STUDENT STAFF:

Fernanda Dobal - Editor-in-Chief
Avi Samarth - Managing Editor
Ariella Chivil - Associate Editor
Robert Gianchetti - Associate Editor
Juliana von Hoeven - Associate Editor
Amanda Johnson - Associate Editor
Claudia Ricciardi - Associate Editor
Carolyn Wu - Associate Editor
Antonio Vazquez - Associate Editor
Kendall Bass - Referee
Chris Lim - Referee
Melanie Redfearn - Referee
Celia Rodgers - Referee

ADVISORS:

Critsina Bicchieri (Director, PPE)
Sumantra Sen (Associate Director, PPE)
Rogers Smith (Professor, Political Science)
KC Tan (Professor, Philosophy)

Cover and Logo Design by Rich Tyson.

Website:
www.sas.upenn.edu/ppe/students/journal.html
05  Letter from the Editor

07  Default Options in Economic Games
    Nikhil Dhingra, Zach Gorn, Andrew Kener

25  Women Entrepreneurship in Tanzania
    Maya R. Perl-Kot

62  The Kyoto Protocol and Brazil
    Erica Reichert

116 Understanding the Strong Basis in Evidence Standard in Ricci vs. DeStefano
    Corey Singer
LETTER FROM THE EDITOR

I am proud to present the students, faculty and friends of the Philosophy, Politics & Economics major at the University of Pennsylvania with the sixth issue of SPICE: Student Perspectives on Institutions, Choices, and Ethics. In this edition, we bring you four articles from PPE graduates that reflect the interdisciplinary approach of the major.

This edition of the journal would not be possible without the work of many dedicated individuals. I would therefore like to thank the student editorial board who worked tirelessly to choose and edit the articles that are presented here. I would also like to thank Dr. Sumantra Sen, whose support makes the journal possible. The papers we included are only a few of several interesting and well-written papers we looked through, and I would like to thank everyone who submitted a paper for consideration. Finally, I would like to thank the staff of Penn Publishing and Monica McVey, who assisted in the publishing of this journal.

We hope that you enjoy these articles as much as we did and that they are successful in provoking thought and engaging you. Whether they are sparking an interest in a new topic for you, or a deeper look at something you are already familiar with, I hope that they stimulate thought and discussion.

Sincerely,

Fernanda Dobal
Editor-in-Chief
Introduction

By now, a large literature has used simple economic games, like the standard one-shot dictator game (see Forsythe et al., 1994), to analyze so-called social preferences. In this game, a “dictator” chooses how to divide an experimental endowment of money between herself and an anonymous, passive “receiver”. If players are selfish, then rational choice predicts they should keep the entire endowment for themselves. Numerous experiments, however, have shown that dictators give away on average more than 20% (Camerer, 2003), even when efforts are made to ensure recipient anonymity (thereby eliminating any concern for reputation, future interaction, or punishment). Various attempts have been made to reconcile this behavior with rational choice by assuming people have social preferences that in some way incorporate concern for others’ payoffs (see, e.g., Fehr and Schmidt, 1999).

Because these games use real-money endowments, typically without time pressures or complicated contexts, it would seem that a player’s actions reflect a clear and reasoned preference. In the present paper, we challenge this notion by presenting evidence of a strong “default pull” effect: Choices are affected by the presence of a subtle default option, even when the default itself is not chosen. We demonstrate this default pull effect in a standard dictator game with seemingly innocuous default options that appear to be a feature of the software presenting the game; i.e., the first in an apparently randomized list of allocation options is checked off by default. Such defaults obviously do not convey normative information about what to choose and, in fact, were often unnoticed by our participants. Although virtually every participant in our experiment denied that the default could have affected their choices, they were more inclined to give away the endowment (on average over 20%) when presented with a generous default option than when presented with a
selfish one. Unlike many previous findings showing that people stick with the default (see Thaler and Sunstein, 2008), the default pull effect sometimes involves participants staying with the default (e.g. when the default is sharing the endowment equally or keeping everything) and sometimes involves moving substantially away from the default (e.g. when the default involves giving everything away). Our findings add to our understanding of default biases by showing that even in a simple environment with clear incentives, people can be unconsciously affected by the presence of a default option.

Before moving to a description of the experiment, we briefly review findings from the literature on default biases, dictator games, and anchoring effects, and then differentiate our default pull from them.

Default Biases

Johnson et al. (1993) examines recent changes made in the insurance laws of Pennsylvania and New Jersey and illustrates how simple framing manipulations (changing default settings) can lead consumers to make inconsistent choices when evaluating the premiums and benefits associated with insurance purchases. Both states introduced a new insurance plan granting the option of lower insurance rates with limited suing rights (tort), but New Jersey set this new plan as the default while Pennsylvania featured the alternative full tort option (higher rates and full suing rights) as the default. 75% of Pennsylvanians retained the full right to sue, while only 20% of New Jersey purchasers actively chose to acquire the same rights. With limited tort as the default, Pennsylvanians could have paid over $200 million dollars less for auto insurance. In some cases, it is likely that defaults carry normative value; purchasing insurance can be a relatively complex task and people may often assume that the default is best for the standard, average consumer. Similarly, Thaler and Sunstein (2008) discuss the ramifications of default options in the context of the Medicare prescription drug program. Participants eligible for both, but who failed to sign up for a plan, were assigned to default plans set by the government. While they could freely
switch to a plan that would better suit their medical needs, many failed to do so.

Korobkin (1998) demonstrates a similar default effect with legal contract default rules. Law student participants provided advice to clients in hypothetical contract negotiation scenarios with the content of the legal default terms manipulated between experimental groups. The results suggest that the choice of legal default terms affects not only the final terms contracting parties will agree upon, but also their actual preferences for these terms. The study shows the effects of status quo and default biases, suggesting that if parties perceive legal default terms as part of the status quo, then their preferences will shift in favor of those terms more than they would if other terms were the legal defaults. Default options have the power not only to influence an agent’s actions, but also an agent’s underlying preferences.

Abadie and Gay (2004) examine default options within the opt-in/opt-out structure of organ donations. The study compares opt-in systems of informed consent, in which one must demonstrate explicit consent to being a donor, with opt-out systems of presumed consent, in which a person is classified as a potential donor when there is no explicit opposition. The survey results show that people retained the provided default option, but that switching the default option from explicit to presumed consent would substantially increase donation rates. Clearly, default options wield powerful influence over decisions of considerable significance.

**Dictator Games**

Dictator games are notoriously sensitive to descriptions and context, and studies often present the game in a context-laden environment. For example, when simply expanding the dictator’s choice set to allow taking from the receiver’s initial endowment, players rarely choose to transfer money and most choose to take from the receiver (Bardsley et al., 2005; List, 2007). An immediate concern is that such studies are susceptible to endowment effects, for the language used implies the assignment of a property right to one of the players (e.g. “you have 10 dollars,” or “player Y has 10 dollars”). Further, the experimental
design of such games often implies some normatively appropriate behavior that can have a significant influence on dictators’ decisions. Krupka and Weber (2008) showed that dictators behave more equitably when the endowment is initially split equally between the two players (“bully” game) than when the entire endowment is given initially to the dictator (“standard” game). The results show that dictators believe it is more socially acceptable to keep the majority of the initial endowment in the standard game, but it is not acceptable when framed in terms of taking from the receiver’s initial endowment, such as in the bully game. To address these concerns, we went through meticulous efforts to construct a purely abstract environment by allowing dictators to merely select the payoff double of their choosing, thereby eradicating the normative implications associated with initial endowments.

The positive amount consistently given away in dictator games makes it difficult to describe dictator preferences in terms of only outcomes. Some studies attempt to rationalize such systematic behavioral inconsistencies by attributing them to the fact that motivation for giving may not be adequately captured by monetary payoffs alone. Hayashi (2008) finds evidence that people who are not willing to create unequal distributions of wealth in their favor are nonetheless willing to permit such randomly generated inequality to persist. He finds that participants appear to exploit omission bias in a self-serving manner, failing to act only when omission is in their self-interest. Dana et al. (2006) finds that people have an “illusory preference” for fairness: They dislike appearing unfair to themselves or others when the causal link between their action and outcome is transparent; however, when transparency is relaxed they tend to act more selfishly, exploiting the uncertainty, or “moral wiggle room” between the distributional fairness of their action and its outcome. Another reason is that people may be compelled to give when in fact they truly prefer a purely self-interested outcome; and, vice-versa, they may be compelled to take more than they would normally if presented with the costless opportunity. In Dana (2006), people are responding to some consistent desire (don’t appear unfair) and thus, the behavior is able
to be rationalized. Our study involves an exploration of a broader hypothesis than many of the aforementioned studies: We think that people are affected by normatively irrelevant contextual factors, regardless of whether such factors are part of the individual’s payoff function.

**Default Pull Effect**

In this section, we aim to disentangle the default pull effect from traditional default, status quo, and anchoring-adjustment effects. Thaler and Sunstein (2008) emphasize that people tend to stick with the default, suggesting that default effects are defined discretely (e.g. opt-in or opt-out) rather than continuously. However, our results showed that participants sometimes retained the default and sometimes moved substantially away from it. Although almost every participant denied the possibility of being affected by the default when surveyed, the data strongly suggest otherwise.

It is important to note that the default pull is not anchoring in the classical sense. Tversky and Kahneman (1974) describe anchoring as a process by which “people make estimates by starting from an initial value that is adjusted to yield the final answer” (p. 1178), and anchors are typically used when people are uncertain about an objectively correct answer to a question. The dictator game, however, does not involve any decision under uncertainty because there is no “right” answer; rather, it is a matter of choice involving personal preference, and preferences are subjective and not typically thought of as being anchored (Simmons et al., 2009; see discussion on self-generated anchors). A dictator game outcome should be solely a function of individual subjective preference; anchors should be irrelevant. Many anchoring studies present a clearly arbitrary and uninformative anchor, such as the number generated by a wheel of fortune (Tversky and Kahneman, 1974) or the first three digits of the participant’s social security number (Ariely et al., 2003). Our experiment, however, features relevant default options that, unlike anchors, can be retained.

In general, default options often carry some normative value. They are typically
suggestions that are meant to actively influence choice with the individual’s best interest in mind (e.g. default funds, default insurance programs; see Thaler and Sunstein, 2008). Some anchors, particularly when relevant to the estimated value, may appear to carry normative value. Our default settings, on the other hand, carry no normative value. We avoided this in a few ways: First, we did not explicitly inform participants of the default options, nor were they actively endorsed or suggested in some way. Second, each default was preselected as a filled-in bubble on a simple computer interface, thereby appearing to be a feature of the game software. Third, the scrambling of the payoff doubles suggested that the top preselected option was randomly generated.

**Experiment**

**Participants**

Participants (n = 50) were all students at the University of Pennsylvania, all of who participated voluntarily in response to an online advertisement for paid decision experiments. All participants gave informed consent prior to the experiment. Four sessions were run, the first with 12 participants, the second with 18, the third with 8, and the fourth with 12 participants.

**Procedures**

We presented each participant with a series of four dictator-type games, three of which had different “default” conditions, or preselected options of wealth allocation between themselves and another anonymous player. Default options were designated as such by being the first and only “preselected” allocation on the list of thirteen wealth allocations. The fourth “no-default” condition simply had no preselected payoff double. We arrayed the payoff doubles along a straightforward, minimalist interface, with the checked off default allocation appearing at the top of the list. By doing this, we avoided the inevitable normative implications of context-laden settings. We ran four sessions with 50 participants in total. In every session, each participant received the four-game series in a different order.
Participants were situated at computer terminals for the entirety of the experiment and were given visual as well as oral instructions. No interaction other than via computer took place. First, participants were randomly assigned identification numbers that designated their role; odd-numbered participants were determined to be dictators (Player X), and even-numbered were receivers (Player Y), however specific role assignments were withheld until after the experiment. We used the strategy method to elicit choices; that is, participants made all decisions under the assumption that they were a dictator (Player X) and after all decisions were made, specific role assignments were revealed. Participants were then instructed that they would play a simple economic game on decision-making with another anonymous and random person in the room and that their payment for participation in the experiment would be determined in part by their and/or others’ decisions. Regardless of role assignment, players were informed that each would receive an equal base pay.

Each participant completed a series of four consecutive dictator games. The series was scrambled so that each player received a different ordering of the games. We ran all $4! = 24$ possible combinations of these four rounds. In each round, thirteen wealth allocations were presented as payoff doubles in a multiple choice format arrayed along the computer interface, and participants were instructed to ensure the allocation of their choosing was selected before clicking “submit” (see Fig. 1 below). The very first allocation at the top of the list was the default setting in three of the four conditions and was designated as such simply by being preselected, or bubbled-in, on the screen (e.g. 10,0 in Fig. 1). The other twelve allocations were all randomized. Ten of these twelve allocations consisted of all the other possible whole-number allocations adding up to ten, and the remaining two allocations were “distractors,” which did not add up to ten. The “distractors” are Pareto-dominated, and thus, should not be chosen, but we included them in an effort to promote independence (along with the scrambled orderings) so that each player would treat each round as a new choice. In the fourth “no-default” condition, there was no pre-selected
option on the screen, so participants simply chose from among the thirteen allocations.

Figure 1: Interface

After participants completed all four rounds, a screen appeared with their calculated monetary compensation (at an exchange rate of $0.25 per point). We revealed how dictator and receiver roles were determined, after which all participants filled out a brief survey with questions relevant to their performance during the experiment.

Results

Dictator Choices

Our results indicated that participants were affected by the default wealth allocation presented in the first game, and that initial effect persisted throughout all four rounds. Statistical analyses compared the mean allocation (and average payout) to the receiver (Y) across all four rounds in each condition. Condition A denotes the group of participants
whose first game had a (10,0) default, Condition B saw a (5,5) default, Condition C saw a (0,10) default, and Condition D saw no-default. The mean allocation to Player Y in Condition A was 1.44, in Condition B was 3.08, in Condition C was 3.58, and in Condition D was 1.56.

Figure 2 shows the mean allocation of dollars given to the receiver over all four rounds, according to the first-round default allocation. We found differences in the mean allocations between Condition A and Condition B that were significant (t=3.487, df=98, p<.001) and differences between Condition A and Condition C that were significant (t=4.512, df=102, p<.0001). Differences in the mean allocations between Condition B and Condition D were also significant (t=3.200, df=98, p=.0019), and differences between Condition C and Condition D were significant (t=4.218, df=98, p<.0001). However, differences in the mean allocations between Condition A and Condition D, and differences between Condition B and Condition C were not significant (t=.283, df=98, p=.777) and (t=.950,
df=98, p=.344), respectively. Furthermore, an ANOVA test returned p<.001, indicating that the means of all four conditions were not equal. Since there are statistically significant differences in the mean allocations to the receiver across all four rounds, and not in just the first round, we interpret this as evidence that the default allocation that participants saw in the first game affected their responses across all four rounds of the game. In addition, we observed that participants allocated fairly consistently over all four rounds, which can be seen visually in Figure 3.1.

Fluctuations in the flatness of each condition across all four rounds can essentially be attributed to random variation, as approximately one participant in each condition distinctly skews the data by varying his/her responses significantly more than the others. Figure 3.2 shows much smoother curves by removing just five data points: two from round 1 and 2 for the (0,10) default condition, and three from round 2 and 3 for the (10, 0), (5,5), and no-default conditions.
Strong positive correlations between round 1 and round 2 ($r = .76$), round 1 and round 3 ($r = .47$), and round 1 and round 4 ($r = .60$) further highlight the fact that there is a strong effect of the participants’ round 1 decision on subsequent rounds. Furthermore, from Figures 2 and 3.1-3.2 we gather that Conditions A and D have similar mean allocations across all four rounds, and Conditions B and C have similar mean allocations, with participants in Conditions B and C allocating nearly $2 more on average than those in Conditions A and D. Our results further demonstrate that participants in these conditions were not simply sticking with the default allocation. Rather, they actively shifted from the default, but still remained within close proximity in terms of wealth distribution. This is where our study parses default bias from our observed “default-pull” effect. Figures 4.1-4.4 help demonstrate this observation:
Figure 4.1

Condition A: (10,0)

Figure 4.2

Condition B: (5,5)

Figure 4.3

Condition C: (0,10)
As can be seen from the above graphs, in Conditions A and D the modal response was (10,0), whereas in Conditions B and C the modal response was (5,5). In Condition B, just as many participants chose (6,4) as they did (5,5), while in no other condition did more than one participant choose (6,4) in the first game. This shows that in Condition B, those who actively strayed from (5,5) nonetheless often remained very close to the default. In Condition C, almost half the participants chose (5,5) in the first game, demonstrating that while they would likely not stick with the “hyper-fair” (0,10) default (only one participant did), they still were much more fair than in Conditions A and D.

Survey Responses

General responses from the survey are shown in Figures 5.1 and 5.2. The most intriguing effect we found from the surveys was that only approximately 5% of participants believe the default allocation affected their choices in that particular game, even while the data indicate that the first default affected not only the first round game, but also the following rounds. This result of two people who believe the default affected that round can essentially be attributed to chance (note: the responses from one participant who answered “no” for the first question of “did you notice the default settings?” and “yes” to “did the default affect your decision in that particular game?” were discarded from the survey data
analysis because those answers are inconsistent). This means that participants were unwilling to admit that the default affected them and/or that the default had a subconscious influence on their decisions. However, participants were willing to acknowledge their degree of selfishness. When asked how generous they felt they were in their allocations on a five-point scale, only 11 percent of participants said “generous” or “very generous” (see figure 5.2). This is somewhat inconsistent with the data given that this group was fairly generous in comparison to other dictator studies, with a total mean allocation of $2.42 to Player Y across all conditions. Thus, people believe that they are, for the most part, selfish in their decisions, but are unwilling to acknowledge that their degree of selfishness might have been influenced by the default options.

![General Survey Results](image)

**Figure 5.1**

**Discussion**

Like many studies in the field of behavioral economics, we find results inconsistent with models that assume people’s preferences are defined solely over payoffs, outcomes, or consequences. But our findings go further. According to a purely consequentialist model, default rules should be theoretically extraneous features on choice; however, we find that presenting participants with default wealth allocations significantly affects decision-making
in dictator games. Our findings show participants giving over $3 on average to a passive and anonymous receiver in two of the default conditions. This begs the question: Can we interpret a significantly positive mean allocation—as much as $3.58 in the (0,10) default condition—as a simple preference for a fair or efficient outcome? Our findings suggest otherwise.

Studies with similar results inconsistent with preferences attempt to rationalize this irrational behavior, attributing such inconsistencies to self-serving biases, such as self-serving omission bias. Hayashi finds that people who do not actively choose unequal wealth distributions are nonetheless willing to let randomly generated inequality persist when it is self-serving. However, our results suggest that such inconsistencies are not able to be rationalized and are thus not necessarily attributable to self-serving biases, for we see effects when the default option represents both equal and “hyper-fair” allocations, such as (5,5) and (0,10). We find the greatest effects with the (0,10) and (5,5) default conditions, and these effects are in the other direction, with people giving significantly more on average. The self-serving bias theory would predict our default setting of (10,0), which is essentially a randomly generated purely self-interested and self-serving option, to result in more selfish behavior; it would also suggest that an initial “hyper-fair” allocation will essentially be disregarded and treated as a standard game, but we observe something different. Moreover, we find no significant differences between the (10,0) default condition and the no-default condition; participants essentially treated each of these games the same.

While we do not see the purely self-interested default of (10,0) affecting outcomes, we do see the default (5,5) and (0,10)—the more generous default allocations—affecting outcomes, but not in a manner that is consistent with the traditional views of default or status quo bias. Our results show that participants do not simply stick with the default or “do nothing” (i.e. leave initial wealth allocations as they were), but rather chose allocations with distributions in the general direction of the initial default allocation. Thus, it appears that the default has a somewhat mysterious and anomalous effect different from a simple
bias toward the present state; similar to a subtle anchoring effect, we call this the “default-pull” effect. Rather than defining default bias by the traditional bright-line test of either doing nothing or doing something, our “default-pull” concept defines the present state along a continuous variable, a continuum that shows how people can be biased toward a similar (but different) state as the default.

One way to view our overall results is that although people may not necessarily consciously care about the causal path that leads to an outcome, causal paths nonetheless affect their choices. As illustrated by our survey responses, of the people who noticed the default setting, about 95% said that it did not affect their choices. Most people did not admit to the possibility of the default setting having any influence on their decision. One participant said, “I noticed that there was a pre-selected option. This did not affect my decision because originally I was going to allocate some points to the other player already.” It is unclear whether this is a statement of true preferences or simply an ex-post rationalization of the “default-pull” effect. Another participant said, “Yes, in the first round of the game there was a default setting…it did have some leverage on my decision making, but not significantly.” Our data tells a different story, however, showing a significant inconsistency between the participants’ subjective responses (which could very well be cheap talk) and their actual actions. This discrepancy strikingly reveals the behavioral component of our study. One participant who was presented with the default condition of (0,10) in round 1 tellingly wrote, “Yes, [I did notice the default] and was trying to make choices so that the probability was close to 50-50,” which is an interesting answer that is consistent with our aggregate data for the (0,10) condition.

While our findings are robust, we hope that they serve to spur more theoretical and empirical research on the dynamics of human decision-making. For instance, a follow-up experiment could be conducted to see if participants, knowing that they are a Player Y, would pay a premium to have the Player X with whom they are matched see a default allocation of (5,5) or (0,10) in the first round. The participant would be shown the data from our
experiment, which indicates a Player Y should pay a premium of up to $2 to have Player X placed in (5,5) or (0,10) default conditions. Thus, there are many more questions to be explored, and some future research should be aimed at discovering the causes and origins of such a “default-pull” effect. And, of course, while disentangling behavioral anomalies is a pursuit in and of itself, the ultimate goal of our research is to enhance the predictive power of economic models.

Some critics argue that selfish preferences do not necessarily mean maximizing monetary payoffs, for some may derive utility from other-regarding behavior. However, this assertion assumes that people will show consistent preferences for some particular outcome, which is not always the case. Ariely et al. (2003) find that people’s initial preferences are not well-defined, even arbitrary, but become articulated in the process of making a decision (Ariely et al., 2003). Participants’ absolute valuations of goods are arbitrary, but relative valuations of different amounts of the good appear orderly, as if supported by demand curves derived from fundamental preferences. This phenomenon they coin “coherent arbitrariness” is consistent with the large body of literature on revealed and constructed preferences, which says that initial preferences or valuations are malleable but then become precisely defined relative to the initial arbitrary point.

Bibliography


Introduction

Traditionally, economic development had a clear negative correlation with rates of entrepreneurship and self-employment in societies: because of a variety of structural and socio-demographic changes associated with economic development, business ownership in the advanced nations of the West is much scarcer than in the developing world. Shifts in sector composition that are at the core of economic development were accompanied by modifications in production, employment and consumption, including a steep increase in the number of available formal jobs (Wennekers et al, 2009) that decrease the need to open independent enterprises as a last-resort occupational solution.

Recent evidence, however, has suggested another trend: a U-shaped relationship between economic development and entrepreneurship, which indicates a return to the activity at the pinnacle of the economic growth spectrum. Since the late 1970’s, several (though not all) developing nations have exhibited a reemergence of entrepreneurship (Wennekers et al, 2009). In general, entrepreneurship in developed nations is opportunity-driven, and related to pull factors: a desire for self-fulfillment, an unmet market need or an ambition for status and freedom (Perunović, 2005). By contrast, entrepreneurs in poorer nations are driven to business out of necessity: a lack of alternatives and insufficient income push individuals to start their own businesses.

Both forms of entrepreneurship have been proven to contribute, albeit in different ways, to economic performance. The former is growth-oriented and linked to innovation,
while the latter assists with poverty alleviation by bringing much needed income and employment to places where no alternatives exist. Women entrepreneurs, as a manifestation of the social and economic phenomenon and a specific research area in its own right, contribute even more to economic performance. In developed and developing economies alike, they have been proven to improve familiar conditions through their increased earnings, giving back to their communities, hiring other women as employees, thus taking them away from unemployment, and empowering their female relatives and acquaintances (Justo et al, 2006).

As mounting evidence has confirmed the macro U-shaped relationship—a resurgence of entrepreneurship as economic growth progresses—one hypothesis is that the same relationship holds internally, within developing nations, such that individuals from lower socioeconomic status pursue businesses because of necessity, while members of richer classes form ventures because of opportunistic and personal motivations. In light of the divergence between push-factors affecting necessity-driven entrepreneurs and pull-factors influencing opportunity-driven ones, the question that remains concerns women, and where they fall along the push-pull pendulum in the developing world context.

To answer this question, and examine the connection between socioeconomic background and the reason for becoming entrepreneurs and the characteristic of businesses, I conducted a case-study analysis of women entrepreneurs from Tanzania and undertook theoretical inquiry into the literature regarding micro and macro trends in women entrepreneurship. Tanzania, a developing country in east Africa has over 40 million residents and is one of the world’s poorest nations, 188th out of 210 according to 2009 World Bank estimates (WDI, 2009), but has a small upper-middle class residing in urban areas. In August 2009, I interviewed 7 women business owners in Tanzania from diverse economic, social and demographic backgrounds. I then analyzed the findings in order to infer conclusions about the relationship between socioeconomic status, the reasons that women in Tanzania pursue entrepreneurial activities and the characteristics
of their ultimate business of choice. The case-study evidence (albeit anecdotal, rather than empirical) suggests that pull factors such as market opportunities, ambition and a desire for self-fulfillment are more ubiquitous among women of middle and upper-middle classes, while push factors such as lack of alternatives, a need for supplemental income and a desire to accommodate domestic responsibilities alongside a work schedule are more common among poorer women in Tanzania. These results confirm the hypothesis that the relation between women entrepreneurs of divergent economic backgrounds within a developing country follows a similar trend to the relation between entrepreneurs from countries with richer economic profiles.

The first chapter reviews the definitions of the fundamental terms and research aspects pertinent to the case study, including entrepreneurship at large, and the unique characteristics of women entrepreneurship. The second chapter explores the literature on the relation between economic growth and rates and types of entrepreneurship, and then draws a hypothetical parallel to the micro-level within a developing country, whereby women from different classes exemplify the same pattern as the macro relationship between nations. The third chapter places the hypothesis in the specific context of Tanzania and provides background data about the country. Finally, it presents interviews with 7 women from Tanzania, explores themes in their enterprises and analyses the differences between them.

The Study of Women Entrepreneurship

In order to understand the nuances and particularities of women business owners, a general overview of entrepreneurship as an economic, social and occupational phenomenon is required.

Definition and Types of Entrepreneurship

Entrepreneurship is defined in many different ways, depending on context and
research objective. Generally, the simple common denominator, which applies to enterprises of varying sizes and characters, is that the entrepreneur be a business owner who creates and heads a new legal entity that engages with other actors in the marketplace by vending a product or offering a service (Hurst and Lusardi, 2002; Cagetti and De Nardi, 2006). For purposes of this study, self-employment—that is, working for independent gain rather than for wages paid by a supervisory authority (Blanchflower and Oswald, 1999) without a registered business—is not considered entrepreneurship, as the study aims to investigate business ownership specifically.

Various disciplines have defined entrepreneurship from their respective academic focal point. The behavioral approach, spearheaded by Casson (1982) sees entrepreneurship as a set of activities toward organization creation. The trait approach, advocated by McLelland (1961) views the concept as a set of personality traits which business-owners share in common (Gartner, 1988), while Schumpeter, who laid the foundation for the study of entrepreneurship as an economic field, believed innovation as the pivotal element in the activity (Rispas, 1995).

For purposes of this discussion, an all inclusive definition of entrepreneurship will be adopted, one that refers to the basic economic activity entrepreneurs engage in, traits and objectives common to entrepreneurs, as well as the manifestation of certain behaviors in aspiring business owners. Since the study involves a comparison between developed and developing countries, where employment and business environments differ significantly, and between entrepreneurs who own vastly different kinds of businesses, the term entrepreneurship will be used to describe business ownership regardless criteria such as size, length of existence or a minimum number of employees (as adopted by Quadrini, 1999).

Differences between those who enter into private business out of personal desire and those who pursue it due to constraints on other opportunities (whether these be self-imposed or objective) have always clearly existed. The discrepancy was institutionalized with the inception of the Global Entrepreneurship Monitor (GEM) survey, which asked
business owner interviewees to self-identify their reason for pursuing entrepreneurship and, thus, distinguished between necessity-driven and opportunity-driven entrepreneurs. The former are affected by push factors such as insufficient income, dissatisfaction with a current position, lack of alternative employment or a need for time-management flexibility, while the latter are motivated by a desire for self-fulfillment, socioeconomic status, wealth or power (Orhan and Scott, 2001).

Regardless of the qualifying criteria for the definition, it has been determined that entrepreneurship as a whole is directly related to many positive social and economic indicators. Opportunity-driven entrepreneurs were found to be an invaluable source of ideas that propel technological and economic advancement in developed and developing economies alike (Acs, 2000). By discovering knowledge or pockets of demand in the market that were previously undetected, entrepreneurs answer consumer needs and improve overall economic efficiency (Boetteke and Coyne, 2003). Entrepreneurs, in general, have been found to boost job creation and, compared to large, established, firms, they contribute a disproportionate share to employment in the market (meaning, entrepreneurs create more jobs, relative to their size, than larger companies) and respond in a more agile manner to changes in the operating environment (Stam, 2008).

**Women Entrepreneurship**

Women entrepreneurship is considered a special segment within the overall study of the economic and social phenomenon, and warrants a specific investigation both because of the defining criteria of the term and the special societal impact of women’s entrepreneurial activities. Women’s participation rate in economic activity, including business ownership, is almost universally lower than men’s. Women’s unemployment rates are higher (WDI, 2009), and they own fewer independent businesses compared to men (GEM, 2009). Studies into the causes of this situation investigate several explanatory dimensions.

On the psychological plane, which examines what distinguishes the personalities of
entrepreneurs from non-entrepreneurs, no significant differences were found between men and women (Scherer, Brodzinski and Wiebe, 1988). Traits common to all entrepreneurs were independence, emotional stability, self-assertiveness (Brandstatter, 1996), increased risk-taking propensity, locus of control, achievement motivation and tolerance for ambiguity (Low and MacMillan, 1988; Shaver and Scott 1991). Irrespective of gender, all entrepreneurs exhibited the traits more dominantly than non-entrepreneurs within their peer comparison groups. Individual psychological explanations, therefore, cannot account for the significant differential between rates of male and female entrepreneurship.

Behavioral aspects and strategic decision-making, measured in terms of success when entrepreneurs were already heading operational ventures with established business processes, were also found to be similar across both genders. Survival rates are similar in female-owned and male-owned firms (Chaganti, 1986), meaning the difference in overall participation rates cannot stem from women entrepreneurs’ lower likelihood of business survival.

Instead, what has emerged as the main explanatory factor for women’s lower entrepreneurial rates is the sociological aspect, which focuses on the “female” socialization experience and its derivatives, including professional and career choices. This aspect concentrates on the psychological macro-trends that shape societal conceptions, culture and values, rather than the decision-making processes of the individual entrepreneur.

According to the sociological approach, several indicators contribute to women’s less entrepreneurial endeavors: educational preparation, career expectations and expectation about the role of one’s gender in the professional working world. These are more negative among women, making them less likely to want to pursue interdependent businesses (Scherer, Adams and Wiebe, 1989). Females are not necessarily less well-educated or well-trained, but have worse perceptions of their self-efficacy (the individual’s confidence in his or her skill and ability to successfully engage in career-related activities) and worse expectations about women owning businesses as a whole (Scherer, Adams and Wiebe, 1989).
In addition, feedback from the external environment regarding an individual’s ability to handle time-constraints, work under pressure or be innovative is a key determinant for entrepreneurs, and women receive less of this positive feedback compared to men (Scott and Twomey, 1988). Other tacit influences also affect women; family considerations and a lack of role models are important restraints for many potential women entrepreneurs (Scott and Twomey, 1988). Since fewer women become entrepreneurs to begin with, the trend is perpetuated: with fewer role models available, fewer women can be inspired by them.

Sociological norms are sometimes also expressed and institutionalized through social policies. For example, women report discrimination and lack of access to credit and investment capital as a main challenge in pursuing entrepreneurial activities (GEM Women Report, 2007). Moreover, an important linkage between employment and entrepreneurship creates a vicious cycle that contributes to women’s lower participation in both. Most entrepreneurs report being employed before forming a venture, often in a similar field (GEM Women Report, 2007), which constitutes professional preparation and allows individuals to develop their entrepreneurial idea. Since women are unemployed more often than men, they are automatically at a disadvantage when trying to start their own businesses.

Cumulatively, these various sociological reasons lead to the current predicament: entrepreneurship is a predominately male phenomenon across almost every country and region, in any class of economic development and at every stage of entrepreneurial activity, whether nascent, new or established (GEM Women Report, 2007). In place of wage labor or entrepreneurial activities, women are driven to self-employment, which leads to their being the vast majority of agricultural self-sustenance workers in low income countries. In Tanzania, for instance, agriculture constitutes 80% of female employment, as opposed to 72% of male (WDI, 2006) and 42% of all women are engaged in agriculture, as opposed to 37% for men (Tanzania National Bureau of Statistics, 2001). While agricultural figures in the developed world are starkly different (in high-income countries agriculture accounts for
less than 3% of employment and more men than women are engaged in it (WDI, 2006),
women in these locations still find themselves at a disadvantage compared to men in their
respective fields, earning less money and being self-employed more often than men.

Women Entrepreneurship and Contributions to Society

A second distinguishing factor of women entrepreneurship from the general
entrepreneurial phenomenon is the former’s amplified positive social impact. From a
strictly financial vantage point, women entrepreneurship leads to increased income and
earnings of individuals who would otherwise not work and, thus, contributes to family-
level conditions and therefore to macroeconomic improvement. The correlation between
women’s education and entrepreneurial success is further translated into strict economic
consequences: additional schooling for girls leads to more success as entrepreneurs and
concrete impact on GDP, income per capita, productivity and national growth rates
(Lawson, 2008).

Due to various psychological and sociological factors, women entrepreneurship has
socioeconomically constructive effects beyond the purely financial, quantifiable, ones. In
general, women define success in broader and more diverse terms than men; ones that
include giving back to the community and the well-being of their families, as well as
making profit. Women describe success by assessing the state of their children, their ability
to choose their own schedule and daily tasks and their reputation (Justo et al, 2006).

Not surprisingly, therefore, women entrepreneurs make “important contributions
to the world economy, particularly in low and middle-income countries” (GEM Women
Report, 2007). Their nuanced reasons for becoming business owners and the intensified
benefits they bring about play a particularly critical role in the context of development.

First, at face-value level, women’s enterprises contribute to household income, and
thereby assist in raising families’ purchasing power and standard of living, which is critical
in developing nations. Additionally, women entrepreneurship assists in poverty alleviation
in indirect ways. When women are economically mobilized through ownership of an independent business, they tend to change the consumption patterns of the household in a more healthy way, compared to men, that is conducive to familiar and child well-being. Women spend their entrepreneurial earning on clothing, food and necessities for the household while men may spend it on alcohol or entertainment (Kantor, 2001).

Women tend to hire other women, which helps ease the cycle of chronic female unemployment in developing nations (Aidis et al, 2005). They tend to pass on their professional know-how and skills, thereby constituting the much-needed role models for future generations, and are able to “build and maintain long-standing relationships and networks to communicate effectively, to organize efficiently, to be fiscally conservative, to be aware of the needs of their environment, and to promote sensitivity to cultural differences” (Jalbert, 2000). Moreover, starting and heading successful independent businesses helps raise women’s sense of self-worth, “making them even more eager to be productive members of society” (Seymour, 2004).

Because of these augmented positive benefits, women entrepreneurship has been a top policy priority for international, regional and domestic development organizations. Many policies and programs, including the ILO’s Women Entrepreneurship Development (WED) and the IFC’s Women in Business, promote it as a catalyst of economic growth. These programs try to address the various underlying causes of lower rates of women entrepreneurship, seeking to change social perceptions, provide access to capital, and run training and educational programs to prepare women before, and support them after, they open their businesses.

The U Relationship

After a general overview of the concept of entrepreneurship, the specifics of women entrepreneurship and the role of both in development, a more in-depth look into the hypothesis regarding the relationship between entrepreneurship and development can be
Entrepreneurship in Developing and Developing Nations

Entrepreneurship is believed to be rooted in human nature; it has been an economic activity central to evolution and historic progress—in the form of creation of new ideas, the accumulation of wealth and the rise of the global economy as we know it today. In general, those nations or regions with higher rates of opportunity entrepreneurism have achieved greater economic growth, while countries or areas where necessity-driven entrepreneurs engaged primarily in agriculture or craftsmanship are now, economically speaking, less advanced (Perunović, 2005).

This historic relationship has translated into a modern-day negative correlation between national economic output and rates of business ownership. The developed nations of the West, which underwent the industrialization process in the 18th and 19th centuries, saw uninterrupted sharp declines in self employment and small business ownership, and a consistent shift toward wage labor, that only slowed (and then reversed) in the 1980’s. Attributed primarily to a move away from agricultural self-sustenance as economies began to develop manufacturing and industrial sectors (Wennekers et al, 2009), this decline persists today: entrepreneurship is correlated with economic underdevelopment. The GEM has found rates of entrepreneurship to be highest among low and middle-income countries as compared to high-income nations (GEM, 2009). This large-scale supply of entrepreneurship at the low end of the macroeconomic spectrum is often composed of agricultural self-sustenance. This is unsurprising, as in rural areas jobs are too scarce, and market wages too low, to incentivize and enable wage labor (Blau, 1985).

What is surprising, and of potential further interest, however, is the relatively recent trend that suggests the correlation between entrepreneurship and development may not be linear, or even consistently negative, after all. Since the late 1970’s, many developed nations have exhibited increases in rates of self-employment, business ownership and
entrepreneurship (Wennekers et al, 2009).

On the macro level, this trend has been associated with several financial, political and regulatory indicators. Business density and entry rates (the number of businesses as a percentage of the population and new firms as a percentage of the working-age population) are higher in developed countries where the overall business environment is better (Klapper et al, 2010).

On the micro level, the question centers on the entrepreneur and his or her motivating forces for entering into business. Entrepreneurs in developed nations—even women entrepreneurs—report various reasons related to opportunity, ambition and self fulfillment for venture formation much more frequently than peers in the developing world, who are generally necessity-driven and influenced by push-factors.

In the developing world, the forces that exert pressure on society at large to be self-employed (and, hence, lead to greater rates of business ownership) have an intensified effect on women. Because of discrimination, men often have priority in receiving the meager number of salaried jobs available in the market (Mueller, 2004). Women across cultures are more subject to self-confidence problems than men, and are expected to meet the requirements of child-rearing and attending to the home in addition to working. Employment is much more problematic for women in these situations and independent businesses (or self-employment) become the only, or the preferred, viable solution. Evidence of this can be found in the narrower gender gap in low-income countries, compared to high-income countries: the 2007 Global Entrepreneurship Monitor survey of women entrepreneurship found that, although males are likelier to own businesses across nations and cultures, the relative differences are smaller in poorer nations (GEM Women Report, 2007), meaning women in developing nations, are more entrepreneurial than the men. This suggests that, in the developing world, women have even fewer employment possibilities than men so are compelled to undertake business ownership instead.

In developed countries, too, women are more necessity-driven than men. Compared
to female peers in the developing world, however, women in developed economies report more opportunistic motivations for starting their own business. When they cite push-factors, they are of a different variety than those of developing world women. Rather than altogether absent employment or professional opportunities, they name less fundamental necessity motivations such as boredom with a position and layoffs. Other times, they are obliged to take over family businesses that no one else is able to direct (Orhan and Scott, 2001). Another key necessity-reason that appears consistently among women in middle or high-income countries, but is absent in developing nations, is women’s limitation on upward mobility and promotion in the workplace: the male-dominated professional world and the resulting glass-ceiling effect for women drive them to seek independence and control by opening their own firms (Orhan and Scott, 2001). Even in places where educational levels and labor-force participation rates are similar, women earn less than men, and entrepreneurship “offers a vehicle for... women to achieve economic parity” (Lerner, Brush and Hisrich, 1997).

Women entrepreneurship at the two ends of the economic development spectrum is therefore distinct. While they have several common characteristics—primarily in the way they assess additional dimension of success—these women differ in their location, relative wealth and, consequently, the reason and characteristics of their businesses.

An Analogy to the Internal U-Shaped Relationship

One interesting implication of the resurgence of entrepreneurship in high-income nations is whether an analogy can be drawn domestically, within developing countries, such that entrepreneurship diminishes among lower middle-class individuals then reemerges in upper-middle and upper class circumstances. According to the analogy, the two types of domestic entrepreneurship will differ both in the preliminary conditions, the motivations for starting a business, the formation process and the final success measures of the venture.

Although, as discussed previously, women almost always face more push factors
than their male counterparts, and particularly so in low income nations, the pressures inflicted on women from different socioeconomic statuses are not identical. Therefore, neither are their motivations for becoming business owners, their occupational and field choices, nor their ability to succeed and grow the businesses.

The overwhelming majority of women in developing nations are poor and, because limited educational or professional background, have a limited range of potential business ideas avenues. Traditional crafts, small-scale retail or food-vending are often the only options available. In addition, being credit constrained and without initial capital, large up-front investments are impossible and they can mostly only pursue ventures in industries with low entry-costs and few barriers (GEM Women Report, 2007).

In contrast, women of the middle or upper-middle class follow occupational-choice models that take qualifications, preferences and goals into account. The initial distribution of wealth is a critical determinant of occupational choice (Banerjee and Newman, 1993), and these women choose professional positions based on the level of their education and training (which is a function of initial wealth), and later branch off of those positions into independent firms. They primarily concentrate on the trade and service sectors (Vandenberg, 2006).

The internal analogy of the U-shaped relationship thus holds on two plains: the reasons for entering into business are different for poor and rich women (as they are different among women of developing and developed nations), and are the characteristics of their businesses (including pre-business professional preparation, choice of sector, size etc’). Based on this psychological and sociological schema of occupational choice, and on the statistics on women’s push-and-pull factors when entering business, an inquiry into the range of different types of women entrepreneurs within the developing world is warranted.

The Internal U-Relationship: the Range of Socio-Economic Backgrounds of Women Entrepreneurs in Tanzania
With an aim to understand the different drivers, characteristics and features of women of different socioeconomic status within the context of a developing country, a qualitative study of 7 women entrepreneurs in Tanzania was conducted in 2009, with demographic indicators representing various societal segments: the rural poor, urban internal migrants, the upper middle class and the country’s social elite. The 7 women interviewed come from different, non-related backgrounds that exemplify the diversity of the female entrepreneurial experience in Tanzania.

**Tanzania’s Economic and Demographic Profile**

Tanzania, with over 42 million people as of 2008, is Africa’s 6th most populous. The average individual income, expressed in purchasing power terms, is $1167, ranking 188th among 210 countries according to the World Bank.

For entrepreneurship, Tanzania is considered a difficult business environment. As Table 1 shows, the business entry rate is low, and indicates low growth in entrepreneurial activity. Registering a business is a complex task involving many steps, and Tanzania’s rank in the ease-of-doing-business indicator is nearly 30 times that of the United States’ (120 compared to 4).

It is no surprise, therefore, that starting a business for women in Tanzania is difficult. For poor rural women, who are usually driven to entrepreneurship by necessity, disadvantages in education and literacy have an acute effect; registration requires competency in reading and an ability to navigate government bureaus and regulations. HIV-Aids also represents a challenge as prevalence rates in Tanzania are high, and higher among women, which can leave them debilitated or responsible for sick relatives’ other family members.

For urban middle-class women who, based on the U relationship hypothesis, would enter into business for more opportunistic reasons, underrepresentation in tertiary education represents a significant drawback, as professional work experience has been proven to be a strong foundation for growth-oriented enterprises, and higher education is
often a prerequisite for gaining such experience.

### Table 1: Tanzania Business Environment

<table>
<thead>
<tr>
<th>Series</th>
<th>Tanzania</th>
<th>World</th>
<th>Low Income</th>
<th>High Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business entry rate (new registrations as % of total)</td>
<td><strong>6.64</strong> (2005)</td>
<td>10.99</td>
<td>18.18</td>
<td>11</td>
</tr>
<tr>
<td>Cost of business start-up procedures (% of GNI per capita)</td>
<td>47.1</td>
<td>64.33</td>
<td>181.54</td>
<td>9</td>
</tr>
<tr>
<td>Ease of doing business index (1=most business-friendly regulations)</td>
<td><strong>126</strong> (2008)</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>New businesses registered (number)</td>
<td>..</td>
<td>2,381,157</td>
<td>21,034</td>
<td>1,352,073</td>
</tr>
<tr>
<td>Start-up procedures to register a business (number)</td>
<td>12</td>
<td>9.06</td>
<td>10.3</td>
<td>7</td>
</tr>
<tr>
<td>Time required to start a business (days)</td>
<td>29</td>
<td>42.76</td>
<td>53.53</td>
<td>23</td>
</tr>
<tr>
<td>Total businesses registered (number)</td>
<td><strong>59,163</strong> (2005)</td>
<td>19,878,084</td>
<td>56,920</td>
<td>12,593,439</td>
</tr>
<tr>
<td>Firms with female participation in ownership (% of firms)</td>
<td><strong>30.93</strong> (2006)</td>
<td>..</td>
<td>..</td>
<td>..</td>
</tr>
<tr>
<td>Female adults with HIV (% of population ages 15+ with HIV)</td>
<td>58.46</td>
<td>32.93</td>
<td>39.16</td>
<td>25</td>
</tr>
<tr>
<td>Ratio of female to male enrollments in tertiary education</td>
<td>47.73</td>
<td>108.26</td>
<td>66.79</td>
<td>123</td>
</tr>
</tbody>
</table>

Source: World Development Indicators, 2007 (except where bold).
Although Tanzania did not participate in the Global Entrepreneurship Monitor, several studies have been conducted regarding gender and business ownership trends in the country. Legislation has been implemented that formally equates male and female rights in business but parity has not been reached; men still constitute the majority of business owners, particularly in the formal sector (Richardson, Howarth and Finnegan, 2004). Women own primarily micro, small and medium sized enterprises (MSMSEs). Several definitions of micro, small and medium sized enterprises exist in Tanzania, varying by organization and defining criterion. The correct categorization of enterprises is important because specific policies in Tanzania target certain classes of businesses; the government’s definition, as outlined in the official policy of the Ministry of Industry and Trade, is given in Table 2.

<table>
<thead>
<tr>
<th>Type of Firm</th>
<th>Employees</th>
<th>Capital Investment in Machinery (TSHs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Micro</td>
<td>1-4</td>
<td>&lt;5m</td>
</tr>
<tr>
<td>Small</td>
<td>5-49</td>
<td>5m-200m</td>
</tr>
<tr>
<td>Medium</td>
<td>50-99</td>
<td>200m-800m</td>
</tr>
<tr>
<td>Large</td>
<td>&gt;100</td>
<td>&gt;800m</td>
</tr>
</tbody>
</table>

Source: Tanzania Ministry of Industry and Trade

The women featured in the study fall under all four of these categories, in different industries and sectors. In order to divide them and study the internal U-shaped relationship hypothesis, interviewees were divided into three groups, primarily based on background (as evidenced in parents’ educational level). First are poor women, who are often pushed into
entrepreneurship, opposite them are women of the upper middle class who are pulled into it because of personal desire, and thirdly women “in the middle” that experience some of both.

Motivations and Business Characteristics of Tanzania Rural/Poor Women

To present the experience of small-scale entrepreneurs from rural, poor areas, Augustina Ameena, Amali Nhuru and Levina Ndusilo were interviewed. Augustina and Amali are in a joint-partnership of a basket-weaving business, while Levina is the owner of a small convenience shop in town. The women live in Njombe, a town of 100,000 residents 1,000km from the capital Dar es Salaam, in the Iringa province. With an additional 50,000 people in the surrounding villages, and two other relatively large towns, Iringa is not Tanzania’s most rural or poorest province, but it has more adults working in agriculture than the national average (67% compared with 62%), and performs poorly on developmental indicators (such as the percentage of households with running water, concrete walls or connection to the electricity grid) compared to the rest of Tanzania (Tanzania Bureau of Statistic, 2001).

Augustina Ameena and Amali Nhuru

Augustina and Amali, 35 and 37 in 2010, have known one another most of their lives, growing up in the outskirts of Njombe. In 2007, they entered in joint partnership into a basket-weaving business (with no official name). Augustina and Amali rent a small dirt-floor room in the edge of Njombe town, where they make bags and baskets from cut and processed bamboo cane. Once weekly they travel to a nearby village to buy the materials, and rent space on a truck to deliver it back to their workspace. They charge between $4 and $16 per item, depending on size and complexity; each can take several days to complete, and may involve various types and shades of bamboo that make the raw materials expensive. Augustina and Amali have gone to trade shows before to market their
products, but sales volume at these events was insufficiently high to justify the expense of the 13-hour trip to Dar es Salaam. The women also found it difficult to leave home for several consecutive days, so the strategy was abandoned in favor of their traditional business model. They take local orders from women around town and sell their products at a stand at the market. Both women had similar motivations for starting the business:

“The revenue that my husband used to bring-in by selling our produce and milk was declining. We needed more money to pay for basic food and clothing and so there was very little I could do aside from try to earn money from weaving baskets. I completed only my primary education so there were no jobs for me in Njombe. But even with a secondary degree, I couldn’t have done much but open my own business. Weaving basket is a traditional craft here, one that many women learn young, so it was the obvious business choice. A partnership was a great opportunity because it enabled us to participate in the initial expenses of registering a business—which are very costly and cumbersome.” (Augustina)

“I had another baby and needed additional income. No other option was more feasible than starting our business: we had no training in other areas, and many of the women around us were also engaged in independent basket weaving. With the business, I was able to keep to my chores and commitments at home: I brought my baby with me to work and could leave whenever I needed, if I needed, knowing that Augustina would finish the orders. The flexibility is very important to me.” (Amali)

**Levina Ndusilo**

Levina was born, raised and married in Njombe. Her parents were farmers in a nearby village, and had very minimal education. They managed to send their children to primary school and all six completed it. Levina married in Njombe and had five children. Although her husband had a relatively stable job—driving a supply truck for Coca Cola—
his salary was not enough to maintain the household, particularly after Levina’s sister became ill and the family took in another three children. At the same time, work around the house abounded; with no additional help, Levina had to take care of eight children, her husband and her sister. In 2000, she opened a small convenience shop in the town’s central road, which vends basic supplies and food products. Because of her husband’s association with Coca-Cola, she is also the central distributor of the product for other shops in the area, so that they have to purchase all soda bottles supplies from her. She keeps a storage space behind her shop, which is just a minute’s walk from home, and hires one employee, which enables her to leave when needed to attend to needs at home.

“The strongest reason for opening the shop was financial. I simply needed the income, and there are no jobs in Njombe for me to apply to. When we took in my sister’s three additional children is when it became evidently clear. Getting the initial capital was difficult, and at first I thought about working as an employee at someone else’s shot. But the pay would have been meager, and the job wouldn’t have awarded me the same flexibility; I would have had to abide by their schedule and demands; having my own allows me to set them.”

Analysis

Women in this category fit the template of necessity-driven entrepreneurs influenced by strong, dominant push forces. Neither one opened the business because of specific aspirations or personal qualities; they were driven to it because of life events that made the need for additional income even more imperative. The women had no alternative employment opportunities, as formal economic activity in their locality is limited, and nowhere to get additional skills that would qualify them for positions (such as secretarial training or education in accounting). In addition, they are charged with domestic responsibilities that limit their flexibility and compel them to find activities they can fit into their schedules. The choice of sector or industry was not a targeted business opportunity;
the basket weaving market is saturated and highly competitive, since many other women engage in the craft, but Augustina and Amali had no other qualification or choice. Although Levina has a comparative advantage through the wholesale of sodas to other vendors, she, too, faces steep competition as the number of competitors selling similar merchandise is high. Overall, all three women agree that they were obliged to enter into business, and that the examples of women around them, who engage in similar types of entrepreneurship, constituted the only source of information for them on how to start, register and run their business. They received no training, assistance or guidance.

Motivations and Business Characteristics of Urban, Upper Middle-Class Women

The experience of the urban, middle (and upper) class women interviewed reflects the existence of an internal U-shaped relationship among women entrepreneurs within Tanzania; the women interviewed reported different reasons, motivations and paths toward opening their businesses. They bear little resemblance to the rural, survivalist entrepreneurs; their background, business planning and aspirations are all starkly different. The women in this category are Emelda Mwamanga, 31, the founder and Editor in Chief of Bang! Magazine, Tanzania’s first lifestyle magazine (See Exhibit 3) and Modesta Mahiga, 29, the founder and Director of Professional Approach, a consulting firm (see Exhibit 4).

Emelda Mwamanga

Emelda was born in 1979 in Dar es Salaam. Her mother is a teacher at an international school and her father, an economist by training, worked in a government-owned industrial auto-parts company until it was privatized and he became the owner and CEO. Emelda went to a private high-school in the capital, then to the University of South Africa, where she studied marketing and human resources administration. After graduating with her bachelor’s degree, Emelda worked for Coca-Cola in Dar es Salaam as a Human Resources Officer for two years. She learned a lot from the role, including how to identify
successful workers and how to communicate in a professional environment. Emelda started Bang! Magazine in 2005 from a corner in her father’s office. The business is described on its website:

Bang! Magazine is a bi-monthly fix for readers fascinated by celebrity culture, successful personalities, human-interest stories, fashion and beauty. Our over 90 thousand bi-monthly readers don’t just read BANG! They live it. Between each and every page, the magazine is full of information to enrich, encourage and enlighten the entire family. BANG! Offers fresh, exciting images and perspectives on the new African and remains the only local general interest and major bi-monthly magazine that covers all of East Africans unbiased.

To symbolize the importance of her family, Emelda named the publishing company, through which the magazine is registered, Relim Entertainment—an acronym of the names of her immediate family members. Emelda used her own savings from her time at Coca Cola, as well as a small $10,000 loan from a private bank, to finance initial expenses. She printed 3,000 copies at a cost of $3,000, and used the rest to pay contributors and employees, including layout, design and marketing. Although distribution channels were virtually nonexistent, as Tanzania had no tradition of vending and reading magazines, even among the urban elite, Bang succeeded and continued to expand its readership. In 2007, it began distributing in neighboring Kenya and Uganda and by 2010 had 10 permanent staff in addition to regular contributors. Emelda’s motivations for starting Bang! were almost entirely pull-related:

“I had a strong interest in entertainment, media, writing and women’s issues. I admired Oprah Winfrey, as a model for a successful black woman who branched out into multiple ventures, and dreamed of importing the business concept to Tanzania. My position at Coca-Cola was lucrative enough, and stable, but I wasn’t feeling fulfilled. I came to work each day thinking about starting a magazine, and
left my office with the same thoughts. Also, I knew that there was a rare market opportunity at play: no one else was offering a similar product, and I would be the first one to capitalize on the idea. Magazines were not a part of life in Tanzania traditionally, but as more people internationalize readership is bound to grow—and I knew that. I had a concrete model of a successful growth-oriented entrepreneur from my father, who was a tremendous influence along the way. Besides, having my own business leads people to respect me a lot more than when I say I’m working for someone else; I can see it in their expression.”

Modesta Mahiga

Modesta was born in 1981, also in Dar as Salaam, to parents who were academics and then working professional at international firms. Modesta’s education reflects her future business venture: eclectic, rigorous and innovative. Modesta went to primary school in Tanzania, boarding school in Kenya, and then moved to Australia with her mother when she worked on her master’s degree in economics. She received her International Bachelorette in Swaziland, interned for PriceWaterhouseCooper Tanzania for 6 months, moved to England to work on her LLB and, after two years of working for a multinational corporation in Tanzania as a human resource manager, moved to Germany to receive an LLM in order to become a lawyer. After finishing her degree and returning to Tanzania permanently, she decided she didn’t want to practice law and, instead, in 2008 founded Professional Approach. On its website, the firm was described as:

“A Tanzanian company founded on the conviction that Tanzania’s human resource can be harnessed, and its organizations energized, to transform our economy”, with a vision “to spearhead Tanzania’s economic development through the injection of skilled labour” and “To offer individuals techniques and opportunities to become effective world-class professionals, driving excellence in the way they think, perform and present themselves.”
With 2 permanent employees and 4 part-time staff, the firm is still in its infancy. Although profit margins are big, at least theoretically, the service is not one customary to Tanzania, and Professional Approach sales volume is not satisfactory. Modesta has to spend the majority of her time promoting and marketing the firm, and “could be making more money as an employee”. She started Professional Approach only with savings and help from family, without any borrowing, and continues to plant profits back into the business and use it as operational capital. Modesta names several reasons for her decision to open a business, all opportunity-related:

“I felt a calling in the venture I decided to pursue. When I returned to Tanzania, I wanted to do public advocacy law, in an effort to drive societal change. But after contemplating about it, I realized that changing society requires changing and adapting norms and behaviors—so I decided to start a firm that would help organizations do that. It was something I felt I had to do, although I know for a fact I would be making more money in my previous job, or as an employee elsewhere, or as a lawyer. The fact my parents were able to support me while I was building the firm was invaluable; otherwise I could have no paid employees or survive myself. The prestige that comes with the venture, too, had an allure: I would be telling managers—men, older executives, high-ranking government officials—what to do and how to run their organizations; climbing the corporate ladder in a company that’s not my own would have never gotten me there.”

Analysis

The entrepreneurial motivations and business characteristics of the upper middle class women clearly differ from those of the rural poor in several dimensions: family background, educational levels, professional preparation, initial capital investment, role models and support in the enterprise’s nascent stage. Both women had well-educated
parents, an international education and prior, professional work experience. They had role-models and examples of entrepreneurs or successful professionals in their immediate family and received the ideas for their ventures from a combination of marketplace exposure and personal interest, beliefs or passions. These women were not pushed toward entrepreneurship because of monetary constraints; on the contrary, they acknowledge that staying employees was safer and more lucrative in the short-term. They invested significant sums into the firms, and continue to make below their earning potential because of aspirations and a desire to “make it” as independents. Thus, they exemplify the same trends that entrepreneurs in developed countries do in the macro U-shaped relationship: an opportunity-driven return to entrepreneurship due to pull-factors.

Motivations and Business Characteristics of Lower Middle Class Women

The last category of women corresponds to the curved area of the U-shaped relationship: they are “middle-of-the-way” entrepreneurs who are neither poor or from a rural area, nor from an affluent household at the top of the domestic economic ladder. They have varied backgrounds and mixed reasons for forming ventures, and similarly diverse business characteristics. The women interviewed are Clara Ibhiya, the owner of Mama’s Flavor, a dried-fruit brand vendor, Dina Bina, founder of Dina’s Flowers decorative business as well as other ventures and Victoria Kisyombe, director of SERO Lease and Finance Ltd., a leasing firm specializing in women.

Clara Ibhiya

Originally from Tabora in the northwest of Tanzania, Clara, 55, moved to the capital early in life and worked for many years as a government employee and then as a clerk in a medium-sized company. She made several attempts at starting her own business, including a piggery, a poultry facility, a hair salon and a tailoring operation. Finally, she received the business idea for Mama’s Flavor from a practice in her home town, where the
women would dry fruits in the sun and later sell it at the local market. With the expansion of the urban middle class in the capital, and the opening of large, upscale supermarkets with significant purchasing power, Clara believed she could find enough clients to keep demand up for products. She leveraged her familiarity with fruit growers in the country and her connections in the city, bought a large-scale drier using savings and money from her children and opened Mama’s Flavor in 2005. She secured contracts with three large supermarket chains and now produces dried mango, pineapple, banana, papaya, jackfruit and indigenous vegetables. After gaining a reputation for itself, Mama’s Flavor became a Tanzanian favorite. Even Dr. Asha-Rose Migiro, Ban Ki-Moon’s Deputy Secretary General at the United Nations and Tanzania’s former foreign minister tried Mamma’s Flavor and was immediately hooked; she places special orders with local affiliates in Tanzania and has Mama’s Flavor baskets sent to New York periodically. When discussing her drivers, Clara mentions several:

“I had always wanted to have my own business, where I would have the freedom to control my hours, and where there was growth potential. Government jobs are stable, but not growth-oriented; there was nothing for me to aspire to. That’s why I made so many attempts—having my own business was something I truly wanted. At the same time, the financial need was also a key incentive. I needed supplemental income because mine was simply not enough. To make ends meet, to send my youngest son to university, I had to earn more. And the idea was to grow a business that would be my main source, not just secondary one.”

Dina Bina

Dina, in her 40’s, was born to parents who were considered financially-stable in the outskirts of Dar es Salaam. Her father was a pastor and her mother, too, completed secondary education. With both of them working, the parents were able to send Dina to the University of Dar es Salaam, where she majored in chemistry and mathematics. She
dreamed of becoming an engineer but her brother convinced her to study marketing and accounting instead, saying engineering is a field for men only. She indeed switched and, after graduating, worked as an accountant for some time, and as a teacher in church school in Kenya and Uganda. When she returned to Tanzania, a sister-in-law pitched a business to her: to open a flower business, utilizing the fact the couple lived in the northern part of the country, where flowers were abundant and cheap. She opened Dina’s Flowers in 1997, with a mission to:

Design, produce, and sell imaginative and attractive flower arrangements, event and venue decorations, and superbly laid out outdoor gardens and landscapes by a well trained and highly motivated work force that is committed to providing top quality customer service.

Dina’s Flowers now offers several services, including fresh and artificial floral arrangements for all occasions, indoor plants-placing and upkeep, landscaping design and implementation, garden and lawn upkeep, event logistics management and tropical flowers growing and supplying. It has 2 retail locations and 20 permanent workers—more during the high-season when Dina contracts laborers. Dina’s path to entrepreneurship was paved with many factors, as she describes:

“I am just not the kind of person who could work for someone else. I have always been independent and free-thinking, working for someone else just gets me into trouble. I am involved in women entrepreneurship associations, with international intergovernmental organizations, with my church, with another business I am opening now—an academy for business and management specifically for women. I am always involving in and leading something. And I always knew I would have to be my own boss. The fact my husband was very supportive was also a huge encouragement; now he even works in the business, as the Chief Financial Officer. At the same time, there was a strong incentive so start the business for
financial reasons. A job as an accountant wouldn’t bring in the same revenue, and allow us to maintain the life we wanted.”

**Victoria Kisyombe**

Victoria is from Mbeya, a city of over 2 million in the southwest of Tanzania. She is a veterinarian by training, having earned her bachelor’s in Tanzania and her master’s degree in England. She worked for the provincial government, treating farm animals around her town until, in 1991, her husband passed away. Her salary alone was not sufficient to support herself and the three children, and Victoria knew she had to find another solution. Her husband’s only inheritance was a cow which, in addition to the family’s use, gave enough milk to sell in the market. Victoria saw the utility in having such a working asset; she received enough additional income to support the family and even save money. She believed that allowing all women to own such assets could be both a business opportunity and a way to help Tanzanian society. She moved to Dar es Salaam, worked in a women’s advocacy NGO that she help found and in 2002 opened SELFINA (Sero Lease and Finance), a financial company that leases equipment to women entrepreneurs around the country. The firm’s mission is:

To be a major provider of microfinance (micro-leasing finance) to Tanzanian women; to achieve operating levels that will ensure sustainability and expansion to cover all regions of Tanzania; and to be a credible and reliable MFI partner to banks and other financial institutions.

Selfina offers leases on a financial basis—meaning the client pays toward ultimate ownership. Victoria offers various types of assets: agricultural equipment (power tillers, maize milling machines, animal feed, mixers, sunflower-oil extraction machines), catering equipment (coolers, freezers, refrigerators, cookers), tailoring equipment (manual and electrical sewing machines, embroidery sewing machines, overlocks, chain stitch machines, handlooms), secretarial equipment (photocopiers, computers, printers, typewriters) or any
other kind of equipment that’s in need. Because it is a whole-sale buyer, it is able to receive better prices for equipment and lend it to customers at a rate that generates sizeable profits but is still better than what the same women would get through a commercial bank. The company has 12 branches and over 100 employees, with Victoria acting as the Managing Director. She cites several reasons that made her pursue her venture:

“I started SELFINA because I needed to; I was on my own and wanted my children to have an education, a good life. But, also, I saw a need I couldn’t ignore—the need of Tanzanian women to have working assets to start and grow their businesses. And implicit in that was a business opportunity to offer a highly-demanded service: many of these women could not get credit from regular financial institutions so supplying it to them would guarantee a large volume for my firm. Really, it was a combination of things that made me open a company—looking at the market, wanting to give back to the women’s community and my own personal situation.”

Analysis

The women in this category represent the middle-class exceptions who choose to pursue entrepreneurial ventures rather than remain employed. Their motivations are combinations of both push and pull factors that, together, lead them to pursue professional independence. On the one hand, they express the typical opportunistic reasons, such as wanting to seize a market demand or fulfill personal objectives. On the other hand, necessity is a significant determinant for their actions. Financial needs and a lack of growth-potential at the workplace play an important role in their decisions. In the macro-level parallel, they represent the curbed bottom of the U-shaped relationship, where the majority of women in similar situations remain employed and a few become entrepreneurs. Indeed, the women in this category reported that the majority of their friends, acquaintances and peers were employed and they were the exceptions to the rule.
Overall analysis and Conclusions

The women in all three categories exemplify the micro-level parallels of the macro U-shaped relationship between entrepreneurship and economic growth. Similar to women entrepreneurs in developed nations, at the higher level of the socioeconomic ladder in Tanzania, formal employment is more readily available, making the opportunity cost of business ownership higher and the need for it less urgent. By contrast, women from lower socioeconomic backgrounds in Tanzania exemplify the phenomena generally attributed to entrepreneurs from countries with weaker economic performance: a lack of occupational alternatives that drives women toward business ownership.

The three classes of women differ on several categorical levels linked to their basic reasons for becoming entrepreneurs: family origin, education, and professional preparation, among others. The case studies found a clear correlation between the women’s family origins and the trajectory of the entrepreneurs’ business: the more educated the parents are, the likelier is the business is to succeed, both because of the initial support parents can furnish and the basic skills they have equipped the women with.

Education of the women themselves was also a critical factor, not only insofar as it determined the level of specialization and market-entry potential for the entrepreneurs, but also in terms of basic understanding of conducting business. The more highly-educated women were able to find market opportunities that fitted their skills and qualifications and, in addition, had an easier time juggling the registration process and the various business management tasks, including dealing with suppliers and keeping records. Professional preparation was directly linked to the women’s educational experiences: women from higher socioeconomic backgrounds had more exposure to professional settings, and were able to utilize their exposure in their strategy and business planning.

The presence of role models was one category in which the differences between interviewees were subtler: women from poorer socioeconomic backgrounds didn’t suffer
from a total absence of role models; rather, the women they knew engaged in similar survivalist business activities. Thus, they did not act as innovation-seeking role models who spurred growth aspirations in the women interviewed but as examples of peers from similar demographic situations. The women of higher socioeconomic status had examples of thriving, innovative peers while the middle category of women exhibited mixed experiences that can be linked to accounts on both other groups. In general, however, all women agreed that females face a harder time as entrepreneurs than men, and that the phenomenon was less common among their female acquaintances or relatives than male ones.

In all these aspects, the relationship between entrepreneurship and economic progress among the Tanzanian women was analogous to the macro U-shaped relationship. In quantitative terms, entrepreneurship was common when economic abilities were low, scarcer among the lower-middle class where economic progress was nascent, and then resurfaced at the peak of socioeconomic development. In qualitative terms, the explanations behind each trend among the women’s categories were similar to those in the country-level context. Family origin, education and professional preparation create similar influences among women in developed countries and women of upper-middle classes in developing nations, such that their entrepreneurial characteristics are similar. A similar kind of resurgence thus exists because of similar reasons. The typical account of entrepreneurs in developing nations refers to necessity-driven entrepreneurs that enter into business because of personal and financial limitations. Thus, it seems that general accounts of entrepreneurship in developing nations are descriptive of the first category of poor women of rural origin.
## Exhibit 1: Countries Ranked by Average Income and Rates of Opportunity Driven Women Entrepreneurs

<table>
<thead>
<tr>
<th>Country</th>
<th>GDP Per Capita (current US$)</th>
<th>Poverty Rank</th>
<th>Female Opp to Necessity Rank</th>
<th>Female Opp to Necessity Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>India</td>
<td>1068</td>
<td>1</td>
<td>17</td>
<td>1.84</td>
</tr>
<tr>
<td>China</td>
<td>3263</td>
<td>2</td>
<td>8</td>
<td>1.11</td>
</tr>
<tr>
<td>Thailand</td>
<td>3869</td>
<td>3</td>
<td>21</td>
<td>2.53</td>
</tr>
<tr>
<td>Peru</td>
<td>4419</td>
<td>4</td>
<td>20</td>
<td>2.45</td>
</tr>
<tr>
<td>Dom. Republic</td>
<td>4655</td>
<td>5</td>
<td>18</td>
<td>1.84</td>
</tr>
<tr>
<td>Colombia</td>
<td>5440</td>
<td>6</td>
<td>5</td>
<td>0.91</td>
</tr>
<tr>
<td>Serbia</td>
<td>6811</td>
<td>7</td>
<td>1</td>
<td>0.32</td>
</tr>
<tr>
<td>Argentina</td>
<td>8235</td>
<td>8</td>
<td>9</td>
<td>1.24</td>
</tr>
<tr>
<td>Brazil</td>
<td>8400</td>
<td>9</td>
<td>10</td>
<td>1.5</td>
</tr>
<tr>
<td>Kazakhstan</td>
<td>8436</td>
<td>10</td>
<td>16</td>
<td>1.83</td>
</tr>
<tr>
<td>Romania</td>
<td>9300</td>
<td>11</td>
<td>22</td>
<td>2.63</td>
</tr>
<tr>
<td>Uruguay</td>
<td>9654</td>
<td>12</td>
<td>4</td>
<td>0.77</td>
</tr>
<tr>
<td>Chile</td>
<td>10112</td>
<td>13</td>
<td>19</td>
<td>2.34</td>
</tr>
<tr>
<td>Turkey</td>
<td>10745</td>
<td>14</td>
<td>2</td>
<td>0.46</td>
</tr>
<tr>
<td>Venezuela</td>
<td>11230</td>
<td>15</td>
<td>13</td>
<td>1.6</td>
</tr>
<tr>
<td>Russia</td>
<td>11339</td>
<td>16</td>
<td>11</td>
<td>1.54</td>
</tr>
<tr>
<td>Latvia</td>
<td>14909</td>
<td>17</td>
<td>7</td>
<td>1.02</td>
</tr>
<tr>
<td>Hungary</td>
<td>15409</td>
<td>18</td>
<td>12</td>
<td>1.57</td>
</tr>
<tr>
<td>Croatia</td>
<td>15636</td>
<td>19</td>
<td>3</td>
<td>0.77</td>
</tr>
<tr>
<td>Puerto Rico</td>
<td>17692</td>
<td>20</td>
<td>30</td>
<td>3.36</td>
</tr>
<tr>
<td>Portugal</td>
<td>22842</td>
<td>21</td>
<td>35</td>
<td>4.9</td>
</tr>
<tr>
<td>Slovenia</td>
<td>26779</td>
<td>22</td>
<td>32</td>
<td>3.83</td>
</tr>
<tr>
<td>Israel</td>
<td>27298</td>
<td>23</td>
<td>15</td>
<td>1.66</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>30863</td>
<td>24</td>
<td>6</td>
<td>0.94</td>
</tr>
<tr>
<td>Greece</td>
<td>31749</td>
<td>25</td>
<td>36</td>
<td>5.38</td>
</tr>
<tr>
<td>Spain</td>
<td>35204</td>
<td>26</td>
<td>27</td>
<td>2.96</td>
</tr>
<tr>
<td>Country</td>
<td>GDP Per Capita (current US$)</td>
<td>Poverty Rank</td>
<td>Female Opp to Necessity Rank</td>
<td>Female Opp to Necessity Ratio</td>
</tr>
<tr>
<td>------------</td>
<td>------------------------------</td>
<td>--------------</td>
<td>-------------------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>Italy</td>
<td>38309</td>
<td>27</td>
<td>26</td>
<td>2.96</td>
</tr>
<tr>
<td>Japan</td>
<td>38443</td>
<td>28</td>
<td>23</td>
<td>2.83</td>
</tr>
<tr>
<td>U.K.</td>
<td>43088</td>
<td>29</td>
<td>24</td>
<td>2.9</td>
</tr>
<tr>
<td>France</td>
<td>45981</td>
<td>30</td>
<td>14</td>
<td>1.62</td>
</tr>
<tr>
<td>Belgium</td>
<td>46486</td>
<td>31</td>
<td>39</td>
<td>9.14</td>
</tr>
<tr>
<td>United States</td>
<td>46716</td>
<td>32</td>
<td>28</td>
<td>3.01</td>
</tr>
<tr>
<td>Austria</td>
<td>49900</td>
<td>33</td>
<td>38</td>
<td>6.84</td>
</tr>
<tr>
<td>Finland</td>
<td>51062</td>
<td>34</td>
<td>31</td>
<td>3.64</td>
</tr>
<tr>
<td>Sweden</td>
<td>52057</td>
<td>35</td>
<td>25</td>
<td>2.93</td>
</tr>
<tr>
<td>Netherlands</td>
<td>52321</td>
<td>36</td>
<td>33</td>
<td>3.88</td>
</tr>
<tr>
<td>Iceland</td>
<td>52557</td>
<td>37</td>
<td>34</td>
<td>4.04</td>
</tr>
<tr>
<td>Denmark</td>
<td>62332</td>
<td>38</td>
<td>41</td>
<td>17.69</td>
</tr>
<tr>
<td>Ireland</td>
<td>63185</td>
<td>39</td>
<td>37</td>
<td>5.99</td>
</tr>
<tr>
<td>Switzerland</td>
<td>64015</td>
<td>40</td>
<td>29</td>
<td>3.3</td>
</tr>
<tr>
<td>Norway</td>
<td>94353</td>
<td>41</td>
<td>40</td>
<td>12.91</td>
</tr>
</tbody>
</table>

* The Puerto Rico figure is for 2001, the last year available in the category.
### Exhibit 2: Tanzania Statistics Comparative Panel Data

<table>
<thead>
<tr>
<th>Category</th>
<th>Tanzania</th>
<th>World</th>
<th>Low Income</th>
<th>High Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population, total (millions)</td>
<td>41</td>
<td>6,614</td>
<td>953</td>
<td>1,061</td>
</tr>
<tr>
<td>Population growth (annual %)</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Surface area (sq. km) (thousands)</td>
<td>947</td>
<td>133,946</td>
<td>19,311</td>
<td>35,150</td>
</tr>
<tr>
<td>GNI, Atlas method (current US$) (billions)</td>
<td>16</td>
<td>52,851</td>
<td>439</td>
<td>39,743</td>
</tr>
<tr>
<td>GNI per capita, Atlas method (current US$)</td>
<td>400</td>
<td>7,990</td>
<td>461</td>
<td>37,460</td>
</tr>
<tr>
<td>GNI, PPP (current international $) (billions)</td>
<td>47</td>
<td>65,815</td>
<td>1,269</td>
<td>38,564</td>
</tr>
<tr>
<td>GNI per capita, PPP (current international $)</td>
<td>1,130</td>
<td>9,950</td>
<td>1,332</td>
<td>36,349</td>
</tr>
<tr>
<td>Life expectancy at birth, total (years)</td>
<td>55</td>
<td>69</td>
<td>59</td>
<td>79</td>
</tr>
<tr>
<td>Fertility rate, total (births per woman)</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Adolescent fertility rate (births per 1,000 women ages 15-19)</td>
<td>130</td>
<td>51</td>
<td>90</td>
<td>20</td>
</tr>
<tr>
<td>Contraceptive prevalence (% of women ages 15-49)</td>
<td>…</td>
<td>60</td>
<td>37</td>
<td>…</td>
</tr>
<tr>
<td>Births attended by skilled health staff (% of total)</td>
<td>…</td>
<td>65</td>
<td>43</td>
<td>99</td>
</tr>
<tr>
<td>Mortality rate, under-5 (per 1,000)</td>
<td>116</td>
<td>68</td>
<td>120</td>
<td>7</td>
</tr>
<tr>
<td>Immunization, measles (% of children ages 12-23 months)</td>
<td>90</td>
<td>82</td>
<td>78</td>
<td>93</td>
</tr>
<tr>
<td>Primary completion rate, total (% of relevant age group)</td>
<td>85</td>
<td>87</td>
<td>65</td>
<td>98</td>
</tr>
<tr>
<td>Ratio of girls to boys in primary and secondary education (%)</td>
<td>..</td>
<td>95</td>
<td>91</td>
<td>99</td>
</tr>
<tr>
<td>Category</td>
<td>Tanzania</td>
<td>World</td>
<td>Low Income</td>
<td>High Income</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
<td>----------</td>
<td>-------</td>
<td>------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Prevalence of HIV, total (of population ages 15-49)</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>GDP (current US$) (billions)</td>
<td>17</td>
<td>54,891</td>
<td>465</td>
<td>40,462</td>
</tr>
<tr>
<td>GDP growth (annual %)</td>
<td>7</td>
<td>4</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Inflation, GDP deflator (annual %)</td>
<td>9</td>
<td>5</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Agriculture, value added (% of GDP)</td>
<td>46</td>
<td>3</td>
<td>25</td>
<td>2</td>
</tr>
<tr>
<td>Industry, value added (% of GDP)</td>
<td>17</td>
<td>28</td>
<td>29</td>
<td>26</td>
</tr>
<tr>
<td>Services, etc., value added (% of GDP)</td>
<td>37</td>
<td>69</td>
<td>46</td>
<td>72</td>
</tr>
<tr>
<td>Exports of goods and services (% of GDP)</td>
<td>21</td>
<td>27</td>
<td>34</td>
<td>26</td>
</tr>
<tr>
<td>Imports of goods and services (% of GDP)</td>
<td>27</td>
<td>27</td>
<td>45</td>
<td>26</td>
</tr>
<tr>
<td>Gross capital formation (% of GDP)</td>
<td>16</td>
<td>22</td>
<td>27</td>
<td>21</td>
</tr>
<tr>
<td>Revenue, excluding grants (% of GDP)</td>
<td>...</td>
<td>27</td>
<td>...</td>
<td>27</td>
</tr>
<tr>
<td>Cash surplus/deficit (% of GDP)</td>
<td>...</td>
<td>(1)</td>
<td>...</td>
<td>(1)</td>
</tr>
<tr>
<td>Time required to start a business (days)</td>
<td>29</td>
<td>43</td>
<td>54</td>
<td>23</td>
</tr>
<tr>
<td>Market capitalization of listed companies (% of GDP)</td>
<td><strong>4.2 (2005)</strong></td>
<td>121</td>
<td>..</td>
<td>124</td>
</tr>
<tr>
<td>Military expenditure (% of GDP)</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Mobile cellular subscriptions (per 100 people)</td>
<td>20</td>
<td>51</td>
<td>16</td>
<td>101</td>
</tr>
<tr>
<td>Internet users (per 100 people)</td>
<td>1</td>
<td>21</td>
<td>4</td>
<td>65</td>
</tr>
</tbody>
</table>
### Bibliography


Introduction

‘All across the world, in every kind of environment and region known to man, increasingly dangerous weather patterns and devastating storms are abruptly putting an end to the long-running debate over whether or not climate change is real. Not only is it real, it’s here, and its effects are giving rise to a frighteningly new global phenomenon: the man-made natural disaster.’

In a speech on April 3, 2006 entitled “Energy Independence and the Safety of Our Planet,” President Barack Obama remarked on the irrefutable importance of mitigating climate change. While scientists have been studying the effects of climate change and warning the public about its detrimental and irreversible effects for decades, it has not been until recently that the world has started to pay attention and has made attempts to limit these effects. One of the most important organizations credited with “popularizing” the issue of climate change and with establishing arguably the most commonly accepted facts about climate is the Intergovernmental Panel on Climate Change, or the IPCC. The IPCC was established in 1988 as a joint initiative of the United Nations Environment Programme (UNEP) and the World Meteorological Organization (WMO). The IPCC focuses on the science of climate change and has the self-proclaimed objective of, “… provid(ing) the world with a clear scientific view on the current state of climate change and its potential environmental and socio-economic consequences.” The IPCC itself does not conduct research, but instead reviews research of scientists from around the world in order
to establish a well-informed consensus as to the current state of climate change in the world. Many pinpoint the success of the IPCC to its dual-nature of being grounded in science, but also of being an intergovernmental body involving the participation of countries that are participants in the United Nations and the WMO. The IPCC’s latest report was “Climate Change 2007,” which was pivotal in furthering the international community’s understanding of climate change and in helping the organization win the Nobel Peace Prize in 2007.

Although initially scientists and green activists were at the forefront of the recent push towards containing the inevitable effects of climate change, individual countries and international groups have also made their own efforts to diminish their reliance on non-renewable sources of energy like oil and coal. One important step was the creation of the United Nations Framework Convention on Climate Change, or the UNFCCC, which was created primarily in response to the IPCC’s Assessment Report in 1990, the organization’s first publication that stressed the necessity of an international coalition to address successfully the environmental situation at hand. The UNFCCC is an international treaty among most of the countries of the world that calls on the international community to begin devising plans for how they will confront climate change together.

One of the major successes of the UNFCCC has been the creation of the Kyoto Protocol, which holds participant countries to individualized emission reduction targets. This represents the first international attempt to mitigate climate change in a legally binding way. While the Kyoto Protocol primarily involves the direct participation of industrialized countries, developing countries are involved indirectly through means such as the Clean Development Mechanism (CDM) projects. These projects represent efforts to reach the emission reduction targets for the industrialized countries, but also to help develop sustainable environmental practices within developing countries. One of the main ways that the CDM projects achieve this goal in developing countries is through the implementation of renewable energy projects, like the creation of wind farms, within these
countries. Throughout this paper, I will be focusing on the developing country of Brazil and assessing the effectiveness of CDM projects in modifying that country’s environmental behavior. While it is obvious that the CDM project initiative of the Kyoto Protocol has prompted an increase in the number of renewable projects present in Brazil, mainly hydroelectricity dams, it is questionable whether the Kyoto Protocol has had any further positive environmental effect on the country.

I will begin by outlining the basics of the Kyoto Protocol and CDM projects. I will then focus specifically on the implementation of CDM projects in Brazil and on Brazil’s environmental policy. What I will find is that the Kyoto Protocol and the CDM projects have not had as positive an impact in Brazil as had been hoped. I will then outline three hypotheses that could potentially explain the inadequacy of progress seen in Brazil. These three hypotheses, broadly stated, are: Brazil lacks the resources necessary to bring about meaningful change in its environmental policy; regardless of resources, Brazil lacks the will to change its environmental behavior; and finally the Kyoto Protocol itself is at fault for the lack of environmental progress seen in Brazil. What I conclude is that a fusion of all three of these hypotheses correctly explains Brazil’s environmental behavior.

Overview of the Kyoto Protocol and the Clean Development Mechanism (CDM) Projects

The Kyoto Protocol was a protocol of the United Nations Framework Convention on Climate Change (UNFCCC), a treaty established in June of 1992. The UNFCCC had always been working towards mitigating the effects of climate change, but it was not until the Kyoto Protocol that the idea of a law-enforced international reaction would be put into place. The UNFCCC clearly states that the goal of the Kyoto Protocol is to create, “legally binding emissions targets that strengthen the world’s reaction to the problem of climate change.”

For each of the industrialized countries, known as “Annex I countries” in the
protocol jargon, the initiative sets legally binding greenhouse gas emission reduction targets applicable to the six principal greenhouse gases: carbon dioxide, methane, nitrous oxide, sulfur hexafluoride, HFCs and PFCs. Each of the Annex I countries is bound to achieve at least a 5.2% reduction of greenhouse gas (GHG) emissions below a baseline year defined as either 1990 or 1995, depending on the greenhouse gas, but most countries commit to more stringent targets. This, more precisely, means that each country must reduce emission levels at least 5.2% below those levels reported in 1990 or 1995.

Developing countries are not officially involved in the Kyoto Protocol, as they are not assigned emission reduction targets. That being said, many of the developing countries signed and ratified the Kyoto Protocol, like Brazil in 2002, and have pledged to try to reduce their greenhouse gas emissions without specific targets being assigned. Furthermore, many theorists like Babiker, Reilly and Jacoby argue that a sort of “ripple effect” from those industrialized nations with reduction targets to those countries without, mainly developing countries, will occur and in a sense force sustainable environmental practices on these other countries;

*Economic trade links among countries will transmit effects of greenhouse-gas control measures adopted by one set of nations, in a ripple effect, to countries that may not have agreed to share the burdens of control. For example, emission restrictions under the Kyoto Protocol will increase the cost to Annex B regions of using carbon-emitting fuels, thereby raising manufacturing costs of their energy-intensive goods, some of which may be exported to developing countries. The restrictions also will lower global demand for carbon-emitting fuels, reducing their international prices.*

Besides this “ripple effect,” there also exist three market-based provisions within the protocol, two of which encourage the participation of developing nations as well, even if in an indirect manner. These three market mechanisms are: Emissions Trading, the Clean Development Mechanism (CDM) and Joint Implementation (JI). These measures are not only a way of including the developing countries, but also they are an attempt to help
industrialized countries meet their targets and to increase the incentives for them to do so.

With Emissions Trading, each country receives allowances for its emissions. The number of allowances received by each country is equal to its emission reduction target, i.e. the level that is equivalent to the country’s specified reduction below 1990 levels. The allowances are distributed to the countries in single units, called assigned amount units, or AAUs. The idea behind granting countries these allowances is that if a country reduces its emissions below its respective targets, it can then sell the extra allowances to another country to apply to its target. The exact details of Emissions Trading are described in Article 17 of the Kyoto Protocol:

The Parties included in Annex B may participate in emissions trading for the purposes of fulfilling their commitments under Article 3. Any such trading shall be supplemental to domestic actions for the purpose of meeting quantified emission limitation and reduction commitments under that Article.

This mechanism has created a very competitive and profitable market, known as the carbon market.

The Joint Implementation mechanism is very similar to the Clean Development Mechanism. Each involves the creation of emission reduction projects; it is the location of these projects, however, that is the difference between JI and CDM projects. Joint Implementation projects are those that are managed by Annex I countries and located in another Annex I country. This leads to the creation of credits, known as Emissions Reduction Units (ERUs), for the managing country, and physical improvement of the environment in the industrialized country where the project is located. The Kyoto Protocol specifically outlines this process in Article 6:

1. For the purpose of meeting its commitments under Article 3, any Party included in Annex I may transfer to, or acquire from, any other such Party emission reduction units resulting from projects aimed at reducing anthropogenic emissions by
sources or enhancing anthropogenic removals by sinks of greenhouse gases in any sector of the economy…

Certain criteria must be met in order for a Joint Implementation project to be deemed legitimate. First of all, both countries involved must give their consent. Furthermore, the project must exemplify “additionality” and must follow the guidelines of the protocol. “Additionality” is the idea that projects must prove that they will be adding environmental benefits that would not otherwise be present. There are four steps to prove “additionality”: identification of alternatives, investment analysis, barrier analysis and common practice analysis.

For the purposes of this paper, I will be focusing on the Clean Development Mechanism (CDM) projects, which are outlined in Article 12 of the Kyoto Protocol:

2. The purpose of the clean development mechanism shall be to assist Parties not included in Annex I in achieving sustainable development and in contributing to the ultimate objective of the Convention, and to assist Parties included in Annex I in achieving compliance with their quantified emission limitation and reduction commitments under Article 3.

This idea was proposed by Brazil at the original negotiations of the Kyoto Protocol; “In the Kyoto Protocol negotiations in 1997, Brazil proposed that differentiated targets for reductions in greenhouse gas emissions should be established, corresponding to each individual country’s historical contribution to the increase in global temperatures.” Under the Clean Development Mechanism system, projects that are established in developing countries can earn credits known as Certified Emissions Reductions (CERs). Each Certified Emissions Reduction is equal to one ton of carbon dioxide, often written as one CO2e (one CO2 equivalent). These reduction credits can then be applied to an industrialized
country’s greenhouse gas emissions reduction target. CDM projects have been in existence since the beginning of 2006, and, to date, there have already been more than 1,650 projects registered, although many thousand more are in the validation process. Overseen by the CDM Executive Board, the registration and issuance of CER’s for each project is an extensive process. The breakdown of the steps can be seen in Figure 1.

One of the main incentives for these projects on the part of the industrialized countries is that it is often cheaper to carry out the projects approved under the Kyoto Protocol in developing countries rather than in industrialized countries. And, in turn, many developing countries are attracted to the idea because it means a possible influx of foreign capital, fostering development in their countries;

…next to the objective of reducing GHG emissions, CDM projects shall also aim at supporting sustainable development in developing countries…a CDM project enables the transfer of a low-carbon technology to a developing country which would be in accordance with that country’s development needs and priorities.

Although the industrialized countries are in charge of implementing the projects, the developing countries are responsible for the creation of a designated national authority (DNA) to oversee the implementation of CDM projects in their country. The DNA is responsible for approving CDM projects, using its country’s personalized criteria for approval, before the CDM Executive Board can review the projects for approval. Most DNAs use an approval process that consists of two steps: a screening of the ideas for projects, followed by a final approval of the project design requests. The DNA also has the responsibility of reporting to the CDM Executive Board periodically and of determining additionality of projects.

There are a number of different ways that CDM projects can be funded and managed: by an Annex I country; by a group of Annex I countries working together; by non-Annex I
countries; by private companies; or by any mix of countries and/or companies. A dominant way that governments and private companies often pool their money is through the creation of an overarching fund, which can then be used to invest in CDM projects. One of the main funds established around the emergence of CDM projects is the World Bank’s Prototype Carbon Fund, or PCF. The PCF was set up as a fund for governments and companies in Europe and Japan to use towards the development of CDM projects. It was established in April of 2000 with three main objectives: the creation of high-quality emissions reductions; knowledge dissemination; and the formation of public-private partnerships. The fund was created initially as a way to attempt to understand and test the procedures for creating a market for emission reductions generated through projects associated with the Kyoto Protocol, and more specifically with CDM and JI projects.

The Prototype Carbon Fund is comprised of “investors who want hands-on experience acquiring emissions-reduction credits under the CDM.” Investors that comprise the PCF include governments of Canada, Finland, Japan, the Netherlands, Norway and Sweden and companies from Europe and Japan, including BP-Amoco, Norway’s Statoil and Mitsubishi. Using contributions made by the companies and governments involved, the PCF invests these funds in projects designed to reduce emissions. The projects, focusing on areas such as biomass and reforestation, are consistent with the guidelines of the Kyoto Protocol and the framework for JI and CDM projects. After the completion of a project, each member of the PCF receives a pro rata share of the emission reduction credits generated from the project.

The PCF seems to focus a large portion of its funds in Latin America, primarily because the countries of Latin America have shown great eagerness to encourage carbon projects, even before the establishment of the Kyoto Protocol. Furthermore, Latin America has made huge strides towards developing the legal structures necessary for these projects to be successful:
‘Latin American has the market orientation necessary for carbon trading, knows how the CDM works and has been extremely proactive in designing and promoting projects,’ says Eduardo Dopazo, an Argentine consultant on the staff of the PCF.

In Latin America, the PCF has already invested in nine projects focusing on renewable energy, energy-efficiency and/or forestry projects. Of their total funds, 30% were invested in Latin American from 2000 to 2003, while 23% were invested in Eastern Europe, 20% in Africa and 12% in East Asia.

The PCF is not the only fund investing heavily in projects in Latin America. A study funded by the European Union estimates that there are a total of 789 CDM projects in Latin America. Of the total number of CDM projects in Latin America, Brazil is responsible for 41% of them, Mexico for 25% and Chile for 7%. Significant numbers of projects are also seen in Argentina, Colombia, Peru, Honduras, Ecuador and Guatemala. Of all the CDM projects in Latin America, projects focusing on renewables seem to dominate, making up 54% of the projects. Following renewables, one sees many projects focused on reducing emissions caused by agricultural practices and landfills. For further breakdown, please see Figures 2 and 3.

Many question whether CDM projects, like those instituted by the PCF, will actually benefit the developing countries of the Latin American region. Some are thoroughly convinced that these projects will be positive for the countries:

‘This project has enormous social benefits,’ says Marta Castillo, head of the office within Colombia’s Environment Ministry that prepares projects under the Kyoto Protocol. ‘It shows how the PCF can stimulate projects in Latin America and benefit poor communities.’

Other possible benefits of CDM projects for countries in Latin America are generating
jobs, restoring the environment, reducing pollution and improving public health. On the other hand, people have raised questions about the potential negative effects of these projects. For example, some CDM projects have the capability of destroying the natural habitats of both people and animals, causing more greenhouse gas emissions than mitigating and not confronting the true environmental problems from which the host country is suffering.

**Effectiveness of Brazil’s Environmental Policies**

In this section, I will focus on Brazil’s specific involvement in CDM projects with the goal of determining the effectiveness of the country’s environmental policy in general. I will first outline the specifics of Brazil’s involvement in the Kyoto Protocol through its CDM projects and discuss why there is such a large number of projects in Brazil compared to the rest of Latin America. Subsequently, I will define a proper measure to judge effectiveness of Brazil’s CDM projects and environmental policies, in general, in order to come to a conclusion as to whether or not CDM projects have been effective in Brazil.

As has been mentioned before, Brazil is the third largest contributor of CDM projects in the world and is the leading contributor in the Latin American region. Of the total expected average annual CER’s from registered projects worldwide, Brazil is estimated to be responsible for 6.52% of them, which is equivalent to an estimated 2,019,987.4 credits. For a further breakdown of CER generation by country, please see Figure 4. According to the UNFCCC website, Brazil has 164 registered projects, excluding a considerable number of projects that are in the process of registration and validation. These projects account for about 8.93% of the total number of registered CDM projects, which is approximately 1,873 projects. For more information about the registered CDM projects by host country, please refer to Figure 5.

According to a study conducted by the United Nations Environment Programme (UNEP) Risoe Centre on Energy, Climate and Sustainable Development in January 2009, the state of São Paulo has the highest concentration of projects among the regions of Brazil.
These projects in São Paulo represent 99 out of an estimated country total of 424 projects, including both those that are registered and those that are unregistered. Other states with notable numbers of CDM projects are Minas Gerais with 66 projects, Rio Grande do Sul with 36 and Santa Catarina with 34 projects. The methodology of methane avoidance is the most prominent in Brazil with 125 of the projects making use of this technology. Methane avoidance is a technology that can take many forms, but generally involves the capturing of methane that otherwise would be released into the atmosphere and transforming captured methane into a more environmentally friendly form, like energy, for example, through its combustion in the cogeneration process. Biomass energy comes in a close second behind methane avoidance with 106 of the projects in Brazil using this clean energy method. Many other projects in Brazil also use hydro (84 projects) and landfill (45 projects).

Being the third largest host country of CDM projects in the world and the largest in Latin America, Brazil has shown a major commitment to CDM project implementation. One might ask why there exists such a large number of projects in this country? First of all, Brazil is the largest country in Latin America and is often looked to as the leader of the region; therefore, it is not surprising that Brazil would also lead the region in its involvement under the Kyoto Protocol. Brazil also has the strongest economy in comparison to the rest of the countries of Latin America. In terms of GDP ranking in 2009, Brazil was ranked 10th in the world, followed by Mexico in 12th place and Argentina in 24th. Another reason for Brazil’s exemplary CDM project implementation is that it is one of the largest emitters of greenhouse gases in the world. For this reason, it is under the most scrutiny among all the Latin American countries to take action to mitigate its negative environmental impact. In a study conducted by the Union of Concerned Scientists in 2006, Brazil was ranked the 17th highest emitter in the world in terms of total emissions, emitting an estimated 377.24 million metric tons of carbon dioxide, which equates to 2.01 tons per capita. Please refer to Figures 6 and 7 for more information.

Despite the considerable number of CDM projects in Brazil, are these projects
proving to be effective? And how is effectiveness actually measured? If I consider effectiveness in a narrow sense to be a reduction of greenhouse gas emissions, Brazil has not seen a reduction in total emissions since it began participating in the Kyoto Protocol’s CDM mechanism in 2002. In 1995, Brazil’s emissions equaled an estimated 289.53 million metric tons of carbon dioxide. This level then increased to 344.93 in 2000 and up to 377.24 million metric tons in 2006. However, even though there is not a decrease in total emissions, there has been an overall reduction in the growth rate of Brazil’s emissions. Between 1995 and 1996, there was a 6.09% increase in emissions. Comparatively, the growth rate decreased to 1.2% from 2000 to 2001 and increased only slightly to 1.78% from 2005 to 2006, but still remains a decrease from the 1995-1996 period. This shows that although the total amount of emissions is still increasing since the initial negotiations of the Kyoto Protocol in 1995, Brazil’s rate of increase of emissions, for the most part, has declined.

Because Brazil is not an Annex I country, it is not legally bound to reduce its emissions in any way under the Kyoto Protocol. But the fact that Brazil is still, even without mandate, taking initiative to halt its growth of emissions through the implementation of CDM projects is noteworthy. Many of these CDM projects, for example, have focused on increasing reliance on renewable energy, a sector that was already existent in Brazil prior to the implementation of CDM projects but that has been greatly strengthened with the creation of the CDM. Brazil is even considered to have one of the cleanest energy sectors in the whole world. Resulting from this relatively clean energy sector, in comparison to the United States, energy emissions in Brazil are 0.50 tons of carbon per person, while in the United States this number is drastically higher at 5.58 tons of carbon per person. Therefore, one can see that while not mandated to cut its emissions under the stipulations of the Kyoto Protocol, Brazil is still playing a significant role through its production of renewable energy.

Part of this excellence in terms of renewable energy production can be attributed to
Brazil’s manufacturing of ethanol from sugarcane. Brazil’s ethanol is one of its renowned technologies, and its production has proven to be a model for the rest of the world. Ethanol proves to be a viable and renewable substitute for gasoline and, because of its renewability, is better for the environment; “Brazilian ethanol from sugarcane can cause a 70-90% reduction in greenhouse gas emissions when compared to gasoline.” For almost three decades, Brazil has been using ethanol as a substitute for fossil fuel, which has drawn the attention of many other countries like the United States. In a meeting with President Luiz Inácio Lula da Silva, President Barack Obama admitted his admiration of Brazilian ethanol production;

‘I have a lot of admiration for the big steps that the Brazilian government has taken in its development of biofuel. It is an investment to which Brazil has been dedicated for a long time. My policy with regards to biofuel is that the United States needs to focus its attention towards finding the same potential for clean energy. We must learn together with Brazil.’

Brazil’s ethanol production from sugarcane is definitely contributing to the fight against climate change, specifically by reducing emissions usually caused by gasoline. Besides ethanol, other renewable technologies that contribute to Brazil’s clean energy sector are biomass energy and hydroelectricity.

Even with its involvement in CDM projects and its ethanol production, is Brazil doing enough to help mitigate its contribution to climate change? Is participation as a host country for CDM projects and its production of ethanol sufficient to qualify Brazil as a serious proponent of environmental improvement? Although the focus of CDM projects is on reduction of emissions, one must look further and assess the environmental implications of these projects more generally. The implicit goal of CDM projects is reduction of emissions, but furthermore, it is to develop sustainable behavior within the country where
the projects are implemented;

\[\text{in accordance with Article 12, the purpose of the clean development mechanism is to assist Parties not included in Annex I to the Convention in achieving sustainable development and in contributing to the ultimate objective of the Convention, and to assist Parties included in Annex I in achieving compliance with their quantified emission limitation and reduction commitments under Article 3 of the Kyoto Protocol.}\]

The renewable sector in Brazil does not seem to be a major problem for the country, as it is one of the cleanest in the world and has the potential for significant growth in the future. However, although making strides in its production of renewable energy, including its production of ethanol, there are two major problems with Brazil’s environmental policies: the use of questionable technologies, like hydropower, and the fact that CDM projects do not focus on the true environmental problems that the country is facing.

The first of Brazil’s problems with its environmental policy is that some of the techniques used to achieve development in its renewable sector are extremely controversial and have negative environmental implications. Prime examples of this are hydroelectric projects; the Kyoto Protocol has received much criticism for its acceptance of such projects as an approved CDM. With hydroelectricity, it is questionable whether the environmental benefits of this mechanism outweigh its environmental destruction. These projects often involve the destruction of natural habitats and the resettlement of a large number of people. For example, in two different dam projects implemented in the 1980’s by the World Bank in Sobradinho and Macadinho, in northeastern and southeastern Brazil, the resultant number of displaced people reached about 70,000. Besides being displaced, those displaced did not even receive adequate compensation for the hardships they experienced from their displacement.

Besides displacing locals, hydroelectricity can also pose a danger to the lives of
inhabitants in the vicinity of the dam;

Imagine people encircled by water from all sides, running for [the] safety of their lives. Many of them climbed over trees. Their houses and other properties washed away. Boribinda, Gopalur, Durrie and Birdih are a few villages among 52 villages in the state of Bihar which were submerged during August 26-27 [1991] due to the ongoing construction of the Chandil dam…

Besides physical destruction to surrounding environment, many also question whether the dams are actually improving the energy matrices, i.e. making the electricity grids “cleaner.” It is estimated that hydroelectric dams have the potential of emitting higher levels of carbon dioxide and methane than power plants that run on fossil fuels:

This is because large amounts of carbon tied up in trees and other plants are released when the reservoir is initially flooded and the plants rot. Then after this first pulse of decay, plant matter settling on the reservoir’s bottom decomposes without oxygen, resulting in a build-up of dissolved methane. This is released into the atmosphere when water passes through the dam’s turbines.

For example, Philip Fearnside of Brazil’s National Institute for Research in the Amazon in Manaus estimates that the emissions from the hydroelectric dam of Curuá-Una in Pará, Brazil in 1990 were three-and-a-half times larger than the emissions that would have been generated had the same amount of electricity been produced using oil.

Besides the dams in Sobradinho and Macadinho described above, hydro is prevalent throughout Brazil, as 76% of its energy is generated from hydro technology. Not only does the Kyoto Protocol allow hydropower to be included as a methodology, Brazil has also made efforts to increase investment in this methodology. President Lula da Silva’s
administration has even changed the power sector regulations to increase the attractiveness of investing in hydropower by decreasing environmental risks to investors; “The new regulations mean that hydropower projects cannot be presented to public tender until after the governmental energy planning agency grants an environmental license.” The guaranteeing of an environmental license means that once up for public tender, the projects will have lower risk of disapproval than those projects put up that may or may not be granted an environmental license. This relaxing of the regulations surrounding the implementation of hydropower projects has led to a large increase in the number of dams in Brazil, thus making hydropower projects the third most implemented type of renewable project in Brazil.

Although the renewable CDM projects seem to be flourishing in Brazil despite their use of questionable renewable techniques, Brazilian CDM projects, for the most part, do not truly address the major environmental problems that the country is facing. While for most countries in the world their energy sector contributes to the majority of greenhouse gas emissions, it is the unsustainable land use and forestry that contribute the most to greenhouse gas emissions in Brazil. In 1994, only 17% of Brazil’s emissions came from energy production, while 81% came from agriculture, land use and forestry. For further breakdown of Brazil’s emissions, please see Figure 8. The country presents an atypical emissions profile that can be categorized as LULUCF, Land Use, Land Use Change and Forestry. Instead of focusing on the real environmental problems at hand in Brazil, i.e. problems caused by agriculture, land use and forestry, Brazilian CDM projects primarily focus on renewable technologies like hydropower.

One of the main environmental problems that Brazil is facing is deforestation; “At the current rate of deforestation, around one-third of the forest in Amazonas (a state in the Amazon region of Brazil) will have been lost by 2050, releasing a colossal 3.5 billion tonnes of carbon dioxide into the atmosphere.” The Brazilian Amazon rainforest lies primarily in the Northern region of the country and is made up of the states of: Acre, Amazonas, Pará,
Roraima, Amapá, Rondônia, Mato Grosso and Tocantins. In a study conducted in 2008 by the Instituto Nacional de Pesquisas Espaciais (Inpe), it was estimated that the area of deforestation that has already occurred in Brazil to date is bigger than the total area of the three states that make up the southern region of Brazil: Paraná, Santa Catarina and Rio Grande do Sul. This area is equal to over 576,000 square kilometers of land. It was further estimated, that in the last five months of 2007, 1,250 square miles of forest suffered from deforestation in Brazil.

Some of the main causes of this deforestation are: the growth of the Brazilian population, logging, cattle herding, the construction of hydroelectric dams, the building of roads, mining and grain production. The main impacts caused by such high rates of deforestation are: the loss of some of the most precious biodiversity in the world, a change in wind patterns and the releasing of huge quantities of greenhouse gas, primarily carbon dioxide, into the atmosphere. While the deforestation is occurring within the boundaries of Brazil, most of its effects have repercussions for the world as a whole. Experts have estimated that if this problem is not dealt with, it will become an even larger problem in the near future, because even if Brazil is not currently experiencing the effects of climate change, it will start to feel them very soon. As Fearnside argues, “…the vastness of the remaining forests means that the impacts of continued clearing are far more serious than the – already severe – impacts resulting from their loss to date.”

Unfortunately, given such high rates of deforestation, one does not see Brazil trying to reduce its reliance on deforestation, and the country has even claimed a right not to be mandated to reduce its deforestation, or to reduce any of its emissions for that matter, in potential future international agreements after the termination of the Kyoto Protocol in 2012;

*In international negotiations, Brazil points out that climate change is driven more by the accumulation of greenhouse gases in the atmosphere than by yearly emissions, primarily because the most important greenhouse gas (carbon dioxide) remains in the*
atmosphere for more than a century on average. Yearly emissions data therefore generally overestimate developing countries’ contributions to climate change, and underestimate that of developed countries. Brazil therefore says that it will not limit its greenhouse gas emissions until the middle of the century (around 2050).

Adding to the country’s refusal to be mandated to reduce its emissions, particularly those related to deforestation, the Kyoto Protocol under its CDM projects does not give any incentive to developing countries to avoid deforestation. For the first commitment period of the Kyoto Protocol culminating in 2012, the avoidance of deforestation was not included as an approved CDM methodology;

*Investment interest in carbon with a view to short-term returns is likely to be limited, given the fact that the agreement over the Kyoto Protocol reached in July 2001 excludes credit for forest maintenance in the Clean Development Mechanism (CDM) during the protocol’s first commitment period (2008-2012). Should this be allowed in future commitment periods, Brazil could potentially gain substantially from CDM projects to reduce deforestation.*

While neither the country of Brazil nor the Kyoto Protocol seem to be prioritizing avoidance of deforestation, one also does not see a significant drive towards reforestation within Brazil. The technique of reforestation, put simply, is the planting of trees in a plot of land where forest has been cleared in an attempt to gain back the lost forest. Although many consider reforestation as a mechanism to undo the effects of deforestation, the advantages that come from reforestation do not nearly right the harm that comes from deforestation;

*But some experts doubt replanted forests provide reliable storage for atmospheric carbon. They say carbon must be stored for at least 100 years to benefit the atmosphere since that’s how long a CO2 molecule remains in the atmosphere before being reabsorbed.*
As the attempt to mitigate the effects of deforestation through reforestation seems to be unsuccessful, one must not consider these two mechanisms as necessarily coupled together. Instead, reforestation should be a methodology instituted to help the environment, while deforestation should be a methodology avoided, as it is destructive to it.

In the case of Brazil specifically, the country does not engage in the positive technique of reforestation and instead engages in the negative one of deforestation. Brazil must attempt to stop this deforestation before its effects are irreversible. Even though developing countries were initially not included in the Kyoto Protocol because it was the general consensus that these countries should have the same opportunities as the industrialized had when developing (for example, the absence of emission reduction targets), this general consensus has shifted somewhat, in part due to Brazil’s uncontrolled rates of deforestation, and it is suggested that these developing countries should play a more active role in this fight to combat climate change.

Although making huge strides in producing renewable energy through initiatives like CDM projects, Brazil needs to focus its attention more on the environmental problems at hand, specifically the deforestation of the Amazon and the usage of questionable technologies like hydropower. Besides helping the industrialized countries achieve their emission reduction targets more easily, the CDM projects are also intended to promote sustainable environmental behavior within the countries in which these projects are implemented. Because of the persistence of serious environmental problems in Brazil without an effort to combat them, the effectiveness of CDM projects in Brazil can be disputed.

Explaining Inadequacies in Brazil’s Environmental Policies

I will now discuss potential reasons why CDM projects in Brazil have not been
successful in addressing the larger environmental issues at work in Brazil in an attempt to come to a conclusion about the projects’ overall effectiveness within Brazil. The CDM projects have helped in reducing the growth of emissions within the country, but they have not helped to deal with Brazil’s more pressing problems of deforestation and usage of the controversial renewable technology of hydropower. In order to address this concern, I will be considering three different hypotheses. Two of the hypotheses will approach the issue at a domestic level. And the third hypothesis will deal with the concern at a more international level.

The first hypothesis focuses on resources. Perhaps Brazil is committed to improving its environmental performance, but because it is a developing country, it does not have the means to realize truly its environmental aspirations. As a number of theories of economic growth and development suggest, resources, particularly capital that can be used for investment, greatly influence a country’s ability to achieve development. The lack of these resources has made this task extremely difficult for developing countries. For example, in his book The Elusive Quest for Growth, William R. Easterly outlines an idea, originally put forth by Evsey Domar, which suggests that in order to achieve economic growth, a country requires more capital; “…the growth rate of GDP was proportional to last year’s investment/GDP. Foreign aid and private finance were to fill the financing gap between saving and the necessary investment to get high growth.” As has been stated here, this investment usually comes in the form of foreign aid;

The difference between the required investment and the country’s own savings is called the financing gap. Private financing is assumed to be unavailable to fill the gap, so donors fill the financing gap with foreign aid to attain target growth. This is a model that promised poor countries growth right away through aid-financed investment. It was aid to investment to growth.
A central problem for developing countries is finding sufficient capital to fund economic development. Without the funds to cover central things like developing economically, it is no wonder that these same countries will not have enough capital to develop environmentally.

Additionally, even if capital became available for environmental development within these countries, change in environmental behavior does not come at a cheap price. Renewable technologies like solar, for example, are extremely expensive alternatives to non-renewable technologies. Furthermore, the rates of return on the implementation of some renewable technologies will not be immediate. A good portion of industrialized countries and definitely in developing countries where resources are scarce will probably veer away from an investment with a long pay-back period and instead choose one where the pay-back period is shorter;

The trouble with mitigating climate change is that the benefits are uncertain and distant. Compared with investments that deliver clear benefits in the near future—such as education in developing countries, for instance, which commonly produces returns of around 10% a year—they do not look worthwhile. Conventional analysis would therefore suggest that those who want to make the planet a better place should invest in school in Malawi rather than in clean energy.

Because of the steep costs associated with technologies that work towards dealing with the problem of climate change and because of their lagging rate of returns, it is almost obvious why implementation of such technologies in developing countries are not these countries’ priority.

Besides a dearth of economic resources to effectively create environmental change, developing countries like Brazil also may not have the political resources to stimulate this change. Complex legislation and overall growth of a society are often the products of
effective institutions within a country. As Dani Rodrik argues in his paper “Institutions for High-Quality Growth: What they are and how to acquire them,” there are certain types of institutions of great importance for the stimulation of economic growth of a society. Rodrik breaks the kinds of institutions that support this growth into five basic categories: “property rights; regulatory institutions; institutions for macroeconomic stabilization; institutions for social insurance; and institutions of conflict management.” In describing these five different types of institutions, Rodrik stresses that there is not one model of implementation of these institutions that can be applied to each country. Various factors must be considered and accounted for in the implementation of any institution; “…there is no single mapping between the market and the set of non-market institutions required to sustain it.” Besides diversity of implementation and type of institutions, Rodrik further argues that the adoption of the right kind of institutions is a stimulant for growth within a society;

Think of institution acquisition/building as the adoption of a new technology…
Let us call this new technology a “market economy”…Adoption of a market economy in this broad sense moves society to a higher production possibilities frontier, and in that sense is equivalent to technical progress in economist’s parlance.

The cruel irony of Rodrik’s idea about institutions and growth is that unlike industrialized countries, developing countries often do not have the leisure of focusing on developing effective institutions. Due to limits on resources, they are instead forced to concentrate on resource extractions and development. Without these effective political institutions, a country is unable to achieve such growth that comes from institution creation, which is often a necessary component for legislative change. Concentrating on the situation of Brazil, implementing changes in environmental policy involves the creation of complex legislation. Brazil may not have the necessary resources to develop institutions to create such economic growth and subsequently such complex environmental legislation.
Our second hypothesis is that Brazil is not committed to its environmental performance at all, regardless of whether or not it has the resources to instigate changes in its policies or not. According to realist perspectives, there are five basic assumptions that states make about the international relations that lead to a semi-pessimistic view of the international system. Theorist Mearsheimer outlines these five assumptions in an article entitled “The False Promise of International Institutions” and they are: the international system is anarchic; states, with their militaries, have the capability of harming and even destroying each other; the intentions of states are unknown to each other; states are driven by the motive to survive; and, finally, states are rational and think strategically about their survival in the international arena. Taken together, these five basic assumptions used to explain the behavior seen in the international setting instill a sense of fear and a drive to fight for survival among the states. Furthermore, states seek to maximize their relative power to other states, usually through building their economies and militaries, a condition that gives states a sense of security in the competitive international setting that realists depict. When making decisions, states use a cost-benefit analysis to determine which action will further their economic and security interests. In this sense, Brazil, in assessing competing demands on resources and time, has determined that protecting the environment is not in its self-interest compared to advancing its economy. Brazil is more concerned with developing economically and, therefore, its environmental goals take a back seat to furthering the economy.

Our final hypothesis approaches the issue of lack of sufficient progress in environmental policy in Brazil with a more international scope, looking at the institution of the Kyoto Protocol as a whole, instead of focusing specifically on Brazil. This idea states that the problem is not that of Brazil, but instead of the Kyoto Protocol. There are two potentially central problems with the international institution of the Kyoto Protocol that might jeopardize the effectiveness of its implementation in countries like Brazil. The first problem with the institution is that there was not enough incentive given to the developing
countries to modify their environmental behavior and, subsequently, the countries did not prioritize adaptation of their environmental behavior towards more conscientious practices. As I have explained at the beginning of this thesis, the Kyoto Protocol only sets legally binding standards for Annex I countries. This inclusion of solely Annex I countries could be the result of many different factors. First, coming to an agreement on how to address the problem of climate change is never an easy task. Besides the sheer number of countries involved, making it hard for consensus to be reached, the developing countries were also not completely on board with the idea of being involved in the aftermath of a problem that they felt they played no part in causing. The principal stance of developing countries in the Kyoto Protocol negotiations was that this was a problem caused by industrialized countries and, therefore, it was their responsibility to fix it. In the Georgetown International Environmental Law Review in 1999, Deborah Cooper points exactly to this point that developing countries feel it would be unfair for them not to be allowed to grow without restriction, as industrialized countries had done during their industrialization period;

*Developed countries profess that stringent controls on all sources of GHGs in China and other developing countries will be necessary to prevent significant global warming. But such measures are viewed as offensive to developing countries like China – tantamount to telling these nations to stay poor. The message of restricted development touted by the develop world is seen as hypocritical since no one told them to halt their own industrialization for the sake of the environment.*

China’s senior energy researcher at the State Planning Commission in Beijing furthers this point about limiting energy use on developing countries by saying, “You try to tell the people of Beijing that they can’t buy a car or an air-conditioner because of the global climate-change issue…It is just as hot in Beijing as it is in Washington.” This popular stance by developing countries caused a tension between developed and developing countries during
the Kyoto negotiations over the issue of whether or not all countries should all be held to the same standard. Furthermore, there exist huge economic disparities among the countries of the world and because of this it is likely that countries with weaker economic situations would be hesitant to invest resources in climate change mitigation and in meeting legally binding targets.

For this reason among many others, the developing countries of the world were not delegated an active role in the Kyoto Protocol. Because of this, strong incentives were not created for the developing countries to reduce their emissions, as was the case for the industrialized countries. Without these strong incentives, developing countries did not feel the pressure to curb their emissions and, therefore, could not be held accountable for questionable environmental practices. The only way that these countries are motivated to curb their emissions by the Kyoto Protocol is through the CDM projects, but this encouragement is limited.

The second problem with the Kyoto Protocol that potentially causes its ineffectiveness in countries like Brazil is that the CDM includes projects that are not ideal given their broader environmental implications for the countries in which the projects are implemented. In analyzing the claim that the Kyoto Protocol has led to the suboptimal result that it includes projects of questionable value, I will first look at more general theories with regard to the design of international institutions. A debate exists in international theory over how international institutions should be designed. One of the main views of how international institutions are designed comes from a tradition known as the rational-choice analysis. Theorists Kormenos et al. subscribe to this rationalist ideology and argue that, “…states use international institutions to further their own goals, and they design institutions accordingly.” Furthermore, rationalists believe that the creation of international institutions is the result of states weighing the costs and benefits of issues and coming to agreement over the best way to further their own self-interests; “…actors have (well-behaved) preferences over various goals; and the pursuit of those goals is guided by their beliefs about
each others’ preferences and their relative costs and benefits of different outcomes; and actors are constrained by their capabilities.” In the specific case of the Kyoto Protocol, this would mean that the states participating in the negotiations of this agreement would have weighed the costs and benefits in order to best address the issue of climate change.

While rational-choice theory provides a somewhat convincing explanation of how international institutions are created, the theory fails to take into account key factors, like the behavior of the institutions after they are created, which are covered by the constructivist view of international institution design. For this reason, in order to explain the Kyoto Protocol’s inclusion of technologies that do not have optimal environmental effects for the countries in which they are implemented, like Brazil, I will utilize the constructivist argument on institutional design. Constructivists Barnett and Finnemore in their paper entitled “The Politics, Power and Pathologies of International Organizations” argue that international institutions, “…can become autonomous sites of authority, independent from the state ‘principals’ who may have created them…” The theorists further claim that the enormity of international institutions can render them inefficient and ineffective;

…the folk wisdom about bureaucracies is that they are inefficient and unresponsive. Bureaucracies are infamous for creating and implementing policies that defy rational logic, for acting in ways that are at odds with their stated missions, and for refusing requests and turning their backs on those to whom they are officially responsible.

Moisés Naím makes a similar claim about the size of international institutions causing problems for effective international foreign policy being achieved in his article “Minilateralism.” Throughout the article, Naím argues that international response is necessary for some of the world’s most pressing issues like climate change, but he also points out that the number of actors involved in addressing these issues must be limited;
We should bring to the table the smallest possible number of countries needed to have the largest possible impact on solving a particular problem. Think of this as minilateralism’s magic number…Same with climate change. There, too, the magic number is about 20: The world’s 20 top polluters account for 75 percent of the planet’s greenhouse gas emissions.

Naím swears that although those left out of negotiations will not like the idea of limiting the number of actors involved, the advantages of the small number, like not being caught in a stalemate, will outweigh the small disadvantages.

Utilizing the constructivist logic of Barnett and Finnemore and Naím’s argument, one can make the claim that the Kyoto Protocol is potentially too large and bureaucratic with 130 countries involved that it often becomes inefficient. Because of this inefficiency and ineffectiveness seen in institutions like the Kyoto Protocol, it would make sense that some corners would be cut in the negotiations process in order to facilitate action actually being taken, instead of no agreement being reached at all. For this reason, the focusing of the Kyoto Protocol on technologies that are easier to implement, such as hydroelectricity, is relatively easy to understand. This emphasis on hydroelectricity is upsetting for true environmentalists since it is, as has been previously stated, a highly controversial technology.

Besides over-emphasis on hydroelectricity, there is little emphasis placed on deforestation, the major environmental problem that Brazil faces. One element of institutional design scholars have focused on is the influence and distribution of voting power, which can lead to suboptimal results like those seen in the Kyoto Protocol. For example, many institutions do not give developing countries a principal role both in its creation and in its maintenance. The Kyoto Protocol is of this type of institution; the design is such that the developing countries are not active participants in the emission reduction scheme. As the developing countries are not involved in the Kyoto Protocol directly, this leads to it being a flawed institution in facilitating real change and real environmental strides
in developing countries like Brazil. Instead, the Kyoto Protocol only focuses on the Annex I countries that are the direct participants in the institution. This lack of emphasis on developing countries has resulted in the exclusion of deforestation as a CDM technology. Potentially, this is a result of the industrialized countries’ framing of the Kyoto Protocol being shaped by their own experiences with environmental problems where renewable energy was their priority, not deforestation.

Analysis of These Explanations

Lack of Resources

To restate, the first hypothesis to be examined is that Brazil does not have the means necessary to achieve change in its environmental policy. This lack of means can be viewed as a lack of financial capital to induce change or a lack of the necessary legislative power to enforce change. Let us first concentrate on a lack of economic means to stimulate change. Looking at the economic standings of those countries producing large improvement in their environmental policy like the United States and Japan, these countries have significantly more economic capacity, which can be directly linked to a capacity to improve environmental behavior. Compared to the United States and Japan, Brazil is still considered a developing country and, therefore, does not have as much flexibility to dedicate its resources to the environment as a fully industrialized country would. Although Brazil’s economy has been applauded in recent years, especially for its miraculous recovery from the financial crisis of 2008, the country still sits in the upper-middle income countries bracket and not in the high-income countries bracket, as designated by the World Bank.

Further utilizing information from the World Bank, Brazil can be compared to the United States and to China. The United States is used to compare Brazil to an industrialized country that is part of the Organization for Economic Co-operation and Development (OECD), which consists of thirty developed countries that are considered high-income
and that are committed to democracy and to the maintenance of a market economy. And China is used to compare Brazil to another developing country that is also targeted by the international community for its high emission rates of greenhouse gases (in the case of China, the highest emissions rates for a developing country). Comparing the GDP’s of these three countries in 2008, Brazil’s GDP is far below the other two at 1.58 trillion (US$), while China’s is 4.33 trillion (US$) and the United States’ 14.09 trillion (US$). In comparing Brazil’s GNI per capita at PPP in 2008 to those of China and the United States, one sees that Brazil sits between China and the United States. Brazil’s GNI per capita at PPP stood at 10,080 (current international $), while China’s was lower with 6,010 (current international $) and the United States’ higher with 46,790 (current international $). Finally, Brazil’s inflation rate in 2008 of 6% was 1% lower than China’s in the same year and 4% higher than that of the United States.

While not achieving the same economic status as the OECD country of the United States, and as China in terms of GDP, Brazil has been a major focal point in the world recently, especially due to its miraculous recovery after the economic crisis of 2008, which left the economies of traditionally “unstoppable” countries like the United States in disarray;

What appeared to be the worst global recession since the 1930’s left Brazil relatively unscathed. It was able to cut interest rates and the real held its value. Brazil turned out to be one of the last countries into the downturn and one of the first out, causing national celebration and not a little surprise, given what had gone before.

The country is achieving a feat it has never accomplished before: democracy, economic growth and low inflation. The country has experienced each of these separately and in groups of two over its history, but never has Brazil had all three working in unison in the way that they do now. For example, starting in 1964 under the country’s newly instated military government, Brazil experienced what was deemed “O Milagre Econômico
Brasileiro,” or the Brazilian Economic Miracle. As can be seen in the Figure 9, during the time that Brazil’s military regime (1964-85) ruled the country, the country saw astounding GDP growth rates, “taxas de crescimento do PIB,” up until the 1980’s. Authors Viola and Mainwaring describe the economic situation of Brazil during this period as follows, “Based on an income concentrating model, the Brazilian economy expanded at one of the fastest rates in the world between 1967 and 1980.” During this time, Brazil experienced unprecedented economic growth, but suffered immensely from extremely high interest rates and from a repressive military government, which was far from democratic.

The current state of Brazil with democracy, economic growth and low inflation can be attributed largely to the exemplary presidency of Lula’s predecessor, Fernando Henrique Cardoso, who was in power from 1995 to 2003. Cardoso is responsible for taking the reigns of the economy and steering it in the direction of stability. Before assuming his role as president in 1995, Cardoso served as Brazil’s finance minister and was mainly responsible for the country’s adoption of the Plano Real. This plan involved the introduction of a new currency, the Real, which uses a floating exchange rate. As former U.S. President Bill Clinton notes in his foreword to Cardoso’s memoir,

*Cardoso’s economic strategy as finance minister, the Plano Real, succeeded in curbing the hyperinflation that was crippling Brazil’s economy. Halting inflation dramatically boosted the real incomes of the poor, created a solid foundation for economic growth, and protected the country from many of the financial crises that have afflicted other developing nations.*

While Cardoso was responsible for the meat of the policies implemented to ignite the wondrous economic growth seen in Brazil today, Lula da Silva has also continued to follow the measures set by Cardoso and has added a few measures of his own.

Looking at hard data from the period before Cardoso’s presidency up to the
current period with Lula as president, one can see overall improvements and prosperity of the Brazilian economy. I will use the year 1990 as illustrative of the time immediately preceding Cardoso’s Real Plan and his time as president. The year 1995 represents the year following the implementation of the plan and Cardoso’s first year as president. I will finally utilize the year 2008, as it provides the most recent data available and will demonstrate the improvements seen in Brazil’s current economy and the economy under President Lula. Comparing the years 1990, 1995 and 2008, Brazil’s GDP (current US$) improved from 461.95 billion to 768.95 billion and finally to 1.58 trillion. As far as Brazil’s inflation rate in the same years, the rate in 1990 of 2,735% diminishes to 94% in 1995 and to 6% in 2008.

While it is true that Brazil’s economy has been flourishing in recent years, this does not mean that the country’s dark economic past cannot be ignored completely. The newly achieved economic stability of Brazil is a condition rarely seen in the history of Latin America, a history wracked by waves of economic crises, soaring inflation rates and the accumulation of insurmountable debts. Furthermore, the region is intrinsically linked and, therefore, the downfall of one of the countries does not bode well for the rest of the region. A perfect example of this linkage in the region is seen in Mexico’s 1982 default on its debt payments to finance the debt it had accumulated after the oil shocks of the 1970’s. This default on the part of Mexico caused mistrust by foreigners of the entire Latin American region, and, subsequently, credit started to disappear. This disappearance of credit led to the Latin American nations accumulating even more massive debts, which in turn caused inflation, capital flights and budget deficits. Mexico’s default caused Brazil’s debt to increase dramatically from $12 billion in 1973 to $70 billion in 1982. Also during this time, Brazil suffered from continually soaring inflation rates and low levels of economic growth. While the country’s economy slowly improved during the subsequent years, the recovery process for Brazil’s economy has been long and arduous. Even in the period between 1990 and 1995, Brazil’s inflation rate still averaged an estimated 764% per year.
Therefore, even with Brazil’s current commendable state of its economy, the country is so closely intertwined with the other countries of Latin America that the permanence of this economic condition is not entirely certain. Although Brazil made a miraculous recovery from the world economic crisis in 2008, it is not clear whether Brazil would be able to make a similar recovery if the crisis were to have started closer to home in Latin America. Furthermore, Brazil has three questionable characteristics of its economic policy, which could serve to be obstacles for its economy in the future, “…its suspicion of free markets; its faith in the wisdom of government intervention in business and finance; and persistently high interest rates.”

Although there is compensation for Brazil through the development of CDM projects, this does not mean that the country does not incur costs through their implementation. In order to curb something like deforestation, Brazil’s major environmental problem, the country would need a significant amount of capital in order to be successful, capital that even the thriving Brazilian economy may not possess. For this reason, the Brazilian government would most likely require assistance from foreign nations in the form of financing in order to be able to give financial incentive to ranchers, loggers and others to halt this deforestation. Brazil has already initiated programs like A Reserva Juma and O Programa Bolsa Floresta no Amazonas that pay Brazilians not to cut down trees, but these measures have not been sufficient to reduce the still significant amount of deforestation occurring within the country. The country has even conceded that it lacks the necessary means to be successful in reducing its rate of deforestation. Brazil has pledged to cut deforestation rates by about 70% in the coming ten years, but with the stipulation that it must receive appropriate compensation to do so. Brazil needs assistance from outside nations and may not be able to achieve meaningful changes in its environmental policies without such assistance; “Developing economies may be technically able to make the sorts of near-term emissions cuts the world needs, but they are not going to pursue them effectively unless they get the right assistance from the world’s wealthier nations.”
Another area in which Brazil may lack the resources necessary in order to change its environmental behavior is in terms of its legislation; “Significant institutional gaps, especially between legislative power and law enforcing authorities, makes enforcing environmental laws in the Amazon region very difficult.” Brazil’s specific environmental legislation can be broken down into two distinct periods: before and after the year 1981. The main distinction between these two periods is based on what exactly constituted pollution. In the time before 1981,

…pollution generally stood for industrial emissions that did not conform to the standards established by law and technical guidelines. At that time, based on the assumption that every production activity caused an impact on the environment, pollutant emissions were fully tolerated, provided that they were kept within preset limits.

While seemingly sufficient prior to 1981, this definition of pollution proved to be too narrow starting in 1981. In the second phase of Brazil’s environmental legislation, pollution was more generally considered anything that has caused damage to the atmosphere, even if created in compliance with emission allowances. Environmental legislation was also dedicated a whole chapter in the Federal Constitution in October of 1988. Within this chapter, thirty-seven articles were presented that dealt with environmental law:

*The constitutions conferred a series of duties on the public authorities, including: (i) preservation and recovery of species and ecosystems; (ii) preservation of the variety and integrity of genetic heritage, and supervision of entities engaged in genetic research and manipulation; (iii) environmental education at all educational levels, and development of public awareness about the need to preserve the environment; (iv) definition of the territorial areas eligible for special protection; and (v) compulsory environmental impact assessment for the setup of any activities that may translate into significant ecologic*
In addition to dedicating a whole chapter in the Federal Constitution to environmental laws, numerous federal agencies were also created in the post-1981 period to enforce environmental legislation. These agencies are all part of the Brazilian Environmental System, SISNAMA, which includes agencies such as: the Brazilian Environmental Council, CONAMA; the Ministry of the Environment; and the Brazilian Institute for the Environment and Renewable Natural Resources, IBAMA.

Although there has been seemingly significant development of environmental legislation, especially with the 1988 Constitution, the effects of this legislation have not been as grand as they may seem on paper. In a study conducted by Prates and Serra entitled, “O impacto dos gastos do governo federal no desmatamento no Estado do Pará,” the authors argue that many of the principal factors causing deforestation can be attributed directly or indirectly to the Brazilian federal government. The authors focus on the state of Pará in their study and analyze the exact effect of government spending on deforestation. What they conclude is that government spending on rural credit and transportation contributes to deforestation, while federal spending on environmental management helps the deforestation problem. According to the argument of Prates and Serra, the Brazilian federal government both mitigates and causes deforestation in the Amazon; “the government’s political expenses have opposite functions and, therefore, nullify each other. In summary, the government has expenses that affect deforestation in the determined region and, at the same time, has other expenses that mitigate deforestation” The authors further point out that the Brazilian government frequently values the benefits that come from deforestation over the costs from the loss of the Amazon. In Fearnside’s paper “Deforestation in Brazilian Amazonia: History, Rates, and Consequences,” he stresses the urgency for the Brazilian federal government to develop environmental legislation that is more than just a “symbolic base.”
While Brazil has shown improvement in both its economy and environmental legislation, it does seem that the country lacks the economic, political and environmental resources necessary to foster successful environmental change like that hoped for under the Kyoto Protocol and the CDM.

Lack of Will

To restate our second hypothesis, I am interested in investigating the claim that Brazil is not committed to its environmental performance at all, disregarding the availability, or unavailability, of resources to make such changes in its policies possible. One sees the bulk of evidence supporting this claim in Brazil’s focus on economic gain, rather than on bettering the environment. For example, looking specifically at the Amazon, Brazil’s neglect of the world’s largest rainforest that is home to some of the most vibrant biodiversity is often attributed to the country’s desire for development over sustainability;

In broader terms, the basic arguments of Brazilian officials (against saving the Amazon) involve matters related to development – the economic importance of the timber industry for the region as the second source of income tax for the Amazonian states, and also providing one hundred thousand jobs.

Cattle herding and logging are major sources of revenue in Brazil, and both industries can potentially involve deforestation. In the case of the logging industry, it is obvious why deforestation is involved. In the cattle herding industry, in order for the industry to expand, free land must be found, or created. One of the principal ways that this industry has facilitated its own growth is through the expansion of its frontier onto the land occupied by the Amazon, which includes the cutting down of trees. While this expansion of both the cattle herding and logging industries has included the destruction of the Amazon, many believe that the Brazilian government turns a blind eye to such devastation because
of the major role that cattle herding and logging play in the strong growth of the Brazilian economy. Furthermore, it seems that with specific respect to deforestation in the Amazon, Brazil’s government simply lacks the will to change; “The fundamental element to reduce the velocity of deforestation and to one day stop it completely is the political will to do so… Above all, the leaders of the country (Brazil) have to have the confidence that government action really will slow down, or even stop, deforestation.”

With its lack of motive to change its behavior in the Amazon, the Brazilian government is failing to realize the importance that the Amazon plays not only in Brazil, but also in the rest of the world;

*The Amazon’s global importance is well established. It acts as a climate regulator, directly affecting rainfall patterns in Brazil and Argentina. Its winds, recent studies say, may even affect rainfall in Europe and North America. The burning and decomposition of trees cut down for development makes Brazil’s chunk of the Amazon responsible for about half of the world’s annual greenhouse-gas emissions from deforestation, says Meg Symington, Amazon director for the World Wildlife Fund in the United States.*

Brazil’s disrespect for the Amazon is having dire effects on the atmospheric makeup of the world. This has led to a tarnishing of the Lula administration’s reputation in the eyes of the international community, as the Amazon is a resource valued by much of the world. In response to critics of Brazil’s treatment of the Amazon, Lula’s administration has responded with the common developing country argument in climate change negotiations that it is not his country’s responsibility to help combat a problem that the industrialized countries are responsible for creating. In a speech in 2005 that makes environmentalists shudder, Lula further argues that the Amazon cannot be a resource left untapped; “…it can’t be treated like was something from another world, untouchable, in which the people don’t have the right to the benefits.” While it may be that outside help, possibly in the form
of financial aid or more stringent legal measures, may be necessary in order to push the Brazilian government to curb its deforestation, many theorists suggest that these measures may not even be enough. They may, “…slow deforestation, (but) they will not solve the problem; there is still too great an economic incentive for people to continue clearing forests and for the government to continue allowing it.”

Besides Brazil’s ineptitude at diminishing its rate of deforestation, possibly due to its valuing of economic development over environmental performance, Brazil has also exemplified its lack of commitment towards improving the environment, in general, in its treatment of ex-environmental minister Marina Silva. Marina Silva was born in the Amazon state of Acre and, therefore, had a personal connection to the Amazon. President Lula appointed her environmental minister in his cabinet in January of 2003. Prior to becoming minister, however, Silva had thoroughly proven her track record both within and outside of the environmental world. Many looked to her as being the potential savior of the Amazon when she assumed her role as environmental minister.

During her time as minister, Silva’s main achievement was the development of the Plano Amazônia Sustentável (PAS), or the Sustainable Plan for the Amazon. The PAS was an initiative by the Brazilian Federal Government calling for cooperation among the states of Acre, Amapá, Amazonas, Maranhão, Mato Grosso, Pará, Rondônia, Roraima and Tocantins. The objective of the plan was to outline a schedule for sustainable development of the Amazon. Due to her major involvement in the development of the plan, Silva expected to be designated the coordinator, or “o coordenador do Conselho Gestor do PAS” of the PAS. To her and most of Brazil’s surprise, President Lula named Mangabeira Unger the coordinator of the PAS on May 8th, 2008. In reaction to this designation of Unger over herself as the coordinator, Silva resigned from her position as environmental minister five days after his appointment.

Although there has been much speculation as to why Silva was not chosen by Lula to be the coordinator, one of the major explanations is that Lula had been pressured by
the more business-minded people of the government not to encourage someone like Silva whose sole objective was to save the environment, and did not truly consider the larger picture of economic development;

Silva’s resignation is the story of a conflict between two very different Brazil’s. In one corner are the farmers, businessmen and ordinary Brazilians who see the country’s vast natural resources as a sure route to economic success. In the other are the environmental activists, indigenous groups and concerned spectators who believe that Brazil’s march to economic greatness will mean the continued devastation of its rainforest and its people.

In her letter of resignation to Lula, Silva even cites the resistance she felt from those in Lula’s administration who were more business lobbyists as part of the reason for her resignation;

In her resignation letter to the president of the Worker’s party, Silva said her decision was an attempt to break with the idea of ‘development based on material growth at any cost, with huge gains for a few and perverse result for the majority’ including ‘the destruction of natural resources.’ She added that ‘political conditions’ had meant that ‘environmental concerns had not been able to take route at the heart of the government.’

In the beginning of her time as environmental minister, Lula had supported Silva’s efforts for bettering the environment. But once his support dwindled, she had no political backer within the administration and she, inevitably, was virtually forced to resign. Silva’s resignation marked a win economically and a loss environmentally.

As if Brazil’s reputation in the international arena had not been tarnished enough by its deforestation of the Amazon, the resignation of Marina Silva also did not bode well for Brazil’s standing in the world, especially in the environmental world; “Marina takes
all of Lula’s environmental credibility with her, credibility which she has brought to his
government over the last five years. Without her, King Lula is completely naked.” Without
Marina Silva as environmental minister, many argue that Brazil’s reputation with respect
to the environment was a complete sham. Around the time of Silva’s resignation, there
were two other resignations of key officials within the Environmental Department of
Lula’s administration. One of the resignations was Bazileu Margarido, President of the
Environmental Agency.

Marina Silva was replaced by current environmental minister Carlos Minc. There
are not high hopes for Silva’s successor, specifically with regard to his addressing of the
situation in the Amazon. The current environmental minister uses the bureaucracy of
Lula’s administration as an excuse for his lack of clear progress towards the protection of
the Amazon. This is a stark change, as Silva used to try to work around the bureaucracy to
fight for the issues she believed in. While this may have been one of the main contributors
to her “forced” resignation, the fire Marina showed during her time as the environmental
minister is something that Minc will never have.

The Brazilian government’s treatment of Marina Silva and its favoring of furthering
the economy over developing sustainable environmental behavior suggest that Brazil lacks
the will to truly modify its environmental behavior in a way consistent with the overarching
goals of the Kyoto Protocol.

**Problems with the Kyoto Protocol**

Our final hypothesis explores the idea that Brazil’s lack of progress with regard
to the environment may not be grounded solely in shortcomings of Brazil, but instead
in failings of the Kyoto Protocol itself. This hypothesis points at two potential problems
with the protocol. First, the Kyoto Protocol includes project types that are not addressing
country-specific environmental problems. And second, it does not give enough incentive
for developing countries to change their environmental behavior.
The Kyoto Protocol’s first potential problem is that it does not include projects that deal with the specific environmental problems within the developing countries being targeted under the CDM scheme. It does not seem that the Kyoto Protocol and the CDM projects in particular take into account the fact that not all countries’ environmental problems lie in the energy sector, as they do not focus on methodologies to help combat these atypical environmental profiles; “Kyoto’s approach has not obviously paid off. Global carbon-dioxide emissions have grown by 25% since the protocol was adopted in 1997. That is partly because the treaty left out big emissions sources such as deforestation.” Theorists like Santilli et al. have argued that the Kyoto Protocol is almost contradicting its principal mission by not targeting the extreme deforestation existent in various developing countries;

The current annual rates of tropical deforestation from Brazil and Indonesia alone would equal four-fifths of the emissions reductions gained by implementing the Kyoto Protocol in its first commitment period, jeopardizing the goal of Protocol to avoid ‘dangerous anthropogenic interference’ with the climate system.

One cannot deny that without the implementation of CDM projects in Brazil, the country’s emissions would be even higher than they currently are, but this does not justify Brazil’s continual reliance on mechanisms such as deforestation going unchecked.

Instead of focusing on the true environmental issues at hand within developing countries, the Kyoto Protocol has instead focused on more easily implemented technologies like hydroelectricity. According to the U.S. Geologic Survey (USGS) power deriving from hydroelectricity is one of the most valuable and commonly used sources of renewable energy in the world, and Brazil is the third largest producer of it. The USGS also outlines the main advantages of hydropower as follows:

Fuel is not burned so there is minimal pollution; water to run the power plant
is provided free by nature; hydropower plays a major role in reducing greenhouse gas emissions; relatively low operations and maintenance costs; the technology is reliable and proven over time; it’s renewable – rainfall renews the water in reservoir, so that fuel is almost always there.

Furthermore, hydropower seems to have a relatively fast return on investment, which is not always the case with some other renewable technologies. This makes it a more attractive technology to implement especially for developing countries that do not usually have the luxury of waiting years to see the return on their investments. While hydropower energy does seem like a very attractive alternative to other “dirtier” forms of energy like coal and oil, there are a number of disadvantages, as have already been discussed, like displacement of inhabitants and destruction of natural habitats. The USGS also points to other potential disadvantages of hydropower such as: “high investment costs; hydrology dependent (precipitation); inundation of land and wildlife habitat; loss or modification of fish habitat; changes in reservoir and stream water quality; and fish entrainment or passage restriction.”

Brazil is a country that does exhibit this “atypical” environmental profile, as its energy sector, due to initiatives like its production of biofuel, is for the most part pretty clean. However, the country does suffer from seemingly uncontrolled deforestation, a point discussed in depth throughout this thesis. This is in large part a problem that Brazil has created for itself, but also the Kyoto Protocol is not designed in a way that facilitates dealing with such problems like deforestation, which is the largest contributor to Brazil’s emissions. As I have mentioned previously, avoidance of deforestation was not even included as an approved CDM mechanism for the first commitment period of the Kyoto Protocol lasting until the year 2012. This lack of inclusion was a decision that was made with the knowledge that deforestation avoidance being an approved CDM mechanism would have greatly helped countries like Brazil. While the Kyoto Protocol does not address the deforestation
problem for developing countries, the problem is addressed for those participant Annex I countries. In Article 3.3, the terms for this approach towards deforestation for these countries are outlined;

3. The net changes in greenhouse gas emissions by sources and removals by sinks resulting from direct human-induced land-use change and forestry activities, limited to afforestation, reforestation and deforestation since 1990, measured as verifiable changes in carbon stocks in each commitment period, shall be used to meet the commitments under the Article of each Party included in Annex I.

Deforestation is also an accepted methodology under the Emissions Trading mechanism for Annex I and Annex B countries, but not for developing countries that have not legally pledged to reduce their emissions. Theorists like Fearnside have argued that Brazil would even benefit greatly from accepting an emission reduction target; “Although not currently favored by Brazil’s Ministry of External Affairs, the country always has the option of accepting national limits on emissions that would allow it to earn much more by emissions trading…”

In an attempt to justify the lack of focus on environmental problems such as deforestation in the Kyoto Protocol, many argue that the monitoring involved in initiatives to diminish deforestation would be too involved a process to be successful; “Schemes to avoid such deforestation were left out of the Kyoto treaty because it was thought, correctly, that it would be too hard to monitor them precisely.” It seems that to address the dire problem of deforestation from which Brazil is suffering will require much more vigorous initiatives than the CDM was willing to address. Although addressing deforestation in developing countries would not be a simple task, the hopes going into the Kyoto Protocol had been that the international community would be the perfect candidates to attempt to conquer such a problem.
Fortunately, there have been subsequent efforts on the international front to develop some kind of legislation to address the deforestation issue that was not confronted under the Kyoto Protocol. One of the most promising initiatives is the UNFCCC’s Reduced Emissions From Avoided Deforestation and Degradation (REDD) scheme that would target deforestation and forest degradation in developing countries. The idea, still in the planning stages, would be to give financial incentive to developing nations to avoid this misuse of land. Contrary to the structure of the Kyoto Protocol, the REDD would give credit to entire nations for progress instead of to individual projects within the countries. As to be expected, there are a lot of unresolved issues surrounding this initiative at its inception, “…including issues such as the deforestation baseline to be used, the role of developing countries that have a low recent rate of deforestation, and the protocols for measurement and validation of emissions reductions.” Despite these issues, the UNFCCC in partnership with the World Bank has started preparations for the eventual implementation of the REDD mechanism in 2012; “Thirty-seven countries have applied and already been approved to participate in the World Bank’s Forest Carbon Partnership Facility (FCPF) for support to prepare for a future REDD Mechanism.” Brazil has also been targeted as a priority country for this REDD Mechanism.

Addressing the second potential problem with the Kyoto Protocol, very much associated with the first, that there is not enough incentive given to developing countries to curb their emissions, I can see that this is reflected in Brazil’s carbon dioxide emissions continuing to increase, especially because of the country’s deforestation problem increasingly becoming out of hand. Even though the Kyoto Protocol did not deal with deforestation problems in developing countries, the creation of the Certified Emission Reductions being awarded under the CDM was an attempt to add incentive for developing countries, and for industrialized countries and private companies, to implement renewable and carbon reduction projects within these countries. While the CER program did lead to the creation of new projects in these developing countries, the majority of project developers opted for
projects that were easier to implement and cheaper, rather than promoting technologies like reforestation, which could not have reversed the effects of deforestation, but at least could have helped rebuild deforested plots of land.

Focusing on the projects chosen in Brazil specifically, the majority of the projects implemented in the country utilize technologies that are more easily registered like methane avoidance and hydroelectricity, rather than technologies like reforestation. Although the process of registration for any project is long and arduous, projects using reforestation or afforestation (the planting of trees on new plots of land) have an even tougher time getting registered. The CDM Executive Board divides CDM projects into three groups: Large-scale CDM Projects, Afforestation and Reforestation CDM Projects and Small-scale CDM Projects. Each type of project includes different processes for validation. Large-scale and small-scale projects have almost identical processes, as usually they are implementing the same type of technology, differing only in the scale of the project. Afforestation and reforestation projects, contrastingly, have a completely different process, as these types of projects involve highly specific technology. Furthermore, one can see the added work in the validation process for afforestation and reforestation projects when comparing the number of steps involved in the process. For both large-scale and small-scale CDM projects, there are 10 steps involved, while there are 14 steps involved in afforestation and reforestation projects. Given the more intensive validation process involved in projects utilizing reforestation and afforestation, it is no wonder that only 0.53% of all of the registered CDM projects in the world are projects utilizing such technology, and 0% of these projects are in Latin America.

The numbers definitely provide clear evidence that the CDM mechanism does not provide sufficient incentive for countries like Brazil to implement technologies like reforestation and afforestation that could make an attempt at counteracting the deforestation rates; “Slowing deforestation in Brazil’s Amazon will not work through marginal incentives such as the CDM.” Furthermore, because developing countries are not direct participants in the Kyoto Protocol, they are not held to legally binding targets, as the industrialized
countries involved in the protocol are. If developing countries like Brazil had been more actively involved and been given binding targets, one would expect to see that Brazil’s deforestation would have been forcibly checked and diminished and, potentially, that afforestation and reforestation would have been implemented technologies to right the damage that had already occurred in forests like the Amazon.

It is unfortunate that the Kyoto Protocol was unable to provide incentive to countries with such high rates of deforestation like Brazil, as there had been many theorized schemes as to how such incentive could have been achieved. In Santilli et al.’s paper “Tropical Deforestation and the Kyoto Protocol,” the authors suggest the creation of a concept they call “compensated reduction.” The authors claim that this method would be, “…a means of both reducing the substantial emissions of carbon from deforestation and facilitating significant developing country participation in the Kyoto Protocol framework.” Santilli et al. envisioned developing countries voluntarily reducing their rates of deforestation below a well-established baseline and receiving credits similar to CER’s issued under the CDM for their progress. Although the authors had hoped for the implementation of such a program during the first commitment period of the Kyoto Protocol, fortunately, the REDD program of the UNFCCC is of a similar nature and may help alleviate the deforestation problem in the near future.

These two faults of the Kyoto Protocol were acknowledged by the international community, which, in reaction, attempted to correct them at the Copenhagen Conference in December of 2009. Going into the Copenhagen Conference, the hopes were that the resultant accord would include legally binding targets for industrialized countries and suggested, but not legally binding, targets for developing countries. This would have been a step forward from the Kyoto Protocol, but still not the ideal of legally binding targets for all countries. Unfortunately, the Copenhagen Conference failed to live up to these aspirations and, furthermore, even to live up to the precedent that the Kyoto Protocol had set. The results of the Copenhagen Conference were a disgrace to the international
community. The bureaucracy of the conference led to next to nothing being accomplished. No real agreement was made. No emission reduction targets were created. The only significant success that came out of Copenhagen, especially with respect to countries facing deforestation problems like Brazil, is that a nearly unanimous consensus was reached that the UNFCCC’s REDD mechanism is the correct path to combating deforestation and forest degradation in the future.

Instead, the conference merely achieved the agreement that the international community, or at least those nations involved in the conference, would not let the world’s overall temperature rise by more than two degrees Celsius. There was no indication of how this target would be met and who would be responsible for reducing what, but merely the agreement that the world as a whole did not want to see this increase in atmospheric temperature. Many attribute the insufficient progress seen at Copenhagen to the inefficiencies of international agreements, especially when the large number of actors involved in the agreements is itself an obstacle. In an interview with the United Kingdom’s Guardian, Prime Minister Brown calls for the creation of an international body with the sole duty of handling environmental negotiations. He argues that the world must learn from the mistakes of Copenhagen and other failed international negotiations and realize that traditional ways of dealing with international issues must either be discarded or improved if progress is to be made on issues such as climate change in the future.

Conclusion

The Kyoto Protocol has no doubt made history by representing the world’s first legally binding international response to climate change. Although focusing on modifying the environmental behavior of industrialized countries, the Kyoto Protocol has also had positive effects on the environmental behavior of developing countries. The main initiative under the Kyoto Protocol to involve developing countries is through CDM projects. While the intrinsic goal of these projects is to reduce emissions, the CDM projects also have the
larger goal of developing sustainable environmental behavior within developing countries.

While CDM projects in Brazil have undoubtedly helped improve the country’s energy matrix, especially shown in the country’s diminishing growth rate of emissions, Brazil’s key environmental problems have scarcely improved with the implementation of these CDM projects. As has been mentioned, Brazil possesses an atypical environmental profile. While the environmental problems for most of the countries in the world lie in their energy sectors, Brazil has a surprisingly clean energy sector. The main environmental problem from which Brazil suffers is deforestation. Unfortunately, besides showing some progress when environmental minister Marina Silva was in power, the Brazilian government has been, for the most part, ineffective at dealing with its deforestation problem. Although the deforestation in Brazil has caused huge international criticism, particularly due to the fact that the Brazilian Amazon is one of the major regulators of the world’s temperature, one does not see the Kyoto Protocol, supposedly an international reaction to the world’s climate change problems, as effective in ameliorating the situation.

In examining Brazil’s environmental situation, I have proposed three potential hypotheses to explain the failure of Brazil to show more progress in dealing with its true environmental problems. The first two hypotheses address potential inadequacies on the part of Brazil, specifically a lack of resources and a lack of will to change its environmental behavior, which could account for the disappointing environmental performance of the country. The third hypothesis moves away from putting fault on Brazil and instead suggests that the Kyoto Protocol and the Clean Development Mechanism projects are designed in a way that does not foster the addressing of the specific environmental problems faced by Brazil and that does not provide adequate incentive to the developing countries, like Brazil, to modify or to even care about their environmental policies.

When examining the hypothesis dealing with a lack of resources, I show that resources, especially capital, are needed in order to attain environmental development. Although Brazil has been experiencing astounding economic growth in the past couple
of years, the capital available for environmental purposes is not directly correlated with this economic growth. Countries, especially developing ones, often do not feel able to use capital to invest in expensive technologies, like renewable technologies, especially when the pay-back period for such investments would be in the long-term; short-term, measurable returns on investment are limited. Furthermore, in order for countries to modify their environmental behavior, they need to have supportive political resources as well. Brazil’s environmental legislation is developing, but not nearly to the point necessary to promote the changes in environmental policy necessary to address the country’s true problems. For all of these reasons, it is evident that a lack of resources is inhibiting Brazil’s ability to address its environmental problems.

While lacking the necessary resources to modify its environmental behavior, I argue that Brazil also lacks the will to change its environmental behavior. This is particularly clear with the practically forced resignation of environmental minister Marina Silva and the economic focus of the government. Silva is internationally renowned for her dedication to the environment and for her authorship of many fundamental pieces of environmental legislation, especially those dealing with the Amazon. Although strong in her knowledge about the environment and in her innovative ideas, Silva was unfortunately too weak politically to remain in power. Had Silva stayed in power, one would have hoped that she would have been able to encourage the Lula government to take more proactive action in the Amazon. This being said, however, as one of the main reasons for her resignation was the resistance she felt within Lula’s administration, I am not confident that had Silva remained environmental minister she would have been able to combat the strong lack of desire on the part of the majority of the administration to address environmental issues.

Within the Brazilian government and the country as a whole, there existed and exists to this day a strong push to promote the economy over the environment and a belief that improving the environment is not in the country’s self-interest when compared to advancing the economy. The economy and the environment are often at odds in Brazil, as some of
the most economically advantageous industries in the country often function at the expense of the environment. Although it is hoped that Brazil will soon realize the importance of mitigating climate change through measures like curbing its deforestation, the country currently seems to lack the drive to make such necessary changes to its environmental behavior.

My final hypothesis for Brazil’s failure to truly modify its environmental policy finds the Kyoto Protocol and the CDM projects also to be at fault. The Kyoto Protocol does not have measures that address situations of deforestation and instead focuses on addressing energy sector environmental problems. Although the CDM projects focusing on the energy sector have, for the most part, helped to improve Brazil’s already clean energy sector, the principal environmental problems that Brazil faces do not lie in its energy sector and instead are problems of LULUCF (Land Use, Land Use Change and Forestry). Besides not addressing the exact problems faced by Brazil, the Kyoto Protocol also includes the highly criticized renewable technology of hydropower as an approved CDM. Hydropower, although probably better than coal or oil, is extremely destructive to surrounding habitats and causes the displacement of settled communities, among other negatives. Finally, by not including developing countries as direct participants of the treaty, the Kyoto Protocol fails to provide developing countries with strong incentives to change their environmental behavior.

As has been demonstrated in this thesis, there are three main reasons that have contributed to Brazil’s mediocre change in its environmental policy, even in the presence of strong international pressure for change and initiatives like the Kyoto Protocol’s Clean Development Mechanism. The resultant behavior of Brazil cannot be explained solely by one of these reasons. True environmental change within Brazil cannot be achieved merely by a true will to change, by the necessary resources or by a strong international institution; taken separately, these three reasons are not sufficient to describe the environmental situation seen in Brazil. But taken together, they sufficiently explain why Brazil’s progress
with regard to the environment has not been more significant; Brazil’s lack of sufficient resources and will to change its environmental behavior coupled with certain failures of the Kyoto Protocol explain Brazil’s failure to demonstrate true environmental change to date. While the international community had hoped that the Copenhagen Conference of 2009 would help to deal with the problems seen in the Kyoto Protocol and to involve the more direct participation of developing countries, the conference was an almost complete failure.

The hope for future international negotiations surrounding climate change is that countries will swallow their national pride and self-interests and be able to take those actions required by the international community for successfully addressing climate change. For Brazil, this would mean that the country would actually pledge to deal with its deforestation problem, notwithstanding the shorter-term effect on the country’s economic activities. This being said, future negotiations would also have to assist developing countries like Brazil that lack necessary resources to achieve the changes that would be required of them. While the CDM projects represent an attempt to supply financial incentive to implement renewable projects, a future treaty would have to involve even more assistance, especially financial, to these developing countries than that seen in the CDM.

What the Brazilian government and the international community are failing to realize is that climate change is only going to get worse and its effects are, for the most part, irreversible;

‘The issue of climate change is one that we ignore at our own peril. There may still be disputes about exactly how much we’re contributing to the warming of the earth’s atmosphere and how much is naturally occurring, but what we can be scientifically certain of is that our continued use of fossil fuels is pushing us to a point of no return. And unless we free ourselves from a dependence on these fossil fuels and chart a new course on energy…we are condemning future generations to global catastrophe.’

No matter the costs to individual actors, the international community must come
together and find a way to negotiate effectively a containment strategy for global warming. For each day without an effective plan, the problem of climate change will continue to worsen.

Bibliography


Cardoso, José Maria. “É a Vitória de uma Visão Conservadora.” O Globo. 15 Maio 2008.


“Geopolítica da Amazônia” Bertha K. Becker


http://cdm.unfccc.int/about/index.html
http://cdm.unfccc.int/Statistics/Registration/AmountOfReducedRegisteredProjPieChart.html
http://cdm.unfccc.int/Statistics/Registration/NumOfRegisteredProjByHostPartiesPieChart.html
http://cdmpipeline.org/index.htm
http://ec.europa.eu/environment/climat/kyoto.htm
http://wbcarbonfinance.org/Router.cfm?Page=PCF&ft=About
http://www.ahk.org.br/cdmbrazil/index.asp
http://www.eoearth.org/article/Kyoto_Protocol
http://www.kyotoprotocol.com/
http://www.oecd.org/pages/0,3417,en_36734052_36734103_1_1_1_1_1,00.html


Phillips, Tom. “‘I’d lost the strength to carry on’: Last week, with the resignation of Marina Silva as Brazil’s environment minister, the Amazon lost its greatest champion. But how did a poor orphan rise to such power, only to walk away from it? And what is the future for her beloved rainforest?” The Guardian. 22 May 2008. Lexis Nexis. University of Pennsylvania. 29 October 2008.


Introduction

“The way to stop discrimination on the basis of race,” Chief Justice John Roberts recently wrote, “is to stop discriminating on the basis on race.” Roberts advances the color-blind view that equal treatment can only be achieved when all racial preferences are eliminated. It directly contrasts with Justice Harry Blackmun’s race-conscious account from thirty years earlier, which argues that, “in order to get beyond racism, we must first take account of race. There is no other way. And in order to treat some persons equally, we must treat them differently.” Since the passage of the landmark 1964 Civil Rights Act, affirmative action efforts have centered on these two competing perspectives—one that advocates measures that discounts race entirely and the other that actively uses race as an integral factor in promoting equality and diversity.

Over the past forty-five years, these two frameworks have collided numerous times. Efforts to increase equality and diversity have been repeatedly challenged in the fields of education, voting, public contracting, and employment. The judiciary has struggled over how to best approach the issue. For instance, the Supreme Court within the last decade has determined that a rigid, point-based university admissions policy classified as an unconstitutional quota system, yet a more holistic policy that allows race to be a significant factor was permissible under the Equal Protection Clause of the Fourteenth Amendment. The answers to dealing with affirmative action are far from clear-cut. However, if we want to live in a world where discrimination no longer exists, there must be policies in place that provide citizens with an equal opportunity including laws that protect minorities from taking employment tests that may be facially neutral but unintentionally discriminate against them. It is this subject where the most recent controversy between the race-
conscious and color-blind orders factions is focused. They each faced answering the difficult
question: Under what circumstances, if any, can an employer’s purpose to avoid liability
for unintentional discrimination excuse what otherwise would be prohibited intentional
discrimination?

In March 2004, the City of New Haven wrestled with this problem. The
examinations that it had administered as promotional tests for the lieutenant and capital
positions for the New Haven Fire Department had racially skewed results. Even though
forty-two percent of the test takers were racial minorities, whites were expected to fill
seventeen of the nineteen available positions based upon the exam scores. Nonwhites were
historically underrepresented in the department despite the fact that they made up nearly
sixty percent of the City’s racial composition. If the City chose to certify the results of the
exam, it risked the possibility of being held liable for unintentionally discriminating against
nonwhites, a provision known as disparate impact that is enumerated in Title VII of the
1964 Civil Rights Act. On the other hand, if the City discarded the test results, it knew
that it would likely face a lawsuit from the seventeen white firefighters and one Hispanic
firefighter who performed well on the exam for intentionally discriminating against them
on the basis of race, a different provision known as disparate treatment under Title VII.
After months of debate, the City voted to throw out the results and was indeed sued for
intentional discrimination.

The district court granted summary judgment in favor of the City in September
2006. Two years later, the Second Circuit Court of Appeals, which then included Sonia
Sotomayor, affirmed the decision. In June 2009, the Supreme Court reversed in a five-to-
four decision and held that the City had violated the disparate treatment provision of Title
VII by discarding the test results. The Court established a new standard, based upon equal
protection case law, that “before an employer can engage in intentional discrimination
for the asserted purpose of avoiding or remedying an unintentional disparate impact, the
employer must show a strong basis in evidence to believe it will be subject to disparate
impact liability if it fails to take the race-conscious, discriminatory action.” Applying the new standard, the Court’s majority argued that the City did not have enough evidence to believe that it would lose a case of disparate impact if it had certified the test results because there was not a strong basis in evidence that the promotional exams were not job-related, consistent with business necessity, or that equally valid, less discriminatory alternatives were available.

Initially designed in Wygant v. Jackson Board of Education (1986) and further developed in Richmond v. J.A. Croson (1989), the strong basis in evidence standard was created to ensure that no state agency could adopt race-conscious policies unless they showed direct evidence—not a generalized assertion—that these measures were narrowly tailored to remedy specific acts of intentional racial discrimination, committed generally if not always by the agency itself. By raising the standard, it became an influential device that the Court has used to limit race-conscious practices. The Court’s introduction of the strong basis in evidence standard to disparate impact liability litigation, rather than equal protection remedial cases, marked a radical departure from its legal origins. The statutory domain was a new legal frontier. The standard still was used to determine whether race-conscious measures were justified. However, how a strong basis in evidence applied to a case where the employer feared losing a disparate impact suit that does not involve remedying intentional discrimination was unclear.

The Court has indicated that it adopted the strong basis in evidence standard in Ricci v. DeStefano (2009) to reconcile the differences between the two central principles of Title VII—disparate impact and disparate treatment. This essay argues, however, that the Court failed to achieve its goal. In reality, the standard provides more questions than answers. How does the standard relate to its legal precedents? If the Court is so certain that New Haven’s actions in Ricci do not pass this standard, what cases would pass this benchmark? What grounds are necessary for a Court to deem that there is a strong basis in evidence for avoiding disparate impact liability? It is these ambiguities that I will examine
further in this essay. I will also explore the future implications that the Supreme Court’s ruling may have on discrimination in the workplace and how Title VII of the 1964 Civil Rights Act will be interpreted.

First, I provide a brief history of how Title VII was developed, and define central terms like disparate impact and disparate treatment that are integral to understanding the Ricci ruling. The second section charts the legal evolution of the strong basis in evidence standard from when the test was first established to its most recent applications to illustrate how the standard has developed over its thirty-year history. Third, I provide a complete background of the facts in Ricci in addition to detailing how the Court’s majority introduced the strong basis in evidence standard. I also summarize the dissenting and two concurring opinions.

In the next section, I propose three different readings that the lower courts could potentially have of this new standard. For each reading, I demonstrate the grounds necessary to fulfill the strong basis in evidence threshold and its related practical consequences for both minorities and employers, including whether it effectively reconciles disparate impact with disparate treatment. The final section looks at how district and circuit courts have applied the principles of Ricci in the last year and examines the long-term viability of the strong basis in evidence standard.

II. Title VII and its Principles

Congress enacted Title VII as a provision of the 1964 Civil Rights Act to avoid employers from using employment practices that disadvantage persons on the basis of race, sex, religion, national origin, or sexual orientation unless the practice is job related and consistent with business necessity. Title VII seeks to combat workplace discrimination by providing citizens with the promise of equal employment opportunities. At first, Title VII was meant as a regulation of interstate commerce and applied only to private employers. However, in 1972, the Equal Employment Opportunity Act extended Title VII’s coverage
to the public sector in accordance with Congress’ authority under Fourteenth Amendment to ensure that “[no] State shall…deny to any person within its jurisdiction the equal protection of the laws.”

The Supreme Court’s initial application of the clause focused on holding employers liable for engaging in intentional discrimination, or disparate treatment. Such discrimination occurs when one person is intentionally treated differently due to their race, sex, national origin, or other characteristic. Disparate treatment is the type of discrimination with which most people are familiar. The prevailing framework that the Court uses to analyze a case of disparate treatment comes from McDonnell Douglas Corp. v. Green (1973). First, a plaintiff must establish a prima facie case of disparate treatment by showing the plaintiff has been disadvantaged in some concrete way. Second, the burden then falls upon the employer to “articulate some legitimate, non-discriminatory reason” for its actions. Third, the burden then shifts back to the plaintiff who must demonstrate that the employer’s defense was “a pretext or discriminatory in its application.”

While the 1964 Civil Rights Act explicitly accounts for intentional discrimination, there is no mention of unintentional discrimination, or disparate impact. According to the Supreme Court in International Brotherhood of Teamsters v. United States (1977), disparate impact refers to “employment practices that are facially neutral in their treatment of different groups but that in fact fall more harshly on one group than another and cannot be justified by business necessity.” The Court first expanded Title VII to include unintentional disparate impacts in the landmark employment discrimination case, Griggs v. Duke Power Co (1971).

Griggs v. Duke Power Co. (1971) dealt with an employer who changed the requirements for positions in their labor department, the plant’s lowest-paying sector. Applicants were now required to have a high school diploma and receive a passing score on an aptitude test. On the surface, these additional qualifications did not appear discriminatory. However, it became evident that there was an adverse impact when “a
markedly disproportionate number of [African-Americans]” were rendered ineligible. In North Carolina, where the plant was located, “thirty-four percent of white males, but only twelve percent of African-American males, had high school diplomas.” Far fewer African-Americans passed the aptitude tests in comparison to whites. The aptitudes measured in the tests, nevertheless, were rarely required to perform the duties of the jobs in the plaintiff’s labor department.

In result, the Griggs Court unanimously concluded that the high school diploma and test requirements violated Title VII of the 1964 Civil Rights Act. They argued that Title VII was intended to include “the consequences of employment practices, not simply the motivation.” Even if a policy appears to be facially neutral, an employer still assumes the risk that they could be held responsible when an adverse impact exists, if there is no legitimate business justification for the measures producing that impact.

The Court specified a two-step process, much like the McDonnell Douglas framework, for determining whether an employer is liable for disparate impact. First, the plaintiff must demonstrate a prima facie case of unintentional discrimination by illustrating that the disputed employment practice, despite appearing facially neutral on the surface, has a disproportionate impact on a protected group such as, African-Americans or women. Second, the employer then has the burden to prove what the Court calls the “touchstone” of disparate impact liability, that the challenged practice was justified as a “business necessity” and “reasonably related” to the job at hand.

In Albemarle Paper Co v. Moody (1975), the Court further refined the scope of what “business necessity” entailed. The justices stated that when the contended practice is a test, the employer is required to establish the test’s validity under specific testing standards that are now laid out by the Equal Employment Opportunity Commission. The burden fell upon the employer to demonstrate that its test fairly and accurately reflected the qualifications of the job. For the next decade, the Griggs and Albemarle precedent had an influential impact on federal and appellate court decisions as the concept of disparate
impact gradually evolved.

Disparate impact faced its first major challenge in Wards Cove Packing Co v. Atonio (1989) when the Court made a drastic shift in its interpretation. The Court significantly lowered the standard for employers so that they no longer had to prove business necessity, just that their hiring practices were reasonably justified. An employment practice would be permissible so long as it “serve[d], in a significant way, the legitimate employment goals of the employer.” Moreover, once a prima facie case of disparate impact was established, the burden shifted to the employer to show justification for the disparate impact. The employer’s burden is one of production and not persuasion. In other words, the employer only has to produce evidence that is sufficient to show that the criteria used in hiring are objectively related to their employment goals. The burden of persuasion, the task of convincing the Court that the criteria are not part of any bona fide occupational qualification, remains with the plaintiff. This relaxing of standards provided fewer incentives for the employers to ensure that their practices did not violate disparate impact. Following Wards Cove, the Supreme Court continued to diminish the scope and effectiveness of federal civil rights protections and generated considerable protest.

In response, Congress passed the 1991 Civil Rights Act. The bill explicitly codified for the first time the concept of disparate impact as an amendment to Title VII of the Civil Rights Act of 1964. It declared that engaging in disparate impact discrimination was an “unlawful employment practice.” Furthermore, the Act revived the “business necessity” and “job-related” requirements that were first laid out in Griggs and Albemarle. Once a prima facie case of disparate impact is established, the burden once again falls upon the employer “to demonstrate that the challenge practice is job related for the position in question and consistent with business necessity.” This time, Congress enacted a third, more stringent step. Should the employer successfully meet the burden, the plaintiff then has the opportunity to respond by identifying “an alternative employment practice” which the employer “refused to adopt.”
The tension between Title VII’s core tenets—disparate impact and disparate treatment—most directly lies in its relationship with the Constitution’s Fourteenth Amendment. The Equal Protection Clause is intended to protect equal opportunities and thus, clearly challenges intentional forms of discrimination. However, it does not explicitly prohibit unintentional discrimination. It is this distinction upon which debate has centered for the past thirty-five years.

The seminal differentiation of disparate impact and equal protection occurred in the Washington v. Davis decision (1976). Two African-Americans, whose applications for a police position were rejected, sued the Washington, D.C. police department for using racially discriminatory hiring practices. The Court ruled against them and stated that an official act will not be held “unconstitutional solely because it has a racially discriminatory impact regardless of whether it reflects a racially discriminatory purpose.” The plaintiffs needed to show a discriminatory motive, also known as intent. Justice White, writing the majority opinion, went on to clarify that “disproportionate impact is not irrelevant, but it is not the sole touchstone of an invidious racial discrimination forbidden by the Constitution.”

Following Washington v. Davis, the Supreme Court has come to accept that disparate treatment principles apply at the constitutional and statutory levels. The terms disparate treatment and equal protection have become “virtually interchangeable in their conceptual relationship to disparate impact” over time. Disparate impact, though, applies only in the statutory context, unless the Equal Protection Clause is read differently than the Court has decided to do.

The more extreme defenders of the Washington decision assume that efforts to avoid disparate impact are incompatible with equal protection. They read the Equal Protection Clause to prohibit all decisions based on race. Concern for unintentional discrimination, on the other hand, necessarily means adopting polices to avoid or correct unequal racial consequences—decisions that are therefore made on racial grounds. To proponents of
color-blind approaches, the Equal Protection Clause is inconsistent with concern for unintentional discrimination. It forbids only intentional discrimination. Should the Court take this position in the future, avoiding or remediying racially disparate impacts could be deemed unconstitutional under the Fourteenth Amendment, or at least be severely diminished.

However, opponents of the Washington decision argue that the two principles are complementary to one another. They assume that the clause is read to mean that all people should be protected equally under the law by remediying past racial injustices and avoiding actions that needlessly perpetuate or exacerbate racial inequalities. Decisions aimed at such remedies and avoidance necessarily involve race. Under this construction, because the Equal Protection Clause guarantees not only equal opportunities, but just if not necessarily equal outcomes as well, the clause includes both intentional as well as unintentional forms of discrimination. These contrasting interpretations of how disparate impact relates to the Equal Protection Clause create the main point of contention between the majority and the dissent in evaluating the strong basis in evidence standard in Ricci.

III. The Legal Evolution of the Strong Basis in Evidence Standard

Before examining the strong basis in evident standard within the context of Ricci, it is important to examine how the test has been elaborated in legal precedents. The standard has a relatively young history—a mere twenty-five years—in comparison to older tests such as strict scrutiny and rational basis. Strong basis in evidence originates in affirmative actions cases dealing with the Equal Protection Clause of the Fourteenth Amendment.

The test was first introduced in Wygant v. Jackson Board of Education (1986), a case that dealt with increasing the percentage of minority personnel in the school system. Based upon a collective bargaining agreement between the Jackson, Michigan Board of Education and the local teacher’s union, a provision was passed to ensure that in the event of layoffs, the percentage of minority personnel laid off would never exceed the current
percentage of minority personnel employed. The school board could also not lay off teachers with the most seniority. The lawsuit occurred when schools in Jackson, Michigan laid off nonminority teachers with seniority, yet retained minority teachers with less seniority. In a five-to-four decision, the Supreme Court ruled that the layoffs violated the Equal Protection Clause, claiming that the public employer’s affirmative actions plan failed to meet the requirements of strict scrutiny. In order to satisfy the strict scrutiny test, a law or policy must be justified by a compelling government interest such as national security or equal protection, narrowly tailored to achieve that interest and the least restrictive means for achieving that interest.

The strong basis in evidence standard makes its appearance when Justice Powell, writing the majority opinion for the Court, argues that employers “must have sufficient evidence to justify the conclusion that there has been prior discrimination.” A legacy of societal injustice is not a strong enough argument to warrant remedial action because the Court claims it is “too amorphous.” More proximate and firm evidence of discrimination by the employer must exist in order to make a remedial action warranted. In result, the Court states that “the trial court must make a factual determination that the employer had a strong basis in evidence for its conclusion that the remedial action was necessary.” A public employer must now fulfill a higher benchmark to use race-conscious remedies.

Following Wygant, the most critical case to flesh out the strong basis in evidence standard is Richmond v. J.A. Croson Co. (1989). In 1983, the City Council of Richmond, VA required companies awarded city construction contracts to subcontract thirty percent of their business to minority business enterprises. When Croson lost its contract because it failed to set aside thirty percent to minorities, the company sued, arguing that Richmond violated the Equal Protection Clause. In Fullilove v. Klutznick (1980), the Court had ruled that Congress had the power to design minority set-aside programs under the Commerce Clause. However, in Richmond, the Court held that states and local governments had less power to do so. Applying strict scrutiny, the City of Richmond’s minority set-aside program
was deemed unconstitutional.

The City of Richmond brought forth four key pieces of evidence to justify their remedial actions. First, minorities faced a significant lack of opportunities in the construction industry due to past public and private racial discrimination, making it exceedingly difficult for a minority business to succeed in the Richmond area. A historically small number of minority businesses in the local contracting industry had been awarded contracts (0.67%) despite minorities making up fifty-percent of the city’s population. Second, African-Americans had also been excluded from skilled construction trade unions and training programs. Third, the plan was consistent with the recent Fullilove decision. Fourth, a legacy of widespread racial discrimination existed on local, state, and national level in the construction industries. The City of Richmond’s enactment of the thirty-percent quota was, therefore, intended to encourage minority entrepreneurship and remedy former discrimination, public as well as private.

In the majority opinion, Justice O’Connor rejects the City’s argument. She claims that “a generalized assertion that there has been past discrimination in an entire industry provides no guidance for a legislative body to determine the precise scope of the injury it seeks to remedy.” It is a slippery slope if one takes into account all past injustices. O’Connor feared that encouraging arguments of this nature would set the standards too low in determining what was sufficient to justify a racial discrimination claim. It would be uncertain as to how far in the past that discrimination could extend, or when something could still be classified as discrimination.

Even though it had supplied four pieces of evidence to support its decision, Richmond had not shown enough proof to provide a strong basis in evidence that the remedial action—the thirty-percent quota in awarding government contracts—was necessary. It lacked “a prima facie case of a constitutional or statutory violation”. According to the Court, the fact that there was a legacy of racial discrimination in the construction industry was not a viable argument to justify the thirty-percent quota. A more proximate
cause of discrimination, including significant discrimination by the City itself, was required. O’Connor points to the fact that there was “no evidence that qualified minority contractors had been passed over for city contracts or subcontracts, either as a group or in any individual case.”

Richmond v. J.A. Croson Co. (1989) marked the first time that the strong basis in evidence standard was applied to a specific case. The Court did not provide very specific guidance as to how the standard should be used, nor as to how much evidence is sufficient to prove that remedial action is justified. Nevertheless, the Court did establish that a strong basis in evidence standard was more demanding than making a generalized assertion of past racial discrimination.

Since Richmond, the strong basis in evidence standard has been used to evaluate remedies for prior discrimination in circumstances when strict scrutiny is applied. In Shaw v. Hunt (1996), the Court applied the strong basis in evidence standard when determining whether a North Carolina voting reapportionment plan represented an appropriate response to voting inequities, or an intentional act of racial gerrymandering that violated the Equal Protection Clause. The deliberate creation of a majority-minority district in North Carolina was meant to increase African-American representation in the state’s political system as well as the likelihood that the state would elect an African-American representative to Congress for the first time since the Reconstruction era. But the Court declared that a compelling state interest did not exist for North Carolina to minimize the consequences of racial bloc voting for African-Americans by consciously creating a district in which they would be a majority. It deemed the reapportionment plan to be unconstitutional.

In order to provide a strong basis in evidence that the state was justified inremedying past discrimination through this district plan, the state of North Carolina relied on passages from two reports written by a historian and a social scientist regarding racial discrimination in the state. This evidence was not sufficient. The state of North Carolina had failed to meet the strong basis in evidence standard because the Court found that “there is little
to suggest that the legislature considered the historical events and social-science data that the reports recount, beyond what individual members may have recalled from personal experience.” The Court, therefore, did not accept that the districting plan was consciously, much less precisely, designed to address the racial inequities documents in the reports.

More recently, the Roberts Court used the strong basis in evidence standard to help justify their argument against the constitutionality of remedial programs to promote racial integration in public schools in Parents Involved v. Seattle School District No. 1 (2007). The Seattle School Board did not provide any evidence about remedying past discrimination and instead focused on forward-looking justifications, such as the promise of lessening racial disparities in education. Justice Clarence Thomas’ concurring opinion argues that future consequences or simply allegations of a racial imbalance “cannot substitute for specific findings of prior discrimination.” He emphasizes the idea of proximate evidence that O’Connor stressed in Richmond. The Seattle school district failed to show a strong basis in evidence that the race-based measure were necessary as remedies for specific acts of discrimination by the district, and it was found in violation of the Fourteenth Amendment.

This benchmark, up until Ricci, had solely been used for constitutional purposes involving the Equal Protection Clause. Strong basis in evidence had been utilized to determine whether a certain action fulfilled a compelling state interest under the strict scrutiny test. The test has been tied to the idea that a state agency adopting race-conscious policies must show direct evidence—not a generalized assertion—that racial measures are narrowly tailored to remedy specific acts of proven racial discrimination, usually if not always discrimination by the state agency itself. Because the test raises the standard for evaluating race-conscious measures, it has been an influential tool that the Supreme Court has used to limit race-conscious decision making over the last three decades.

By applying the standard with the same intent but within a statutory context, Ricci represents a radical departure in legal precedent. It marks the first time that the standard has been applied to disparate impact liability. For both equal protection and disparate impact
cases, the strong basis in evidence standard is intended to determine whether race-conscious measures are justified. However, prior to Ricci, the standard was only used to ensure that there are no racial measures for remedial purposes unless specific evidence of intentional discrimination has been shown; it is unclear what it means when strong basis in evidence is applied to a city that fears losing a disparate impact suit not involving intentional discrimination. The set of criteria that one would have to show must be different, but how? It is this question that Justice Ginsburg posed to the Court’s majority, and that I seek to answer later on in this essay.

IV. Case Background

Following an affirming summary judgment from both the district court and the Second Circuit Court of Appeals, the Supreme Court granted certiorari to Ricci v. DeStefano and placed the case on the docket for their 2008-2009 term. On June 29, 2009, they reversed the Second Circuit’s ruling in a landmark 5-4 decision. The Court held that the City of New Haven had violated Title VII of the 1964 Civil Rights Act.

Facts of the Case

In 2003, the City of New Haven fire department administered examinations to determine the most qualified candidates to be promoted for lieutenant and captain positions. The examination was comprised of a written component that accounted for sixty percent of the final score, and an oral section that was worth forty percent. This formula was agreed upon in a contract between the City and the local firefighters’ union.

The City hired Industrial/Organizational Solutions, Inc. (IOS), a company that specialized in designing promotional exams for fire and police departments. They were responsible for developing, testing, and administering the examinations. In designing the test, the company identified the key characteristics, knowledge, and skills necessary to fulfill
the lieutenant and captain positions. They interviewed current lieutenants and captains and asked them to complete a questionnaire as part of their research. When testing the validity of the examination, IOS deliberately oversampled minority firefighters to protect against unintentionally favoring white candidates.

Following the administration of the examination, the New Haven Civil Service Board was required to certify the results. They then used a system called the “rule of three” where a single candidate is chosen from the top three scorers on the ranked list in order to choose who is actually promoted. Since these promotional lists remain active for two years and therefore are infrequent, the stakes for taking this examination were quite high.

A total of 118 New Haven firefighters took the examinations in November and December 2003. For the lieutenant examination, there were seventy-seven candidates—43 whites, 19 blacks, and 15 Hispanics. Only thirty-four candidates passed—25 whites, 6 blacks, and 3 Hispanics. Using the rule of three, the eight vacant lieutenant positions would be filled by the top ten candidates. All ten were white. For the captain examination, there were forty-one candidates—25 whites, 8 blacks, and 8 Hispanics. Only twenty-two candidates passed—16 whites, 3 blacks, and 3 Hispanics. Using the rule of three, the seven vacant captain positions would be filled by the top nine candidates. This included seven whites and two Hispanics.

Given the test results, the City feared that the examinations unintentionally discriminated against minority candidates. The City’s director of human relations noted how a statistically significant racial disparity existed on both exams. Under Title VII of the 1964 Civil Rights Act as amended by the Civil Rights Act of 1991, an employer might be held liable for violating disparate impact, defined as “policies or practices that are not intended to discriminate but in fact have a disproportionately adverse effect on minorities.”

The Civil Service Board first met in January 2004 to discuss certifying the test results. Over the course of five meetings, they heard testimony from firefighters and expert witnesses who supported certifying the test as well as those who opposed its certification.
An industrial-organizational psychologist criticized the weighting formula as arbitrary and advocated the use of performance assessment centers to measure the most qualified candidates more accurately. Others, like Frank Ricci, a dyslexic New Haven firefighter who took the exam, argued that test questions accurately measured the requirements to fulfill the position. The representative from the IOS stated that “in [his] professional opinion, [the examinations were] facially neutral. There’s nothing in those examinations…that should cause somebody to think that one group would perform differently than another group.”

After the fifth meeting, the CSB voted to not certify the test results.

The Civil Service Board’s decision to not certify the test results is the proximate cause of what prompted the lawsuit. The plaintiffs—seventeen white firefighters and one Hispanic who passed the examinations but were denied a chance to be promoted on the basis of their exam results—filed suit against the City of New Haven for violating Title VII of the 1964 Civil Rights Act and the Equal Protection Clause of the Fourteenth Amendment. They contended that the City had engaged in an act of intentional discrimination against whites when the CSB refused to certify the test results.

Both parties filed for summary judgment. The district court awarded summary judgment in favor of the City. They concluded that under Title VII, the respondents’ “motivation to avoid making promotions based on a test with a racially disparate impact…does not, as a matter of law, constitute discriminatory intent.” As for the Equal Protection claim, the district court claimed that the City did not act “based on race” since “all applicants took the same test, and the result was the same for all because the test results were discarded and nobody was promoted.” The Second Circuit affirmed the decision of the district court, adopting the previous court’s reasoning.

The Supreme Court’s Majority Opinion

Justice Kennedy wrote the majority opinion. He claims that the City of New Haven made an explicitly race-based decision when they decided to discard the test results,
and it therefore violated the prohibition against disparate treatment under Title VII absent some “valid” defense. Kennedy argues that “the City rejected the test results solely because the higher scoring candidates were white.” Rather than examining further whether the City’s behavior was intentionally discriminatory within the meaning of the law, Kennedy chooses to focus on the question of “whether the purpose to avoid disparate impact liability excuses what otherwise would be prohibited disparate treatment discrimination.” In doing so, he attempts to reconcile what he believes are two conflicting principles, as well as to provide lower courts with guidance.

Kennedy rejects both the plaintiff and the defendant’s claims in the case. He criticizes the plaintiff’s argument that posits “avoiding unintentional discrimination cannot justify intentional discrimination” as being “overly simplistic and too restrictive of Title VII’s purpose,” given that Congress had added prohibitions of certain disparate impacts in the 1991 Civil Rights Act as an amendment to Title VII. The defendants, along with the US Government, contend that “an employer’s good-faith belief that its actions are necessary to comply with Title VII’s disparate impact provision should be enough to justify” that the City should not be held liable for disparate treatment because it discarded the exams. In doing so, the Court instead adopts a new test—the “strong basis in evidence” standard—for when employers may use race conscious means to prevent disparate impacts. The majority rules that it is permissible for an employer to use the avoidance of disparate impact liability as a defense when engaging in intentional discrimination if and only if there is a “strong basis in evidence to believe it will be subject to disparate impact liability if it fails to take the race-conscious, discriminatory action.”

According to Kennedy, introducing this standard into Title VII litigation allows courts to reconcile the tension between disparate impact and disparate treatment by accepting “violations of one in the name of compliance with the other only in certain, narrow circumstances.” He further states that “the standard leaves ample room for employers’ voluntary compliance efforts” while at the same time it “appropriately constrains employers’
discretion in making race-based decisions.” Employers should not have to concede actual violations, just have strong basis in evidence that they will be held liable for disparate impact.

Kennedy then applies this new standard to the present case and rules that the City has failed to show with a strong basis in evidence that discarding the test results would pose serious risks of being held liable for disparate impact. An employer can no longer administer an exam and then choose to disregard it based upon an adverse racial impact when there is evidence that the exam is job-related, consistent with business necessity, and equally valid, less discriminatory alternatives do not exist. The Court stresses that the strong basis in evidence standard only applies after the selection process is underway. To avoid disparate impact liability, it is crucial that an exam’s validity is evaluated before an exam is administered.

As we should see, in light of the evidence on New Haven’s conduct, the majority implicitly appears to reject the possibility that failure to search for alternative exams that might be equally effective, job-related, and more racially inclusive would provide a strong basis in evidence of liability to lose a disparate impact case. A selection procedure must be narrowly tailored to predict how a candidate would perform the duties of a particular position, as mentioned in Ricci. The opinion does not note, however that a selection procedure could be done that includes finding an exam that both measures appropriate qualifications well and is as racially inclusive as possible. By not mentioning a search for racially inclusive tests as a component in demonstrating a strong basis in evidence of a disparate impact claim, Kennedy discounts its importance and subsequently ratchets down what an employer needs to show to avoid disparate impact liability. Even if a test produces racially skewed results, employers did not have to pay attention to racial inclusiveness in the first place, because it was not one of the requirements that Kennedy set forth to prove that they had reason to fear disparate impact liability. Only under a sympathetic reading of the standard would racial inclusiveness be considered a requirement.

Since the City did not show that it had a strong basis in evidence to fear loss of
a Title VII disparate impact lawsuit in the Court’s eyes, the majority refused to make any ruling on the constitutional claims. The plaintiffs were awarded summary judgment. The case was reversed and remanded.

Scalia’s Concurrence

Justice Scalia filed a concurring opinion where he predicted that this ruling “merely postpones the evil day that the Court will have to confront the question: Whether, or to what extent, are the disparate impact provisions of Title VII of the Civil Rights Act of 1964 [as amended in the Civil Rights Act of 1991] consistent with the Constitution’s guarantee of equal protection?” The tension lies in the fact that the Equal Protection Clause and disparate treatment provisions of Title VII, according to Scalia, presuppose that judgments based upon race are prohibited, while disparate impact explicitly assumes race-conscious decision-making. He is suggesting that disparate impact, as amended to Title VII in 1991, will eventually be held unconstitutional under his reading of the Equal Protection Clause.

Alito’s Concurrence

Justice Alito filed a concurring opinion as well. In it, he claims that two questions must be decided when evaluating whether an employment practice is legitimate: first, an objective question of “whether the reason given by the employer is one that is legitimate under Title VII;” and second, a subjective question of the employer’s intent. Alito argues that because the City did not have a legitimate reason not to certify the results of its original tests, it is not necessary to examine their intent in refusing to do so. He disagrees with the majority in how they interpret the first question. Alito makes the case that the City chose to discard the test for illegitimate political reasons, as an effort to please African-American constituents. He chronicles how an important African-American leader exerted strong political pressure, using his personal ties with the New Haven mayor to push the CSB to vote against certifying the test results. Given these motives for suspending its test, not any
discovery of its inefficacy, Alito believes that a reasonable jury would not have found the City liable for disparate impact. His argument, therefore, does presume to know the City’s intent, even though he says otherwise in the opinion.

**Ginsburg’s Dissent**

Justice Ginsburg, siding with the respondents and the US Government, maintains that the City of New Haven had “good cause that it would be vulnerable to a Title VII disparate impact suit if it relied on those results.” She criticizes the majority for adopting the new strong basis in evidence standard and for not considering that there was “substantial evidence of multiple flaws in the test.” She also defends the rulings made by the district court and Second Circuit by emphasizing that their actions were “race-neutral.” Therefore, Ginsburg anticipates that “the Court’s order…will not have staying power.”

Ginsburg recounts how there has been a history of racial discrimination in municipal employment including New Haven. Even though African-Americans and Hispanics composed thirty percent of the New Haven population in the 1970s, they made up fewer than four percent of the City’s firefighters. In 2003, with the racial composition of New Haven unchanged, African-Americans and Hispanics remained only sixteen percent of the City’s firefighters. In upper level positions, the racial disparities were even greater.

She suggests that Title VII was set forth to advance two complementary goals: “ending workplace discrimination and promoting genuinely equal opportunity.” However, Ginsburg feels that the majority’s ruling has placed these two objectives at odds with one another. She advocates a reading of disparate impact and disparate treatment that are complementary, not conflicting as the majority sees them. She understands her “good cause” standard as significantly less demanding than what the Court appears to require, but still significant.

Ginsburg concentrates her dissent on attacking the strong basis in evidence standard that the majority has implemented. She contends that the Court finds documented failings
on the part of the City insufficient to show vulnerability to a disparate impact lawsuit but does not adequately explain why. The dissent faults the Court for “denying respondents any chance to satisfy the newly announced strong basis in evidence standard.” It is ordinary protocol to allow the Court to devise a new principle, but permit the lower courts to interpret its meaning and then litigate to demonstrate whether the lower courts have indeed interpreted it correctly.

Finally, Ginsburg responds to Alito’s concurring opinion. She dismisses his argument, claiming that there is “scant evidence that [the CSB’s] motivation was anything other than to comply with Title VII’s disparate impact provision.” Ginsburg demonstrates how the board heard from both sides, and that there is little evidence to believe that political maneuvering of any kind took place. She concludes by emphasizing that “New Haven might well have avoided [this unfortunate situation] had it utilized a better selection process in the first place.”

V. The Three Readings of the Strong-Basis-In-Evidence Standard

It is at best uncertain how the lower courts will interpret “strong basis in evidence” now that the Court has applied it to the statutory domain. For starters, the strong basis in evidence standard is surely a more rigorous standard than the good-faith standard that the dissent advocates. Employers can no longer just reasonably believe that they will be held liable for disparate impact. Instead, they now must demonstrate a strong likelihood that a court will rule against them in a disparate impact case. The new standard, therefore, asks employers to fulfill a much more demanding criterion.

Nevertheless, the way in which the Court has adopted strong basis in evidence appears to have provided more questions than answers in seeking to reconcile disparate impact and disparate treatment. It is not clear how the standard’s legal precedents make it relevant to this case. If the Court is so certain that Ricci does not pass this standard, what
cases would actually pass this benchmark? What grounds are necessary for a Court to deem that there is a strong basis in evidence? Justice Ginsburg frames this uncertainty as a central part of her dissent.

I propose three different readings that the lower courts could potentially have of this standard and question whether each reading effectively reconciles disparate impact with disparate treatment. The first construction of the strong basis in evidence benchmark would effectively equate a disparate impact violation with a disparate treatment violation, as Ginsburg suggests the majority was doing. Under this demanding interpretation, evidence of negligence alone is insufficient to prove disparate impact liability. As a result, an employer must prove to a court that, had they used the desired examination, they would probably have lost an intentional discrimination, disparate treatment case—not just a disparate impact case—if a suit was brought. Employers would essentially be required to show an actual constitutional violation against themselves. In this reading, the Court aspires to make it as difficult for the employer as possible to promote race conscious practices. If lower courts choose to apply this reading, then the disparate impact provisions of Title VII will be severely undermined at the expense of disparate treatment, significantly limiting the authority and duty of governments and other employers to avoid policies with disparate impacts that unintentionally increase racial inequalities.

A second approach that lower courts could take to the strong basis in evidence standard would be a more expansive one. Uncertain as how to apply the standard to disparate impact liability, lower courts may be left with the dissent’s good-faith argument as a reasonable equivalent. An employer’s legitimate belief that it would be held liable for disparate impact would in practice be treated as sufficient to fulfill the strong basis in evidence requirement. Lower courts might find this reading to be an appealing way to minimize the Court’s recent condemnation of race-conscious measures. This interpretation would allow the City of New Haven to be justified in discarding the test results and thus, be awarded summary judgment. The “good-faith” reading would be a victory for advocates of
diversity in the workplace—but it would involve effectively ignoring the majority decision of the Supreme Court in Ricci.

Neither of the two readings is one that the Ricci majority would recognize as its strong basis in evidence standard, yet they are possibilities because just what the Court attempts to establish in this case is unclear. Because of this uncertainty, I contend that the third reading that the lower courts may take is one where they simply do not know what to do with the strong basis in evidence standard because its meaning is too ambiguous. According to this construction, the Supreme Court failed to provide sufficient guidance on what criteria are necessary to pass the test and thus, lower courts will be unable to make any valid conclusions. The standard, under this reading, fails to achieve what it sets out to accomplish: reconciling the tension between disparate impact and disparate treatment. Lower court judges cannot even apply the rule because they do not understand what has specifically been resolved. Rather than adopting and thereby enforcing the Court’s standard, they will choose to avoid it until the Court offers further criteria to help delineate what constitutes a viable strong basis in evidence argument.

The Most Demanding Reading of the Strong Basis in Evidence Standard

In the majority opinion, Kennedy states that an employer must demonstrate a strong basis in evidence that it would lose a case for disparate impact if it did not take further acts on racial grounds. This can be achieved through a three-step process laid out by the 1991 Civil Rights Act. First, the employer must show a prima facie case of disparate impact. Second, they must demonstrate that the exam used was not job-related or consistent with business necessity. And third, the employer must show that it had equally valid, less discriminatory alternatives available to them.

a) The Equal Employment Opportunity Commission Guidelines

In making these showings, employers are not without established guidelines,
but the majority opinion in Ricci appears to call these guides into question. In 1978, the federal Equal Employment Opportunity Commission laid out specific guidelines to assist employers in designing nondiscriminatory exams, called the Uniform Guidelines of Employee Selection Procedure. These guidelines primarily concern the issue of validity, which answers the question of whether a test is actually measuring what it is supposed to be measuring. Chief Justice Burger held in Griggs v. Duke Power (1971) that the EEOC standards were entitled to “great deference.” Following Albemarle Paper Co. v. Moody (1975), courts have been required to evaluate employers’ tests against these professional standards. Judges often preferred that validity testing be conducted through an outside company so that there is no apparent bias in the measurements. The EEOC guidelines identify three ways that employers can demonstrate that their examinations fulfill the “business justification” standard set forth in landmark Griggs decision. All three methods emphasize the importance of a systematic statistical analysis to prove that an exam is adequately reflecting the qualifications for a particular position.

The first and most common method that employers can use to evaluate their test’s validity is criterion validation, which looks at the strength of the relationship between test scores and levels of job performance through statistical analysis. One type of criterion validation, known as predictive validation, involves administering a test to applicants but then not using the results in the hiring decision. Later on, an employer can match the performance of those employees hired with their tests results to determine whether the test accurately predicted job performance. A much easier approach, known as concurrent validation, is to give tests to current employees and correlate their tests scores with their job performance.

Content validation, a second method, measures whether the technical knowledge and skill set that are integral to the job are accurately reflected in the content of the exam. It is most often used to determine the validity of achievement tests, which assesses how well a candidate has mastered a particular area or subject. In order to persuade a court of
content validation, an employer must demonstrate that it undertook a rigorous job analysis to understand the specific qualifications that a candidate must possess to succeed in the position, and then applied that analysis directly in writing the content of the exam.

The third and most complex validation method is construct validation, often used for aptitude tests. This technique evaluates whether a test correctly assesses, not achieved knowledge or skills, but rather behaviors and traits that are important to job performance. Since many of the constructs tested, like the capacity to lead and ability to work with others, are fairly subjective or intangible, construct validation often poses a much greater challenge for employers to prove. An employer must conduct an even more rigorous job analysis than in the content validation method to determine the desirable work behaviors and attributes of successful job performance.

b) The Flaws in the New Haven Promotional Exams

In the Ricci dissent, Ginsburg uses this framework to conclude that the test that New Haven used had a “selection process [that] was flawed and not justified by business necessity.” There are several reasons that she contends that the tests used were problematic. The test failed to fulfill any of the three methods of exam validity laid out in the EEOC’s Uniform Guidelines. Ginsburg focuses her criticism of the New Haven promotional exams on the contention that heavy reliance “on written tests to select fire officers is a questionable practice.” First, the City of New Haven failed to conduct any validation study when the IOS designed the captain and lieutenant examinations. Second, she questions why there was no discussion of whether the decades-old agreement that exams should be weighted 60/40 for the written and oral section accurately determined the most qualified candidates or why no alternatives to the formula were ever explored. Even if the City had only used written and oral components, the exam would have better reflected an applicant’s qualifications by weighting the oral sections greater than the written component. An oral exam shows that an applicant can apply the knowledge to real-life scenarios, a skill that is vital to being a
successful firefighter, more so than a written exam. A lawsuit following Ricci deals directly with the weighting formula issue, which I will discuss later.

Third, if the City had exercised due diligence, then it would or should have known that the use of assessment centers represented a “more reliable and less discriminatory” alternative that was available at the time it used the IOS tests. She points out that an expert witness in Ricci advocated for the use of assessment centers, where real-life situations could be simulated, because they better reflected the traits necessary for the position than a candidate’s responses on a written or an oral exam. A 1996 study found that two-thirds of surveyed municipalities have moved away from paper-and-pencil exams to using real-life simulations to address better how a candidate would perform in a real-world situation. The fact that the City of New Haven neglected even to search for others forms of testing was therefore a significant oversight that contributed to why their exams were so severely flawed. Had the City of New Haven conducted more thorough job analyses and ensured the exam’s validity, they could have easily avoided years of costly litigation. The neglect of available alternatives provided further good cause to believe that the City might be held liable.

Fourth, the IOS failed to seek input from the New Haven Fire Department concerning the exam before it was administered to ensure that it reflected the necessary characteristics and knowledge of the positions. Looking at the evidence that Ginsburg lays out in the aggregate, it is clear that the City did not follow the EEOC guidelines, as required under Albermarle, and its promotional exams were not clearly shown to be job-related and consistent with business necessity. Therefore, the City had good cause to think that they would have been held liable for disparate impact had they not discarded the test results.

c) Kennedy Discounts The Exams’ Flaws

However, as noted in the Ricci majority opinion, Kennedy contends that the promotional exams were sufficient in meeting the business justification requirements, even
though they did not evaluate the validity of their exams very thoroughly. According to Kennedy, using a good-faith defense has a slippery slope and would be read too expansively to be a desirable standard. It could potentially “amount to a de facto quota system.” Kennedy fears that this defense would encourage race-based decisions like racial balancing.

In result, the Court’s majority rejects the claim that the City had a strong basis in evidence to believe they would be held liable for a disparate impact suit, even though the City of New Haven had demonstrated that passage rates of non-whites on the exams fell well below the eighty percent satisfactory standard that the Equal Employment Opportunity Commission has set forth to implement disparate impact, and despite testimony indicating the flaws of the test and New Haven’s test selection process.

First, Kennedy claims that “there was no genuine dispute that the examinations were job-related and consistent with business necessity,” without alluding to rigorous statistical validation studies as required by the EEOC standards. The IOS gave testimony to the New Haven Civil Service Board about “details steps” to develop and administer the exams. Moreover, the IOS conducted “painstaking analysis” of the positions to ensure that the test accurately reflected the job qualifications of the positions. The IOS had made sure that minorities were overrepresented when they were testing the exams. The New Haven Civil Service Board also approved the source material from which the IOS designed the exams. Kennedy also points that the IOS had detailed information prepared to prove the exam’s validity, but the City did not take the IOS up on its offer.

Second, he contends that the City “lacked a strong basis in evidence of an equally valid, less discriminatory testing alternative.” The City did not provide sufficient evidence to prove that changing the composite-score calculation, for instance from 60/40 for the written and oral examinations to 30/70, would be an equally valid method to find the most qualified candidates. In fact, Kennedy argues that modifying the weighting formula in favor of a racial balance or adopting a different interpretation of the “rule of three” after the fact would have violated disparate treatment under Title VII, just as much as refusing to certify
the test results.

He also dismisses the use of practical assessment centers as viable options, arguing that they too lacked enough justification to be a viable alternative.

By asserting that the City of New Haven did not have a strong basis in evidence to avoid disparate impact liability, Kennedy accepts that tests that create unintentional, unequal racial consequences are often permissible. Employers may well be justified in using tests that have disparate impacts and do not follow the EEOC guidelines for exam validity. Kennedy largely ignores the facts that Ginsburg mentions in the dissent about the exam’s flaws.

Most importantly, he does not clearly delineate what the sufficient threshold is for an employer to believe that it would be held liable for disparate impact. If the City did their due diligence and did enough to show that their exams were justified without having to ensure the validity of their exams comprehensively, then Kennedy might be implying that the only alternative left to meet the strong basis in evidence standard is to show that the City’s actions were intentional racial discrimination in the first place. Vulnerability to disparate impact lawsuits would then be equivalent to vulnerability to disparate treatment lawsuits, effectively eliminating the 1991 disparate impact provisions added to Title VII of the 1964 Civil Rights Act.

d) The Most Demanding Reading

The most demanding reading of the strong basis in evidence standard would then be that the City of New Haven must show that had it certified the initial examinations instead of discarding them, the City would have probably been held liable for violating disparate treatment. The City would need to demonstrate a strong basis in evidence that the initial test would cause it to lose a case accusing it of intentionally discriminating against a certain group. Since Kennedy rejects proof of negligence as an adequate means of showing a strong basis in evidence, then an employer is left with nothing else but proving intent.
A reading such as this would be closely aligned to the way strong basis in evidence has been read previously in strict scrutiny cases. As I mentioned earlier, the standard was initially designed to ensure that state agencies implementing race-conscious practices were narrowly tailored to remedy specific acts of intentional racial discrimination. It essentially made it more difficult for agencies to employ legitimate race-conscious measures. A demanding construction of the strong basis in evidence standard might, therefore, be seen as an extension of the standard’s original goal to cases of disparate impact.

There appear to be two potential ways that the City of New Haven might have passed the strong basis in evidence standard using this restrictive reading of Kennedy’s opinion—but both amount to showing intentional discrimination, not just the negligence that creates disparate impact liability. First, the City could have consciously decided not to use less discriminatory alternatives like assessment centers, even though using less discriminatory alternatives would have decreased the adverse racial impact and provided a more equitable employment opportunity. The City could have shown evidence that it knew assessment centers were equally valid, less discriminatory alternatives to the current examinations and even though that it had this knowledge, it willingly ignored these facts so that it could produce a result that favored white firefighters. Had the City shown it adopted policies no more valid than those it rejected, and ones that it knew would create a racial imbalance, it would have succeeded in achieving the strong basis in evidence standard using this framework.

Another way the City could have met Kennedy’s apparent criteria was by showing that the City purposely chose not to revise the two-decades-old contract with the firefighters union, which laid out that the 60/40 weighting formula for the written and oral portions of the exam, even knowing alternatives existed that would not disadvantage minority firefighters. Despite the fact that City was aware that a neighboring city, Bridgeport, had changed its formula so that the oral component was given primacy over the written component because it better “addressed the sort of real-life scenarios fire officers encounter
on the job,” and Bridgeport now had minorities that were fairly represented in its exam results, the City of New Haven did not emulate Bridgeport. It might therefore have been expressing a preference for white firefighters to remain overrepresented in its upper level positions. By setting out the parameters early on with the IOS, the IOS could not explore other equally valid, less discriminatory alternatives to the 60/40 weighting formula or more importantly, whether or not that weighting formula best reflected the qualifications for the job.

While it would be possible to pass the strong basis in evidence threshold in these ways, I would predict that most, if not all, cases would fail to meet this reading of strong basis in evidence. Cities may well be reluctant to provide evidence that actually incriminates them. It is paradoxical, verging on absurd, to impose a standard that essentially asks employers to show a high probability that a disparate treatment case can be successfully brought against them. They would not only be forced to produce evidence that casts their agencies in a pejorative light, but also willingly tarnish their own reputation.

It is unclear as to what type of physical evidence is necessary today to demonstrate that the City had intended to discriminate. I have provided some examples above, but these are simply conjectures. We no longer live in a world where explicit racism is an accepted cultural norm. Written documentation that clearly enumerates that an employer has a specific racial preference is much harder to come by today than seventy or eighty years ago. Discovering an employer’s intent is often speculative at best.

e) Consequences of the Most Demanding Reading

Although it may be too soon to tell how race-conscious measures will actually be affected by the implementation of the strong basis in evidence standard, the most demanding construction would almost certainly have a negative effect, should the lower courts adopt that position. If the main objective of Title VII is to one day live in a world where workplace discrimination no longer exists, then the most demanding reading of
the strong basis in evidence standard runs counter to that aim. It increases the bar that employers must fulfill to show they had a legitimate belief that they would have been held liable for disparate impact, and it legally permits employment tests with unintended adverse racial impacts. In result, the new standard threatens to invalidate the long-accepted EEOC guidelines as well as opens the door to more tests with disparate impacts.

i. The Potential Invalidation of the EEOC Guidelines

By arguing that the City’s exams were sufficient, Kennedy appears to show that an exam can be job related and consistent with business necessity without having to comply with the testing guidelines. If so, it would be a shocking break from precedent. Ricci could have potentially invalidated, or at least severely diminished, the importance of these forty-year-old EEOC guidelines. Employers would no longer have any incentive to spend money on validation studies. They could also be insulated from disparate-impact suits. Future tests that have substantial disparate impacts and are not statistically proven to be job related could potentially be viewed as acceptable in the eyes of the Ricci Court. Without any guidelines for employers to use, it would be much harder for employers to prove that they will lose a disparate impact suit. There would no longer be a generally accepted method for employers to adopt to ensure that they have avoided disparate impact liability. Instead, they would be at the whims of the court to decide which evidence is sufficient to demonstrate a strong basis in evidence of unintentional discrimination. Disregarding the EEOC guidelines would therefore threaten employers’ efforts to encourage diversity as well as allow tests with unintended racial consequences to persist.

The viability of the EEOC guidelines as a source to demonstrate a strong basis in evidence remains an unanswered question. How the Court responds to similar litigation in the future could very well shift the way employers and lawyers regard the EEOC guidelines in measuring test validity. While I do recognize that Kennedy may be hinting at invalidating the EEOC guidelines, the most recent literature on the strong basis in evidence standard
for both cases of disparate impact and equal protection have emphasized the importance of statistical analysis. Legal scholars have focused on using the three types of validity, as laid out by the EEOC, to prepare employers should they face a suit resembling Ricci.

ii. Opens the Door to More Tests with Disparate Impacts

The promise of promoting diversity in the workplace is to correct for a legacy of discrimination and changes the cultural norms to accept these minorities as equals. However, the Court’s imposition of the strong basis in evidence standard in Ricci threatens the success of this goal. By allowing for more tests with unintended, unequal racial impacts to exist, employers do not have to do their due diligence and can be less cautious about the exams that they use without fearing a disparate impact suit. It also raises the bar that employers must fulfill in order to demonstrate disparate impact liability, making it more difficult for an employer to show that they would lose. In result, employers are likely to be more hesitant in discarding examination results, or similar means, to remedy disparate impacts.

Unless the Ricci opinion radically alters the lower courts’ belief in these ways, the likelihood remains that if the City had certified the test results, it might well have lost a disparate impact suit for having an adverse racial impact. Ricci already makes clear that by not certifying the test results, the City violated disparate treatment. In light of these situations, employers are not left with the “ample room for voluntary compliance” as Kennedy believes it does and undermines the promotion of equal employment opportunities. The new standard significantly lessens the authority and duty of governments to avoid policies with disparate impacts.

Most frustratingly, by permitting more tests with disparate impacts, the Court’s new standard does not resolve the tension between disparate impact and disparate treatment that Kennedy seeks to achieve. On the contrary, it diminishes disparate impact, making it a much weaker protection than disparate treatment, and places their two goals at odds with one another, when both principles together are critical to fighting discrimination in the workplace. Therefore, in its most rigorous reading, the strong basis in evidence standard
does not effectively reconcile disparate impact and disparate treatment. Instead, the standard unjustly imbalances the two principles and discourages employers from making more race-conscious decisions in the workplace.

The “Good-Faith” Reading of the Strong Basis in Evidence Standard

The Court’s introduction of the strong basis in evidence standard to statutory provisions was a radical shift in the way the standard was originally understood. Since Kennedy fails to spell out precisely how an employer can prove with a strong basis in evidence that they would lose a case of disparate impact, lower courts could be left with little more than the dissent’s good-faith reasoning as the minimal requirement to fulfill the standard. In this reading of strong basis in evidence, lower courts may adopt the good-faith argument as a reasonable equivalent.

The strong basis in evidence benchmark could then be fulfilled so long as an employer demonstrate that it has a legitimate belief—a reasonable basis to fear—that it would have been held liable for disparate impact. It would be sufficient proof to show a “reasonable basis” of disparate impact liability to constitute a “strong basis in evidence.” Applying this interpretation, Ricci would satisfy the strong basis in evidence standard because the City of New Haven had good cause to believe that a court would have found the City guilty of violating disparate impact.

The City could establish a reasonable basis by showing more than just a mere statistical disparity. It would also have to show a good cause to believe that it would not meet the “business necessity” requirement and thus, had reason to in fear disparate impact liability. As I have already shown, Ginsburg’s dissent contends that the examinations themselves had several deficiencies and were of questionable character. The weighting formula that the City used was antiquated, and the City neglected the use of available alternatives such as assessment centers, which were “more reliable and less discriminatory in operation,” and so forth. Taken together, the City had good cause to believe that they
could not fulfill the “business necessity” requirement and would be held liable for disparate impact if the nonwhite firefighters brought suit. The City would fulfill the strong basis in evidence standard and therefore, would have been justified in discarding the test results.

This looser strong basis in evidence standard is akin to a paradigm used in equal protection cases called the rational basis test. It is the default standard of review, and the lowest level of scrutiny. The rational basis test determines whether a governmental action was a reasonable means—not arbitrary—to achieving an end. Furthermore, the action must be rationally related to furthering a legitimate government interest. Many of the terms used to describe the rational basis test are analogous to the dissent’s good-faith argument. At its core, a “good-faith” reading of the strong basis in evidence standard is merely a question of reasonableness. Just as the narrowest reading might be a statutory proxy for a strict scrutiny test, it appears the same is true for an expansive reading and a rational basis test.

The relaxed reading of the strong basis in evidence standard provides a more concrete answer as to what can and cannot pass the threshold than the more demanding construction. By framing the standard in this manner, there is greater flexibility about what can meet this test, so this standard is more feasible for employers to achieve. An employer is not required to incriminate itself in order to achieve the standard, nor to have to worry about what constituted intent. Although it still lays the burden of proof on the employer, the benchmark is not nearly set as high as the more demanding construction.

Consequences of the “Good-Faith” Reading

Should the lower courts adopt the “good-faith” reading of the strong basis in evidence standard, employers would not have the lose-lose situation that they would experience under the more demanding reading. Employers could more freely shift to inclusive race-conscious measures without the fear of reverse discrimination lawsuits so long as they could demonstrate a legitimate good cause to fear a disparate impact liability. The EEOC guidelines would remain intact. More importantly, on a practical as well as a symbolic level, incentives would still exist for employers to promote diversity in the
workplace. Employers would not be restricted in making race-conscious decisions, as they would be under the most demanding reading of the standard.

Although lower courts would be effectively ignoring the Ricci majority opinion by embracing this reading, they would be able to reconcile the tension between disparate impact and disparate treatment. By lessening the amount of scrutiny to prove disparate impact liability, employers would be allowed to be more inclusive in the types of candidates that they choose to hire. Employers could successfully satisfy the two complementary pillars of Title VII that Ginsburg discusses in her dissent: “ending workplace discrimination and promoting genuinely equal opportunity.” The good-faith test, aligning itself with the expansive view of disparate impact embodied in the 1991 amendments to the 1964 Civil Rights Act, recognizes that both principles—disparate impact and disparate treatment—are vital complementary components to eradicating discrimination at the workplace.

Yet, it is important to note that disparate treatment remains the stronger provision because it is explicitly protected under the Fourteenth Amendment, where disparate impact is not, at least as the Court has interpreted the Equal Protection Clause. There are two ways to strengthen disparate impact so that it is perceived as fully equal to disparate treatment. Judges could reinterpret the Equal Protection Clause expansively to include disparate impact. Or, a more challenging option would be to amend the Constitution to enumerate disparate impact as a provision of the Fourteenth Amendment. These possibilities will be explored further when discussing the future of the strong basis in evidence standard.

The Ambiguous Reading of the Strong Basis in Evidence Standard

The final interpretation that lower courts may have to Ricci’s strong basis in evidence benchmark is one based upon uncertainty. The legal evolution of the strong basis in evidence standard, as well as Ricci itself, fails to provide specific guidance about what
criteria are necessary to fulfill the standard or how it should even be applied. Without a
good understanding of what strong basis in evidence entails, judges will have difficulty
being able to apply the standard. Therefore, lower courts could potentially read the strong
basis in evidence standard as being too vague to have any legitimate legal meaning.

Besides confirming that the strong basis in evidence is a higher threshold than
the dissent’s good-faith argument, the Court does not fully flesh out what is meant by
instituting a more rigorous standard. It is unknown what cases would actually meet the
strong basis in evidence standard. The Court does not enumerate a specific step-by-step
approach, like the McDonnell-Douglas burden-shifting framework for disparate treatment
and the related Griggs test for disparate impact, as to how an employer can successfully
reach the strong basis in evidence threshold.

Instead, the majority makes broad, descriptive claims such as the standard “leaves
ample room for employers’ voluntary compliance efforts” and “constrains employers
discretion in making race-based decisions” without detailing any specific criteria. If no
explicit steps exist to pass the standard, it then begs the question of why the Court is so
certain that Ricci does not meet the strong basis in evidence standard. It is foolish to think
that a judge can make a ruling applying the strong basis in evidence standard when no
criteria exist for how to meet its threshold.

Kennedy failed to offer reasonable evidence as to why they believed that the strong
basis in evidence standard could make the jump from Equal Protection cases dealing with
absolute racial preferences to a disparate impact case in a statutory context. He assumes
that the “tension between eliminating segregation and discrimination” are the same as “the
interplay between the disparate impact and disparate treatment provisions of Title VII,” but
this connection is grossly oversimplified for the reasons that Ginsburg confronts. Since the
Court’s application of strong basis in evidence in Ricci is such an unwarranted departure
from the standard’s legal evolution, it is unlikely that lower court judges will be able to use
the standard effectively for disparate impact cases.
Consequences of the Ambiguous Reading

If this line of reasoning holds true in application, then employers too will be unsure whether they should encourage race-conscious practices with the risk that they may or may not be held liable or should they just hedge their bets. They will also sit in a holding pattern in the interim. The applicability of the EEOC guidelines will also remain in flux if lower courts resist ruling on the strong basis in evidence standard. The one consequence that is for certain should lower courts adopt an ambiguous reading is that discriminatory employment practices will persist. Tests with disparate impacts will go unchallenged and will continue to flourish. This effect, much like in the most demanding reading, undermine the aims of Title VII and reduce its importance.

Therefore, the ambiguous reading of the strong basis in evidence standard most clearly represents the critical failure of the standard itself. Strong basis in evidence is designed, first and foremost, to reconcile the tension between disparate impact and disparate treatment. Yet, if lower court judges do not even have a strong grasp of how to apply the standard appropriately, then the tension will persist and hotly contested issues will remain unresolved. Until the Court provides further clarity regarding what constitutes a passable strong basis in evidence, I suspect that a majority of lower courts will hold off on adopting the Court’s standard and thereby will exacerbate further discrimination in the workplace.

VI. The Future of the Strong-Basis-in-Evidence Standard

Because it has only been a year since the Ricci decision was released, it may be too soon to draw any conclusions regarding which reading the lower courts will take on the strong basis in evidence standard. Nevertheless, there have been a few instances of legal activity within the past year that have sought to apply the principles of Ricci and provide
further clarification on enumerating what is necessary to fulfill the strong basis in evidence threshold. A great cloud of ambiguity still remains over how the standard works, but courts have now tried to apply the Court’s reasoning. For now, the lower court’s rulings seem to suggest two different directions for evaluating the standard—ones that follow the spirit of Ricci and ones that try to limit its potential influence. In this section, I will examine how recent court decisions on the district and circuit level have interpreted Ricci, as well as what the potential long-term implications of the strong basis in evidence standard may be.

Recent Court Decisions


Less than a month after the Supreme Court issued the Ricci ruling, a federal court in New York held that the New York City Fire Department’s written entrance examinations unintentionally discriminated against African-American and Hispanic applicants and thus, violated Title VII of the 1964 Civil Rights Act. Of the approximately 3,1000 African-American applicants and 4,200 Hispanic applications who sat for the test, the City only hired 184 African-American firefighters and 461 Hispanic firefighters. The City’s selection process prevented roughly a thousand African-American and Hispanic candidates from joining the firefighter department. The Eastern District concluded that the City’s promotional tests had a statistically significant racial disparity between whites and nonwhites, and the tests did not establish that it was job-related or consistent with business necessity.

However, Judge Nicholas G. Garaufis was quick to point out that the Ricci does not directly apply to this case. Where the Ricci Court ruled that New Haven would likely have not been liable for disparate impact, the City of New York feared an actual disparate impact suit. Judge Garaufis does believe that Ricci’s discussion of the importance of validity testing is relevant. Even though the Supreme Court may have potentially invalidated the EEOC guidelines with Ricci, Garaufis makes a strong effort to reinforce the importance of these guidelines by explicitly applying them. He declares in his opinion that “municipalities
must take adequate measures to ensure that their civil service examinations reliably test the relevant knowledge, skills and abilities that will determine which applicants will best perform their specific duties.”

He criticizes the City of New York for taking even fewer steps than the City of New Haven had in validating their tests. Garaufis then uses the criteria that the EEOC has laid out to thoroughly detail how the City of New York’s examinations were poorly constructed. He points out several flaws in the tests, such as how they failed to demonstrate that the reading level was appropriate. They did not test for the important abilities of a firefighter. The City also failed to prove that the exams administered actually tested the abilities they intended to test. The City had imposed arbitrary pass/fail scores that were unrelated to the qualifications for the job.

Five months later, the federal district court declared that the City of New York’s firefighters examinations had not only unintentionally discriminated, but also intentionally discriminated against minority candidates and therefore violated the disparate treatment provisions of Title VII and the Equal Protection Clause. Judge Garaufis writes that it was not a “one-time mistake or the product of benign neglect. It was part of a pattern, practice, and policy of intentional discrimination against black applicants that has deep historical antecedents and uniquely disabling effects.” He criticizes the City for failing to do its due diligence in devising tests that were nondiscriminatory in nature and argues that there was a strong basis in evidence that the City knew its exams were discriminatory, yet failed to take sufficient remedial action.

What was significant in this ruling was that Garaufis determined that the City intentionally discriminated using only the statistics about the number of minority applicants hired compared to the percentages in the applicant pool and the city’s population, along with evidence of failure to fulfill due diligence as defined by the EEOC guidelines. There was no other evidence of any explicit discriminatory statement to establish intent. This decision to focus only on the numbers and decision-making processes as constituting “intent” might
signal a shift away from the courts’ past reliance on clear, smoking-gun statements of intent. Moreover, the district court’s ruling has insulated the City from a reverse discrimination claim should the City choose to revise their firefighter exams as Judge Garaufis ordered, because they have been provided with the strong basis in evidence that the tests were discriminatory. Strong basis in evidence allows them to have a legitimate built-in defense to prove why remedial action was necessary. Although this case found Ricci not to be controlling, the Eastern District of New York seems to indicate that employers still need to illustrate that their examinations have undergone rigorous statistical analysis and are in fact valid. Judge Garaufis was so emphatic about stressing the EEOC guidelines to establish the validity of the examinations that it suggests that there might be some resistance should the Court decide to invalidate the EEOC guidelines. The City of New York case, therefore, indicates that lower courts are willing to go against the principles laid out in Ricci and openly reject the de-emphasis of EEOC guidelines compliance that Kennedy appears to have advocated in the majority opinion.


NAACP v. North Hudson Regional concerns a disparate impact challenge to a consolidated municipal fire department whose residency requirements for hiring had, allegedly, a disparate impact against African-American applicants. Eliminating the residency requirements was intended to achieve a more desirable racial distribution of candidates for hiring. In February 2009, the New Jersey district court concluded that the Regional department did unintentionally discriminate. They granted the NAACP’s motion for a preliminary injunction and refused to allow the Regional to hire new candidates until it expanded its residency requirements. However, in light of the Ricci decision that came after the February 2009 decision, the US Court of Appeals for the Third Circuit remanded the case back to the district to reconsider the preliminary injunction applying the principles from Ricci to its factual and legal analysis.

On April 23, 2010, Judge Debevoise for the New Jersey District Court lifted
the ban against hiring firefighters and argued that “the residency requirements in this case furthers legitimate business goals in a significant way” in accordance with the Ricci decision. He began his opinion with a full discussion of Ricci and how it relates to the current case. When the plaintiffs cited US v. The City of New York (2009) to defend their position, Debevoise denied its applicability and stated his disagreement with the Eastern District’s ruling. He acknowledged that the plaintiff’s claims are factually different from Ricci, but believed that Ricci must still play a factor “because the [present] case implicates the tension between disparate impact and disparate treatment.” Debevoise contended that Ricci should be used any time there is a conflict between these two principles, as the Court’s majority decision suggests. He took full license in adapting the standard to reflect his own views. According to the district court’s ruling, “the strong basis in evidence standard applies whether, in order to cure alleged discriminatory impact, the challenged action is initiated by the employer, such as the NHRFR, or whether the employer is ordered by a court to take a challenge action.”

Applying Ricci, the district court must determine whether a strong basis in evidence existed to discard the residency requirements, the remedial action at hand, in order to avoid violating the disparate impact provisions of Title VII. Debevoise explicitly notes that the Supreme Court failed to provide detailed guidance in determining how the strong basis in evidence standard should be applied. He reformulates the Ricci standard to mean “whether the plaintiff’s have established a strong basis in evidence that they are likely to succeed on the merits of their disparate impact case.” The court finds quite clearly that statistical evidence exists showing a racial disparity, establishing a prima facie case of disparate impact. Only two of 323 employees of the NHRFR are African-American. Looking at proportions comparing the number of African-Americans employed in state and local government jobs in each county, it was indisputable that African-Americans were significantly underrepresented in the NHRFR’s work force and was likely attributed to the use of residency requirements.
The burden now shifts to the NHRFR to prove that the residency requirements have a valid business necessity defense, as laid out in Ricci. The NHRFR argue that the residency requirements were necessary for three reasons: first, to avoid a risk of suit by a groups of Hispanics; second, to follow the rules of a previous settlement where the NHRFR agreed to reach and attract additional qualified applicants of Hispanic/Latino origin; and third, to correct past allegations of discrimination and a historic under-representation of Hispanics in the NHRFR. Expanding the residency requirement would, in effect, contradict the principles underlying this previous settlement, which was intended to increase the number of qualified Hispanic applicants. The fear of diluting the accomplishments of this settlement agreement provides the NHRFR with a business justification for resisting the expansion of its residency requirements. The Court also finds that “averting future liability to other groups of Hispanics is another substantial business justification, since the expense of additional litigation could cost the NHRFR significant sums of money.”

NHRFR also provides two additional reasons for why there is a business justification for residency requirements. First, because these municipalities have a large Hispanic population, it is vital that Hispanics are well represented in the NHRFR’s workforce so that “the NHRFR’s protective services workers can better communicate with the [Spanish-speaking] population that they serve.” Second, because residency requirements increase the probability that a firefighter will live in the community, it will ensure that more firefighters will be able to report to work more quickly in the case of an emergency. These justifications taken together, according to Debevoise, indicate that the NAACP is not likely to succeed on the merits of their claims and therefore, have not established with a strong basis in evidence that discarding the residency requirements avoided disparate impact liability.

NAACP v. North Hudson marks the first time that the strong basis in evidence standard has been applied to a hiring practice outside a civil service test. What makes the case so interesting is how the introduction of the strong basis in evidence standard has forced two minorities groups—both of whom wish to increase their representation and have the
opportunity to equal employment—to be pitted against one another. Where the Hispanics have used the standard to their advantage, the African-Americans in the community now must live with the residency requirements that quite clearly have a disparate impact.

An alternative that Judge Garaufis fails to look at is one that accommodates for Hispanics while still restricting the use of residency requirements. In this instance, the regional fire department can still actively attract Hispanic applicants within the specified municipalities that the previous lawsuit pertains to and at the same time make it more accessible for African-American candidates to apply as well. A plan such as this would have improved equal opportunities for both minority groups.

The case illustrates that applying strong basis in evidence to efforts to avert disparate impact liability does not comply with the goals of Title VII. Rather, it allows discriminatory employment practices to continue and compels judges to choose between which minority group is more “justified” in having race-conscious practices. NAACP v. North Hudson is thus consistent with the principles set forth in Ricci and suggests that lower courts are willing to adopt the high court’s new standard.


In the aftermath of the Supreme Court’s Ricci decision, litigation has continued over the contested New Haven exams. Michael Briscoe, an African-American firefighter who took the promotional test in 2003, filed suit after the Court’s ruling, claiming that the 60/40 weighting formula for the written over the oral component had a disparate impact on African-American candidates. Briscoe belatedly discovered that he was the top scorer on the oral section of the exam for the lieutenant position, yet scored only twenty-fourth overall due to his lower score on the written portion and would not be among the group that was eligible for promotion.

In April 2010, Connecticut’s federal district court dismissed Briscoe’s claims in accordance with the Supreme Court’s ruling in Ricci. Even though Briscoe argues that
it is unfair to apply Ricci to his case, Judge Charles S. Haight, Jr. states that the Supreme Court specifically “foreclosed subsequent disparate impact suits, such as Briscoe’s, against the City based on the 2003 exam” by introducing the strong basis in evidence standard and then applying it to Ricci. He mentions that had the Ricci majority adopted the dissent’s approach of remanding the case in light of the strong basis in evidence standard, then Briscoe could have intervened to raise his claims over the weighting formula. But since the Supreme Court did not remand the case, Briscoe “cannot circumvent that decision by filing another lawsuit with respect to the same exams.” Judge Haight reaffirms the Supreme Court’s decision that there was not a strong basis in evidence that the City of New Haven would lose a disparate impact suit had they certified the tests, and thus rules to dismiss the case.

Briscoe v. New Haven is another instance of the lower courts following the spirit of Ricci. Since the district court was tightly constricted by the decision of the higher court, the ruling is not particularly surprising and does not reveal anything particularly novel about using the strong basis in evidence standard. Although, had Judge Haight ruled in favor of Briscoe, then there would have been an interesting divide between the district court and the Supreme Court on applying the standard. This case does bring closure to the extensive litigation over the City of New Haven’s 2003 promotional exams.

d) Further Cases Remanded in light of Ricci

Like the NAACP v. North Hudson case, several other cases have been remanded in order to reconsider the ruling now that Supreme Court has ruled on Ricci and introduced the strong basis in evidence standard. On the same day as the Supreme Court issued the Ricci decision, the Court also vacated judgment and remanded a similar case back to the lower court, Oakley v. Memphis (2009). The case concerns the Memphis Police Department who decided to cancel its promotional exam because too few minority candidates score well enough to be promoted and thus, Memphis wanted to avoid disparate impact liability. Just as in Ricci, the plaintiffs brought suit alleging that the Memphis Police Department
intentionally discriminated against them by using race as the basis for canceling the valid promotional process. The Sixth Circuit Court of Appeals affirmed the district court’s dismissal of the plaintiff’s complaint on the grounds that the City’s desire to avoid a potential lawsuit was a legitimate nondiscriminatory intent. Now that the Supreme Court has remanded, the Sixth Circuit Court of Appeals will have to reconsider their decision applying the principles of Ricci to its factual and legal analysis as the New Jersey district did in NAACP. Nearly a year later, they have yet to release a new decision.

The Second Circuit Court of Appeals also released a decision in April 2010 remanding a case, Bridgeport Guardians, Inc. v. Delmonte (2010), to the lower court to reconsider given the Ricci decision. Bridgeport deals with an interim modification order set in place so that the Bridgeport Police Department could adopt race-conscious promotional and hiring practices, and reduce tests with disparate impacts. The plaintiffs argue that the order violates Title VII and the Fourteenth Amendment. Because the Supreme Court issued Ricci after the order was appealed, the district court could not review the merits of the order using the strong basis in evidence standard from Ricci. The Circuit Court declined to exercise its discretionary power and remanded the case to the district level to reconsider the order in light of the Ricci decision.

Over the next year, I would expect many higher courts to follow suit and remand cases to apply the principles from Ricci. The spotlight on the development of the strong basis in evidence standard now turns to the lower courts. I suspect that judges will examine the disparate impact doctrine with further scrutiny, but the ways in which it will be applied, I predict, will vastly differ. Since it is so early in the post-Ricci litigation period, judges will likely want to leave their mark in shaping the standard as seen in the cases that I have discussed. Judge Debevoise disagreed with the US v. City of New York ruling and raised his concerns explicitly in the NAACP v. North Hudson opinion. I believe that we will continue to see a heated discussion amongst judges in how to understand the new standard until the Supreme Court clarifies its meaning.
The Long-Term Viability of the Strong Basis in Evidence Standard

Strong basis in evidence has already had an effect on recent court decisions. But, will the standard only have a short-term effect or will it have more far-reaching implications? I argue that it has the potential to do both. The standard itself may have a brief shelf life, but its long-term influence could dramatically change the way the disparate impact doctrine is perceived. Examining the constitutionality of disparate impact helps to resolve this discrepancy.

Although the Court chooses not to confront the constitutional issue over disparate impact, the Ricci decision is rooted in these constitutional concerns and “reaches further than its reasoning would suggest.” The Court imports the strong basis in evidence standard from equal protection case law, but does not clearly flesh out whether disparate impact should conform to equal protection norms or whether the doctrine actually violates the Equal Protection Clause. The primary divergence between the majority and the dissent rests on how disparate impact should be read. The majority’s ruling is consistent with the prevailing trends that the Court has taken over the past two decades towards restricting the doctrine’s influence. On the other hand, Ginsburg’s dissent reflects the expansive view of disparate impact as set forth in the 1991 amendments to the 1964 Civil Rights Act in an effort to combat Wards Cove.

Ginsburg predicts in the dissent that the strong basis in evidence standard of liability to lose a disparate impact suit “will not have staying power.” I believe that with due time, she will be correct. The standard will be invalidated through one of two possible trajectories. Should the Court experience an unexpected ideological shift in the near future toward more liberal-minded justices, the Court will likely overturn Ricci and nullify the standard. They will embrace a more expansive conception of disparate impact that is viewed as complementary to the Equal Protection Clause, and in result, employers will have a greater incentive to promote voluntary compliance efforts. While this option is viable
possibility, its success will depend on the timing of vacancies on the Court.

Assuming that the Court’s ideological composition will remain the same, the more likely path for the strong basis in evidence standard is for the Court to continue building upon the Ricci precedent. The new standard provides a foundation so that the Court can support more radical change over disparate impact in the future by further narrowing the doctrine’s scope to fit the majority’s interpretation of the Equal Protection Clause. Should the Court read the Equal Protection Clause to prohibit race-conscious decision-making, then disparate impact, which implicitly assumes such practices, is in inherent conflict. As Justice Scalia suggests in the concurrent opinion, the Court is likely to deem Title VII’s disparate impact provision as unconstitutional with the Equal Protection Clause, or at least severely limit the provision.

Overturning disparate impact would be an aggressive act for the Court to take, but an unsurprising one given its recent ideological trajectory. Decisions like Parents Involved v. Seattle School District 1 (2007) and Gratz v. Bollinger (2003) show a Court that is eager to thwart affirmative action policies. The Ricci ruling also reveals a Court that is willing to set aside established precedent, namely Griggs and Albermarle, and apply a new standard from a different area of the law in order to reformulate the way disparate impact is used. Clearly, the Court has signaled with Ricci that they have more work to do interpreting disparate impact. It appears that they may have purposely crafted Ricci as a skipping stone to eventually ruling disparate impact as unconstitutional. Thus, it is reasonable to conclude that regardless of the Court’s ideological makeup, the strong basis in evidence standard is likely to be an impermanent standard.

Another Congressional Fix?

Now that the Court has weakened disparate impact by introducing the strong basis in evidence standard, it seems as though the fate of disparate impact, at least in the short-term, has again fallen into the hands of the legislature. Congress came to the rescue
and officially codified disparate impact as an amendment to the 1964 Civil Rights Act in 1991 after Wards Cove significantly undercut the doctrine two years prior. Today, with comfortable Democratic majorities in both houses of Congress and a Democratic President, it would seem that conditions would be more favorable to strengthen disparate impact than in 1991. In fact, the 2009 Lilly Ledbetter Fair Pay Act, the first bill that Congress enacted when Obama became President and dealt with the statue of limitations for filing an equal-pay lawsuit regarding pay discrimination, was intended to directly counteract a 2007 Supreme Court decision. That said, the issues and principles in Ricci is a far more complicated and politically-charged than the simple rule overturned by the Ledbetter Fair Pay Act. While the majority of the 111th Congressional session has dealt with resolving the recent financial crisis and reforming health care, it is hopeful that the 112th session, following this year’s mid-term elections, will tackle the Ricci decision and bolster disparate impact yet again.

If Congress does enact legislation that strengthens the disparate impact doctrine, then I do hope that the Supreme Court will respect this legislation and choose not to hear a case that challenges it. Nevertheless, action from Congress may just be the smoking gun that provokes the Court, as Scalia contended, to address head on the constitutional question of “whether, or to what extent, are the disparate impact provisions of Title VII… consistent with the Constitutions’ guarantee of equal protection?” Since the Court seems to indicate that disparate impact is unconstitutional in Ricci, it is possible that Congress may be doing more harm than good in the long-term by enacting legislation if the Court chooses to actively pursue this issue.

Should a Congressional attempt to restore disparate impact back to its original 1991 definition fail, only two other options remain, and neither has strong prospects. The easier alternative of the two would be to have a radical change in the ideological composition of the Supreme Court to make way for justices who could reinterpret the Equal Protection Clause to include disparate impact. Unless one of the conservative-leaning justices surprisingly
dies or retires in the very near future, this path does not seem promising. The justices who are most likely to retire from the Court are liberal leaning and Obama’s appointments, as a Democrat himself, are not expected to change the ideological composition of the Court.

The more challenging option would be to amend the Constitution to enumerate disparate impact as a provision of the Fourteenth Amendment. In order to do so, the amendment would have to face the extremely difficult burden of either being approved by three-fourths of the state legislatures or ratifying conventions in three-fourths of the states, as per Article V of the Constitution. While this alternative would be the best defense for protecting against conservative efforts to limit to right to prohibit unintentional forms of discrimination, I do not anticipate that an amendment codifying the disparate impact doctrine would generate the necessary political support and media coverage to be successfully enacted as an amendment to the Constitution. Therefore, while I would encourage Congress to take action on the issue, it is likely that the power in developing the strong basis in evidence standard will lie with the courts.

VII. Conclusion

After a six-year legal struggle, the thirteen white firefighters and one Hispanic firefighter who sued after the City of New Haven had thrown out the test results were officially promoted to either lieutenant or captain on December 10, 2009. Now that the dust has settled on Ricci, what can employers and judges alike take away from the case? In short, Ricci has made matters much more complicated. It is uncertain what evidence is sufficient for an employer to demonstrate a strong basis in evidence standard that they would lose a disparate impact suit, and judges are left with little guidance to determine what classifies as appropriate in fulfilling the standard and what does not. Furthermore, if the Court was so certain that the City of New Haven’s examinations were job-related and consistent with business necessity even though they did not follow the EEOC guidelines, they seem to suggest that employers can willingly ignore the EEOC guidelines. But, they
fail to fully elaborate on the issue, puzzling both employers and judges.

I expect that further judicial wrangling from the district level up to the Supreme Court itself will persist in hopes of resolving the questions that the Court created when they implemented the strong basis in evidence standard to disparate impact. The Court’s majority intended to apply the standard to reconcile disparate impact and disparate treatment, two seemingly conflicting principles of Title VII according to Kennedy, for once and for all. I do not think the Court achieved this goal. In actuality, the Court diminishes the disparate impact doctrine at the expense of disparate treatment by allowing tests with unintended racial consequences. They have locked in discriminatory employment practices and made voluntary compliance efforts with Title VII significantly more difficult.

I look forward to a time in the near future when the strong basis in evidence standard is overturned and disparate impact reverts back to its initial construction either due to a change in the Court’s ideological composition or from Congressional action. Employers should have an incentive to promote diversity in the workplace. Citizens should have a right to prohibit measures that unintentionally discriminate against minorities. The Supreme Court’s newly established precedent threatens this right. I fear that the Ricci decision is just the beginning of further limitations of the disparate impact provisions of Title VII that will lead up to the frightful day when the Court holds disparate impact as unconstitutional with the Equal Protection Clause. Hopefully, the nation will never encounter such a crippling setback in affirmative action efforts, and policies to promote diversity in the workplace will flourish.