. For example, if the company’s most powerful and legitimate stakeholder wants the firm to employ tactics that could be harmful to the firm or to society, such as the destruction of the environment, inhumane employee treatment or even human slavery, what should a corporate executive do? In Ramirez’s model the needs and concerns

---

of stakeholders that match two attributes, should be addressed first, but that does not necessarily indicate that they will always have to be met by corporate executives. For example, if a stakeholder wants to implement an idea that would cause harm to the firm’s image, such as one of the three extreme examples provided above, a manager has the right under Ramirez’s model to modify or even negate the stakeholder’s claims if he feels that they are not rational. According to Ramirez, it is important to acknowledge that, “these attributes are variable over time, are socially constructed rather than objectively real, and the resulting influence may or may not be willfully or consciously exercised.”9 Hence, a manager has the flexibility to be rational in his decision-making and implementation of a stakeholder’s needs and desires.

According to the typology of the model, if stakeholders possess two attributes, they are grouped under the heading Expectant stakeholders. Among these, those having power and legitimacy are called Dominant while those that have legitimacy and urgency are deemed Dependent. The most poignant description in the model is that of stakeholders that possess power and urgency; these are called Dangerous stakeholders. Ramirez recognizes that stakeholders who possess only power and urgency may want the firm to engage in activities that are hazardous to the firm’s wellbeing and image, and are thus considered “dangerous” to the firm. After all, they have tremendous influence in the firm, and salience in the minds of managers. However, according to Freeman’s original stakeholder model, stakes change depending on the strategic issue under consideration. It is unlikely that a manager would abide by a “dangerous” stakeholder with dangerous implications, because this would not be advantageous to the strategic planning being implemented at the time.

By putting this useful and successful tool to practice in the workplace, managers can decide the best paths to take when stakeholders’ interest collide. The analysis is structured, but flexible enough for managers to use their own personal judgment and acumen during turbulent times for a firm. Nowhere in the stakeholder analysis does it claim that shareholders take priority above all other stakeholders in a firm, as a shareholder theorist such as Friedman would articulate. Instead, the model views the shareholder as being equally weighted to all other stakeholders in the firms, and if a shareholder falls into a category that takes priority over another at the time, then his or her needs should be tended to first. One caveat to this

---

rule is that since shareholders play such an important role in the decision-making processes and management of the firm, they will often take priority, but the key point is not all the time. Stakeholder analysis presents many opportunities for managers to tend to the needs of other stakeholders before shareholders.

Ed Freeman claims that when comparing the stakeholder and shareholder models, one should not view them as completely contradictory theories. Instead, Freeman states that the stakeholder approach can be implemented to produce an end result as profitable as or even more profitable than the shareholder approach. For Krishna, this objective is the highest level of consciousness and self-awareness that transcends the physical world and physical senses. In the context of management and business, this “supreme state” refers to a higher level of leadership consciousness and firm value. Thus, if firms have the capability to “balance the interests of the firm,” as companies such as P&G claim to do in their mission statement, they could utilize acts of CSR to further the interests of the firm while still benefiting their communities.

Krishna would view the theory through a holistic approach, seeing it as a way to bring out the best in each individual who is important to the wellbeing of the overall organization. Analogous to the “Feminist Standpoint Theory” in Freeman’s Stakeholder Approach, Krishna claims that actions must be taken with high priority in social awareness and relationship management, because these are the factors that drive profits within a firm. This view returns to Krishna’s central message that if managers concentrate on their actions and all those that influence their actions, positive effects will inevitably ensue.

For example, Proctor & Gamble’s mission statement states: “We will provide branded products and services of superior quality and value that improve the lives of the world’s consumers. As a result, consumers will reward us with leadership sales, profit, and value creation, allowing our people, our shareholders, and the communities

---

in which we live and work to prosper.”¹³ P&G believes in satisfying all individuals valuable to the firm, not just their shareholders. As a result, consumers reward them with sales, profit, and value creation that allow the communities in which its employees live and work to prosper.¹⁴ Thus, P&G’s primary focus is social responsibility both within the firm and in its local and global communities: P&G contributes to the economic and social well-being of their employees, shareholders and the local communities in which they operate. On a larger scope, the company is involved in regional, national and international development. Achieving a desirable rate of return for its shareholders is one end goal for companies such as P&G in order to fund its social mission.¹⁵

The late London Business School professor Sumantra Ghoshal believed that the corporate scandals of a few years ago in the U.S. were offshoots of the shareholder theory of capitalism that were being proposed by economists and emphasized at business schools. He believed that corporations are not solely profit-generating robots reacting to market forces; “they are managed by and for humans, and have a symbiotic relationship with the world around them.”¹⁶ Similarly, social responsibility does not have to be the “cloak” that Friedman once described, but instead, could be the key to greater value generation for all firms, especially those who have been weighed down by their prominent shareholder focus. The quest, says Prahalad, is to “develop a capitalism that puts the individual at the center of the universe,”¹⁷ thus positioning suppliers, employees, customers, and other stakeholders of the firm first so that they can benefit shareholders.

III. Contrast between Sun Tzu’s The Art of War and Krishna’s Bhagavad Gita

All warfare is based on deception. Hold out baits to the enemy. Feign disorder and crush him.

Sun Tzu, The Art of War

¹⁴ Ibid.
¹⁷ Ibid., 84.
Be fearless and pure; never waver in your determination or your dedication to the spiritual life. Cultivate vigor, patience, will, purity; avoid malice and pride. Then, you will achieve your destiny.

Krishna, *The Bhagavad Gita*

While the *Bhagavad Gita* and *Art of War* both address a leader’s decision-making and resolve during a time of great tribulation, their original purposes differ immensely. The *Art of War* was written to be a manual for strategy in battle while the *Gita* was composed to present war as just one of the manifestations of struggle that individuals have to overcome in their daily lives. Even the titles of the texts themselves reveal the differences in their intent. The *Bhagavad Gita* literally means “Song of the Blessed One,” alluding to the teachings of Krishna, who is a sage that has reached a higher level of enlightenment.\(^{18}\) The *Art of War* refers to the beauty and craft of decision-making and organization in the circumstance of war.

In the times of Sun Tzu, warfare was prevalent and consisted of burning cities, reading signs of favorable and unfavorable weather patterns, interpreting the disturbance of birds as the enemy advanced, and motivating troops to fight with resolve and valor. If generals strayed from their strategy and their end focus, they failed. At times, even deception was needed to achieve victory. For Friedman, the business environment was as ruthless and unyielding. In his opinion, any firm whose motives strayed from increasing shareholder’s profits had no reason to exist.

In the *Gita*, whether or not Arjuna should fight is a secondary question. The real question that he faces and that managers should acknowledge from the text is: how should individuals live their daily lives.\(^{19}\) For Krishna, battle was more a spiritual quest than a physical display of brute strength and domination. When Arjuna consults his guide on the battlefield, Krishna emphasizes that he should not look towards the opposing army as his friends and family, because he will then get attached to his relationships with these individuals instead of his duties. It is what Gandhi called “renunciation of the fruits of action”\(^{20}\) that propels Arjuna to shed his weaknesses and accomplish his tasks with emphasis placed on the purity of his actions over the end result. Similarly, for Freeman, it is this desire to focus on the

---


intermediate steps of a goal that prompted him to form his six principles of
the Doctrine of Fair Contracts. Freeman hopes to redefine the purpose of
the firm with these principles, by revealing that an emphasis on stakeholders
brings new dimensions of decision-making and leadership that a purely
stockholder-centric approach lacks.

The gravitation towards *The Art of War* as the managerial leadership
text of choice in the past seems logical. After all, for years, people have
regarded Sun Tzu’s *Art of War* as a manual of intricate strategy with tactics
for victory that can easily be implemented in the workplace. The *Gita*, on
the other hand, has been seen as purely a religious text. Only today are
more leaders admiring the *Gita* as an ancient text of spiritual wellbeing and
guidance. “To me, most aspects of Hindu philosophy such as the *Gita*, the
*Vedas* or the *Upanishads* are nothing but scriptures of knowledge and lessons
of leadership,” says Dipak Jain, dean of Northwestern University’s Kellogg
School of Management.

“The Art of War brings leadership that is unidimensional and
completely focused on victory. To me victory with a huge social cost is not
a victory,” says Jain. “You have to ask yourself is life just about winning, or
winning subject to constraints?”

The key perception today is that executives are seeing the corporate
world less as a physical battlefield that justifies sly tactics and deception,
and more as an internal challenge of leadership ability, accountability
and organizational values. The basic idea proposed by Sun Tzu remains
the same: “he whose ranks are united in purpose will be victorious.”

Furthermore, today’s CEO is more a cooperating value-generator than a
money-mongering warrior, and his “troops” entail all stakeholders, not just
those who are attached to wealth generation. The holistic approach used
by some managers today is “paying dividends” in many forms, and if more
corporate executives conducted business operations in this manner, more
firms would experience the long-term value generation associated with CSR.

IV. Shift from Shareholder to Stakeholder Focus

By pursuing his own interest [an individual] frequently promotes
that of the society more effectually than when he really intends to

21 H. Jeff Smith, “The Shareholders vs. Stakeholders Debate,” MIT Sloan Man-
agement Review 44.4 (Summer 2003): 87.
22 Dipak Jain, Telephone interview, 19 Nov. 2006.
23 Ibid.
promote it. I have never known much good done by those who affected to trade for the public good.

Adam Smith, The Wealth of Nations (1776)

In recent years, business ethics and the stakeholder theory has become a popular area of study in academic literature and implementation in the corporate setting. According to Wharton professor Thomas Donaldson, “The focus on corporate stakeholders has become the symbol of the modern effort to redefine the corporation.”

An important element of the shift to the stakeholder approach has been the effort to promoting corporate social responsibility both inside and outside of the workplace. In articulating his orthodox view of free market capitalism with the 1970 article “The Social Responsibility of Business is to Increase its Profits,” economist Milton Friedman balked at capitalists who claimed that business is not concerned “merely” with profit but also with promoting desirable “social ends.” Even today, noted economists such as Arthur Laffer, known as the “father of supply-side economics,” claim that “there is no positive correlation between CSR and business profitability” based on an analysis of 28 companies that were among Business Ethics magazine’s “Top 100 Corporate Citizens” every year from 2000-2004.

However, proponents of CSR including Proctor & Gamble CEO Alan Lafley and Dipak Jain, dean of the Kellogg School of Management, strongly disagree. Lafley would argue that Friedman’s view could not apply to modern-day firms and today’s free market capital system because there are numerous external factors that play a role in a firm’s decision-making today that did not exist when Friedman devised his argument decades ago. For example, firms no longer utilize CSR as a “cloak” for short-sightedness but rather as a strategic tool to enhance their image, and in doing so, their profits. Corporate social responsibility reflects a shift in society’s priorities: civil society has increasingly turned to the private sector to address critical needs including disease prevention, provision of potable water supplies and

educational opportunities.

This shift has also reflected claims on large corporations. In Lafley’s opinion, Friedman’s view can be viewed as a modern-day interpretation of Sun Tzu’s exploitation of an enemy’s weaknesses in order to further one’s own quest towards a supreme goal. In today’s society, this mentality towards running a firm will not suffice. Lafley would propose that his company’s rapid global growth can be attributed to its firm stakeholder model, and this is evidence that caring about more than just shareholders means good business overall, even for businesses who do not use CSR just to attract and retain customers and employees.

Jain cites the Bhagavad Gita as the text that has often guided his actions in satisfying all stakeholders of his organization, the Kellogg School of Management, including the students, faculty, alumni and media. “Shareholders are primarily focused on outcome,” says Jain. “The Bhagavad Gita teaches us that you don’t control the outcome of your actions, you only control the efforts. Thus, if you focus on your efforts, you have to focus on all stakeholders and not just the shareholders.”

For both Lafley and Jain, profits are the means to value generation, and not the end in itself. This sense of “enlightened capitalism” has been one of the main reasons that managers have transitioned from The Art of War to the Bhagavad Gita. Times have changed since it was considered wise for companies to pursue only their own economic self-interests, and managers now feel the need to balance relationships within the firm. It is no surprise that the management guides linked to both the shareholder and stakeholder approach have shifted as well.

V. Responses to Critics of the Stakeholder Theory and Corporate Social Responsibility

Critics of corporate social responsibility such as David Vogel, professor at the University of California at Berkeley, claim that CSR only makes “business sense” for firms in niche markets, or those that have a “need” to pursue CSR to repair their reputations that have been sullied in some way or prevent it from becoming a competitive disadvantage. For instance, Vogel would claim that based on empirical evidence, it does not make sense for Proctor & Gamble, the world’s largest consumer-products company, which does not cater to a niche market and has not had any significant internal crisis or media blitz in the past few years, to invest heavily in CSR projects. However, this is simply not the case. Vogel’s view that CSR provides advantages in terms of public relations and image

29 Dipak Jain, Telephone interview, 19 Nov. 2006.
enhancement falls short because there are other significant reasons why firms should pursue CSR. These reasons include opportunities for firms to tap into untouched international markets, prospects for the firm to stay on top of cutting-edge technology and open avenues for product development, and the pursuit of long-term value generation and sustainability for the firm. Each of these objectives has been a direct offshoot of CSR activities for major firms, such as P&G which will be described later in the section. None of these ideas have been incorporated into Vogel’s limited definition of the benefits of the CSR, and thus his model is incomplete.30

In response to critics like Vogel who argue that CSR has its shortcomings, Steven D. Lydenberg, chief investment officer of Domini Social Investments, a prominent socially responsible investing firm based in Providence, RI proposes that CSR should be a core element of corporate management. In his recent book entitled Corporations and the Public Interest: Guiding the Invisible Hand, Lydenberg proposes many strategies that executives should employ to integrate CSR into the fabric of a company. For Lydenberg, “CSR is ‘a major secular development, driven by a long-term reevaluation of the role of corporations in society.’”31 Lydenberg’s main argument is that corporate social responsibility is an excellent and viable method of creating long-term wealth for companies such as P&G.

Lydenberg cites that a significant development on the horizon, is guidance from the International Organization for Standardization (ISO) for implementing CSR programs. This standard will be called the ISO 26000 standard for social responsibility, and is expected to come out in 2008. This standard will contain guidelines, not requirements, and therefore will be used to be a certification criterion like the ISO 9000 standard for quality. It will continue to add value to firms who have already been employing CSR, and will reinforce the decision to implement CSR for those who have been thinking about making the appropriate decisions, but have not made the move yet. Lydenberg believes that companies that are concerned with sustainability issues will adopt the standard and ask their suppliers and vendors to follow suit.32

For public companies such as P&G, the complete and wide-ranging disclosure of social and environmental performance “has already passed


31 Ibid.

32 Ibid.
the tipping point,” according to Lydenberg. So far, P&G and more than 700 other firms have published sustainability reports using part or all of the Global Reporting Initiative’s guidelines for reporting, in which companies disclose a “triple bottom line” of economic, social, and environmental performance. Lydenberg cites that such disclosure helps educate consumers and investors, who in turn help steer companies to the public interest. By making their CSR activities known publicly, these companies boost their public image, an extra step towards gaining a foothold on competitors and tapping into further markets.

This idea is prevalent in P&G’s ventures in CSR. Corporate social responsibility does not have to reap rewards solely for firms that utilize it as part of their identity, or for firms that do not want it to become a source of competitive disadvantage, as Vogel states. There are many firms that have the capability, time and capital to invest in CSR, and thus should make an effort at some level. According to P&G spokesperson Terry Loftus, the company states that their efforts overseas are a clear link to their approach to understanding diversity, which in turn is a “fundamental business strategy.”

Vogel’s claim that the “international impact of CSR on brands has been negligible” is simply not valid today. For example, P&G offers over 300 brands in more than 80 countries worldwide, and its success “depends entirely on our ability to understand these diverse consumers’ needs,” according to a company spokesperson. A main reason why the company has donated a significant amount of financial support and supplies to help disadvantaged youth in Vietnam, combat childhood malnutrition in India, and provide earthquake relief in Turkey, is to gain the insight and customer base needed to continue growing worldwide. P&G continues to invest in CSR year after year even though they already have such a reputable international image. CSR “makes sense” by such a brand name firm not only because it is at risk of being targeted by activists, as Vogel claims, but due to the fact that

every year such actions increases the company’s scope and awareness of the international community, allowing it to tap into markets worldwide.

Lynenberg states that when companies recognize the need for CSR, they realize the immense potential of “long-term wealth” generation. “Corporations create long-term wealth,” he writes in Corporations and the Public Interest, “when, in addition to generating productivity gains, they preserve natural resources for future generations, create value in their relationships with their stakeholders, and do not externalize costs onto society.” This notion of long-term wealth leads to future sustainability along with profits in the short-term, a point that is missing in the arguments of Vogel and other critics who claim that CSR is only profitable for select companies.

In the article, “Two views of virtue” published in the December 2005 issue of CFO magazine, Vogel himself claims that he will only change his argument about CSR when “articles about companies in the mainstream business press regularly or even occasionally mention some aspect of CSR in terms of...understanding a company’s past or future performance.” This issue is directly addressed by Lynenberg’s proposal of long-term value creation, which supports addressing stakeholder needs in order to evaluate future value generation in comparison to past wealth, expansion and other value-generating opportunities for the firm. Vogel believes CSR will always remain “marginal and secondary,” to firms, but this is simply not the case in today’s corporate world. Without question, CSR is becoming part of the global corporate culture; today, more than 2,000 U.S. companies publish annual corporate social responsibility reports along with their financial statements compared to a few hundred just two years ago.

An article published in The Economist magazine claims that “it would be a challenge to find a recent annual report of any big international company that justifies the firm’s existence merely in terms of profit, rather than ‘service to the community’.” According to the article, “The good company” deputy editor Clive Crook claims that signs of the “victory of CSR advocates” can be seen in the speeches of top executives and the diligent reporting of CSR efforts in their published accounts. In addition, corporate

38 Ibid.
39 Ibid.
social responsibility is now an industry in its own right, and a flourishing profession as well. Consultancies have sprung up to advise companies on how to do CSR, and how to let it be known that they are doing it. In a survey of the 1,500 delegates, which included mostly business leaders, attending the World Economic Forum in Davos, Switzerland in 2004, fewer than one in five respondents stated that profitability was the most important measure of corporate success. While a little under five percent of the constituency named CSR as the single most important criterion, it should be noted that an additional 24% said that the reputation and integrity of the brand, to which good corporate citizenship is a critical element, mattered the most to them.41

Critics of this model also assert that there many other cost-effective ways for firms to create “long-term wealth” than to just employ CSR. Also, a critic may ask how it is possible for all firms to be able to abide by a “triple bottom line” when many are merely trying to maintain the single bottom line of profit generation for its shareholders. Other critics perceive corporate social responsibility as incurring unnecessary costs and being a waste of time for companies whose image does not rely upon it.

The answer to these arguments lies in the fact that there are many levels to corporate social responsibility that managers often fail to realize. Comparable to Krishna’s ideology, corporate social responsibility could start at the individual level, ensuring that relations between managers and employees are open and good-natured. In addition, CSR is one means that is highly effective and can be cost-efficient for all firms if they employ it to address the unique needs of their individual firm, as P&G has done, according to Lyndenberg.42

Taking small steps to incorporating CSR in the workplace is the key for the modern corporate executive. In 1999, the animal rights organization In Defense of Animals launched a vicious campaign against P&G to stop the company’s techniques of animal abuse due to product testing that occurred each year. They labeled their global campaign “P&G Kills” and stated that the company refused to switch to more humane and reliable alternatives due to extra costs involved.42 This campaign and boycott against P&G hurt the firm’s image tremendously. P&G decided to take action and invest in more humane measures, but they did not stop there. Today, they are known worldwide as a global leader in providing animal alternatives and were

recognized by the Humane Society of the United States for their efforts to find alternative methods to test their products.

P&G could have restricted its CSR initiatives to modifying its animal testing practices, as Vogel would have recommended to repair its image, but P&G continued to take larger leaps to enhance its community and to be on the cutting edge of environmentally sound technology. P&G took an aggressive offensive-minded approach, rather than just settling with the “primarily defensive” display of corporate virtue as Vogel advocates. P&G’s risky investment to be on the forefront of animal rights issues has enabled the company to tap into a large market of ecologically-cognizant consumers that was previously accommodated by smaller brands.

In 2002, Vancouver-based Standberg Consulting conducted a study entitled “The Future of Corporate Social Responsibility,” in which it created a CSR continuum open to all firms, not just those in limited and narrow positions. The consulting group intended to pinpoint any firm on the continuum, based on the scope, or lack thereof, of their CSR agenda. The firm identified five levels of CSR: CSR Lite, CSR Compliant, CSR Strategic, CSR Integrated and Deep CSR. CSR Lite includes firms that are primarily concerned with responsiveness to complaints, not CSR, and thus their business model will not be changed. This category also includes firms that have been forced into compliance with CSR objectives, but do not show any commitment to social or environmental progress.43

The consulting firm stated that all companies need to make the leap to at least CSR Compliant, in which firms keep abreast of emerging standards of CSR and ensure they are compliant with those standards. The next level is analogous to Vogel’s viewpoint on CSR. CSR Strategic includes firms that develop business strategies within one or two aspects of CSR around which they can develop a competitive advantage and have a significant impact. The last two levels of CSR include firms that are the most committed to incorporating CSR within their business plans. While it is may not be in the best interest of every firm to operate at the last two levels of CSR, more firms should make the leap from the CSR Lite category for the benefit of the firm and society as a whole. All firms do not have to be heroic by pursuing societal and environmental improvements with alacrity, but as P&G has shown, a gradual focus on CSR efforts can make a tremendous difference for a firm.44

44 Ibid.
VI. *Bhagavad Gita* on Leadership Character: Sattvic vs. Rajasic

A sattvic person performs actions with a firm belief in his or her responsibilities and without attachments to results.

Krishna, *The Bhagavad Gita*

Throughout the *Bhagavad Gita*, Krishna and Arjuna discuss the virtues and vices of human character. Today, these same personal qualities can be seen driving decision-making in the workplace. In chapter 7 of the *Gita*, Krishna describes two main types of human character: sattvic and rajasic. Sattvic leaders are driven by the harmony between the self and the surroundings and are repelled from provoking confrontation or disparity. In addition, they do not hesitate to accept suffering if such hardships cause peace and harmony for the people in their organizations.

Applying Krishna’s philosophy, Proctor & Gamble announced in 2005 that it will launch its Children’s Safe Drinking Water initiative in order to help distribute its low-cost PUR Purifier of Water technology to developing countries around the world. P&G aims to transform contaminated water sources to meet World Health Organization standards for safe drinking water. P&G has already invested millions to provide safe and clean drinking water to children in Africa and around the world, because such sustainability is instrumental to the company’s values and policies. Executives at P&G know that such decisions do not directly affect potential consumers of their products; nevertheless, they continue to maintain their humanitarian focus because corporate social responsibility has been part of their business identity for years. It is this commitment to their community that has helped P&G’s brand shine for decades.

“Sattvic leadership radiates calmness, purity, and righteousness. It embodies happiness and knowledge,” says Krishna. Leadership character should thus incorporate integrity and a focus on people and competency. The central leadership message of the *Gita* directs people to become satvic leaders who put the interest, needs and wants of others ahead of themselves. When leaders act with purely selfish motives, they tend to emit negative energy to those within their firm. This negative energy can have

grave ramifications in the workplace. Krishna calls this type of leadership character rajasic. Rajasic leaders “create chaos and darkness in their organizations”\(^49\) by their egotistical desires.

“If you look at the world of business today, the biggest challenge we face is on the front of leadership,” says Jain.\(^50\) This theme is clearly relevant for company executives who are more inclined to put company profits and the shareholders’ interests over the overall value of the corporation. Financial incentives for business leaders are often based on aggressive financial goals that focus less on customers and employees than on the bottom line of profit generation. Unfortunately, company value is often equated with purely shareholder’s profits and according to the teachings of the Gita, this should not be the case.

VII. Future Implications of CSR

CSR has become a subject of great importance when companies assess their strategic plans. Many of the world’s leading corporations are already integrating socially responsible solutions into their business models, while making great strides towards social and environmental responsibility. Companies are finding that “corporate value and social virtue not only are mutually exclusive but can be mutually reinforcing.”\(^51\) While many companies acknowledge that the benefits of CSR are great today, those that are not currently engaging in CSR should realize that the future implications and amplified effects of CSR are significant.

With the dramatic shift in corporate attitudes toward CSR that has occurred in the past decade, more companies will utilize CSR as an offensive rather than defensive tactic employed to protect oneself from negative publicity. Many companies today engage in CSR, “because they believe that being socially responsible is good for their business -- whether it’s through building their brand reputation, mitigating risk or improving employee retention and productivity, just to name a few benefits,” says James Viray, director of the State Department’s Office of International Labor and Corporate Social Responsibility. Other companies engage in CSR, he added,


\(^{50}\) Dipak Jain, Telephone interview, 19 Nov. 2006.

“because they just believe that it’s the right thing to do,” while “others may do it for a combination of those reasons.” Companies will progressively turn to CSR as employees, customers and government bodies continue to demand for businesses to maintain and be more open about their business practices.

Companies with similar green goals in mind may collaborate to create positive branding for all parties involved. A recent collaboration between Greenpeace and Coca-Cola demonstrates that organizations with similar goals in mind for a project involving CSR, can collaborate to extend the scope of their actions. To alleviate the egregious impact that Coca-Cola’s over nine million coolers and vending machines had on the global climate, both organizations worked together to create a more sustainable energy technology. Due to the overwhelming success of this project, Coca-Cola became more open to cobranding efforts and joint advertising campaigns. Likewise, CSR has opened the door for many full-service sustainability consulting firms like DOMANI, which helps to “engage multinational clients with stakeholders and NGOs for mutual benefit.” Firms like DOMANI understand that engaging in common green projects is a common denominator for companies to work together and utilize synergies to boost stakeholder value. They have helped firms achieve “new bottom line value through precise, market driven sustainability strategies.”

Along with the advent of new firms focused on helping companies engage in constructive CSR activities, the potential for job creation in the field of CSR has surged, as larger companies are beginning to add CSR departments or include CSR functions within particular divisions including communications, investor relations, finance and operations. CSR can be seen as an emerging profession due to the complexity of roles and responsibilities of a CSR practitioner. For companies in the emerging stage of establishing a CSR agenda, managers with experiences in socially responsible projects will be essential. Once considered a budgetary burden for many companies, CSR

The Crisis and Emancipation of the Modern Corporate Executive

now holds tremendous long-term potential for value creation.

Companies may decide to adopt the “bottom of the pyramid” strategy, based on the notion that low-income markets present excellent opportunities for investment by global corporations. This idea suggests that developing nations represent 4 billion people who live on less than $2 a day, on average. DuPont is a prime example of a company that has been able to generate value by participating in developing nation markets. In 2003, it launched its “BOP” pilot programs in Latin America. DuPont researchers screened ideas to meet the triple bottom line – each idea, “must make a positive contribution to society, be good for the environment, and be an attractive business opportunity.” DuPont brings together government officials and NGOs to help in product development to alleviate starvation. Such projects have included hybrid corn seeds that are more robust than conventional seeds, and an affordable product that fuses soy protein and chicken. The company hopes to expand its successful model to other developing regions of the world, and has demonstrated that, “it is possible to integrate sustainability into one’s business model as a driver of value creation.”

Engagement in CSR must not be commonplace and generic, but needs to have a unique angle to have a profound impact. In a June 2007 panel at Georgetown University entitled, “Principles, Profits, and Politics: The Rise of Corporate Social Responsibility,” former U.S. Senator Jim Talent stated that most corporations are already actively engaged in creating value beyond profit generation, but just need to communicate more clearly the specific ways that they are creating value for particular groups of stakeholders. Findings from a national survey by Fleishman-Hillard and the National Consumers League were released at the panel, showing that that a majority of Americans believe that the most important thing a company can do to be viewed as socially responsible is to “treat their employees well.” Hence, CSR does not have to entail huge donations of money or the establishment of multi-millionaire dollar centers of learning, to leave a lasting impression on its stakeholders and the public.

Business executives must assess the potential affects of CSR on overall value creation for their firm. Current and future drivers of corporate

---

57 Ibid.
social responsibility include: ethical consumerism, globalization and market forces, social awareness and government regulations. In his response to *The Economist* Survey of CSR published on January 22, 2005, Jeremy Moon, director of the Nottingham University Business School’s International Centre for Corporate Social Responsibility states that, “the proper business of business includes investment in the social conditions for its own success and in its legitimacy. This is a complement rather than an alternative to other vehicles for business success.”

**VIII. CSR and Public Policy**

Governments are beginning to view CSR as a cost-effective and efficient means to enhance sustainable development strategies. Social responsibility is now being seen as an integral component of national competitiveness strategies. During the 2003 e-conference “Public Policy for Corporate Social Responsibility,” hosted by the World Bank Institute Series on Corporate Responsibility, Accountability and Sustainable Competitiveness, scholars from around the world shared ideas on ways the public and private sectors can work in conjunction to promote CSR activities. Participants agreed that the public sector must work to create a business environment that promotes CSR, as well as implement the appropriate regulatory structures.

Government involvement in CSR must be transparent and pervasive. Governments around the world must clarify their expectations of business with regard to CSR, and develop ways to measure their responses to these expectations. The public sector may decide to include tax exemptions for companies that build social capital within their communities. Governments may also decide to encourage CSR within small and medium-sized enterprises, which lack the impetus to participate in such activities on their own.

While CSR will continue to be attractive at the national level due to its beneficial effects on a nation’s economic competitiveness, there will need to be a great emphasis on defining CSR at a local level. Since CSR is viewed differently around the world depending on cultural and historical conditions, approaches to CSR by different countries, “would depend on the orientation

---


of the government and on the level of business influence on policy.” More research and consideration needs to be placed on research on country-level agendas that affect the role of business in society and the redefinition of the sectors of public and private policies. Responsible business activity will inevitably create a favorable environment for conducting business and promoting social development both locally and nationally.

IX. Conclusion

In the text of the Bhagavad Gita lies the “crisis and emancipation” of the modern corporate executive. Like Prince Arjuna, today’s business leader has to overcome both an internal and external dilemma: whether to ceaselessly pursue the bottom line for the sole benefit of shareholders or whether to broaden the scope of the firm by balancing the relationships of all stakeholders who are impacted by the firm. The challenge for today’s business executive is to place individuals before profits, so that employees and customers can inevitably benefit shareholders. The Gita requests leaders to accept this challenge and be the sattvic frame of reference for ethics and morality that the corporate world has lacked for many years.

As the world moves forward to combat international issues during our global green revolution, companies are expected to be good corporate citizens in their communities not only by governments and shareholders, but by their customers and communities. For firms planning ahead, building partnerships and cultivating a culture of social entrepreneurship “both within the company and around the brand,” will inevitably have a positive impact on stakeholders. Creating a sustainable future will involve making conscious choices about business strategies, and a sustainable future is one in which a company’s social and ecological consequence will be of critical importance in the boardroom. In order for company executives to succeed in their CSR ventures, they must understand how the intersection of business and society is evolving, and must shape these changes in ways that deliver business value that is viable for the firm and for society.

References


Drive for Show, Putt for Dough: Rates of Return to Golf Skills, Events Played, and Age on the PGA Tour

John R. Watkins, Jr.¹
Davidson College

Abstract

The “winner take all” structure of professional golf, where there is substantial incentive to perform at one’s highest ability, invites investigation into the rates of return to golf skills on the PGA Tour. The biggest earner on tour in 2006, Tiger Woods, won $9,941,563 on the PGA Tour, while the 227th earner made just $65,494. Because professional golfers are thought to produce earnings by means of skill in each area of the game, I shall employ a production function to model the marginal effect of golf skills on earnings. I intend to determine if the advancement of golf technology in the last 12 years has affected the impact of each golf statistic on earnings. It turns out that the marginal effect of driving distance and driving accuracy is negligible, but players with exceptional ability in other areas of the game of golf are handsomely rewarded. Age also has a negligible marginal effect on earnings, while events played does have a noticeable effect. Greens in regulation, putting, and the short game are important in determining the magnitude of player’s earnings.

¹ John R. Watkins, Jr. is a third year student concentrating in economics and history at the Davidson College. He would like to thank Dr. Kelly Chaston for her guidance. John may be reached at jowatkins@davidson.edu.
I. Introduction

Professional golf attracts fans that pay to watch golf at the highest level. Unlike team sports such as baseball or football, however, tournament officials determine player compensation according to a structure of remuneration akin to a winner-take-all format. It is worth noting for this analysis that tournament purses vary in size according to the magnitude of the event. Events that are historically more prestigious tend to have a large television viewership and a strong field of players. These events reward each player who “makes the cut” with more money than smaller tournaments are able to do. To make the cut is to shoot a low enough score so that you are invited to play the final two rounds of the tournament, usually held on Saturday and Sunday. The champion of a major PGA Tour event, such as the Players Championship, receives more money for his performance in that event than the champion of a lesser-known event such as the John Deere Classic. Besides this convenient compensation scheme, the PGA Tour warrants analysis for another reason. In 2006, Tiger Woods was the leading money winner, making $9,941,563. Conversely, the 227th earner on tour in 2006 made just $64,494. I intend to determine the factors of earnings production on the PGA Tour which prompt this striking range between the most and least successful players on tour. In fact, $9,941,563 represents only a fraction of Tiger Woods’ income including his endorsements. A player who regularly finishes among the top 10 in PGA Tour events can expect lucrative endorsement deals, which only increases the incentive to perform at one’s peak. The rewards of outstanding performance on the PGA Tour exceed the sum of players’ tournament earnings.

In this study, I shall relate certain factors of golf production to money earnings on the PGA Tour. Golfers produce earnings, fame and recognition, and enjoyment by golf success. Of these, earnings is the dependent variable in this analysis because it is the only variable of the group which can be found in existing literature. The object of the study is to determine the skills with the greatest associated marginal effects. In other words, I hope to discover the skills in which an improvement is associated with the highest returns.

I shall model the factors affecting earnings production as a function of data relating to events, age, driving distance, driving accuracy, greens in regulation, putting, and sand saves. I shall use OLS regression to isolate the effect that each type of playing style produces. If the data demonstrate that performance in statistical categories helps to produce earnings on the PGA Tour, we can surmise the skills that distinguish Tiger Woods and other elite players on the PGA Tour from the players who earn five-digit salaries.
This analysis should provide us with knowledge about which skills are most valuable on the PGA Tour.

II. Data Analysis

I gathered the data from ESPN.com and PGATOUR.com. ESPN provided a table with basic golf statistics for players ranked from 1st to 227th in tour earnings. There were a few gaps in the data for players who play primarily on the European Tour and for a handful of PGA Tour members toward the bottom of the rankings, but I filled that part of the data out using PGATOUR.com. I failed to obtain the statistics of six players ranked above 227 on the PGA Tour. Observations of statistics were consistent between those sources for players whose statistics were provided on both sources. The consistency between the two sources affirms the accuracy of the data. The data sample I shall use for this analysis includes most of the players for whom ESPN included performance statistics. Beyond the 227th player on the money list, ESPN has only recorded a handful of players’ statistics. The data set represents the magnitude in variability in earnings on the PGA Tour. The players included in the data set have all played at least eight events in 2006. I have excluded players who participated in just a few events because it is unlikely their tour performance is an accurate reflection of PGA Tour performance as a whole. Players who participated in one or two events received a sponsor’s exemption in most cases, signifying they should not be considered regular tour players. It happens that the PGA Tour publishes a list of exempted players for the PGA Tour season, all of whom are included in the analysis. An exempted player is entitled to play in the events of his choosing. I considered including the Nationwide and LPGA Tours, but their data is more limited than the data on the PGA Tour.

The explanatory variables are natural choices for the things which they are meant to represent. At least one sports economist has employed each variable in his analysis; I found most of the variables I chose in multiple empirical models. A more thorough discussion of the existing economic literature about golf can be found in Section II, entitled “Model / Ex Ante Predictions.”

Earnings are a sufficient measure of success in professional golf competition. Although scoring average may be a more pure measure of golf skills, the metric of earnings captures the prestige of a tournament along with the performance of the observed golfer. Generally speaking, the size of the purse reflects the prestige of the tournament, because the biggest tournament purses attract the best golfers. Thus, players who perform well in prestigious events against a relatively strong field are rewarded accordingly.
Scoring average indicates a player’s proficiency at golf, but can vary with the difficulty of the courses he played or the strength of the field. A player who competes in tournaments held at the easiest courses on tour may have a lower scoring average than a better player who enters tournaments at the hardest venues. Moreover, a player may do very well against weaker fields, but perform poorly in major events against particularly strong fields. The earnings statistic captures the strength of the tournament field and removes the influence of course difficulty. Because everyone competes on the same course in a tournament, the earnings metric captures the quality of a player’s performance regardless of the difficulty of the course. Tournament purses are public information, thus individual player’s earnings for 2006 are readily available.

Age is a proxy for experience with, or exposure to, the game of golf, although it is hardly a perfect measure because players pick up the game at different stages of their life. Moreover, some players develop their ability more rapidly than others, and so it is likely that age benefits some players more than others. Events is quite simply the number of events in which a player competed during the 2006 season.

The rest of the explanatory variables vary in nature and in the degree to which they can be known. Putting performance on individual holes is a discrete statistic (a player can take 0, 1, 2, 3, etc. putts). Thus, putting average is a precise statistic of a player’s putts per hole. Driving distance, on the other hand, is vulnerable to measurement error and is only recorded on some holes. Driving accuracy, greens in regulation, and sand saves are all percent measurements of success rates. I have rounded the percentage frequency variables to one decimal place, except for means and standard deviations, which I have rounded to two decimal places.

The explanatory variables are, for the most part, good proxies for the part of golf which they are meant to represent. Sand save percentage\(^2\) is the most dubious explanatory variable in its ability to represent some larger aspect of a player’s ability. Sand save percentage is supposed to represent a player’s overall short game ability, in that it is the frequency with which a player is able to hit the ball out of the bunker and make the subsequent putt. If a player has failed to convert a sand save, he has required more than two shots to escape from the bunker and finish the hole. The short game is considered those shots within 90 or so yards of the green but not yet on the green. Unfortunately, a skilled short game player is not necessarily a strong player from the bunker, and vice-versa. Generally speaking, however, a good

bunker player tends to have similar ability in his overall short game. Another problem with the sand save percentage is that it incorporates a player’s putting ability having put the ball on the green from the bunker. A player may be skilled from the bunker, but be an abysmal putter. This player would struggle in sand save percentage, even though his bunker game is strong. Conversely, a great putter can have a high sand save percentage because he is able to simply put the ball anywhere on the green and make the putt. I have chosen the statistic over another short game statistic, scrambling. Scrambling is the percentage frequency with which a player is able to get “up and down,” that is, to hit the ball onto the green from within 90 yards and make the subsequent putt. I decided not use this statistic as a proxy for short game ability because a player who two-putts from just off the edge of the green has completed an “up and down.” In reality, he has done nothing more than two-putt, but he has just increased his scrambling statistic. The sand save metric precludes this two-putt from contributing to a player’s short-game metric.

Earnings range from $65,494 to $9,941,563, with a mean of $1,117,264 and a standard deviation of $117,391 (Table 1). The magnitude of the standard deviation suggests that a seemingly insignificant difference in skill creates an enormous discrepancy in compensation, although that is likely the product of a few very good players who dominate the purses at the most prestigious events. The most highly correlated variable with earnings is greens in regulation, suggesting the PGA Tour puts a premium on having the ability to hit good approach shots into the green.

The average event participation for tour players in 2006 was 24.66, with a standard deviation of 5.96. One player only played in eight events, while the most active player participated in 36. Remarkably, the correlation coefficient between events and earnings is only .04, suggesting greater participation does not at all produce more earnings.

The average age on tour is 36.21 years old, with a standard deviation of 6.59 years. This comes as no surprise, as golfers tend to lose their ability to compete at the high level of the PGA Tour by their mid-40s. It is also of note that players become eligible to play on the Champions Tour at 50, considerably reducing the incentive for older players to continue to try and compete on the PGA Tour. Members of the Champions Tour are excluded from this analysis. Not surprisingly, driving distance has a fairly negative correlation with age, having a Pearson Correlation Coefficient of -0.50. The youngest PGA Tour player is 23. A likely explanation for the high minimum in golf is the common practice of professional golfers playing in college before they try to earn membership to the Tour.

Average driving distance is 288.85 yards, with the minimum and
maximum being 265 and 319 yards, respectively. The standard deviation of driving distance is 8.56 yards.

PGA Tour players hit 63.90% of fairways, with a standard deviation of 5.46%. The most accurate driver of the golf ball hit 78.4% of fairways, while the most erratic driver only hit 49.8%.

Players hit slightly more greens in regulation than they did fairways, averaging 64.94% with a standard deviation of 2.79%. It is of note that the standard deviation of greens hit in regulation is just under half of the standard deviation of fairways hit in regulation, suggesting that players who tend to drive the ball more erratically off the tee still manage to hit greens as consistently as the straight drivers of the golf ball. In other words, players may use the fairway, the rough, or even the woods as a path to the hole, but every tour player tends to put the ball on the green in close to the same number of shots. The most accurate iron player (the player who hits the most greens in regulation), Zach Johnson, hit 74.1% of greens, while the least accurate iron player, Aaron Baddeley, only hit 56.9%. There is a Pearson correlation coefficient of 0.38 between driving distance and driving accuracy. The correlation coefficient is significant at $\alpha = .01$ and indicates that the most accurate drivers of the golf ball are also the more consistent iron players. The correlation between driving distance and driving accuracy supports the notion that players tend to have strong “long games” (i.e., they are either good at driving and approach shots or bad at driving and approach shots), rather than be simply strong drivers or strong iron players.

Players averaged 1.7802 putts per hole, which implies that the average player took well under the par number of putts on each hole. That this average is well under the expected or par putts per hole does not only indicate that tour players are good putters, but also that they hit approach shots close enough to the hole that they can take only one putt to finish the hole. The standard deviation of putts per hole is .0253, while the minimum and maximum observations of putts per hole are 1.712 and 1.851, respectively. Of note, Aaron Baddeley made up for his weak iron play with the lowest putting average on tour. Conversely, Zach Johnson had the very worst putting average. One explanation for the reversal of these two players is that, because Zach Johnson hits so many greens in regulation, he frequently faces long putts. Conversely, Aaron Baddeley hits very few greens in regulation but hits the ball to a short distance from the hole. Thus, he leaves himself many relatively short putts. Generally speaking, having a low average number of putts per hole indicates a player is either talented at positioning the ball near the hole on his approach shot so that his first putt is relatively short or he makes relatively longer putts more frequently than other
The average rate of sand save conversion was 48.66% with a standard deviation of 6.47%. The minimum and maximum sand save conversion rate is 16.7% and 63.8%. The sand save conversion rate is our best measure of short game proficiency. It has a large range, reflecting a large variation in short game ability across the tour. It has a correlation with earnings of 0.2433.

It may seem curious that I have excluded the variable “scoring average.” The existing literature calls for it, except for Patrick James Rishe in his 2001 article, entitled “Differing Rates of Return to Performance: A Comparison of the PGA and Senior Golf Tours.” I have excluded the variable because it does not reflect any one golf skill, but rather a player’s overall ability. While scoring average has immense explanatory power, it does not represent a specific golf skill. There are certain variables other than scoring average, such as cuts made, top 10 finishes, and wins, which are obviously correlated with earnings, but they would likely capture the mental side of golf. Of course, the mental aspect of the game could explain a good deal of variation in earnings, but the literature which I have used as a foundation for my choice in variables largely ignores the mental part of the game. Age is included to capture the change in one’s mental maturity that comes with experience.3

III. Literature Review

In the context of previous empirical work dealing with this subject, this analysis is in many ways a replication of the work Ronald L. Moy and Thomas Liaw did in their 1998 article, “Determinants of Professional Golf Earnings.” They included all of the variables I shall include in my analysis with the exception of age, a variable for which Gerald W. Scully’s 2002 article, “The Distribution of Performance and Earnings in a Prize Economy,” provided the basis. I selected the independent variable and dependent variables, with the exception of scoring average, based on the empirical work they did in 1998 on PGA Tour results from 1994. In 2000, Stephen Shmanske created a similar model to Moy and Liaw’s for data from 1998, so similar that my analysis could be considered a replication of his work as well. I shall try to determine if the advancement of golf technology has significantly affected player performance on the PGA Tour. Driving distance is the most likely statistic to have changed dramatically between 1994 and 2006. I intend to determine if the greater driving distance created by more advanced

technology has caused driving distance to have a more significant effect on tour winnings in 2006 than it did in 1994.

Each of the existing models employs some of the same variables, but a couple of the models use variables that, to the casual observer, would seem unconventional. In Patrick Rishe’s 2001 article “Differing Rates of Return to Performance: A Comparison of the PGA and Senior Golf Tours,” he introduces a few innovative variables, birdie conversions, bounce back, and scrambling among more conventional proxies for golf talent. Stephen Shmanske, Ronald L. Moy and Liaw, Gerald W. Scully, and Patrick James Rishe’s variable choices all provide scholarly authority with respect to the explanatory variables I have chosen to represent player skill (Table 1). Since their studies, however, technology has improved to the extent that there may be a noticeable change in the returns to driving statistics. I shall see if driving, both driving distance and driving accuracy, has had any noticeable effect on player performance overall. Driving distance is the most likely statistic to have changed dramatically since Moy and Liaw’s study in 1994. The introduction of graphite technology in golf clubs preceded their study, but golfers have enjoyed more and more advanced technology in the last 12 years. I hope to add to the empirical studies since Moy and Liaw’s investigation in 1998 by drawing something from each of the authors I mentioned.

The primary study on which my empirical model is based is Stephen Shmanske’s model of the relationship between skills and earnings, published in 2000. Shmanske examined both the PGA and LPGA Tours in order to determine why such a large gap exists between purses on those tours and the skills that had the greatest effect on earnings. Shmanske’s empirical model varies only in that he has selected a special variable to represent putting proficiency. He considered total putts a misleading measure because a player who hits few greens in regulation is able to play an extra approach shot from close range and set himself up with a shorter, more easily holed, first putt. His point is well taken, but Moy and Liaw’s results suggest that Shmanske underestimates the power of a statistic which averages a player’s putts per hole.

Both Shmanske and Moy and Liaw compared the PGA and LPGA tours by examining the relationship between skill and earnings. My empirical model is as much a replication of their empirical model as it is a replication of Shmanske’s model, except I shall exclude one variable Moy and Liaw include: scoring average. Since scoring average is the result of the level of skill a player exhibits in all the other statistical categories, I have left it out. Rishe is the authority behind my decision to leave out scoring average. Moy
and Liaw published their empirical work in 1998 based on PGA Tour results from 1994. Aside from the authority they give my variable choice, their results should come in handy because technology has changed significantly since 1994. They found ln(DDIS) and DACCUR to be statistically insignificant. ln(DDIS) has a p-value of .220, while DACCUR has a p-value of .350. Conversely, their greens in regulation, putting, and sand statistics were all significant at $\alpha = .01$, with significant large parameter estimates.\(^4\) If technology has affected the game, we might expect those variables to be significant in 2006. Contrasting the results of their 1994 data and my 2006 data could demonstrate a transformation in the relationship between different skills and earnings.

In “Differing Rates of Return to Performance: A Comparison of the PGA and Senior Golf Tours,” Patrick James Rishe sought to determine “whether the earnings gap between Professional Golf Association and Senior Tour golfers is due to differences in average skill levels or the rates of return to these skills.”\(^5\) He concluded that the earnings differential across tours is a consequence of different rates of return to performance. The relevant part of Rishe’s paper is his empirical model, which has driving accuracy, greens in regulation, sand saves, driving distance, birdie conversions, bounce back, and scrambling as explaining the variation in the natural log of $Y$, the earnings per event for a given golfer. Birdie conversions, bounce back, and scrambling, as far as previous literature is concerned, are all innovative explanatory variables. The birdie conversions statistic is the percentage of times a player, having reached the green in regulation, is able to convert a birdie or better. Rishe has intended bounce back as a measure of resiliency. He defined it as, “the percentage of times a player, after shooting over par on the previous hole, comes back to shoot under par on the next hole.” He expected that erratic players, who frequently make bogeys followed by birdies and consequently have high bounce back statistics, will finish on the lower end of the money list. He defined scrambling as the percentage of times a player can make a par or better having failed to reach the green in regulation.

Rishe’s bounce back ex ante hypothesis turned out to be significant at the .11 level. The rest of Rishe’s coefficients had their expected sign and were significant at the .05 level, with the exception of driving distance, which was only significant at the .11 level. Conspicuously absent from his


empirical model, however, is some measure of putting ability. This omission could cause bias in the coefficients if there is correlation between it and the included variables. There is likely a correlation between sand saves and putts and scrambling and putts. Because of this correlation, excluding a putting variable could be detrimental to the accuracy of the model.

Gerald W. Scully links prize money on the PGA Tour to scoring average in his paper entitled “The Distribution of Performance and Earnings in a Prize Economy.” Scully does not write purely about the determinants of golfer winnings, but he does dedicate a section of his paper to that topic. Scully takes issue with Shmanske’s model specification from 1992. Scully contends that one extra average yard of driving distance gains a player nothing in earnings. He argues that performance measures such as driving distance affect scoring average and scoring average affects tournament earnings. Scully’s article is of particular relevance to mine because he considered age a determinant of earnings. He believes that the difficulty of replicating a golf swing and the pressure of competing at a high level give an advantage to older player, i.e., players with more experience. He also notes that age becomes a detriment because strength and eye-hand coordination recede with time.

Shmanske, Moy and Liaw, Rishe, and Scully provide the theoretical framework for the relationship between golf skills and earnings. I shall incorporate to some degree each of their contributions to this study. Shmanske’s use of performance measures, along with Moy and Liaw’s older data of the same statistics, will provide the backbone of the model specification. Rishe has a number of proxies worth consideration, and Scully provides the authority to include age into the model.

IV. Theoretical Model Exposition
At least five economists have considered the effect of golf skills on golf earnings. Earnings on the PGA Tour can be modeled as a production function. Shmanske, Moy and Liaw created empirical models which had one variable acting as a proxy for each part of the game: driving, approach shots, short game, and putting. I shall include all of the variables which they employed, with the additional variable of age. Scully was the first to use that variable.

Production is typically conceived as a function of labor and capital. The production function is relevant to modeling earnings in professional golf because professional golfers produce earnings. Production of golf earnings is, for the most part, a function of labor. Just as some workers are more capable of producing a good in an ordinary production function, so too is
there variance in the abilities of golfers to produce earnings on the PGA Tour. We can include capital in the production function in the sense that a player’s physical attributes might represent the capital stock he possesses. There is no evidence in previous literature and little anecdotal evidence to suggest that any particular physical attribute produces earnings. Some observers of the game believe that being “too tall” is a disadvantage, but actual evidence of that claim is scanty. I have included one explanatory variable that could be considered a measure of human capital investment. As a player ages, we expect his level of human capital to increase in proportion to the amount of time he has spent learning how shoot low scores. At a certain point, a player risks loss in hand-eye coordination and physical strength. As it pertains to a production function, this could be conceived of as capital depreciation.

Each skill for which I have included a proxy variable has a detectable effect on earnings. Driving distance, driving accuracy, greens in regulation, and sand save percentage should all be positive, because driving distance, driving accuracy, greens in regulation, and sand saves should all have a positive effect on earnings. The ex ante hypothesis for PUTT, the putting statistic, is negative. Common sense suggests that a player’s earnings will decrease as his putting statistic increases. Since the best putters have the lowest average putts, we expect the player with the lowest putting average to have the highest earnings. The ex ante hypothesis has the potential to fall short in explaining the variation in earnings in the case of a player who hits relatively few greens in regulation and consequently has a low putting statistic. Such a player is usually just off the green and is able to give himself a short first putt. He is not necessarily a particularly good putter; rather, his putting statistic is lower as a consequence of having more short putts than a player who hits relatively more greens in regulation.

Once again, I shall use Ordinary Least Squares Regression in order to estimate the true value of each of my explanatory variables. This technique will give me the best (smallest variance), linear, unbiased, estimate. The following equation represents the earnings return to golf skills I expect.

$$EARN = a + b_1 EVENT + b_2 AGE + b_3 AGE^2 + b_4 DD + b_5 DA + b_6 GIR + b_7 PUTT + b_8 SS + e$$

The statistics represented are event, age, driving distance, driving accuracy, greens in regulation, putting, and sand saves, respectively. EVENT is the number of tournaments in which a player competed during the 2006 season, AGE is the age of the player, DD is the average driving distance a

---

player achieved during the season, DA is the percentage of times a player
hit the fairway on par 4s and 5s over the course of the season, GIR is the
percentage of the greens a player hit in no more than two less than par of a
given hole, PUTT is the putting average a player compiled over the course
of the year, and SS is the percentage frequency with which a player landed a
ball on the green from the bunker and holed out the subsequent putt.

Now I shall describe the marginal effects which I expect each
variable to have on earnings, beginning with EVENT (Table 2).

\[ \frac{\partial \text{EARN}}{\partial \text{EVENT}} = b_1; b_1 > 0. \]

\( \beta_1 \) is the marginal effect of the number of events in which a
player competes on his earnings. It is expected to be greater than zero,
corresponding to the notion that the more events in which a player competes,
the more earnings he can produce. As a player competes in more events he
increases the number of opportunities he has to earn. Moreover, competing
in more tournaments increases a player’s experience over the course of that
season. Consequently, we can expect a positive \( \beta_1 \) based on the notion that
players tend to “get the rust off” when they play in more events.

\[ \frac{\partial \text{EARN}}{\partial \text{AGE}} = b_2 + b_3 \text{age}; b_2 > 0; b_3 < 0. \]

It is difficult to develop a consistent ex ante hypothesis of earnings
with respect to age across an entire sample of tour players because there is a
good deal of variance in skill at any given age. Players improve at different
rates at each point in their careers. Some players improve by leaps and
bounds within a year of joining the tour, while others require 10 years on
tour before their first PGA Tour win. I have squared the AGE term because
I expect greater age to contribute to greater earnings to a point, but to lower
earnings after that point. The loss of strength and hand-eye coordination that
affects players after their peak causes decreasing earnings with increasing
age.\(^8\)

\[ \frac{\partial \text{EARN}}{\partial \text{DD}} = b_4; b_4 > 0. \]

\( \beta_4 \) is the marginal effect of driving distance on earnings. Driving
distance is expected to have a positive effect on earnings.\(^9\) As Shmanske

---

\(^8\) Scully, Gerald W. “The Distribution of Performance and Earnings in a Prize

\(^9\) Shmanske, Stephen. “Gender, Skill, and Earnings in Professional Golf.” *Journal
does, I shall use a linear functional form for it. Anecdotal evidence suggests, however, that as a player’s driving distance increases, he benefits less from one average extra yard of driving distance. An interview from Tiger Woods shows that longer hitters benefit less from a marginal increase in driving distance. When asked why he was not playing more aggressively by trying to hit the ball farther, Woods responded that he did not know what advantage that would allow him. He claimed that, even on the longest whole of the course, he had only a short distance to the whole and that he endangered his position for no apparent reward.10 Tiger is an exceptionally long driver of the golf ball, but his answer reflects a recurring sentiment among professional golfers. When Woods hits a 3-wood on the longest hole on a golf course (a long hole, to be sure), he must not benefit a great deal from extra distance. This anecdotal evidence suggests that the improvement in technology that has occurred since Moy and Liaw’s data were recorded in 1994 will not change the results Moy and Liaw found in their study. If driving distance is statistically significant in the correct direction, I expect it to have hardly any real effect on earnings.

\[ \frac{\partial EARN}{\partial DA} = b_5; b_5 > 0. \]

\( \beta_5 \) is the marginal effect of driving accuracy on earnings. I expect driving accuracy to have a positive effect on earnings.11 Greater driving accuracy can only improve a player’s earnings, because it means he has more shots from a more desirable position, the fairway.

\[ \frac{\partial EARN}{\partial GIR} = b_6; b_6 > 0. \]

\( \beta_6 \) is the marginal effect of greens in regulation on earnings. I expect greens in regulation to have a positive on earnings.12 Once again, I am uncertain about the proper functional form. Previous literature has chosen a linear form, but I believe that a player who hits greens with a relatively greater frequency benefits more from a one percentage point increase in greens in regulation than a player who hits greens with a relatively smaller frequency. Since the player who hits relatively more greens in regulation is more familiar with the situation, he has a higher likelihood of succeeding in

12 Ibid.
that situation. A more sound explanation requires an example. Let us take two players, one with a higher GIR and one with a lower GIR. The one with the higher GIR hits a green in regulation, as he is accustomed to doing. The one with the lower GIR fails to hit the GIR, and requires an extra shot to put the ball closer to the hole than the higher GIR player did. This situation is typical; the higher GIR player will be more accustomed to having to hit longer putts than the player with the lower GIR. When the lower GIR player does hit a green in regulation and faces the longer putt, he will not have refined the skill and will not succeed as often as the higher GIR player. To make matters worse for the lower GIR player, he requires a broader skill set to make the same score on a hole as the higher GIR player. Because he misses the green more frequently than the higher GIR player, he is forced to draw from a skill set the higher GIR player need not worry about as often.

\[
(7) \quad \frac{\partial EARN}{\partial PUTT} = b_7; b_7 < 0.
\]

\(b_7\) is the marginal effect of putting average on earnings. A lower putting average should increase earnings; putting average has a negative relationship with earnings. As demonstrated with Aaron Baddeley and Zach Johnson, the putting statistic reveals more than a player’s putting proficiency. Putting is inexorably linked with GIR. Generally speaking, we might expect a player with a higher GIR to have a higher average putting statistic than a player with a lower GIR. As previously discussed, this is because the player with the higher GIR makes fewer of his first putts, simply because they are longer, not necessarily because he is a worse putter. Interacting the terms would reflect their interrelatedness, but the literature does not interact them and we would learn more by being able to observe each variable’s marginal effect.

\[
(8) \quad \frac{\partial EARN}{\partial SS} = b_8; b_8 > 0.
\]

\(b_8\) is the marginal effect of sand save percentage on earnings. Certainly, increased frequency of sand saves will positively affect earnings. A player’s ability from the sand, and more broadly, from anywhere around the green, should produce returns to earnings at the same rate. That is, a one unit increase in sand save percentage is no more beneficial to a player with a low sand save percentage than a player with a high one, or vice versa.

---


I have provided an explanatory variable to represent each major skill in the game of golf in order to determine the competence of a player in that area. Economic literature suggests that each of these skills has an effect on production of earnings. I have included age because I believe that golf skills increase up to a certain age, where they actually begin to decrease. I expect events played, age, driving distance, driving accuracy, greens in regulation, putting average, and sand saves to have a statistically and economically significant marginal impact upon earnings.

V. Empirical Model and Estimation

The theoretical framework on which my original model was founded yielded an unimpressive adjusted $R^2$ and prompted me to reconsider the functional form of a number of variables. After revisiting the literature, I decided that the linear functional form fails to sufficiently represent returns to skill.

\[
\ln(\text{EARN}) = a + b_1 \ln(\text{EVNT}) + b_2 \text{AGE} + b_3 \text{AGE}^2 + b_4 \text{DDDA} + b_5 \ln(\text{GIR}) + b_6 \ln(\text{PUTT}) + b_7 \ln(\text{SS})
\] (9)

It is different from my original specification in that it has predominantly log-log forms, with an interaction term for driving distance and driving accuracy.

The Ordinary Least Squares Regression produced $t$-values for each parameter estimate (Table 3). One column of the table, labeled Significant?, indicates whether the hypothesis test for that variable was statistically significant. The results are based on one of the following hypothesis tests, depending on the expected sign. In other words, all hypothesis tests of individual significance were one-tailed and with respect to zero.

\[
H_0 : b \leq 0; \quad OR \quad H_0 : b \geq 0; \quad OR \quad H_0 : b < 0.
\]

(10)

The model represented above has an $R^2$ of 0.5048 and an adjusted $R^2$ of 0.4885. Naturally, the adjusted $R^2$ is somewhat lower because it is adjusted for the intercept term and seven independent variables I have included in the model. Even so, the coefficients of determination above are substantially higher than the $R^2$s that the first specification of the model yielded. In that specification, all functional forms but two were linear (Equation 1).


Moreover, the first specification did not call for an interaction term of DD and DA. It had an $R^2$ of 0.3926 and an adjusted $R^2$ of 0.3697.

The AIC is another tool that aids in comparison between models, especially those employing different specifications. It adjusts the RSS for the sample size and number of independent variables, allowing comparison between models with different specifications. It is 498.947 in the log-log model, which is substantially lower than 6709.459, the AIC that the first specification produced. The improved specification of the model explains the vast improvement between the original and final models. The F-statistic for the final model is 31.02, which exceeds the F-critical of 2.10. The F-critical has 7 d.f. in the numerator and $\infty$ d.f. in the denominator, based on 221 observations. The magnitude of the F-statistic indicates that we have evidence to reject the null that $b_1 + b_2 + \ldots + b_7 = 0$. Therefore, all of our variables jointly have an affect on ln(EARN).

The parameter estimates in the original model for driving distance and driving accuracy caused me to reconsider the specification of the model. Neither was significant, and both parameters had the opposite marginal affect from what I expected. When prompted to consider the variables further, however, I realized that they have little meaning when considered ceteris paribus. When a player steps up to the tee, he considers both driving distance and driving accuracy together, not as distinct entities. For example, when a player wishes to hit the ball relatively far, he has to compromise accuracy. When he is intent on hitting the fairway off the tee, he must compromise distance. The new parameter estimate, which consists of the interaction of driving distance and driving accuracy, is both statistically and economically insignificant for an $\alpha$ of .05. The parameter estimate for DDDA is .000076, meaning a one percent increase in DDDA produces a .000076 percent increase in earnings, i.e., a negligible sum. This could indicate that the tee shot is relatively unimportant, as long as the player puts himself in position to hit the green on his approach shot.

Initially failing the Ramsey RESET Test with a linear model prompted me to reconsider the functional form. The Ramsey RESET is a blunt instrument for detecting omitted variable bias or some other specification error. I believed the variables sufficiently represented every golf skill, so I concluded only a problem with functional forms could cause my specification error. Literature suggests that the linear functional form most accurately represents the true functional form of golf production, but economic theory of production suggests that logs are more appropriate. Having failed the Ramsey RESET Test, I resorted to a log-log functional form. The model employing the log-log form passed a modified Ramsey
RESET Test which excluded yhat3. Inclusion of yhat2 and yhat4 indicates that the model, should it pass, has some level of robustness. It turns out that the Ramsey RESET Test without yhat3 yielded an F-statistic of 0.38 which is not significant at any level of $\alpha$. The critical F with 2 d.f. for the numerator and $\infty$ d.f. for the denominator is 4.61. Because 0.38 < 4.61, we do not have evidence to reject the null hypothesis that F is significantly different from 0. The model failed the traditional Ramsey RESET Test with yhat2, yhat3, and yhat4 included. The inclusion of those variables created substantial multicollinearity and altered parameter estimates (Appendix 1).

VIFs for the parameter estimates were all fairly low with the exception of AGE and AGE$^2$. ln(EVENT), DDDA, ln(GIR), ln(PUTT), and ln(SS) all have VIFs less than 1.5, with ln(EVENT), ln(PUTT), and ln(SS), all within .1 of 1. AGE and AGE$^2$ each have VIFs of about 105, suggesting they are highly multicollinear with one another. Given AGE$^2$ is a function of AGE, we would naturally expect substantial multicollinearity between them. However, theory behind modeling AGE as a polynomial is strong enough such that we should be willing to accept the multicollinearity. Even so, it is distressing that the functional form we have selected for AGE has turned out to be empirically wrong. AGE and AGE$^2$ are significant in the wrong direction. It is possible that the polynomial form is not appropriate after all given the small variance in ages of PGA Tour players. For now, I shall leave AGE as a polynomial because it does not noticeably cause other variables to be biased or reduce the adjusted R$^2$ of the model.

As we might expect, the correlation between AGE and AGE$^2$ is strong, 0.995. The next highest simple correlation coefficient is 0.523, between ln(GIR) and DDDA. The correlation between these variables indirectly supports the hypothesis that DD and DA should be an interactive term. If DDDA is a measure of overall driving proficiency, we might expect it to be correlated with GIR, because both are measures of ability in what as known as the “long game.” The long game is generally defined as those shots taken from outside of 90 yards from the green. Because they are so highly correlated, we have some reason to believe that DDDA is a good measure of overall driving ability. The correlation coefficient between ln(PUTT) and ln(SS) is another one worthy of our attention. That correlation coefficient is -0.236. The explanation for this is that when a player successfully converts a save from the sand, he has taken only one putt. Consequently, his putting average will decrease. We might expect players who have a high sand save conversion percentage to have a relatively low putting average, because each time they convert a sand save, they only actually take one putt. None of these correlations was as pronounced with linear functional forms. Finally, the
correlation between DDDA and ln(EVENT) is -.204. This correlation is not high, but it is higher than almost every other correlation. There is no obvious reason for the correlation, but it is low enough such that we can dismiss it.

Driving accuracy was significant in the wrong direction when I tried to isolate its effect in the first specification. The significance in the wrong direction might be attributed to the unimportance of putting the ball in the fairway on the tee shot. It may be that players who are too cautious and play very timidly off the tee fail to put themselves in a good position to hit the green in regulation. The simple correlation statistic certainly supports this hypothesis. We might expect serious multicollinearity between driving accuracy and greens in regulation, because they are both measures of accuracy. They, however, have a simple correlation of only .38. Therefore, players who do hit the green in regulation are accurate enough to hit the fairway with some regularity, but hitting the fairway does not necessarily lead to hitting the green in regulation. The evidence suggests that hitting the fairway is not important to earnings success.

Before discussing the extent to which there is autocorrelation in the model, it is worth noting that I sorted the observations in alphabetical order by players’ last name. Our null and alternative hypothesis is:

\[
H_0 : r \leq 0; \\
H_a : r > 0.
\]  

The autocorrelation procedure yields a DW-d statistic of 1.98. Since 1.98 > 1.83, the critical DW-d statistic (k = 7, n > 100), we do not have evidence to reject the null hypothesis. We have no evidence of serial correlation in the model. Therefore, Ordinary Least Squares regression is still the most appropriate way to regress skills from earnings.

The results of the White Test indicate we do not have evidence for heteroskedasticity.

\[
H_0 : Homoskedasticity; \\
H_a : Heteroskedasticity.
\]

The White Test yields R^2 of .0896, .0917, .1593, respectively when \(X_1, X_1^2,\) and \(X_1X_2\) are used to model the error term. When we multiply these by the degrees of freedom, it is apparent that the chi-square critical value far exceeds each of our statistics. We fail to reject the null hypothesis at \(\alpha = .10\), even if we compare our chi-square statistics to a chi-square critical value with only 100 degrees of freedom.

Having passed the White Test, for which I did not have to specify a factor of proportionality, I proceeded to the Park Test. Only one variable,
\[
\ln(\text{PUTT}), \text{ yields a significant } t\text{-value in the Park Test. Since I passed the White Test, I do not believe the cost of weighting the results by } \ln(\text{PUTT}) \text{ is worth the correction I might be able to make to the error term.}^{17}
\]

In order to determine the marginal effect of each variable on earnings, we must consider the meaning of the variables and their functional forms. First, it is worth noting that I have excluded slope values in X-Y space, \( \frac{\partial Y}{\partial X} \), of each parameter, because the marginal effects keeping the original functional form (e.g., \( \frac{\partial Y}{\partial \ln X} \)), reveal more about the impact of an explanatory variable on the independent variable. All of the values I have provided as marginal effects are in the space of their functional forms. For example, the marginal effect of a one percent increase in PUTT is a -25.45138 percent increase in EARN. Put another way, the marginal effect of \( \ln(\text{PUTT}) \) on \( \ln(\text{EARN}) \) is -25.45138. The marginal affect of a percent change in X on a percent change in Y is simply \( \beta \) for all log-log forms. Let us take another example, \( \ln(\text{EARN}) \). The marginal impact of a one percent increase in the number of events played produces a .88608 percent increase in earnings (Table 3). Considering that a one percent increase could be thousands of dollars, this figure is economically significant statistic. The number of events only needs to be increased by one percent in order to produce a noticeable increase in earnings. \( \ln(\text{GIR}) \) and \( \ln(\text{SS}) \) are expected to produce a 12.96432 and 1.23497 percent increase in earnings. Age is a more complicated case to interpret because the partial derivative of age is represented in two terms. The marginal effect of age on earnings is,

\[
\frac{\partial Y}{\partial X} = .15676 + .00195 \times 2X
\]

At the mean of age, we expect a .298 percent increase in earnings for a one year increase in age.

Finally, the marginal effect of the interaction term DDDA is -0.00007649. We would expect a tiny number for the DDDA interaction term because average driving distance has a range of only 54 yards and DA is only a percentage. In any case, DDDA is not statistically significant, so we need not worry it is in the wrong direction.

For log-log forms, the elasticity for each individual variable is simply \( \beta \). Thus, elasticity is constant. That is, it is the same at the mean as everywhere else (Table 4). The elasticity for age at the mean is a bit

17 I did run a Weighted Least Squares Regression using \( \ln(\text{PUTT}) \) as the factor of proportionality. The parameter estimates were largely unaffected, leading me to believe there is negligible heteroskedasticity in the model.
challenging to calculate (Appendix 2).

The main lesson of actually running the OLS regression is that returns to golf skills on the PGA Tour is not demonstrated best by a linear relationship between golf skills and earnings. Rather, a much sounder specification employs a log-log relationship. Moreover, it seems that an interaction term between DD and DA resembles most closely the considerations of a professional golfer when he steps onto the tee and makes a decision about how he wants to approach the golf shot. The interaction of those terms is not statistically significant, but it does improve the overall explanatory power of the model. However reasonable our theoretical grounds, it seems that driving distance and driving accuracy are all but irrelevant on the PGA Tour. Within reason, a player need not worry too much about what he does off the tee, because it hardly has any impact at all upon his earnings.

VI. Conclusion

Given what I have found, I would like to do a study that includes events played, greens in regulation, a putting statistic, and a short game metric but excludes DDDA, AGE, and AGE2. Such an adjustment could increase the adjusted $R^2$, and more importantly, given this empirical analysis, it might be theoretically sound as well. It seems that players’ driving ability is so similar at the highest level of the game that a marginal increase in driving proficiency has a negligible effect on earnings. Moreover, whatever power a player loses as he gets older is either not important or is made up by that player’s experience. The statistical insignificance of DDDA leads one to believe that the change in golf technology, since Moy and Liaw’s data were collected in 1994, has not affected the outcome of golf tournaments. Such an outcome is hardly surprising. The average tour player drives the ball far enough that he can put himself in an excellent scoring position without the advantage of the few extra yards that new technology brings. With added distance comes the added risk of driving the ball into the rough, or worse. The cost associated with risk of driving the ball into the rough is greater than the reward of gaining a few extra yards. Moreover, players who are able to drive the ball farther than their opponents also tend to miss the fairway more frequently. The simple correlation coefficient between DD and DA is -.59; there is a strong, negative correlation between driving distance and driving accuracy. Just as Moy and Liaw found ln(DDIS) and DACCUR to be statistically insignificant in 1998, I found the interaction of those variables to be insignificant in 2006. We can confidently say that the most significant determinants of earnings on the PGA Tour are still greens in regulation and
Appendix

Functional Form

\[ \text{EARN} = a + b_1 \text{EVENT} + b_2 \text{AGE} + b_3 \text{AGE}^2 + b_4 \text{DDA} + b_5 \text{GIR} + b_6 \text{PUTT} + b_7 \text{SS} + e \]

Description of Data Set

The data have been taken from ESPN.com. The data are a cross-section of the sum of earnings in 2006 and observations of each variable listed below. I have drawn some of the lower earners’ data from PGATOUR.com. The observational unit is one PGA Tour player.

Variable Definitions

\( \ln(\text{EARN}) = \) The natural log of the observed \( i \)th player’s money earnings on the PGA Tour during 2006 season.

\( \ln(\text{EVENT}) = \) The number of tournaments in which the \( i \)th player has participated, including events in which he missed the cut to play all four days.

\( \text{AGE} = \) The age of the \( i \)th player.

\( \text{AGE}^2 = \) The squared age of the \( i \)th player.

\( \text{DDDA} = \) The interaction of the \( i \)th player’s average driving distance on par 4 and par 5 holes and the percentage frequency with which the \( i \)th player hits the fairway.

\( \ln(\text{GIR}) = \) The natural log of the percentage frequency with which the \( i \)th player green in regulation. A player hits a green in regulation when he hits the ball onto the green par for the hole minus two. If a player should hit a green in fewer shots than the number of shots that is considered a green in regulation, it is also considered a green in regulation.

\( \ln(\text{PUTT}) = \) Having hit the ball onto the green, the natural log of the average number of putts the \( i \)th player requires to finish a hole.

\( \ln(\text{SS}) = \) The natural log of the percentage frequency with which the \( i \)th player finishes a hole from the bunker in two shots or fewer.
### Tables

#### Table 1: Existing Models for Returns to Skill on PGA Tour

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moy and Liaw</td>
<td>(\ln(\text{EARN}) = a + b_1 \ln(\text{DIS}) + b_2 \text{DACCUR} + b_3 \text{GIR} + b_4 \ln(\text{PUTT}) + b_5 \text{SAND} + b_6 \text{SCORE} + e)</td>
</tr>
<tr>
<td>Shmankse</td>
<td>(Y_{98} = a + b_1 \text{TOTPUTT} + b_2 \text{EVENTS} + b_3 \text{EVENTS98} + b_4 \text{SCOREAVE} + b_5 \text{DRIVDIST} + b_6 \text{DRIVACC} + b_7 \text{TOTDRIV} + b_8 \text{GIR} + b_9 \text{PUTTPER} + b_{10} \text{PUTTPRED} + b_{11} \text{SANDSAVE} + b_{12} \text{WINPER} + e)</td>
</tr>
<tr>
<td>Scully</td>
<td>(\ln \text{PRIZE} = a + b_1 \ln \text{SA} + b_2 \ln \text{EVENTS} + b_3 \text{AGE} + e)</td>
</tr>
<tr>
<td>Rishe</td>
<td>(\ln Y = b_0 + b_1 \text{DA} + b_2 \text{GIR} + b_3 \text{SS} + b_4 \text{DD} + b_5 \text{BC} + b_6 \text{BB} + b_7 \text{SCRAM} + e)</td>
</tr>
</tbody>
</table>

#### Table 2: Ex Ante Hypotheses

<table>
<thead>
<tr>
<th>Explanatory Variable</th>
<th>Marginal Effect of X on EARN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Events</td>
<td>(\frac{\partial \text{EARN}}{\partial \text{EVNT}} = b_1; b_1 &gt; 0).</td>
</tr>
<tr>
<td>Age</td>
<td>(\frac{\partial \text{EARN}}{\partial \text{AGE}} = b_2 + 2b_3).</td>
</tr>
<tr>
<td>Driving Distance</td>
<td>(\frac{\partial \text{EARN}}{\partial \text{DD}} = b_4; b_4 &gt; 0).</td>
</tr>
<tr>
<td>Driving Accuracy</td>
<td>(\frac{\partial \text{EARN}}{\partial \text{DA}} = b_5; b_5 &gt; 0).</td>
</tr>
<tr>
<td>Greens in Regulation</td>
<td>(\frac{\partial \text{EARN}}{\partial \text{GIR}} = b_6; b_6 &gt; 0).</td>
</tr>
<tr>
<td>Putting Average</td>
<td>(\frac{\partial \text{EARN}}{\partial \text{PUTT}} = b_7; b_7 &lt; 0).</td>
</tr>
<tr>
<td>Sand Saves</td>
<td>(\frac{\partial \text{EARN}}{\partial \text{SS}} = b_8; b_8 &gt; 0).</td>
</tr>
</tbody>
</table>
### Table 3: Parameter Estimates and T-Scores

<table>
<thead>
<tr>
<th>Explanatory Variable Name</th>
<th>Parameter Estimate (Marginal Effect, except for age)</th>
<th>T-Score</th>
<th>Correct Direction?</th>
<th>Significant?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-29.10987</td>
<td>4.93</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>ln(EVENT)</td>
<td>0.88608</td>
<td>4.88</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>AGE</td>
<td>-0.15676</td>
<td>2.03</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>(\text{AGE}^2)</td>
<td>0.00195</td>
<td>1.86</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>DD</td>
<td>-0.00007649</td>
<td>1.67</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>ln(GIR)</td>
<td>12.96432</td>
<td>9.47</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>ln(PUTT)</td>
<td>-25.45138</td>
<td>7.02</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>ln(SS)</td>
<td>1.23497</td>
<td>3.52</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

### Table 4: Elasticity Calculations at the Mean

<table>
<thead>
<tr>
<th>Variable</th>
<th>Elasticity</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(EVENT)</td>
<td>0.88606</td>
</tr>
<tr>
<td>age</td>
<td>-0.56248</td>
</tr>
<tr>
<td>ln(DDDA)</td>
<td>-0.10469</td>
</tr>
<tr>
<td>ln(GIR)</td>
<td>1.43414</td>
</tr>
<tr>
<td>ln(PUTT)</td>
<td>1.08171</td>
</tr>
<tr>
<td>ln(SS)</td>
<td>1.08088</td>
</tr>
</tbody>
</table>

### Table 5: Simple Correlation Coefficients

<table>
<thead>
<tr>
<th>AGE</th>
<th>DD</th>
<th>DA</th>
<th>GIR</th>
<th>PUTT</th>
<th>SS</th>
<th>EVENT</th>
<th>EARN</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE</td>
<td>1</td>
<td>-0.49583</td>
<td>0.27781</td>
<td>-0.02197</td>
<td>0.04035</td>
<td>-0.04876</td>
<td>-0.13106</td>
</tr>
<tr>
<td>DD</td>
<td>-0.49583</td>
<td>1</td>
<td>-0.5998</td>
<td>0.14876</td>
<td>0.06366</td>
<td>-0.15149</td>
<td>0.08518</td>
</tr>
<tr>
<td>DA</td>
<td>0.27781</td>
<td>-0.5998</td>
<td>1</td>
<td>0.3839</td>
<td>-0.0006</td>
<td>0.00248</td>
<td>-0.17702</td>
</tr>
<tr>
<td>GIR</td>
<td>-0.02197</td>
<td>0.14876</td>
<td>0.3839</td>
<td>1</td>
<td>0.06611</td>
<td>-0.03842</td>
<td>-0.00519</td>
</tr>
<tr>
<td>PUTT</td>
<td>0.04035</td>
<td>0.06366</td>
<td>-0.0006</td>
<td>0.06611</td>
<td>1</td>
<td>-0.26356</td>
<td>-0.0111</td>
</tr>
<tr>
<td>SS</td>
<td>-0.04876</td>
<td>-0.15149</td>
<td>0.00248</td>
<td>-0.03842</td>
<td>-0.26356</td>
<td>1</td>
<td>0.07846</td>
</tr>
<tr>
<td>EVENT</td>
<td>-0.13106</td>
<td>0.08518</td>
<td>-0.17702</td>
<td>-0.00519</td>
<td>-0.0111</td>
<td>0.07846</td>
<td>1</td>
</tr>
</tbody>
</table>

### Table 6: Means and S.D.s

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEARN</td>
<td>13.46416</td>
<td>1.02777</td>
</tr>
<tr>
<td>IEVENT</td>
<td>3.16972</td>
<td>0.28402</td>
</tr>
<tr>
<td>AGE</td>
<td>36.21267</td>
<td>6.59373</td>
</tr>
<tr>
<td>AGE2</td>
<td>1355</td>
<td>483.7311</td>
</tr>
<tr>
<td>DDDA</td>
<td>18429</td>
<td>1323</td>
</tr>
<tr>
<td>IGIR</td>
<td>4.17256</td>
<td>0.04333</td>
</tr>
<tr>
<td>IPPUTT</td>
<td>0.57662</td>
<td>0.01422</td>
</tr>
<tr>
<td>ISS</td>
<td>3.87496</td>
<td>0.14682</td>
</tr>
<tr>
<td>wIPUTT</td>
<td>1.73528</td>
<td>0.04279</td>
</tr>
</tbody>
</table>
Mathematical Appendix

Appendix 1

\[
\text{EARN} = 896.160 - 23.573 * \text{EVENT} + 4.166 * \text{AGE} - 0.052 * \text{AGE}^2 + 0.002 * \text{DDDA} - 344.713 * \text{GIR} + 677.115 * \text{IPUTT} - 32.804 * \text{ISS} + 2.067 * \text{yhat}2 - 0.015 * \text{yhat}3 + 0 * \text{yhat}4.
\]

Appendix 2

\[
\frac{\partial \ln Y}{\partial X} = \frac{dY}{dX} = (b_1 + 2b_2 X) * Y
\]

\[
\text{Elasticity} = \frac{\partial Y}{\partial X} * \frac{X}{Y} = ((b_1 + 2b_2 X) * Y) * \frac{X}{Y} = (b_1 + 2b_2 X) * X
\]
References

Determinants of New York City Residential Rental Prices

Larissa Marco
Washington University in St. Louis

Abstract

Economists and financial analysts have invested a tremendous amount of energy in an attempt to explain housing sale and rental prices. Many different approaches have been taken to try to understand residential prices, taking into account qualities such as neighborhood characteristics and housing units’ physical characteristics. This paper considers the physical and demographic characteristics of New York City neighborhoods to better understand residential rental prices in this region. The empirical analysis reveals that premiums are charged for rental properties based on location in the following order: Staten Island, The Bronx, Queens, Brooklyn, and Manhattan, with Manhattan commanding the largest premium. High rental prices are also correlated with high median household incomes and high crime rates. Finally, increases in rental prices are correlated with increases in rent-regulated or rent-subsidized housing. At a certain point, however, the level of rent-regulated housing reaches its maximum and rental prices fall thereafter. Ultimately, this study offers a bundle of attributes that may be used to understand housing prices and establishes the attributes most relevant to the pricing equation. With this study, renters, landlords, and real estate developers can better understand the real estate market and make more informed decisions as players in the market.

1 Larissa Marco is a fourth year Economics student at Washington University in St. Louis. She would like to thank Robert Parks for his support, guidance, and dedication, all of which were instrumental to the writing of this paper. She would also like to thank her friends and family for their love and encouragement. Please direct any questions or comments to larissa.marco@gmail.com.
I. Introduction

The real estate market is a highly complex and challenging one to understand. “Housing is a multidimensional good differentiated into a bundle of attributes that vary in both quantity and quality.”2 Because each piece of property is unique, it is often difficult to identify the appropriate variables that will explain residential sale and rental prices. This paper’s objective is twofold: to ascertain the bundle of attributes that can be used to understand housing prices in New York City and to determine the attributes within this bundle that are most relevant to the pricing equation.

II. Literature Review

Economists and financial analysts have invested a tremendous amount of energy and effort in an attempt to explain residential sale and rental prices. One approach that has been used to model residential prices evaluates properties according to their neighborhood characteristics. These characteristics include quality of schooling systems, level of noise pollution, air quality, proximity to parks, proximity to bodies of water and quality of transportation systems. Other research has tried to explain residential prices on the basis of housing units’ physical characteristics, which include number of bedrooms, number of bathrooms, square footage and age. Additional surveys model residential prices as functions of apartment amenities such as indoor pools, gyms, and covered parking. Palmon, Smith and Sopranzetti (2004) offer an alternative approach in “Clustering in Real Estate Prices.”3 The authors look at transactions themselves as the units of analysis, investigating the impacts that listing price, closing price, number of days on the market and number of properties listed by a given real estate agent have on the resulting real estate prices.

A review of the available literature on real estate prices reveals the many lenses through which these prices can be studied. In a housing market such as New York City, the high wages and cultural attractions available in NYC translate into a particularly high demand for space in the region, making the wide variance in NYC residential prices all the more interesting. An analysis of these prices must therefore reflect New York City’s unique characteristics. Perhaps a more thorough analysis could come from a study that incorporates all sets of potentially relevant data; future research could

Determinants of New York City Residential Rental Prices

involve such an analysis. This paper, however, will be limited to a select set of physical and demographic attributes of New York City neighborhoods in the hopes of solving the mystery of NYC residential rental prices.

Compelling results from a study of the determinants of residential rental prices can have useful real-world implications. With knowledge of the variables relevant to the rental price equation, landlords can learn how to extract maximum rents from their tenants. A refined rental price equation can also help landlords and developers refine their estimates of the expected returns on their investments during their cost/benefit analysis of each residential project. From the consumer side, prospective tenants can break down rental asking prices into their relevant components and negotiate along these lines. Furthermore, pricing information can give economists and financial analysts a better understanding of the real estate market and its supply and demand functions.

III. Data

The data for this research were obtained from The Furman Center for Real Estate and Urban Policy: State of New York City’s Housing and Neighborhoods 2005, this publication represents the results of the NYU Furman Center’s annual collection of data on housing and demographics in New York City. The data are cross-sectional, drawing from the 59 community districts within the five boroughs of New York City. The unit of observation for this data is a community district.

The dependent variable is Median Monthly Rent, which was measured in 2005. Data on this variable were obtained from the New York City Housing and Vacancy Survey, excluding subsidized units. Hence, the rents are ‘gross’ rather than ‘net.’ Using median monthly rents relieves the model of potential bias; median rents are less influenced by the highest and lowest rent levels than are mean rents.

Variables

Table 1: Names of Variables

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Unit of Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mmr</td>
<td>Median Monthly Rent</td>
<td>(in 2005 dollars)</td>
</tr>
<tr>
<td>Brook</td>
<td>Brooklyn</td>
<td></td>
</tr>
<tr>
<td>crimerate</td>
<td>Felony Crime Rate</td>
<td>(per 1,000 residents)</td>
</tr>
<tr>
<td>Idr</td>
<td>Income Diversity Ratio</td>
<td></td>
</tr>
<tr>
<td>Manhatt</td>
<td>Manhattan</td>
<td></td>
</tr>
<tr>
<td>medhouinc</td>
<td>Median Household Income</td>
<td>(in 2004 dollars)</td>
</tr>
<tr>
<td>Bronx</td>
<td>Bronx</td>
<td></td>
</tr>
<tr>
<td>numhouse</td>
<td>Number of Housing Units</td>
<td></td>
</tr>
<tr>
<td>percentimig</td>
<td>% Immigrant Households</td>
<td></td>
</tr>
<tr>
<td>Percrent</td>
<td>% Rent-Regulated/Other subsidized</td>
<td></td>
</tr>
<tr>
<td>Pop</td>
<td>Population (2000)</td>
<td></td>
</tr>
<tr>
<td>Povrate</td>
<td>Poverty Rate</td>
<td></td>
</tr>
<tr>
<td>Queens</td>
<td>Queens</td>
<td></td>
</tr>
<tr>
<td>Racdiv</td>
<td>Racial Diversity Index</td>
<td></td>
</tr>
<tr>
<td>Rentalvac</td>
<td>Rental Vacancy Rate</td>
<td></td>
</tr>
<tr>
<td>Staten</td>
<td>Staten Island</td>
<td></td>
</tr>
</tbody>
</table>

Independent Variables for Monthly Rent Regression: Bronx, Brooklyn, Crime Rate, Income Diversity Ratio, Manhattan, Median Household Income, Number of Housing Units, Percentage of Immigrant Households, Percentage of Public Housing, Percentage of Rent-Regulated Housings, Population,
**Poverty Rate, Queens, Racial Diversity Index, Rental Vacancy Rate and Staten Island.**

*Bronx, brook, Manhattan, Queens, and Staten Island* are dummy variables to indicate the borough associated with each observation. Each dummy takes on a value of 1 when the observation is located in that borough. *Crime Rate* is the number of crimes per 1,000 people living in the area as reported by the New York City Police Department. The crime rate data reflect reports of burglary, larceny and motor vehicle theft, murder, rape, robbery, and assault. The rates are calculated for each community district through a population weighting formula.

*Income Diversity Ratio* represents the 80th percentile income divided by the 20th percentile income. A larger number for this variable indicates a larger range of incomes for a given community. *Median Household Income* is measured for all members of a household who are fifteen years of age and older measured in 2004 dollars.

*Number of Housing Units* in a given area includes houses, apartments, and any other spaces intended for occupancy as separate living quarters. *Percentage of Public Housing* is calculated by dividing the total number of public housing units by the number of housing units in the area. A public housing unit is a unit that is owned and maintained by the NYC Housing Authority. *Percentage of Immigrant Households* is calculated against total households in a given area. *Percentage of Rent-Regulated Housing* includes housing units that are rent-stabilized, rent-controlled, or city-owned. *Population* reflects estimates as reported by the United States Census Bureau.

*Poverty Rate* is calculated by measuring household income data against a yardstick of poverty, which is adjusted for the number of dependents in a household, number of members of household and age of head of household. *Racial Diversity Index* measures the probability of two heads of a household being of different race. A higher number indicates a neighborhood that is more racially diverse. *Rental Vacancy Rate* reflects the number of vacant rental units divided by the total number of rental units in the area.
Table 2: Expected Signs of Independent Variables

<table>
<thead>
<tr>
<th>Name</th>
<th>Expected Sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bronx</td>
<td>?</td>
</tr>
<tr>
<td>Brooklyn</td>
<td>?</td>
</tr>
<tr>
<td>Crime Rate</td>
<td>?</td>
</tr>
<tr>
<td>Income Diversity Ratio</td>
<td>?</td>
</tr>
<tr>
<td>Manhattan</td>
<td>?</td>
</tr>
<tr>
<td>Median Household Income</td>
<td>+</td>
</tr>
<tr>
<td>Number of Housing Units</td>
<td>-</td>
</tr>
<tr>
<td>Percentage of Immigrant Households</td>
<td>-</td>
</tr>
<tr>
<td>Percentage of Public Housing</td>
<td>-</td>
</tr>
<tr>
<td>Percentage of Rent-Regulated Housing</td>
<td>-</td>
</tr>
<tr>
<td>Population</td>
<td>+</td>
</tr>
<tr>
<td>Poverty Rate</td>
<td>-</td>
</tr>
<tr>
<td>Queens</td>
<td>?</td>
</tr>
<tr>
<td>Racial Diversity Index</td>
<td>-</td>
</tr>
<tr>
<td>Rental Vacancy Rate</td>
<td>-</td>
</tr>
<tr>
<td>Staten Island</td>
<td>?</td>
</tr>
</tbody>
</table>

IV. Explanation of Expected Signs

Prior to running the regressions, the direction of the effect of the boroughs is uncertain, although it is expected that the prices of land vary across locations even within a given city. The signs of Crime Rate and Income Diversity Ratio are also indeterminate. An increase in the crime rate may decrease the rental prices being asked and paid for housing units in that area because crime is undesirable. However, crime may be committed more in high-income areas because it is most lucrative in these areas.

Median Household Income and Population are expected to have positive signs. It can be assumed that an increase in the median household income of a community will drive rental prices up. Wealthy renters can not only afford more expensive properties but may also prefer them as luxuries and status symbols. Increases in population will also drive up rental prices because they increase demand.

---

It can be hypothesized that the *Number of Housing Units*, *Percentage of Immigrant Households*, *Percentage of Public Housing*, *Percentage of Rent-Regulated Housing*, *Poverty Rate*, and *Rental Vacancy Rate* will all have negative coefficients. Areas with a higher number of housing units will have lower rents according to the laws of supply and demand. It can be hypothesized that there is a negative relationship between *Percentage of Immigrant Households* and *Racial Diversity Index* and rental prices. According to David Harris, “[a]mong nonpoor whites, at least 70 percent of movers select new neighborhoods that are in nonpoor areas where no more than 30 percent of residents are black. The probability of moving to a neighborhood that is at least 60 percent black is almost zero for this group.”

High levels of public housing are typically undesirable. Since rent controls lower rents, increases in the number of rent-regulated units are expected to lower rents.

**Table 3: Descriptive Statistics as Common Sample of Variables**

<table>
<thead>
<tr>
<th>Borough</th>
<th>Mean</th>
<th>Median</th>
<th>Maximum</th>
<th>Minimum</th>
<th>Std. Dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Jarque-Bera</th>
<th>Probability</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRONX</td>
<td>0.189655</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.395452</td>
<td>1.583278</td>
<td>3.50677</td>
<td>24.85275</td>
<td>0.00003</td>
<td>58</td>
</tr>
<tr>
<td>BROOK</td>
<td>0.310345</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.466675</td>
<td>0.819892</td>
<td>1.672222</td>
<td>10.75872</td>
<td>0.004611</td>
<td>58</td>
</tr>
<tr>
<td>CRIMERATE</td>
<td>36.42241</td>
<td>29.4</td>
<td>226.9</td>
<td>13.1</td>
<td>31.37464</td>
<td>4.517984</td>
<td>25.85549</td>
<td>1459.72</td>
<td>0</td>
<td>58</td>
</tr>
<tr>
<td>HOMERATE</td>
<td>31.23103</td>
<td>28.35</td>
<td>73.7</td>
<td>2.2</td>
<td>18.50228</td>
<td>0.687601</td>
<td>2.741312</td>
<td>4.732077</td>
<td>0.093852</td>
<td>58</td>
</tr>
<tr>
<td>IDR</td>
<td>5.553448</td>
<td>5.4</td>
<td>8.3</td>
<td>3.5</td>
<td>1.062966</td>
<td>0.430242</td>
<td>2.809263</td>
<td>1.877312</td>
<td>0.391153</td>
<td>58</td>
</tr>
<tr>
<td>MANHATT</td>
<td>0.206997</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.408619</td>
<td>1.447136</td>
<td>3.094203</td>
<td>20.26541</td>
<td>0.00004</td>
<td>58</td>
</tr>
<tr>
<td>MEDIHOUINC</td>
<td>40796.03</td>
<td>39150</td>
<td>76010</td>
<td>15544</td>
<td>5880.7</td>
<td>0.529162</td>
<td>2.654835</td>
<td>2.994708</td>
<td>0.223721</td>
<td>58</td>
</tr>
<tr>
<td>MMR</td>
<td>936.3621</td>
<td>900</td>
<td>1640</td>
<td>600</td>
<td>241.5379</td>
<td>1.651107</td>
<td>5.193751</td>
<td>37.98315</td>
<td>0</td>
<td>58</td>
</tr>
<tr>
<td>NUMHOUSE</td>
<td>58720.97</td>
<td>52033.5</td>
<td>135015</td>
<td>40426</td>
<td>16404.27</td>
<td>2.106314</td>
<td>8.206526</td>
<td>108.3975</td>
<td>0</td>
<td>58</td>
</tr>
<tr>
<td>PERCENTIMIG</td>
<td>33.64138</td>
<td>36.6</td>
<td>60.9</td>
<td>10.9</td>
<td>13.97894</td>
<td>0.027752</td>
<td>1.91547</td>
<td>2.849994</td>
<td>0.249059</td>
<td>58</td>
</tr>
<tr>
<td>PERCPUBLIC</td>
<td>8.808621</td>
<td>4.4</td>
<td>45.6</td>
<td>0</td>
<td>10.16534</td>
<td>1.352254</td>
<td>4.724076</td>
<td>24.85977</td>
<td>0.000004</td>
<td>58</td>
</tr>
<tr>
<td>PERCENT</td>
<td>55.09826</td>
<td>58.35</td>
<td>94.6</td>
<td>6</td>
<td>20.64179</td>
<td>0.296197</td>
<td>2.427811</td>
<td>1.6393</td>
<td>0.440586</td>
<td>58</td>
</tr>
<tr>
<td>POP</td>
<td>135682.7</td>
<td>127793.5</td>
<td>242952</td>
<td>34420</td>
<td>46199.91</td>
<td>0.125836</td>
<td>2.642286</td>
<td>0.461279</td>
<td>0.794026</td>
<td>58</td>
</tr>
<tr>
<td>POVRATE</td>
<td>18.97799</td>
<td>15.65</td>
<td>45.6</td>
<td>4.4</td>
<td>10.54978</td>
<td>0.023351</td>
<td>3.389098</td>
<td>10.47807</td>
<td>0.000369</td>
<td>58</td>
</tr>
<tr>
<td>QUEENS</td>
<td>0.241379</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.431657</td>
<td>1.208734</td>
<td>2.461033</td>
<td>18.42573</td>
<td>0.000064</td>
<td>58</td>
</tr>
<tr>
<td>RACDIV</td>
<td>0.696952</td>
<td>0.705</td>
<td>0.96</td>
<td>0.26</td>
<td>0.184786</td>
<td>0.368448</td>
<td>2.196199</td>
<td>2.87369</td>
<td>0.237675</td>
<td>58</td>
</tr>
<tr>
<td>RENTALVAC</td>
<td>2.846552</td>
<td>2.7</td>
<td>6.3</td>
<td>0</td>
<td>1.614758</td>
<td>0.14625</td>
<td>2.322838</td>
<td>1.314921</td>
<td>0.518165</td>
<td>58</td>
</tr>
<tr>
<td>STATEN</td>
<td>0.051724</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.223404</td>
<td>4.048199</td>
<td>17.38788</td>
<td>658.6929</td>
<td>0</td>
<td>58</td>
</tr>
</tbody>
</table>

A brief review of the descriptive statistics reveals that Brooklyn has the greatest representation in this sample. Staten Island, on the other hand, is least represented. An ideal study would feature equal representation of community districts from each borough. Variance in representation may have implications for the results of this analysis. Thus, this inequality must be accepted as a limitation of this study.

V. Models

The goal of this paper is to identify the independent variables that

---

6 Harris, David R. “‘Property Values Drop When Blacks Move In Because…’: Racial and Socioeconomic Determinants of Neighborhood Desirability.” *American Sociological Review* (64.3, 1999) 461-479.
explain median monthly rents of New York City residential properties and to assess the robustness of these variables as explanatory instruments. An OLS method was used to regress median monthly rent of 2005 (Median Monthly Rent) against 15 independent variables.

### Table 4: Models

<table>
<thead>
<tr>
<th>OLS Regressions</th>
<th>Equation 1</th>
<th>Equation 2</th>
<th>Equation 3</th>
<th>Equation 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variable: Median Monthly Rent (Median Monthly Rent) for 2005</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>367.149</td>
<td>5.781196</td>
<td>6.099663</td>
<td>6.213655</td>
</tr>
<tr>
<td>BRONX</td>
<td>-101.561</td>
<td>-0.10323</td>
<td>-0.12406</td>
<td>-0.12395</td>
</tr>
<tr>
<td>BROOKLYN</td>
<td>-91.6838</td>
<td>-0.12274</td>
<td>-0.10937</td>
<td>-0.11455</td>
</tr>
<tr>
<td>STATEN ISLAND</td>
<td>-319.851</td>
<td>-0.27219</td>
<td>-0.30777</td>
<td>-0.31885</td>
</tr>
<tr>
<td>QUEENS</td>
<td>-97.8713</td>
<td>-0.12705</td>
<td>-0.11571</td>
<td>-0.09738</td>
</tr>
<tr>
<td>CRIME RATE</td>
<td>1.549827</td>
<td>0.001499</td>
<td>0.001481</td>
<td>0.001387</td>
</tr>
<tr>
<td>INCOME DIVERSITY RATIO</td>
<td>7.725554</td>
<td>0.01996</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MEDIAN HOUSEHOLD INCOME</td>
<td>0.010915</td>
<td>1.34E-05</td>
<td>1.14E-05</td>
<td>1.09E-05</td>
</tr>
<tr>
<td>NUMBER OF HOUSING UNITS</td>
<td>0.002816</td>
<td>6.34E-07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERCENTAGE OF IMMIGRANT HOUSEHOLDS</td>
<td>1.459037</td>
<td>0.002175</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERCENTAGE OF PUBLIC HOUSING</td>
<td>1.492536</td>
<td>-0.0002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERCENTAGE OF RENT-REGULATED HOUSING</td>
<td>-0.20112</td>
<td>0.011469</td>
<td>0.009567</td>
<td>0.009304</td>
</tr>
<tr>
<td>PERCENTAGE OF RENT-REGULATED HOUSING^2</td>
<td>-0.00011</td>
<td>-9.14E-05</td>
<td>-8.93E-05</td>
<td></td>
</tr>
<tr>
<td>POPULATION</td>
<td>-0.00093</td>
<td>-2.94E-07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>POVERTY RATE</td>
<td>-0.28494</td>
<td>0.001893</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RACIAL DIVERSITY INDEX</td>
<td>53.52097</td>
<td>0.155106</td>
<td>0.126587</td>
<td></td>
</tr>
<tr>
<td>RENTAL VACANCY RATE</td>
<td>-3.19515</td>
<td>-0.01131</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.875718</td>
<td>0.898504</td>
<td>0.881725</td>
<td>0.87396</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.831332</td>
<td>0.858896</td>
<td>0.859548</td>
<td>0.853382</td>
</tr>
<tr>
<td>S.E. of regression</td>
<td>99.19781</td>
<td>0.084959</td>
<td>0.084762</td>
<td>0.086603</td>
</tr>
<tr>
<td>Sum squared resid</td>
<td>413288.6</td>
<td>0.295938</td>
<td>0.344863</td>
<td>0.367503</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-339.571</td>
<td>70.76494</td>
<td>66.32808</td>
<td>64.48408</td>
</tr>
<tr>
<td>Durbin-Watson stat</td>
<td>2.12305</td>
<td>2.287611</td>
<td>2.038685</td>
<td>2.050118</td>
</tr>
<tr>
<td>Mean dependent var</td>
<td>936.3621</td>
<td>6.814738</td>
<td>6.814738</td>
<td>6.814738</td>
</tr>
<tr>
<td>S.D. dependent var</td>
<td>241.5379</td>
<td>0.226172</td>
<td>0.226172</td>
<td>0.226172</td>
</tr>
<tr>
<td>Akaike info criterion</td>
<td>12.26106</td>
<td>-1.85396</td>
<td>-1.94235</td>
<td>-1.91324</td>
</tr>
<tr>
<td>Schwarz criterion</td>
<td>12.82946</td>
<td>-1.25004</td>
<td>-1.5871</td>
<td>-1.59352</td>
</tr>
<tr>
<td>F-statistic</td>
<td>19.72942</td>
<td>22.68486</td>
<td>39.75932</td>
<td>42.47073</td>
</tr>
</tbody>
</table>
Prob(F-statistic) | 0 | 0 | 0 | 0 | 0
Ramsey RESET Test:
| F-statistic Probability | 0.010812 | 0.13001 | 0.065406 | **0.301108**
| Log likelihood Probability | 0.00141 | 0.04812 | 0.03211 | **0.227363**
| Italics = Significant at 1% level | Bold = Significant at 5% level |

VI. Discussion of Models

EQUATION 1

For the first regression, Manhattan is removed from the set of independent variables and used as a baseline due to the fact that the data involved five classifications in terms of geographical location and six descriptors. Research from “Why is Manhattan So Expensive? Regulation and the Rise in Housing Prices” by Glaeser, Gyourko and Saks (2005) supports the use of Manhattan as a baseline.7 According to the paper, data on condominium sales in the five boroughs suggest two distinct housing markets: the first market being that of Manhattan, which is associated with high values for Median Monthly Rent, and the second being that of the remaining boroughs, which are associated with lower values for Median Monthly Rent.

The data make this dual market distinction readily apparent:

<p>| Table 5: Median Monthly Rent by Borough |</p>
<table>
<thead>
<tr>
<th>Borough</th>
<th>Median Monthly Rent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bronx</td>
<td>768</td>
</tr>
<tr>
<td>Brooklyn</td>
<td>850</td>
</tr>
<tr>
<td>Manhattan</td>
<td><strong>1186</strong></td>
</tr>
<tr>
<td>Queens</td>
<td>950</td>
</tr>
<tr>
<td>Staten Island</td>
<td>850</td>
</tr>
</tbody>
</table>

EQUATION 2

The second regression incorporates modifications of the functional forms of the variables to determine whether additional independent variables are relevant to the Median Monthly Rent regression. Squaring all of the terms produces a statistically significant combination of Percentage of Rent-Regulated Housing and Percentage of Rent-Regulated Housing^2.

---

Because of this significance, the second equation includes a squared term for *Percentage of Rent-Regulated Housing*. The second regression is in log-linear form for two reasons. First, the literature addressing the subject of housing prices supports the use of log-linear form. Furthermore, only this form passed the Ramsey test for goodness of fit.

**EQUATION 3**

Having identified an appropriate model for the data in terms of the functional form of the dependent and independent variables, Equation 3 reflects the elimination of statistically insignificant variables *Income Diversity Ratio, Number of Housing Units, Percentage of Immigrant Households, Percentage of Public Housing, Population, Poverty Rate* and *Rental Vacancy Rate*. The elimination of these variables is supported by the WALD test, which tests for a statistically significant relationship between independent and dependent variables.

**EQUATION 4**

In the fourth regression, *Racial Diversity Index* is eliminated on account of the fact it is significant at the 10% level and the model is operating at the 95% confidence level. Further support for the elimination of this variable comes from the fact that a report entitled, “Does Federally Subsidized Housing Depress Neighborhood Values,” suggests that *Median Household Income* captures some of the effects *Racial Diversity Index* may have on New York City rental prices. The report states that public housing developments are commonly sited in neighborhoods comprising low-income families with the highest poverty rates and highest proportion of non-Hispanic Black and Hispanic populations. This being the case, the statistically significant inclusion of *Median Household Income* may be a sufficient stand-in for the other variables omitted for their statistical insignificance (namely, *Poverty Rate, Percentage of Public Housing and Racial Diversity Index*).

**VII. Interpretation of Results**

*Bronx, Brooklyn, Queens, and Staten Island* all command lower

---

Determination of New York City Residential Rental Prices

As the Median Household Income of residents in a particular community district increases, so too do the median monthly rents charged. Renters with greater amounts of disposable income can afford to pay higher rents and may even look for more expensive properties for the status and luxuries that these properties offer their occupants.

Together, Percentage of Rent-Regulated Housing and Percentage of Rent-Regulated Housing^2 produce a cap; initial increases in the percentage of properties that are rent-regulated increase the median monthly rent, but only to a point. Before reaching a peak, increases in rent-regulated properties limit the supply of non-subsidized housing units. This limitation puts upward pressure on the prices of remaining residential units. When enough rental units become regulated, increases in Percentage of Rent-Regulated Housing make the lower rents associated with subsidized housing the predominant rental values in the market.

The crime rate data used in this study reflect New York City Police Department reports of burglary, larceny and motor vehicle theft, murder, rape, robbery and assault. While data indicating the breakdown of reported crimes would be useful, the positive coefficient assigned to this variable links crime to rental areas. The association of high rental prices with high rates of crime suggests that more crimes are committed in these areas on account of their relative value as targets. Based on the positive relationship between rents and median household income, it can be argued that higher crime rates are also associated with higher income areas.

VIII. Implications

Judging from the results of the final model for this data, landlords can ascertain how to price their rental properties and renters can break down the prices they are charged to understand how certain attributes are valued. More specifically, premiums are charged based on location in the following order: Staten Island, The Bronx, Queens, Brooklyn, and Manhattan, with Manhattan commanding the largest premium. Prospective renters with higher median household incomes will typically pay higher median monthly rents and live in areas with higher crime rates. Districts where the percentage of subsidized housing is below 51.06% will see higher rents as subsidized housing increases to a certain point, after which rent-regulated prices will come to dominate.

With an understanding of the relationships between these variables and monthly rents, renters can make more informed decisions about where to reside according to their willingness to pay for housing and negotiate along the relevant dimensions of the pricing equation. An understanding of how to
break down rents into their relevant components also helps landlords and real estate developers approximate the expected returns on their investments with more accuracy. Knowing that the borough, median household income, crime rate and percentage of subsidized housing in a district influence monthly rents, resources can flow to investments that will best suit investors’ strategic desires and expectations. A matching of expectation and returns of this kind can ultimately reduce volatility in the real estate market as developers and landlords’ secure their ability to collect the rents necessary for subsistence and profitability.

With regard to economic efficiency, the results of this study make public the inputs of the residential pricing formula. With more information available to the general public, transactions costs should decrease, and consumers and sellers should be able to make more informed decisions and realize an increase in social welfare.

**IX. Areas of Future Research**

There are multiple aspects of this project that could be further developed in future studies. Among these is a review of data on the sales market to see whether the factors relevant to rental prices are similarly relevant to sales prices. Any discrepancies between these two sets of prices could be explored in an attempt to model the difference between the sale and rental markets. It would also be interesting to see whether this model would be transferable to rental prices of other cities.

It is also possible to look at the data used in this project over a period of time. Such a panel study could reveal the relative importance of these independent variables over time. Perhaps some understanding of their relationship with general changes and trends in the real estate market could be developed.

Most interesting, however, is the prospect of conducting a comprehensive study of New York City rental pricing that would incorporate every kind of attribute. From neighborhood characteristics to apartment amenities to individual unit dimensions and features, there is an incredibly wide range of variables that may be relevant to this pricing question. Including data that represent all of these variables would illustrate which bundles of attributes and which attributes within those bundles are most significant in explaining the rental prices of New York City residences. This comprehensive data set would then yield results that would paint the most detailed portrait of rental prices as a function of their many different determinants.
Appendix A

Discussion of Equation 1

In Equation 1, very few of the variables result in compelling p-values. However, the Ramsey RESET test, which tests the model for misspecification of functional form, suggests that the regression is correctly specified. The regression passes the White Test, indicating that the model does not suffer from heteroskedasticity, corroborating the assumption that the error term has a constant variance.

To determine whether borough (as an indicator for geographical location) is significant in explaining housing prices or rents empirically, a WALD Test is run. This test would reveal if the coefficients for the borough dummy variables were simultaneously equal to zero and could consequently be omitted. The poor results for the WALD test lend empirical support for the inclusion of these dummies.

Theory also supports the notion that borough is significant in explaining housing rental prices. Oftentimes residential units of seemingly identical qualities demand much higher rents in one region than in another. Cheshire and Sheppard (1998) discusses this notion of rental disparity in “Estimating the Demand for Housing, Land, and Neighborhood Characteristics.”9 Their analysis of residential prices operates on the premise that rental prices are largely location-dependent.

---

**Equation 1: Rental price regressed on all of the variables (with Manhattan as a baseline)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>367.149</td>
<td>310.7906</td>
<td>1.181339</td>
<td>0.2441</td>
</tr>
<tr>
<td>BRONX</td>
<td>-101.561</td>
<td>60.70606</td>
<td>-1.67299</td>
<td>0.1018</td>
</tr>
<tr>
<td>BROOK</td>
<td>-91.6838</td>
<td>60.67182</td>
<td>-1.51114</td>
<td>0.1382</td>
</tr>
<tr>
<td>STATEN</td>
<td>-319.851</td>
<td>92.32771</td>
<td>-3.4643</td>
<td><strong>0.0012</strong></td>
</tr>
<tr>
<td>QUEENS</td>
<td>-97.8713</td>
<td>80.984</td>
<td>-1.20853</td>
<td>0.2336</td>
</tr>
<tr>
<td>CRIMERATE</td>
<td>1.549827</td>
<td>0.668143</td>
<td>2.319603</td>
<td><strong>0.0253</strong></td>
</tr>
<tr>
<td>IDR</td>
<td>7.725554</td>
<td>20.94101</td>
<td>0.36892</td>
<td>0.714</td>
</tr>
<tr>
<td>MEDHOUINC</td>
<td>0.010915</td>
<td>0.002844</td>
<td>3.837305</td>
<td><strong>0.0004</strong></td>
</tr>
<tr>
<td>NUMHOUSE</td>
<td>0.002816</td>
<td>0.001961</td>
<td>1.436259</td>
<td>0.1583</td>
</tr>
<tr>
<td>PERCENTIMIG</td>
<td>1.459037</td>
<td>1.863143</td>
<td>0.783105</td>
<td>0.438</td>
</tr>
<tr>
<td>PERCPUBLIC</td>
<td>1.492536</td>
<td>2.353582</td>
<td>0.634155</td>
<td>0.5294</td>
</tr>
<tr>
<td>PERCRENT</td>
<td>-0.20112</td>
<td>1.061202</td>
<td>-0.18952</td>
<td>0.8506</td>
</tr>
<tr>
<td>POP</td>
<td>-0.00093</td>
<td>0.000705</td>
<td>-1.32145</td>
<td>0.1935</td>
</tr>
<tr>
<td>POVRATE</td>
<td>-0.28494</td>
<td>3.532521</td>
<td>-0.08066</td>
<td>0.9361</td>
</tr>
<tr>
<td>RACDIV</td>
<td>53.52097</td>
<td>89.40825</td>
<td>0.598613</td>
<td>0.5526</td>
</tr>
<tr>
<td>RENTALVAC</td>
<td>-3.19515</td>
<td>12.69394</td>
<td>-0.25171</td>
<td>0.8025</td>
</tr>
</tbody>
</table>

R-squared       | 0.875718    | Mean dependent var | 936.3621 |
Adjusted R-squared | 0.831332 | S.D. dependent var | 241.5379 |
S.E. of regression | 99.19781  | Akaike info criterion | 12.26106 |
Sum squared resid | 413288.6 | Schwarz criterion | 12.82946 |
Log likelihood    | -339.571   | F-statistic | 19.72942 |
Durbin-Watson stat | 2.12305  | Prob(F-statistic) | 0        |

*Bold* indicates statistical significance at the 5% level.
### WALD Test on Equation 1:

<table>
<thead>
<tr>
<th>Test</th>
<th>Value</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wald Test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equation: REGRESSION1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Null Hypothesis:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(2)=0</td>
<td>3.226243</td>
<td>0.021373</td>
</tr>
<tr>
<td>C(3)=0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(4)=0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C(5)=0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-statistic</td>
<td>3.226243</td>
<td>0.021373</td>
</tr>
<tr>
<td>Chi-square</td>
<td>12.90497</td>
<td>0.011750</td>
</tr>
</tbody>
</table>

### RAMSEY Test on Equation 1:

<table>
<thead>
<tr>
<th>Test</th>
<th>Value</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ramsey RESET Test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-statistic</td>
<td>5.080364</td>
<td>0.010812</td>
</tr>
<tr>
<td>Log likelihood ratio</td>
<td>13.12847</td>
<td>0.00141</td>
</tr>
</tbody>
</table>

### WHITE Test on Equation 1:

<table>
<thead>
<tr>
<th>Test</th>
<th>Value</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Heteroskedasticity Test</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F-statistic</td>
<td>1.062293</td>
<td>0.432327</td>
</tr>
<tr>
<td>Obs*R-squared</td>
<td>27.32767</td>
<td>0.392311</td>
</tr>
</tbody>
</table>
Discussion of Equation 2

Taking the log of Median Monthly Rent significantly reduces the skew-factor and improves the normality of the variable, as measured by the Jarque-Bera Test. Using log-linear form produces better R-squared, Akaike and Schwarz values, all of which measure the goodness of fit of the model. The equation also produces coefficient estimates of greater statistical significance and passes the Ramsey and White tests.

Normalizing Median Monthly Rent changed the distribution from Median Monthly Rent:

![Histogram of Median Monthly Rent](image1)

**Figure 2: Un-normalized Median Monthly Rent**

To log(Median Monthly Rent):

![Histogram of log(Median Monthly Rent)](image2)

**Figure: To log(Median Monthly Rent)**

<table>
<thead>
<tr>
<th>Series: MMR</th>
<th>Sample 1 59</th>
<th>Observations 59</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>934.0508</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>900.0000</td>
<td></td>
</tr>
<tr>
<td>Maximum</td>
<td>1640.0000</td>
<td></td>
</tr>
<tr>
<td>Minimum</td>
<td>600.0000</td>
<td></td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>240.1038</td>
<td></td>
</tr>
<tr>
<td>Skewness</td>
<td>1.677584</td>
<td></td>
</tr>
<tr>
<td>Kurtosis</td>
<td>5.292097</td>
<td></td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>40.58921</td>
<td>0.000000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Series: LOGMMR</th>
<th>Sample 1 59</th>
<th>Observations 59</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>6.812533</td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>6.802395</td>
<td></td>
</tr>
<tr>
<td>Maximum</td>
<td>7.402452</td>
<td></td>
</tr>
<tr>
<td>Minimum</td>
<td>6.396930</td>
<td></td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.224853</td>
<td></td>
</tr>
<tr>
<td>Skewness</td>
<td>1.110127</td>
<td></td>
</tr>
<tr>
<td>Kurtosis</td>
<td>4.120632</td>
<td></td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>15.20563</td>
<td>0.000499</td>
</tr>
</tbody>
</table>
**Equation 2**: Log(rental price) regressed on all of the variables with the addition of percrent^2 (with Manhattan as a baseline)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>5.781196</td>
<td>0.280541</td>
<td>20.60729</td>
<td>0</td>
</tr>
<tr>
<td>BRONX</td>
<td>-0.10323</td>
<td>0.052344</td>
<td>-1.97211</td>
<td>0.0554</td>
</tr>
<tr>
<td>BROOK</td>
<td>-0.12274</td>
<td>0.055243</td>
<td>-2.22179</td>
<td>0.0319</td>
</tr>
<tr>
<td>STATEN</td>
<td>-0.27219</td>
<td>0.079078</td>
<td>-3.442</td>
<td>0.0013</td>
</tr>
<tr>
<td>QUEENS</td>
<td>-0.12705</td>
<td>0.07171</td>
<td>-1.77166</td>
<td>0.0839</td>
</tr>
<tr>
<td>CRIMERATE</td>
<td>0.001499</td>
<td>0.000574</td>
<td>2.61272</td>
<td>0.0125</td>
</tr>
<tr>
<td>IDR</td>
<td>0.01996</td>
<td>0.01796</td>
<td>1.11130</td>
<td>0.2729</td>
</tr>
<tr>
<td>MEDHOUINC</td>
<td>1.34E-05</td>
<td>2.54E-06</td>
<td>5.28462</td>
<td>0</td>
</tr>
<tr>
<td>NUMHOUSE</td>
<td>6.34E-07</td>
<td>1.78E-06</td>
<td>0.357063</td>
<td>0.7229</td>
</tr>
<tr>
<td>PERCENTIMIG</td>
<td>0.002175</td>
<td>0.001597</td>
<td>1.362029</td>
<td>0.1806</td>
</tr>
<tr>
<td>PERCPUBLIC</td>
<td>-0.0002</td>
<td>0.002105</td>
<td>-0.0949</td>
<td>0.9249</td>
</tr>
<tr>
<td>PERCRENT</td>
<td>0.011469</td>
<td>0.003396</td>
<td>3.376777</td>
<td>0.0016</td>
</tr>
<tr>
<td>PERCRENT^2</td>
<td>-0.00011</td>
<td>3.13E-05</td>
<td>-3.59481</td>
<td>0.0009</td>
</tr>
<tr>
<td>POP</td>
<td>-2.94E-07</td>
<td>6.20E-07</td>
<td>-0.47364</td>
<td>0.6383</td>
</tr>
<tr>
<td>POVRATE</td>
<td>0.001893</td>
<td>0.003133</td>
<td>0.604191</td>
<td>0.549</td>
</tr>
<tr>
<td>RACDIV</td>
<td>0.155106</td>
<td>0.076635</td>
<td>2.023946</td>
<td>0.0495</td>
</tr>
<tr>
<td>RENTALVAC</td>
<td>-0.01131</td>
<td>0.01123</td>
<td>-1.0072</td>
<td>0.3197</td>
</tr>
</tbody>
</table>

R-squared: 0.898504  Mean dependent var: 6.814738
Adjusted R-squared: 0.858896  S.D. dependent var: 0.226172
S.E. of regression: 0.084959  Akaike info criterion: -1.85396
Sum squared resid: 0.295938  Schwarz criterion: -1.25004
Log likelihood: 70.76494  F-statistic: 22.68486
Durbin-Watson stat: 2.287611  Prob(F-statistic): 0

*Bold* indicates statistical significance at the 5% level.

*Bold and italicized* indicates statistical significance at the 10% level.
**RAMSEY Test on Equation 2:**

<table>
<thead>
<tr>
<th>Ramsey RESET Test:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>F-statistic</td>
<td>2.150691</td>
</tr>
<tr>
<td>Log likelihood ratio</td>
<td>6.068129</td>
</tr>
<tr>
<td>Probability of F-statistic</td>
<td>0.13001</td>
</tr>
<tr>
<td>Probability of Log likelihood ratio</td>
<td>0.04812</td>
</tr>
</tbody>
</table>

**WHITE Test on Equation 2:**

<table>
<thead>
<tr>
<th>White Heteroskedasticity Test:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>F-statistic</td>
<td>1.134003</td>
</tr>
<tr>
<td>Obs*R-squared</td>
<td>29.29569</td>
</tr>
<tr>
<td>Probability of F-statistic</td>
<td>0.36725</td>
</tr>
<tr>
<td>Probability of Obs*R-squared</td>
<td>0.346709</td>
</tr>
</tbody>
</table>

**WALD Test on Equation 2:**

<table>
<thead>
<tr>
<th>Wald Test:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Equation: Untitled</td>
<td></td>
</tr>
</tbody>
</table>

Null Hypothesis:  
C(7)=0  
C(9)=0  
C(10)=0  
C(11)=0  
C(14)=0  
C(15)=0  
C(17)=0

| F-statistic                              | 0.9683 |
| Probability                              | 0.466883|
| Chi-square                               | 6.778098|
| Probability of Chi-square                | 0.452344|

**Discussion of Equation 3**

After eliminating *Income Diversity Ratio, Number of Housing Units, Percentage of Immigrant Households, Percentage of Public Housing, Population, Poverty Rate and Rental Vacancy Rate*, all of the remaining variables become significant at the 5% level except for *Racial Diversity Index*, which is significant at the 10% level. The Ramsey Test results, however, strongly indicate misspecification of some type, which calls for testing for multicollinearity (as measured by the Variance Inflation Factor—VIF) and heteroskedasticity (as measured by the White Test). The
Determinants of New York City Residential Rental Prices

results from the VIFs identify *Percentage of Rent-Regulated Housing* and *Percentage of Rent-Regulated Housing^2* as being somewhat collinear, which indicates that there is a high degree of linear correlation between the two variables. On account of the fact that the two variables are highly statistically significant and that they are derived from the same data, the multicollinearity is understandable. The variables are therefore kept in the regression. The regression passes the White Test but fails the Ramsey Test, suggesting that the regression is incorrectly specified.

**Equation 3: Equation 2 having omitted those whose elimination was supported by WALD**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>6.099663</td>
<td>0.121152</td>
<td>50.34712</td>
<td>0.0000</td>
</tr>
<tr>
<td>BRONX</td>
<td>-0.12406</td>
<td>0.044692</td>
<td>-2.77594</td>
<td>0.0078</td>
</tr>
<tr>
<td>BROOK</td>
<td>-0.10937</td>
<td>0.04129</td>
<td>-2.64876</td>
<td>0.0109</td>
</tr>
<tr>
<td>STATEN</td>
<td>-0.30777</td>
<td>0.069253</td>
<td>-4.4441</td>
<td>0.0001</td>
</tr>
<tr>
<td>QUEENS</td>
<td>-0.11571</td>
<td>0.042964</td>
<td>-2.69317</td>
<td>0.0097</td>
</tr>
<tr>
<td>CRIMERATE</td>
<td>0.001481</td>
<td>0.000418</td>
<td>3.545903</td>
<td>0.0009</td>
</tr>
<tr>
<td>MEDHOUINC</td>
<td>1.14E-05</td>
<td>9.62E-07</td>
<td>11.83738</td>
<td>0.0013</td>
</tr>
<tr>
<td>PERCRENT</td>
<td>0.009567</td>
<td>0.002792</td>
<td>3.425969</td>
<td>0.0007</td>
</tr>
<tr>
<td>PERCRENT^2</td>
<td>-9.14E-05</td>
<td>2.53E-05</td>
<td>-3.61691</td>
<td>0.0822</td>
</tr>
<tr>
<td>RACDIV</td>
<td>0.126587</td>
<td>0.07131</td>
<td>1.775179</td>
<td>0.0475</td>
</tr>
</tbody>
</table>

R-squared 0.881725 Mean dependent var 6.814738
Adjusted R-squared 0.859548 S.D. dependent var 0.226172
S.E. of regression 0.084762 Akaike info criterion -1.94235
Sum squared resid 0.344863 Schwarz criterion -1.5871
Log likelihood 66.32808 F-statistic 39.75932
Durbin-Watson stat 2.038685 Prob(F-statistic) 0
**Discussion of Equation 4**

Eliminating *Racial Diversity Index* from the regression yields coefficient estimates for the independent variables that are all statistically significant at the 1% level save *Queens*, which is significant at the 3% level. The regression passes the Ramsey and White Tests, confirming that there is neither misspecification nor heteroskedasticity.

The regression is then tested for dummy interactions to eliminate the potential of each dummy having its own distinct slope and intercept. The resulting coefficient estimates are statistically insignificant and their omission from the regression is supported by the WALD Test. The interaction results confirm that there is no difference in slope between each borough—the only difference is in the intercept.
Equation 4: Equation 3 with the elimination of racdiv

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>6.213655</td>
<td>0.104966</td>
<td>59.19697</td>
<td>0</td>
</tr>
<tr>
<td>BRONX</td>
<td>-0.12395</td>
<td>0.045662</td>
<td>-2.71448</td>
<td>0.0091</td>
</tr>
<tr>
<td>BROOK</td>
<td>-0.11455</td>
<td>0.042081</td>
<td>-2.72215</td>
<td>0.009</td>
</tr>
<tr>
<td>STATEN</td>
<td>-0.31885</td>
<td>0.070469</td>
<td>-4.52465</td>
<td>0</td>
</tr>
<tr>
<td>QUEENS</td>
<td>-0.09738</td>
<td>0.042609</td>
<td>-2.28529</td>
<td>0.0267</td>
</tr>
<tr>
<td>CRIMERATE</td>
<td>0.001387</td>
<td>0.000423</td>
<td>3.276336</td>
<td>0.0019</td>
</tr>
<tr>
<td>MEDHOUINC</td>
<td>1.09E-05</td>
<td>9.46E-07</td>
<td>11.54418</td>
<td>0</td>
</tr>
<tr>
<td>PERCRENT</td>
<td>0.009304</td>
<td>0.002849</td>
<td>3.265603</td>
<td>0.002</td>
</tr>
<tr>
<td>PERCRENT^2</td>
<td>-8.93E-05</td>
<td>2.58E-05</td>
<td>-3.46195</td>
<td>0.0011</td>
</tr>
</tbody>
</table>

R-squared 0.87396  Mean dependent var 6.814738
Adjusted R-squared 0.853382  S.D. dependent var 0.226172
S.E. of regression 0.086603  Akaike info criterion -1.91324
Sum squared resid 0.367503  Schwarz criterion -1.59352
Log likelihood 64.48408  F-statistic 42.47073
Durbin-Watson stat 2.050118  Prob(F-statistic) 0

RAMSEY Test on Equation 4:

<table>
<thead>
<tr>
<th>Test</th>
<th>F-statistic</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ramsey RESET Test:</td>
<td>1.231469</td>
<td>0.301108</td>
</tr>
<tr>
<td>Log likelihood ratio</td>
<td>2.962411</td>
<td>0.227363</td>
</tr>
</tbody>
</table>

WHITE Test on Equation 4:

<table>
<thead>
<tr>
<th>Test</th>
<th>F-statistic</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>White Heteroskedasticity Test:</td>
<td>1.181531</td>
<td>0.325902</td>
</tr>
<tr>
<td>Obs*R-squared</td>
<td>12.77724</td>
<td>0.308134</td>
</tr>
</tbody>
</table>
Discussion of Crime Rate

Crime Rate’s distribution is highly skewed:

![Crime Rate distribution graph](image)

**Figure 5: Crime Rate Descriptive Statistics**

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Series</td>
<td>CRIMERATE</td>
</tr>
<tr>
<td>Sample</td>
<td>159</td>
</tr>
<tr>
<td>Observations</td>
<td>58</td>
</tr>
<tr>
<td>Mean</td>
<td>36.42241</td>
</tr>
<tr>
<td>Median</td>
<td>29.40000</td>
</tr>
<tr>
<td>Maximum</td>
<td>226.9000</td>
</tr>
<tr>
<td>Minimum</td>
<td>13.10000</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>31.37464</td>
</tr>
<tr>
<td>Skewness</td>
<td>4.517984</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>25.85549</td>
</tr>
<tr>
<td>Jarque-Bera</td>
<td>1459.720</td>
</tr>
<tr>
<td>Probability</td>
<td>0.000000</td>
</tr>
</tbody>
</table>

The mean is 30, while the maximum is 226.9. Three outliers may be driving the non-normality of this variable: 106.4, 124.9 and 226.9. These outliers are related to observations 1, 4, and 5, all of which fall within the borough of Manhattan. Omitting these three observations, however, produces insignificant results. Normalizing *Crime Rate* by using $\log(\text{Crime Rate})$ in place of *Crime Rate* does not improve the outcome either. As a result, *Crime Rate* remains in its original form.
References


Harris, David R. “‘Property Values Drop When Blacks Move In Because…’: Racial and Socioeconomic Determinants of Neighborhood Desirability.” American Sociological Review 64.3 (1999): 461-479.


Counterfeit Goods and Their Potential Financing of International Terrorism

Zachary A. Pollinger\textsuperscript{1}
Harvard University

Abstract

The production and sales of counterfeit products is an increasingly serious problem in the international trade market. Beyond issues of intellectual property theft and consumer safety, there is the notion that counterfeit sales directly fund terrorist organizations. I use a number of economic controls to analyze the effect of two proxies of annual counterfeit sales on two measures of international terrorism. My findings suggest inconclusive results as the effect of counterfeiting differs significantly between the two terrorism indexes.

\textsuperscript{1} Zachary Pollinger is a third year honors economics concentrator at Harvard University. He would like to thank Ms. Allison Rone for her excellent suggestions and guidance with the research and writing of this paper. He can be reached at zpolling@fas.harvard.edu.
I. Introduction

Murder, rape, and robbery: these are the crimes that routinely grab headlines and attract the attention of media outlets, government authorities, and concerned citizens. Far less publicized are intellectual property crimes which often go largely unnoticed despite representing one of the most serious and pervasive crime industries in the world. One significant source of intellectual property crime is the counterfeiting of branded or luxury goods. According to the World Customs Organization, the international sales of counterfeit goods comprise a $600 billion industry representing between 5-7% of total world trade.² Counterfeiting has become such a problem in the United States that Congress has even gone so far as to pass the “Stop Counterfeiting in Manufactured Goods Act” in 2006. In this paper, I will discuss the history and implications of the international criminal counterfeiting industry and focus on the larger issue of whether or not this market segment contributes to international terrorism. I will begin with a brief review of the many facets of this topic, continue to do an independent regression analysis of the correlation between counterfeiting and terrorism, and conclude by discussing the various difficulties and limitations in this area of study.

II. Economic Costs of Counterfeiting

Some types of counterfeit products are relatively common and widely-known in society. Luxury goods, for instance, are recognized as widely counterfeited due to the prevalence of fake designer handbags and wallets on sale in major US cities, and on display on the arms of many a faux-fashionista. Beyond luxury goods, counterfeits also account for a significant portion of pharmaceutical products, automotive and aviation parts, software, and video and music goods.³ While 10% of pharmaceutical drugs and luxury goods are thought to counterfeit products, as much as one-third of software and music CDs are said to be fake.⁴

The harmful effects of counterfeit products vary widely and generally relate to the category of good. Pharmaceutical products represent the most dangerous type of counterfeit good, as false medicines are produced

Counterfeit Goods

without regard to safety of production methods or composition. End products may contain accurate ingredients in incorrect proportions or may be based on entirely wrong chemical compositions. The counterfeiting of medication can be a deadly crime and routinely results in fatalities of innocent consumers. This problem is especially prevalent in developing countries where false medications may outnumber genuine products three to one. All told, the costs of counterfeit medications amount to $17 billion worth of losses for the legitimate pharmaceutical industry and an incalculable loss of life.5

Transportation parts and accessories are an equally dangerous market for counterfeiters to target. Fake auto parts, spare airplane pieces, and even spacecraft accessories may be of a far lower quality than parts subject to the exacting standards imposed by legitimate industries. In addition to the potential loss of life, fakes of these goods cost the automobile industry an estimated $12 billion, cost airline manufacturers sales of over half a million products, and lead to a loss of an additional $1 billion.6

Counterfeit consumer products such as software, music, and movies do not pose any perceptible public health or safety concern; however, they still do represent a serious problem for legitimate firms. Worldwide, 43% of user software is thought to be of illegal origin, with losses to legitimate firms in excess of $11 billion. Exactly one-third of music CDs are counterfeit products, with losses to the recording industry of up to $5 billion. Finally, an approximate $1.2 billion loss is incurred annually by motion picture companies as a result of the sales of bootleg DVDs.7

Still, the most visible goods in the counterfeit market are without doubt luxury products. European and American clothing and accessory brands are constant targets of counterfeiters seeking to capitalize on insatiable consumer demand for status goods. Most notably, French and Italian designers such as Louis Vuitton or Gucci face significant competition from nearly identical fakes. LVMH, the parent company of Louis Vuitton, is said to employ 40 lawyers and 250 independent investigators and to spend in excess of $20 million each year in efforts to fight increasing levels of

6 Ibid.
7 Ibid.
counterfeiting of their products.\(^8\)

While these societal and economic costs of counterfeit products are largely incontrovertible, one final effect of this crime industry is less definite: its support of international terrorism. Anti-counterfeiting organizations and luxury goods manufacturers are quick to suggest that counterfeit product revenues are directly funding terrorism. There is, however, only a small amount of hard data in support of this claim.

### III. Literature Review

A great deal of research has been commissioned on the links between counterfeiting and terrorism. Most studies, however, are qualitative as opposed to quantitative and focus largely on anecdotal evidence in drawing connections between the two. An example of one such statistic is Interpol’s 2004 seizure of $1.2 million worth of counterfeit German brake pads. Later investigations of the products revealed that their proceeds were earmarked for supporters of the Lebanese terror organization Hizbollah. According to Interpol Secretary General Ronald K. Noble, “Linking the Hizbollah to counterfeit brake parts shows not only the link between terrorist financing and intellectual property crime, but also how intellectual property crime is not a victimless one – the potential danger to the public from this sort of criminal activity is too serious for governments and law enforcement to ignore.”\(^9\)

More specious connections to Hizbollah have been reported with counterfeiters based in Los Angeles County. Authorities have found case-specific evidence of these connections (in the form of Hizbollah flags, tattoos, and pamphlets) in the homes and on the persons of numerous convicted counterfeiters. In another incident, a woman found to be a retailer of counterfeit cigarettes was arrested in an airport en route to Lebanon with $230,000 cash strapped to her body. While the reported reason for her trip was “vacation,” authorities believed her to be funneling money to Hizbollah


Some specific terrorist incidents also appear to have been funded by counterfeit operations. The FBI has compiled evidence that the terrorists who bombed the World Trade Center in 1993 financed their activities with counterfeit textile sales from a store located on Broadway in New York City. Three years later, the FBI confiscated 100,000 counterfeit products manufactured for sale at the summer Olympics. This operation funded an organization run by Sheik Omar Abdel Rahman, who was later sentenced to 240 years in prison for plotting to bomb historic landmarks in New York. Interpol has found that Chechen rebels fund their operations through the sale of pirated CDs and that paramilitary groups in Northern Ireland fund their operations by counterfeiting DVDs. Even Al Qaeda has been linked to the counterfeit industry through the sales of fake perfumes and shampoos. Finally, according to New York City police commissioner Raymond Kelly, the sale of pirated CDs was responsible for funding the 2004 bombing of a Madrid train – an incident that resulted in the deaths of 191 people.

The reason that terrorist organizations would fund themselves through counterfeiting is simple: fast, easy, plentiful cash. Counterfeiting and piracy are extremely easy industries to enter and would help terrorists maintain a certain level of anonymity. In addition, the profits from counterfeit sales significantly outweigh those of other illegal products. While the sales of cocaine might yield an entrepreneurial criminal a 100% profit margin, sales of pirated Windows software would earn a savvy counterfeiter profits of up to 900%.

IV. Research Question, Regression Model, and Hypothesis

While it is relatively easy to provide historical data citing an observational link between counterfeiting and terrorism, it is much less straightforward to analyze the aggregate effects of the counterfeit industry on international terrorist crimes in general. For this reason, this analysis will conduct an inquiry into the purported causal link between the two. The


method of this inquiry will be an OLS regression with a dependent variable that measures the number of international terrorism incidents. Linear predictors will include the factor of interest – a measure of international counterfeiting – as well as a number of control variables.

The regression can be modeled formally as:

\[ T_t = \beta C_t + \delta X_t + \epsilon_t \]

where \( T \) represents international terrorist incidents in year \( t \), \( C \) reflects the included measure of counterfeiting, and \( X \) is a matrix of additional control variables including a number of economic indicators. Log transformations of the terrorism and counterfeiting variables will also be used where appropriate to capture the elasticity between the two.

I predict that despite the many observational links between counterfeiting and terrorism, the empirical analysis will fail to provide any conclusive evidence of a causal link between the two. A lack of readily-available, reliable data makes it necessary to use proxies for the counterfeiting factor, and this shortcoming in measuring the variable of interest may inhibit the ability of the model to capture the true causal association.

V. Data Sources, Assumptions, and Potential Biases

Data on the dependent variable – number of international terrorist incidents – will come from two sources: the US Department of State and the RAND-MIPT Terrorism Incident Database, and will include values from years 1992-2005. Both sources are utilized to determine whether the subtle differences in data (mainly attributable to differences in definitions of the term “terrorist incident”) significantly alter the relationship between the dependent and independent factors of interest. In addition, the inclusion of both factors will hopefully hedge against potential biases associated with one particular agency’s motivation for reporting incidents. For instance, the RAND Database might be pressured to overreport incidents by donor organizations looking to attract attention to the problem of terrorism, and the Department of State might feel compelled to underreport incidents in an effort to show improvement resulting from new anti-terrorism initiatives.

Data on levels of international counterfeiting is, by comparison, much more difficult to procure. Since the counterfeiting industry is illegal and unregulated, precise statistics on number of counterfeits produced or sold simply do not exist. As a proxy for total number of counterfeits, this regression analysis takes into account the number of counterfeit products seized by the US Customs and Border Protection, as reported by end-of-year intellectual property rights seizure reports. This measure is, however, far
from ideal. Customs seizures would constitute an accurate proxy only if the border patrol captured an equal percentage of total fakes being imported into the country each year. Unfortunately, changes in policing strategies, import policies, and factors such as funding and political pressure likely cause fluctuations in the seizure rate, causing the proxy to be an imprecise measure. Ultimately, with no better measure of counterfeiting available, this regression must rest on the assumption of a stable capture rate.

As a check of the general validity of the seizures proxy, one additional factor is substituted for the independent variable of interest. This measure – referred to as world trade share – takes into account a very rough estimate of the aggregate yearly value of the counterfeit trade. The data for this control is generated from annual figures of aggregate world trade value (as reported by the World Trade Organization) and a 6% estimate of the trade share attributable to counterfeit goods (simply the mean of the 5-7% figure discussed earlier). While this control is also imprecise, it may provide interesting confirmation or refutation of the results obtained with the seizure statistics. World trade value figures are also more reliable than counterfeit values reported by companies suffering from intellectual property theft (such as LVMH) in that they avoid an almost certain overreporting bias.

Controls in the regression include GDP per capita and consumption per capita in the United States, since the seizures proxy is created from US Customs data. Figures on GDP and consumption per capita are also included for China (including Hong Kong) since China is the source of approximately 75% of annually traded bootleg products (see also: Appendix A, Figure 1). The two measures of economic controls are obtained from reports of the IMF. Finally, time-fixed effects are used to capture external variables that vary across years.

VI. General Findings and Regression Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Min</th>
<th>Max</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOS Terrorist Incidents</td>
<td>334.917</td>
<td>81.300</td>
<td>205</td>
<td>440</td>
<td>12</td>
</tr>
<tr>
<td>RAND Terrorist Incidents</td>
<td>251.571</td>
<td>83.910</td>
<td>106</td>
<td>395</td>
<td>14</td>
</tr>
<tr>
<td>US Customs Seizures</td>
<td>3844.214</td>
<td>2179.844</td>
<td>1253</td>
<td>8022</td>
<td>14</td>
</tr>
</tbody>
</table>
A first interesting feature of the data is the marked difference between the Department of State (DOS) and Rand/MIPT (RAND) terrorism statistics. As shown in Table 1, the DOS index categorizes far more incidents as terrorism. This distinction between the two is the first of many differences which will be discussed.

Figure 2 – Scatterplot of Seizures vs. DOS

Figure 3 – Scatterplot of Seizures vs. RAND

Figure 4 – Scatterplot of Seizures vs. lnDOS

Figure 5 – Scatterplot of Seizures vs. lnRAND

Figure 6 – Scatterplot of Seizures vs. DOS After Removing the 2004 Outlier
An initial regression of seizures on the two indexes yields interesting yet somewhat ambiguous results. Figures 2 through 5 show scatterplots of observed terrorism vs. seizures, along with simple regression lines and 95% confidence intervals of their slopes. Figures 2 and 4 suggest very weak or no association between number of customs seizures and terrorist incidents recorded by DOS. By contrast, Figures 3 and 5 seem to suggest that increases in seizures correlate to more terrorism, as recorded by RAND. Two additional important features of the data are revealed by the scatterplots. First, as seen in Figure 2, one large outlier (in the top right corner) seems to be significantly skewing the regression. This point represents the year 2004 in which recorded DOS incidents more than tripled from 208 to 651. Rather than an actual change in levels of terrorism, this discrepancy reflects a change in the Department of State’s definition of “terrorist incident” to a much broader measure. After removing the outlier, there is a negative correlation between the variables, as shown in Figure 6. To avoid an errors-in-variable measurement bias, the 2004 outlier is removed from all regressions. Second, in each case, there is strong evidence for a linear relationship between the variables, allowing for formal regressions ignoring the possibility of alternate functional forms. The one relationship most nearly parabolic – \( \ln\text{Rand} \) and \( \ln\text{Seizures} \) (as shown in Figure 5) – is in fact determined to be best modeled by a linear regression after testing a model inclusive of the regressor \( \ln\text{Seizures}^2 \). The coefficient on this term is insignificant, and the adjusted R\(^2\) of the regression decreases.

Results of the simple linear regression of seizures on the two indexes of terrorism are shown in Table 2. R\(^2\) values vary widely between the four regressions, ranging from .339 to .873. While it is encouraging to see relatively high values, it is worrisome that a vast inconsistency exists between the regressions on DOS incidents and those on RAND. It seems highly suspect that the included regressors would account for between 34% and 57% of the variation in terrorist incidents recorded by the DOS index, but a much greater 84% to 88% of the variation in incidents recorded by RAND.

---

Table 2 – Regression Statistics of Effects of Seizures on International Terrorist Incidents

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1) DOSt Terrorist Incidents</th>
<th>(2) RAND Terrorist Incidents</th>
<th>(3) ln DOS Terrorist Incidents</th>
<th>(4) ln RAND Terrorist Incidents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customs seizures</td>
<td>-.048 (.038)</td>
<td>.055*** (.014)</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>ln Customs Seizures</td>
<td>-----</td>
<td>-----</td>
<td>-.442 (.541)</td>
<td>.768*** (.275)</td>
</tr>
<tr>
<td>US per capita real GDP</td>
<td>96.079 (139.431)</td>
<td>-78.642* (45.317)</td>
<td>.272 (.457)</td>
<td>-.541*** (.209)</td>
</tr>
<tr>
<td>China per capita real GDP</td>
<td>-318.319 (383.843)</td>
<td>39.036 (129.210)</td>
<td>-1.268 (1.324)</td>
<td>.462 (.594)</td>
</tr>
<tr>
<td>US consumption per capita</td>
<td>-5.698 (33.007)</td>
<td>-21.091 (14.671)</td>
<td>.017 (.137)</td>
<td>-.113* (.060)</td>
</tr>
<tr>
<td>China consumption per capita</td>
<td>130.893 (88.365)</td>
<td>50.667 (43.340)</td>
<td>.817 (.534)</td>
<td>.416*** (.161)</td>
</tr>
</tbody>
</table>

Year indicators? Yes Yes Yes Yes
Country indicators? No No No No
R² .566 .842 .339 .873

Notes: Heteroskedastically-robust standard errors are shown in parentheses below each coefficient. Significance at the 99% level is denoted by β***, and at the 90% level by β*.

The coefficients on the main independent variable are somewhat surprising. In regression (1), customs seizures are shown to have a negative (though insignificant) correlation to DOS incidents. By contrast, seizures have a positive correlation to RAND incidents, which is significant at the 99% level. Despite this significance, the coefficient is quite small, and the finding relatively difficult to interpret in a real world sense. Much more relevant are the results of regressions (3) and (4). The insignificant coefficient on Seizures in regression (3) shows that a 1% increase in levels of seizures corresponds to a .4% reduction in international terrorist incidents (as recorded by DOS). The coefficient in regression (4) reveals a .7% increase in RAND incidents for every 1% increase in Customs seizures (again significant at the 99% level.) While it may technically be appropriate to claim that the
regression results evidence a positive relationship between counterfeiting and terrorism – considering that the coefficients on the seizures proxy are only significant when positive – such a claim is not entirely straightforward. Instead, the distinction should be made that counterfeiting seems to increase RAND-defined terrorism, while decreasing or having no significant effect on DOS-defined incidents. Ultimately, this ambiguity makes it more or less impossible to put forth a confident claim about the effect of counterfeiting on aggregate international terrorism.

Results of the control variables in the model are equally unconvincing. There is a great discrepancy between the signs of correlations of economic factors to terrorism when comparing the two indexes. Ultimately, most of these coefficients are insignificant, and it cannot be claimed that they are statistically significantly different from zero.

Table 3 – Regression Statistics of Effects of World Trade Share on International Terrorist Incidents

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1) DOS Terrorist Incidents</th>
<th>(2) RAND Terrorist Incidents</th>
<th>(3) ln DOS Terrorist Incidents</th>
<th>(4) ln RAND Terrorist Incidents</th>
</tr>
</thead>
<tbody>
<tr>
<td>World trade share</td>
<td>.001 (.001)</td>
<td>-.0008 (.0009)</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>In World trade share</td>
<td>-----</td>
<td>-----</td>
<td>2.075 (1.651)</td>
<td>-1.576* (.926)</td>
</tr>
<tr>
<td>Economic controls?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year indicators?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Country indicators?</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.496</td>
<td>.755</td>
<td>.439</td>
<td>.847</td>
</tr>
</tbody>
</table>

Notes: Heteroskedastically-robust standard errors are shown in parentheses below each coefficient. Significance at the 90% level is denoted by $\beta^*$. The inclusion of world trade share as a proxy for counterfeiting does very little to clear up the ambiguous results obtained using Table 2. Table 3 in fact shows opposite results – a negative correlation between counterfeiting (through the world trade share proxy) and RAND terrorism, and a positive correlation between counterfeiting and DOS terrorism. However, only regression (4) bears a significant coefficient, and evidences a 1.5% decrease
in RAND terrorism for a 1% increase in the estimated aggregate value of international counterfeit products. As in the regressions of Table 2, a very large portion of the variation in terrorist incidents is accounted for. The included regressors in Table 3 account for between 43% and 50% of the variation in DOS incidents, and between 75% and 85% of the variation in RAND incidents.

VII. Limitations to the Regression Model

A number of significant problems arise in this regression analysis that limit the extent to which results can believably be generalized to the real supply of international terrorism. As mentioned earlier, the lack of reliable data is the first serious shortcoming of the model. Data on the dependent factor – international terrorism – is surprisingly subjective and varies greatly depending on its source. US government data employs questionable methodology that leads to tripling of terrorism rates in response to changing definitions of the term “terrorist incident.” RAND index data is more conservative, using a generally stricter definition of terrorism, but may be subject to a number of oversights itself. In 2004, for instance, when the DOS statistics reported 295 terrorist incidents taking place in the war-stricken Kashmir region, RAND failed to find a single act of terrorism in the province. The distinction boils down to a debate over whether the insurgency in the region – manifested in recurring violence against unarmed non-combatants – is or is not international in scope. While DOS interprets the violence as contributing to Indian-Pakistani conflict and posing a threat to US and international security, RAND views the incidents as isolated within the region and inherently removed from international influence.\(^\text{15}\)

While an errors-in-variable bias calls into question the validity of data on the dependent factor, data used for the independent variable of counterfeiting is subject to error simply in that the construction of the model may not be entirely accurate. The principle assumption on which the regression rests – that US Customs captures an equal percentage of total counterfeits each year – is likely not accurate. Rather, unmeasured factors such as changes in policing strategies, pressures from anti-counterfeiting groups, and even sheer luck most likely have significant bearing on Customs seizures for a given year. The world trade share estimate is similarly flawed in that the oversimplification of assigning 6% of annual world trade to counterfeit products is not entirely likely. In either case, the independent

variable of interest is simply a proxy rather than a direct measure of the level of counterfeiting. Beyond the issues of data accuracy, this makes the analysis much less meaningful in terms of interpretation of the direct effect of the counterfeiting industry on terrorism.

In addition, one alternate theory might further confound the interpretation of the counterfeiting coefficients (under the seizures proxy). An argument could be made that counterfeits are produced in roughly equal quantities each year, and hence, increases in Customs seizures reduces the net sales and profits of the counterfeit industry. Under this theory, a negative—not positive—coefficient on Seizures would indicate the financing of terrorism through pirated goods. In reality, this hypothesis is probably not true. However, the most accurate measure of counterfeit goods sold may still somewhat benefit from the subtraction of captured products that will never make it to market.

A potentially large omitted variable bias constitutes the second main threat to the internal validity of the regression model. After the issues with data accuracy, the fact that few controls are included represents a main shortcoming of the analysis. Despite the fact that R² values are rather high, a number of important predictors are undoubtedly absent from the model. An example of one such important factor which can most likely never be fully quantified is social attitudes toward right of law and crime. This variable would undoubtedly influence counterfeits (on both the supply and demand sides) and also serve as a predictor of terrorism.

A final, more general limitation to this regression model rests on the question of whether a phenomenon such as terrorism can ever really be sufficiently explained. Deliberate acts of large-scale violence can partially be attributed to factors such as economic conditions, political strife, religion, and societal upbringing, but in another sense can only be linked to entirely irrational psychological aspects and behaviors unable to be explained by data compilation and OLS. Future studies might do better to avoid the psychological and focus instead on the purely economic variables that can be fully analyzed and accounted for. If there were a way to obtain detailed information on a factor such as the aggregate annual revenues of militant or terrorist organizations, then a regression might be able to link counterfeit profits to the funds backing terror groups. Unfortunately, with such limited perspective on the financial dealings of terrorists, no such analysis can currently be conducted. Very likely, the best solution lies somewhere between the realm of purely observational analysis and the type of study in this paper which searches for a link between the counterfeit trade and actual terrorism events.
VIII. Conclusion

The regression results of this analysis leave open much room for debate over the effect of the counterfeiting industry on international terrorism. Table 2 shows that if RAND figures are accurate, then counterfeits may indeed correlate with terrorism. If, however, DOS statistics are correct, then such a positive correlation may not exist.

This worrisome observed contradiction between indexes is unfortunately representative of data biases in general. Felix Salmon argues that the vast majority of counterfeiting statistics are fabricated, often in an attempt to pander to influential lobbying forces. According to Salmon, statistics reported by New York City Comptroller William C. Thompson Jr. make the far-fetched implication that one-third of all expenditures in the city go toward counterfeit goods on the black market. Regarding the value of annual counterfeits and the 5-7% figure concerning world trade share, Salmon reports that Peter Lowe, the assistant director of the International Chamber of Commerce’s Counterfeiting Intelligence Bureau, regards the elusive initial source of the statistics as “lost in the mists of time.”

Ultimately, very little is certain in assessing the magnitude of the counterfeiting industry and the proposed connection between the industry and international terrorism. To create a stronger model with which to analyze the effect of any policy changes, obtaining more accurate, in-depth data is obviously of the greatest importance. Levitt and Venkatesh used detailed financial information provided by the incarcerated leader of a Chicago gang in assessing the sources of revenue and expenditures for the organization. This information revealed a great deal about gang financing, drug trafficking, and inner-city violence. With the sophistication and high level of organization associated with some terrorist groups, it is likely that similarly detailed financial records for these factions do exist. Though it might be more difficult to obtain cooperation from an incarcerated terrorist leader than from a gang leader, it would nevertheless be enlightening to see accounting statements of the sources of funding for an organized terrorist group.

Finally, it should be debated what exactly is the appropriate

methodology for evaluating the effect of the counterfeiting industry on terrorism. This study utilizes OLS regression in an attempt to link changes in counterfeit sales to yearly fluctuations in terrorist incidents. One factor not accounted for in any of the regressions is the potential importance of lags. The implicit assumption behind Tables 2 and 3 is that counterfeiting revenues are earmarked for the funding of terrorist attacks in the same calendar year. However, there may be a number of reasons why funds earned in one year finance incidents only after a certain amount of time. For example, it may take a year to launder illegal revenue to make it available for use in executing an attack, or several years just to plan such an attack. If this is the case, then the original regression:

\[ T_t = \beta C_t + \delta X_t + \epsilon_t \]

should be modified to read:

\[ T_t = \beta C_{t-y} + \delta X_t + \epsilon_t \]

so that terrorist incidents in year t are regressed as a function of counterfeiting profits after a lag of y years.

More generally, it should be questioned whether OLS is an appropriate method of study at all. The supply of international terrorism is fueled by funding from a wide array of sources. While revenue derived from the sales of counterfeit products may fund some terrorist incidents, it is inconclusive whether fluctuations in the overall levels of bootleg product sales will be a significant enough predictor to establish causation in the international terrorism market. Therefore, it may be in the best interest of law enforcement officials to focus on accounting figures and financial information of terrorist groups in reviewing incidents on a case-by-case basis rather than at an aggregate level. Policymakers in turn may be best advised to combat intellectual property theft as a crime in its own right and hope for a secondary benefit of reductions to terrorism potentially financed by the counterfeit trade.
Appendix A

Figure 1 – Countries of Origin of Counterfeit Products (2005)

Source: US Customs and Border Protection, L.A. Strategic Trade Center, 11/3/05

Appendix B

Figure 7 – Counterfeiting and Terrorism Incidents over Time
Figure 8 – Top Counterfeit Commodities Seized (2005)

Appendix C

References


