ECONPress

of Northeastern University

ECONPress is a publication for undergraduate compositions in economics. We publish twice a year during each fall and spring semester. ECONPress invites the highest quality submissions from undergraduate students in various economics related disciplines. It provides a forum for the undergraduate economics community to engage in active discussion and debate about the topics, theories, and applications they've learned in the classroom.

Special Thanks

Edward J. Meehan
Marcel Landesmann
Prof. Steven Morrison

Edited and published at Northeastern University

Editorial Offices:
ECONPress
360 Huntington Avenue
301 Lake Hall
Boston, MA 02115-5000

ISSN:
2158-8864 (Print)
2158-8872 (Online)

Electronic Mail:
EditorInChief@ECONPress.org

Web Site:
www.ECONPress.org

ECONPress
of Northeastern University

Spring 2012 *Volume II, Number II*

Letter from the Editors

The greatest insights into our world are often not revolutionary, but evolutionary. Progress is not made in leaps and bounds but by small, determined steps. The authors that we present you in this edition of ECONPress have all taken well-founded topics, approached them with new questions and set into motion a process of researching answers that have not been given before.

We at ECONPress have not developed the research within, but instead provide an outlet for such work to be distributed and recognized. Non-published research is akin to a tree fall not heard. We seek to make the work heard.

Submitting to ECONPress is perhaps the easiest part of the entire research process and the labor of "rethinking our world" the most difficult. Yet all of the authors who submitted to ECONPress should be commended in their confidence and determination because as undergraduates we have not traditionally had such a way to share our work before. Each submission requires a modest confidence that one's work is quality research - quality that each one of us undergraduates possess. The authors and research here are distinguished for hard work, unique insights, economic understanding, and adept prose - qualities that all other undergraduates like you possess and that are waiting to be shared.

Let the research presented within inspire and instigate a debate in your own mind. Put that question to the test and build upon the work of those that have come before us. "Rethink your world" and share it with all of us. Such is a process to continue throughout your lifetime, and we at ECONPress hope to be but a small part.

Thank you for reading,

Alec Loudenback	*Victor Martinelli*
Robert Dent	*Pooja Trivedi*
Michael Onore	*Tala Borno*

The Editorial Board of ECONPress

Voting with Bidirectional Elimination

Matthew Stephen Cook
Stanford University

Abstract

Two important criteria for judging the quality of a voting algorithm are strategy-proofness and Condorcet efficiency. While, according to the Gibbard-Satterthwaite theorem, we can expect no voting mechanism to be fully strategy proof, many Condorcet methods are highly susceptible to compromising, burying, and bullet voting. In this paper I propose a new algorithm which I call "Bidirectional Elimination," a composite of Instant Runoff Voting and the Coombs Method, which offers the benefit of greater resistance to tactical voting while nearly always electing the Condorcet winner when one exists. A sophisticated program was tailor-made and used to test IRV, the Coombs Method, and Bidirectional Elimination on tens of billions of social preference profiles in combinations of up to ten voters and ten candidates. I offer mathematical proofs showing the new algorithm meets the Condorcet criterion for up to 4 voters and N candidates, or M voters and up to 3 candidates; beyond this, program results show Bidirectional Elimination offers a significant advantage over both IRV and Coombs in approaching Condorcet efficiency.

Acknowledgments

I am grateful to Jonathan Levin for his instructive insights, engaging discussions, and thoughtful guidance. A special thanks to Jaehyun Park for lending his world class coding talents in creating the program used on tens of billions of samples. Discussions with Jonathan Zhang were enlightening: It was during conversation with him that the concept of Bidirectional Elimination was born. Jon also independently created his own proof of Condorcet efficiency for the [3 candidates, M voters] case. Tyler Mullen showed me the first set of conditions under which my algorithm would not elect the Condorcet winner. This led me to design the specifications for the program Jaehyun coded. Thanks to the four of you for your support of this stimulating project.

1. Introduction

For decades, the Academy of Motion Picture Arts and Sciences used a plurality vote to determine Best Picture: Over five thousand voters would submit a single vote among just a few nominees, and the movie with the most votes would win. In 2010, the Academy replaced plurality voting with a new system called Instant Runoff Voting (IRV). Instead of submitting a single choice, voters were asked to submit a preferential ranking of ten nominees. After all ballots were cast, the mechanism would systematically eliminate candidates with the fewest first-place votes until one winner remained.

IRV offered several advantages over a simple plurality vote. The plurality mechanism could have easily elected candidates many voters strongly disliked, and it also allowed for the *spoiler effect*—the negative effect a weaker candidate has over a stronger candidate by stealing away votes. To compensate for the spoiler effect, voters in the plurality elections were able to strategize by failing to report actual first choices; voters could *compromise* by voting for the candidate they thought realistically had the best chance of winning. IRV eliminated the spoiler effect and reduced the chance of electing candidates whom a majority disliked.

In his February 15, 2010 article for *The New Yorker,* Hendrik Hertzberg explains the effects of IRV in the context of that year's Academy Awards:

> *"This scheme, [known as] instant-runoff voting, doesn't necessarily get you the movie (or the candidate) with the most committed supporters, but it does get you a winner that a majority can at least countenance. It favors consensus. Now here's why it may also favor The Hurt Locker. A lot of people like Avatar, obviously, but a lot don't—too cold, too formulaic, too computerized, too derivative. . . . Avatar is polarizing. So is James Cameron. He may have fattened the bank accounts of a sizable bloc of Academy members—some three thousand people drew Avatar paychecks—but that doesn't mean that they all long to recrown him king of the world. . . . These factors could push Avatar toward the bottom of many a ranked-choice ballot.*
>
> *On the other hand, few people who have seen The Hurt Locker—a real Iraq War story, not a sci-fi allegory—actively dislike it, and many profoundly admire it. Its underlying ethos is that war is hell, but it does not demonize the soldiers it portrays, whose job is to defuse bombs, not drop them. Even Republicans (and there are a few in Hollywood) think it's good. It will likely be the second or third preference of voters whose first choice is one of the other 'small' films that have been nominated."*

As a process, voting requires design. According to Allan Gibbard and Mark Satterthwaite, no voting mechanism is fully strategy proof, and according to Kenneth Arrow, no voting mechanism meets all reasonable criteria of fairness. A theoretical framework must be constructed to evaluate methods of aggregating individual interests into a collective decision. This theoretical framework, dating back to the work of French philosopher and mathematician Condorcet, is called *social choice theory*.

Within this framework, criteria exist for evaluating the justness of a voting mechanism. The purpose of this paper is to define important criteria, explain their importance, and evaluate a new voting mechanism using these criteria. The mechanism I will analyze is called Bidirectional Elimination and is conducted in a manner similar to IRV. The criteria I will use are the Condorcet criterion and strategy proofness. I will primarily analyze the mechanism based on the Condorcet criterion, using mathematical proofs and experimental data from computer simulations. I will discuss strategy proofness briefly, and recommend an experimental methodology one could use for a rigorous analysis.

The paper is organized as follows. Section 2 reviews criteria, background, terminology, and theorems related to social choice theory that are necessary for analyzing a voting mechanism. Section 3 dissects Instant Runoff Voting, the Coombs Method, and Condorcet Methods. Section 4 introduces Bidirectional Elimination, runs the algorithm step by step on a given set of preferences, and analyzes based on certain criteria presented in section 2. Section 5 offers logic proofs related to Condorcet efficiency. Section 6 presents and analyzes experimental results from a computer simulation. Section 7 offers concluding remarks.

Results from the computer simulation show that Bidirectional Elimination comes very close to meeting Condorcet efficiency. Given the susceptibility of most Condorcet methods to tactical voting, and given that failure to elect the Condorcet winner occurs less than .6% of the time for up to 10 voters and 10 candidates using Bidirectional Elimination, I find this new algorithm may offer a significant advantage over Condorcet methods, and certainly a strict advantage over Instant Runoff Voting and over the similar Coombs Method. The importance: Bidirectional Elimination has the potential to improve justice in democratic systems.

2. Background and Terminology

The first key term is *Condorcet winner*. The Condorcet winner is the candidate who, when paired against any other individual candidate, will always capture a majority vote. A voting method that always elects the Condorcet winner when one exists is called *Condorcet efficient,* or is said to meet the *Condorcet criterion.* The *Condorcet loser* is one who always loses such pairwise runoffs. A system that never elects the Condorcet loser is said to meet the *Condorcet loser criterion.* A *Condorcet method* is a mechanism that will always elect the Condorcet winner. A Condorcet winner does not always exist. Sometimes there will be a tie, or a *cycle*; the latter is an example of *Condorcet's paradox.* For example, three voters with preferences (A>B>C), (B>C>A), and (C>A>B) cause a cycle; in this case, the group chooses A over B, B over C, and C over A, so there is no clear winner.

A second criterion for evaluating voting systems is *strategy proofness.* A voting system is said to be strategy proof if, given full information over everyone else's preferences, no individual can improve his outcome by misreporting his own preferences. Such misreporting is called *tactical voting*. According to the *Gibbard-Satterthwaite theorem,* no preferential voting system is fully strategy proof.

Various types of tactical voting can occur. The first is called *push-over voting,* in which a voter ranks a weak alternative higher, but not in the hopes of getting that alternative elected. The second is called *compromising*, which occurs when voters dishonestly rank an alternative higher than their true alternative in hopes of getting it elected. The third strategy, called *burying,* occurs when a voter dishonestly ranks a candidate lower in hopes of seeing it defeated. The last form is called *bullet voting,* which means voting for only one candidate in a preferential system when submitting a list of ranked preferences is an option.

Failure to meet certain criteria creates susceptibility to strategic voting. Among these criteria is the *monotonicity* criterion, which states that a candidate X must not be harmed—that is, changed from being a winner to a loser—if X is raised on some ballots without changing the orders of the other candidates. Conversely, the *later-no-harm* criterion states that giving a more positive ranking, or simple an additional ranking, to a less preferred candidate must not cause a more preferred candidate to lose.

3. Relevant Voting Methods

Instant Runoff Voting

Instant Runoff Voting is a system using ranked preferences to elect a single winner. The procedure works as follows. Voters submit ranked preferences, and the candidate with the fewest first choice votes is systematically eliminated until only the winner remains. In some cases, candidates will be tied for the position of "fewest first choice votes." If this happens, special tiebreaker rules must be constructed. IRV satisfies the later-no-harm criterion and the Condorcet loser criterion but fails monotonicity, independence of irrelevant alternatives, and the Condorcet criterion. IRV is susceptible to push-over voting.

Coombs Method

The Coombs Method is a system similar to IRV using ranked preferences to elect a single winner. Instead of systematically eliminating the candidate with the fewest first choice votes, Coombs systematically eliminates the candidate with the most last choice votes until one remains. Similarly, a special tiebreaker must be devised to deal with cases in which candidates are tied for the position of "most last choice votes." The Coombs Method satisfies the Condorcet loser criterion but fails monotonicity, independence of irrelevant alternatives, and the Condorcet criterion. The Coombs Method is susceptible to compromising and burying.

Condorcet Methods

Condorcet methods, methods that will always elect the Condorcet winner, include Copeland's method, the Kemeny-Young Method, Minimax, Nanson's method, ranked pairs, and the Schulze method. No Condorcet method satisfies the later-no-harm criterion or independence of irrelevant alternatives. The methods vary in monotonicity.

Condorcet methods are generally quite susceptible to tactical voting—in particular, compromising, burying, and bullet voting.

4. Bidirectional Elimination

I now propose a new voting mechanism using ranked preferences. I call this new mechanism "Bidirectional Elimination," as it is a hybrid of both Instant Runoff Voting and the Coombs Method. The algorithm works as follows. (Note: The tiebreaker procedure described in stages 1 and 2 is important in preventing premature elimination of the Condorcet winner.)

The Bidirectional Elimination Algorithm

Stage 1: Use regular IRV to elect potential winner X. If in any round of elimination two or more candidates share the least number of first choice votes, a special tiebreaker is performed. To perform the tiebreaker, create a set of "potential losers" consisting of all candidates sharing the least number of first choice votes in the leftmost column. Then move to the next column to the right and reiterate this form of IRV among potential losers only, possibly narrowing the set of potential losers, until one is eliminated; then continue with regular IRV. If potential losers ever tie in the rightmost column containing potential losers, all potential losers are eliminated then. Under certain tiebreaker conditions, no winner will be elected; all candidates are eliminated.

Stage 2: Use the Coombs Method to elect a potential winner Y. If in any round of elimination two or more candidates share the highest number of last choice votes, a special tiebreaker is performed. To perform the tiebreaker, create a set of "potential losers" consisting of all the candidates sharing the highest number of last choice votes in the rightmost column. Then move to the next column to the left and reiterate this form of the Coombs Method among potential losers only, possibly narrowing the set of potential losers, until one is eliminated; then continue with the regular Coombs Method. If potential losers are ever tied in the leftmost column containing any potential losers, all potential losers are eliminated at once. Under certain tiebreaker conditions, no winner will be elected; all candidates are eliminated.

Stage 3: If X and Y are the same candidate, X = Y= winner; if X and Y are different candidates, a simple pairwise runoff between X and Y determines the winner; if only one of the stages elects a potential winner, the potential winner is the final winner; if neither former stage elects a potential winner, there is no final winner.

Non-Monotonic

Bidirectional Elimination fails the monotonicity criterion and is therefore not strategy-proof. The social preference profiles below are examples susceptible to tactical voting.

# Voters	True Preferences
4	A > B > C
3	B > C > A
2	C > A > B

IRV and Coombs would each elect A in this case, so Bidirectional Elimination also elects A. However, a voter with preferences (B>C>A) could simply state falsely that his preferences were (C>A>B). Preferences would then look like this:

# Voters	Reported Preferences
4	A > B > C
2	B > C > A
3	C > A > B

Now, IRV elects C; Coombs elects A; and C wins with Bidirectional Elimination when compared pairwise with A. By misreporting preferences, a voter was able to improve his outcome, now getting his second choice instead of his third choice.

Visualization

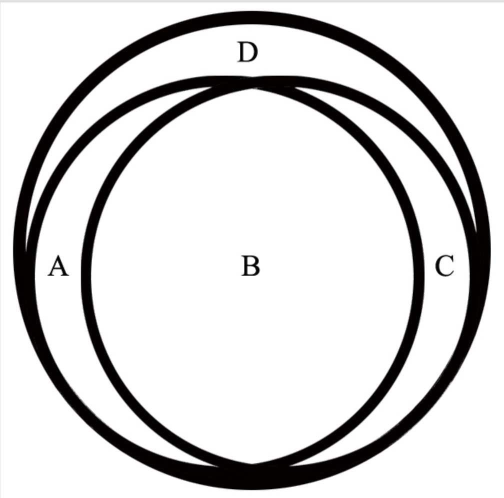
The circle below, not to scale, represents an arbitrary space for N>3 candidates and M>4 voters containing permutations of voter preferences.

A + B + C + D: Permutations of preferences such that candidate A is a Condorcet winner.

A + B + C: Bidirectional Elimination elects candidate A.

A + B: Instant Runoff Voting elects candidate A.

B + C: Coombs Method elects candidate A.



A: Instant Runoff Voting elects candidate A, but Coombs Method does not.

B: Both Instant Runoff Voting and Coombs Method elect candidate A.

C: Coombs Method elects candidate A, but Instant Runoff Voting does not.

D: Bidirectional Elimination fails to elect candidate A.

When No Condorcet Winner Exists

While Bidirectional Elimination fails to meet the Condorcet criterion, it comes very close, as shown in section 6. However, as the number of N candidates or M voters increases, the number of Condorcet winners generally decreases.

When no Condorcet winner exists, there must be new criteria for evaluating the justness of a voting algorithm. No Condorcet winner exists in the presence of a tie or a cycle.

When votes are tied—such as in the simplest case, when two voters disagree over which of two candidates to elect—there is no mathematically "just" way of choosing between candidates. Bidirectional Elimination does not choose a winner in cases of ultimate ties. This is not a bad thing; de-

pending on the nature of the collective decision, either no winner should be elected, or an alternative method should be used to arbitrarily select a winner.

Electoral cycles can be similar to ties in that no mathematically just winner exists. The classic example of Condorcet's paradox—three voters with profiles (A>B>C), (B>C>A), and (C>A>B) has no clear winner (this is an example of a circular tie). But what if votes are not as symmetrical? What if we add a fourth voter with preferences (A>B>C)? Now, A is strictly preferred over B, which is strictly preferred over C; but A and C are tied in a pairwise runoff. This is no longer a circular tie. While there may still be no clear winner, electing an arbitrary candidate seems less than ideal: One could argue that B should be eliminated since it is strictly dominated by A; and one could debate whether C should be eliminated, since while tied with A, C is still strictly dominated by B. Establishing criteria for justice is more difficult given an asymmetric cycle like this one.

A *Condorcet completion method* is required to elect a winner when ambiguities like these arise. One such completion method uses the "minimax rule" developed by Simpson and Kramer. This rule gives a score to each candidate as follows. Candidate A_1 is paired against $A_{2\text{-}N}$. For each pairwise comparison, a score is given to A_1 equal to the number of votes his opponent has, minus the number of his own votes. The maximum of these scores is recorded as W_1 (there are various ways of calculating a W score; this one, called using *margins*, is the simplest) The minimum of ($W_{1\ldots N}$) corresponds to the winner.

Andrew Caplin and Barry Nalebuff proposed a system based on Simpson and Kramer's minimax using a "64%-majority rule" that can in fact eliminate all electoral cycles given a restriction on individual preferences, and on the distribution of preferences.

Bidirectional Elimination may not eliminate cycles, but the mechanism does satisfy other pleasant, though informal, criteria. The IRV stage of the mechanism will rarely eliminate a candidate whom voters find appealing, while the Coombs stage will rarely elect a candidate whom voters find unappealing, and between the two stages, the better candidate is chosen.

5. Logic Proofs for IRV, Coombs Method, and Bidirectional Elimination

IRV, Coombs Method, and Bidirectional Elimination meet the Condorcet criterion in certain specific combinations of N candidates and M voters. In this section I examine for which sets of [N candidates, M voters] each voting system meets the criterion.

For consistency and ease of notation, candidate "A" refers to the Condorcet winner.

We will ignore two trivial cases, [1 candidate, M voters] and [N candidates, 1 voter], in which voters have no choice between candidates, or one voter has the ultimate choice (a situation similar to dictatorship). Tie and cyclical cases are not discussed here. I am primarily concerned with electing the Condorcet winner when one exists.

Logic Proofs: Instant Runoff Voting

Instant Runoff Voting, the weakest of the three voting mechanisms studied here, meets the Condorcet criterion for up to 2 voters and N candidates, or M voters and up to 2 candidates.

N candidates, 2 voters

When only two voters are present, a Condorcet winner can exist only when a tie does not exist; that is, when the voters agree on the candidate to be chosen. Both candidates have therefore ranked this candidate (candidate A) as their first choice. All other N-1 candidates will be eliminated at once, leaving A as the winner from IRV.

2 candidates, M voters

When only two candidates are present, a Condorcet winner exists when either M is odd, or M is even but the two candidates are not tied in their votes received. In both cases, candidate A captures more than 50% of the voters. If this is the case, the non-A candidate will be eliminated in the first and only iteration of IRV.

Logic Proofs: Coombs Method

The Coombs method, which meets the Condorcet criterion in a larger set of candidate/voter combinations, works consistently for up to 4 voters and N candidates, or M voters and up to 2 candidates.

N candidates, 2 voters

When only two voters are present, a Condorcet winner can exist only when a tie does not exist; that is, when the voters agree on the candidate to be chosen. Both candidates have therefore ranked this candidate (candidate A) as their first choice, and the losing candidates (candidate non-A) as their second and last choice. All other N-1 candidates will be eliminated systematically until only A remains from the Coombs Method.

N candidates, 3 voters

To preserve the existence of a Condorcet winner under these conditions, only one of the three voters may rank a given non-A candidate above A; otherwise, at least two-thirds of the voters would prefer a non-A candidate over A.

When columns of preferences are drawn, while N>1 (a winner as not yet been found, as multiple candidates remain), it follows that A can appear in the rightmost column at most once. If A were to appear in the rightmost column more than once, at least one non-A would be preferred in majority over A. This is impossible, as A is the Condorcet winner.

If A does appear in the rightmost column once, either (1) the same non-A is now ranked in last-place by two voters, in which case this non-A is eliminated, or (2) three different candidates appear in the last column, in which case A cannot be eliminated in the tiebreaker, because one of the other two potential losers will be eliminated first. This is because as we move leftward through columns, A cannot appear before another candidate: At least three voters rank every other candidate below A.

If A appears in the rightmost column zero times, it will not be eliminated in this round of the Coombs Method.

Candidate A can never be eliminated as the algorithm iterates. Under [N candidates, 3 voters] conditions, Candidate A can never be eliminated using the Coombs method. A will always win.

N candidates, 4 voters

This proof is directly parallel to the [N candidates, 3 voters] proof. To preserve the existence of a Condorcet winner under these conditions, only one of the four voters may rank a given non-A candidate above A; otherwise, a tie could exist, or a majority of the voters could prefer a non-A candidate over A.

When columns of preferences are drawn, while N>1 (a winner as not yet been found, as multiple candidates remain), it follows that A can appear in the rightmost column at most once. If A were to appear in the rightmost column more than once, at least one non-A would be preferred in majority over A. This is impossible, as A is the Condorcet winner.

If A does appear in the rightmost column once, either (1) a non-A is eliminated right away, or (2) all different candidates appear in the last column, in which case A cannot be eliminated in the tiebreaker, because one of the other three potential losers must be eliminated first. This is because as we move leftward through columns, A cannot appear before another candidate: At least three voters rank every other candidate below A.

If A appears in the rightmost column zero times, it will not be eliminated in this round of the Coombs Method.

Candidate A can never be eliminated as the algorithm iterates. Under [N candidates, 4 voters], Candidate A can never be eliminated using the Coombs method. A will always win.

Logic Proofs: Bidirectional Elimination

Bidirectional Elimination meets the Condorcet criterion for up to 4 voters and N candidates, or M voters and up to 3 candidates.

Since Bidirectional Elimination utilizes IRV and the Coombs Method in electing two potential winners who eventually face off in a pairwise election, when IRV or the Coombs Method elects a Condorcet winner, so will Bidirectional Elimination. Thus from the proofs above, we have already established that Bidirectional Elimination meets the Condorcet criterion for up to 4 voters and N candidates, or M voters and up to 2 candidates.

Additionally, Bidirectional Elimination meets the Condorcet criterion for 3 candidates and M voters, which is more than either IRV or the Coombs Method can accomplish alone.

3 candidates, M voters

Here we use an example with a total number of voters equal to $X_1+X_2 \ldots +X_6$ choosing between candidates A, B, and C. A is given as the stable Condorcet winner. We wish to prove that Bidirectional Elimination will always elect candidate A given the [3 candidates, M voters] condition. Proving that Bidirectional Elimination will never eliminate A is the same as proving that Bidirectional Elimination will always elect A if and only if a candidate is eliminated each round. Since one candidate is eliminated each round, I will prove Bidirectional Elimination will never eliminate A—in order to show that Bidirectional Elimination will always elect A.

The table below represents all possible permutations of voter preferences.

Since we are given that A is the stable Condorcet winner, to satisfy the stable Condorcet condition, A must have more votes than either B or C in pairwise counts. The following two statements must be true:

$$X_3 + X_4 + X_6 < X_1 + X_2 + X_5$$

(1, implies A > B)

Number of Ballots	First Choice	Second Choice	Third Choice
X_1	*A*	*B*	*C*
X_2	*A*	*C*	*B*
X_3	*B*	*A*	*C*
X_4	*B*	*C*	*A*
X_5	*C*	*A*	*B*
X_6	*C*	*B*	*A*

$$X_4 + X_5 + X_6 < X_1 + X_2 + X_3$$

(2, implies A > C)

If the Condorcet condition is to be satisfied by Bidirectional Elimination, then we must show that A cannot be eliminated by both IRV and the Coombs Method. For this to be true, if the stable Condorcet winner A is eliminated in one of these two mechanisms, the other mechanism must *not* eliminate it.

Candidate A could be eliminated in the IRV stage in one of three ways:

IRV Case 1: The first round eliminates B; the second round eliminates A.

IRV Case 2: The first round eliminates C; the second round eliminates A.

IRV Case 3: The first round eliminates A.

IRV Case 1 is impossible: If B is eliminated in the first round, only A and C remain in the final pairwise runoff. By (2), A would win this runoff and C would be eliminated, not A.

IRV Case 2 is impossible: If C is eliminated in the first round, only A and B remain in the final pairwise runoff. By (1), A would win this runoff and B would be eliminated, not A.

Let us continue unraveling IRV Case 3. If first-round IRV eliminates A, then we know:

$$X_1 + X_2 < X_3 + X_4 \quad (3)$$

$$X_1 + X_2 < X_5 + X_6 \quad (4)$$

Given these conditions, can candidate A now also lose in a Coombs election? A loss could happen in one of three ways:

Coombs Case 1: The first round eliminates B; the second round eliminates A.

Coombs Case 2: The first round eliminates C; the second round eliminates A.

Coombs Case 3: The first round eliminates A.

IRV Case 1 is impossible: If B is eliminated in the first round, only A and C remain in the final pairwise runoff. By (2), A would win this runoff and C would be eliminated, not A.

IRV Case 2 is impossible: If C is eliminated in the first round, only A and B remain in the final pairwise runoff. By (1), A would win this runoff and B would be eliminated, not A.

This leaves us with only Coombs Case 3—the first round eliminates A—which would require the following to hold:

$$X_1 + X_3 < X_4 + X_6 \quad (5)$$

$$X_2 + X_5 < X_4 + X_6 \quad (6)$$

From (1) and (3), we can show algebraically,

$$X_6 < X_5 \quad (7)$$

From (7) and (6), we can show algebraically,

$$X_2 < X_4 \quad (8)$$

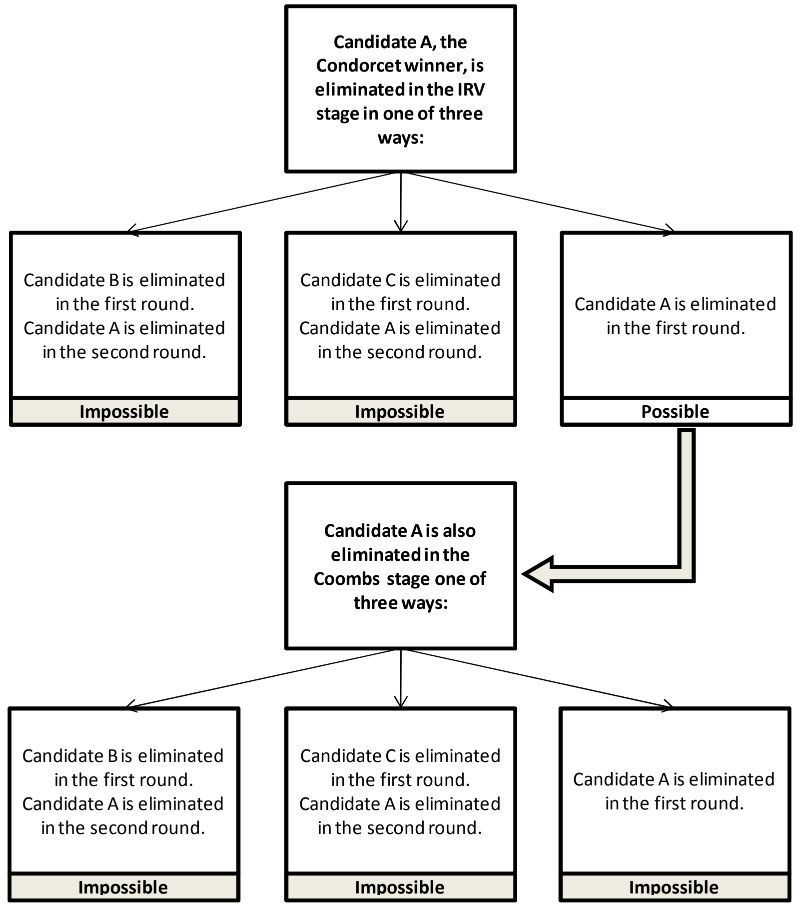
From (6) and (1), we can show algebraically,

$$X_3 < X_1$$

From (4) and (2), we can show algebraically,

$$X_4 < X_3 \quad (10)$$

From (5) and (10), we can show algebraically,



$$X_1 < X_6 \quad (11)$$

From (5) and (2), we can show algebraically,

$$X_5 < X_2 \quad (12)$$

The following chain can be derived from numbers (7) through (12) above:

$$\mathbf{X_5} < X_2 < X_4 < X_3 < X_1 < X_6 < \mathbf{X_5}$$

Since X_5 cannot appear on both sides of the inequality, Coombs Case 3 must also be impossible. We have just shown that it is impossible for candidate A to be eliminated by IRV, and also by the Coombs Method. It is therefore impossible for A to be eliminated by Bidirectional Elimination under these conditions. Bidirectional Elimination always elects the Condorcet winner given the [3 candidates, M voters] condition.

The proof can be visualized in a tree (above).

6. Experimentation via Computer Simulation

A Program to Simulate Voting

Jaehyun Park generously wrote a program according to my specifications to test the results of IRV, the Coombs Method, and Bidirectional Elimination for various social preference profiles. The simulation takes in three inputs: N candidates, M voters, and X cases. Here we define “social preference profile" as the set of ranked preferences for all voters $V_{1...M}$. When a user enters X=0, the program will test all possible permutations of social preference profiles exactly once and output results. (For example, for 2 voters and 3 candidates, V_1 has 6 possible rankings, and V_2 has 6 possible rankings, for a total of 36 possible social preference profiles. Entering X=0 tests each one of these.) When a user enters X>0, the program will create X random social preference profiles and perform the algorithms on each one. Since these profiles are random when X>0, it is possible, though rare as M and N increase, for duplicate profiles to be tested. This duplication is generally expected of Monte Carlo experiments, and in this experiment does not significantly affect results (only by thousandths of a percent).

Outputs of the simulation include:

Q, number of total cases, or social preference profiles, tested. Q is equal to X when X>0, or $(N!)^M$ when X=0. Individual preferences were restricted to include only sets in which each voter gave a ranking to every candidate. That is, if options A through E were available, each voter included options A through E in his individual preference profile, omitting no option from his ordering. Removing this domain restriction would expand the total number of possible social preference profiles to $(N! + N!/2! + N!/3! \ldots + 1)^M$, making the program unfeasible.

R, number of cases with A as Condorcet winner. The program labels one of the candidates as A and tests preference permutations to determine if A is a Condorcet winner.

R/Q, percentage of cases with A as Condorcet winner. This calculation performs R/Q, the number of times that A is a Condorcet winner divided by the number of cases tested.

*N*R, number of cases with a Condorcet winner.* By symmetry, we can multiply R by the number of candidates to determine for how many permutations a Condorcet winner exists.

*(N*R)/Q, percentage of cases with a Condorcet winner.* Dividing N*R by the total number of cases tells us the percentage of cases for which a Condorcet winner exists.

S, the IRV success figure. This is equal to the number of times that A is the Condorcet winner, and Instant Runoff Voting elects A.

S/R, the IRV success ratio. Dividing S by the number of cases where A is the Condorcet winner returns a valuable percentage.

T, the Coombs Method success figure. This is equal to the number of times that A is the Condorcet winner, and the Coombs Method elects A.

T/R, the Coombs Method success ratio. Dividing R by the number of cases where A is the Condorcet winner returns a valuable percentage.

U, the Bidirectional Elimination success figure. This is equal to the number of times that A is the Condorcet winner, and Bidirectional Elimination elects A.

U/R, the Bidirectional Elimination success ratio. Dividing U by the number of cases where A is the Condorcet winner returns a valuable percentage.

The program reports the most accurate results when X=0, and all cases are tested exactly once. However, since Q grows astronomically as N or M increase, the program can take years to run beyond a [N=5 candidates, M=5 voters] simulation. We approximate by setting X equal to a large number—in this case, one billion—when necessary to shorten the simulation. When $(N!)^M < 1$ billion, we test all possible samples

Table 1: Number of Cases Tested (Q)

	Candidates									
Voters	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	1	2	6	24	120	720	5040	40320	362880	3628800
2	1	4	36	576	14400	518400	25401600	1000000000	1000000000	1000000000
3	1	8	216	13824	1728000	373248000	1000000000	1000000000	1000000000	1000000000
4	1	16	1296	331776	207360000	1000000000	1000000000	1000000000	1000000000	1000000000
5	1	32	7776	7962624	1000000000	1000000000	1000000000	1000000000	1000000000	1000000000
6	1	64	46656	191102976	1000000000	1000000000	1000000000	1000000000	1000000000	1000000000
7	1	128	279936	1000000000	1000000000	1000000000	1000000000	1000000000	1000000000	1000000000
8	1	256	1679616	1000000000	1000000000	1000000000	1000000000	1000000000	1000000000	1000000000
9	1	512	10077696	1000000000	1000000000	1000000000	1000000000	1000000000	1000000000	1000000000
10	1	1024	60466176	1000000000	1000000000	1000000000	1000000000	1000000000	1000000000	1000000000

Number of cases to test equals = ((number of candidates)!)^(number of voters)
Cases tested capped at 1 billion.

Table 2: Percentage of Cases where A is a Condorcet Winner (R/Q)

	Candidates									
Voters	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	100.000%	50.000%	33.333%	25.000%	20.000%	16.667%	14.286%	12.500%	11.111%	10.000%
2	100.000%	25.000%	11.111%	6.250%	4.000%	2.778%	2.041%	1.559%	1.221%	0.998%
3	100.000%	50.000%	31.481%	22.222%	16.800%	13.296%	10.866%	9.111%	7.786%	6.759%
4	100.000%	31.250%	14.815%	8.550%	5.535%	3.862%	2.839%	2.176%	1.712%	1.389%
5	100.000%	50.000%	31.019%	21.528%	16.001%	12.474%	10.054%	8.318%	7.024%	6.030%
6	100.000%	34.375%	16.958%	10.002%	6.571%	4.623%	3.425%	2.633%	2.091%	1.695%
7	100.000%	50.000%	30.833%	21.248%	15.694%	12.148%	9.736%	8.006%	6.727%	5.745%
8	100.000%	36.328%	18.397%	11.034%	7.295%	5.180%	3.857%	2.985%	2.377%	1.935%
9	100.000%	50.000%	30.734%	21.099%	15.526%	11.977%	9.562%	7.841%	6.570%	5.595%
10	100.000%	37.695%	19.447%	11.781%	7.876%	5.620%	4.208%	3.262%	2.604%	2.125%

Table 3: Percentage of Cases where a Condorcet Winner Exists (N*R/Q)

	Candidates									
Voters	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%
2	100.000%	50.000%	33.333%	25.000%	20.000%	16.667%	14.286%	12.473%	10.988%	9.983%
3	100.000%	100.000%	94.444%	88.889%	84.000%	79.778%	76.060%	72.884%	70.071%	67.587%
4	100.000%	62.500%	44.444%	34.201%	27.675%	23.172%	19.875%	17.410%	15.406%	13.889%
5	100.000%	100.000%	93.056%	86.111%	80.005%	74.845%	70.375%	66.543%	63.217%	60.304%
6	100.000%	68.750%	50.874%	40.008%	32.857%	27.738%	23.974%	21.066%	18.821%	16.955%
7	100.000%	100.000%	92.498%	84.992%	78.469%	72.888%	68.153%	64.051%	60.540%	57.452%
8	100.000%	72.656%	55.190%	44.136%	36.474%	31.083%	27.001%	23.881%	21.396%	19.346%
9	100.000%	100.000%	92.202%	84.395%	77.632%	71.863%	66.935%	62.729%	59.132%	55.949%
10	100.000%	75.391%	58.340%	47.122%	39.382%	33.723%	29.459%	26.098%	23.435%	21.252%

and set X=0; when $(N!)^M > 1$ billion, we set X=1,000,000,000. This way, the program will never test more than a billion samples.

Although for a [N=10 candidates, M=10 voters] simulation one billion trials represent an extremely small fraction of the total possible permutations, (exactly 2.53×10^{-52}), the results are still closely accurate and change in the thousandths place with repeat simulations.

Tables from Sample Runs

The tables of results from Park's voting simulator appear below.

Table 1 reports the number of social preference profiles tested. When the number is less than a billion, all possible social

Table 4: IRV Success Ratio (S/R)

Percentage of Cases where a Condorcet Winner Exists, and IRV Elects the Condorcet Winner

	Candidates									
Voters	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%
2	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%
3	100.000%	100.000%	100.000%	98.438%	96.429%	94.359%	92.349%	90.466%	88.720%	87.070%
4	100.000%	100.000%	100.000%	100.000%	99.729%	99.241%	98.560%	97.839%	97.019%	96.168%
5	100.000%	100.000%	97.512%	95.758%	93.947%	92.078%	90.172%	88.326%	86.560%	84.800%
6	100.000%	100.000%	100.000%	99.970%	99.768%	99.313%	98.668%	97.883%	96.989%	96.030%
7	100.000%	100.000%	99.676%	98.360%	96.815%	95.143%	93.403%	91.644%	89.938%	88.212%
8	100.000%	100.000%	99.275%	98.377%	97.563%	96.748%	95.890%	94.986%	94.099%	93.001%
9	100.000%	100.000%	98.332%	96.598%	94.941%	93.309%	91.677%	90.066%	88.497%	86.938%
10	100.000%	100.000%	99.893%	99.613%	99.031%	98.309%	97.493%	96.594%	95.621%	94.627%

Table 5: Coombs Method Success Ratio (T/R)

Percentage of Cases where a Condorcet Winner Exists, and the Coombs Method elects the Condorcet Winner

	Candidates									
Voters	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%
2	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%
3	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%
4	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%
5	100.000%	100.000%	100.000%	95.758%	97.755%	96.672%	95.820%	95.161%	94.637%	94.183%
6	100.000%	100.000%	100.000%	100.000%	99.862%	99.630%	99.374%	99.143%	98.944%	98.779%
7	100.000%	100.000%	98.054%	96.663%	95.709%	94.863%	93.999%	93.130%	92.286%	91.516%
8	100.000%	100.000%	100.000%	99.716%	99.450%	99.255%	99.060%	98.847%	98.593%	98.305%
9	100.000%	100.000%	98.698%	97.232%	95.775%	94.569%	93.613%	92.827%	92.118%	91.431%
10	100.000%	100.000%	99.429%	99.034%	98.565%	98.082%	97.661%	97.320%	97.039%	96.765%

Table 6: Bidirectional Elimination Success Ratio (U/R)

Percentage of Cases where a Condorcet Winner Exists, and Bidirectional Elimination Elects the Condorcet Winner

	Candidates									
Voters	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%
2	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%
3	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%
4	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%
5	100.000%	100.000%	100.000%	100.000%	99.979%	99.947%	99.905%	99.855%	99.798%	99.731%
6	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	99.999%	99.998%
7	100.000%	100.000%	100.000%	99.994%	99.968%	99.925%	99.865%	99.787%	99.689%	99.577%
8	100.000%	100.000%	100.000%	100.000%	99.999%	99.998%	99.995%	99.993%	99.990%	99.983%
9	100.000%	100.000%	100.000%	99.983%	99.942%	99.881%	99.804%	99.713%	99.609%	99.485%
10	100.000%	100.000%	100.000%	100.000%	99.999%	99.995%	99.990%	99.981%	99.971%	99.956%

preference profiles were tried.

Table 2 divides the number of social preference profiles for which A is the Condorcet winner by the total number of social preference profiles tested. The result is the percentage of cases for which candidate A is the Condorcet winner.

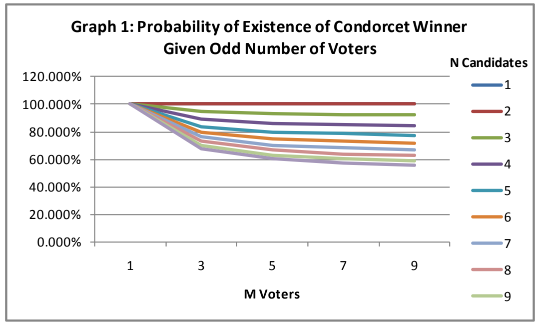
Multiplying Table 2 by number of candidates tells us how often a Condorcet winner exists.

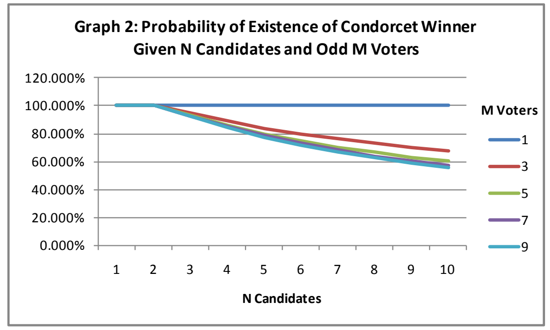
Table 4 tells us how often an existing Condorcet winner will be elected by IRV.

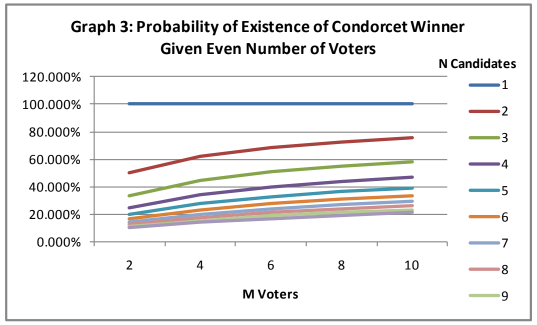
Table 5 tells us how often an existing Condorcet winner will be elected by the Coombs Method.

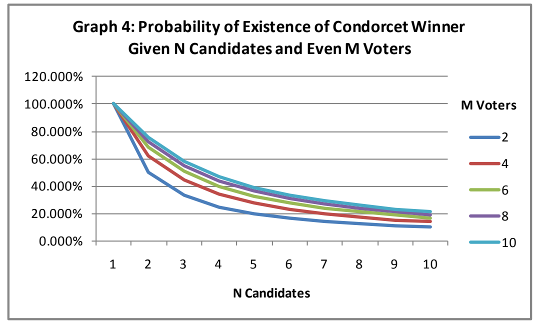
Table 6 tells us how often an existing Condorcet winner will be elected by Bid. Elim.

Analysis of Results









Existence of a Condorcet Winner

Table 3 reveals interesting trends relating to the existence of Condorcet winners for various combinations of N candidates and M voters.

In the row with 2 voters, we observe that the result (percentage of cases where a Condorcet winner exists) is equal to 1/N. Probabilistically this makes sense: With two voters, a Condorcet winner can exist only when both voters rank the same candidate in first place; that is, when the second voter chooses the same candidate as the first voter, which happens with a probability of 1/N.

For two candidates and an odd number of voters, a Condorcet winner will always exist because even splits or ties are not possible.

Except for the dictatorship case when M=1, the addition of candidates lowers the probability of the existence of a Condorcet winner.

The results become quite interesting when we split Table 3 into odd and even voters:

Observe the major difference between Tables 3A and 3B, or Graphs 1 and 3. For an odd number of voters, the probability of the existence of a Condorcet winner starts high and decreases with the addition of voters. For an even number of voters, the same probability starts low and increases with the addition of voters. This interesting result leaves room for further study; probabilistic analysis could lead to a theorem that would answer these questions:

Why does p, the probability of the existence of a Condorcet winner given a specific even number of voters and arbitrary N candidates, increase with more voters? Why does q, the probability of the existence of a Condorcet winner given a specific odd number of voters and arbitrary N candidates, decrease with more voters?

Why does p follow a linear pattern as N increases, while q bows steeply?

As M approaches infinity, what do p and q approach?

Table 3A: Percentage of Cases where a Condorcet Winner Exists (N*R/Q); Odd Voters Only

	Candidates									
Voters	1	2	3	4	5	6	7	8	9	10
1	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%	100.000%
3	100.000%	100.000%	94.444%	88.889%	84.000%	79.778%	76.060%	72.884%	70.071%	67.587%
5	100.000%	100.000%	93.056%	86.111%	80.005%	74.845%	70.375%	66.543%	63.217%	60.304%
7	100.000%	100.000%	92.498%	84.992%	78.469%	72.888%	68.153%	64.051%	60.540%	57.452%
9	100.000%	100.000%	92.202%	84.395%	77.632%	71.863%	66.935%	62.729%	59.132%	55.949%

Table 3B: Percentage of Cases where a Condorcet Winner Exists (N*R/Q); Even Voters Only

	Candidates									
Voters	1	2	3	4	5	6	7	8	9	10
2	100.000%	50.000%	33.333%	25.000%	20.000%	16.667%	14.286%	12.473%	10.988%	9.983%
4	100.000%	62.500%	44.444%	34.201%	27.675%	23.172%	19.875%	17.410%	15.406%	13.889%
6	100.000%	68.750%	50.874%	40.008%	32.857%	27.738%	23.974%	21.066%	18.821%	16.955%
8	100.000%	72.656%	55.190%	44.136%	36.474%	31.083%	27.001%	23.881%	21.396%	19.346%
10	100.000%	75.391%	58.340%	47.122%	39.382%	33.723%	29.459%	26.098%	23.435%	21.252%

Approaching Condorcet Efficiency

Tables 4, 5, and 6 show success ratios for IRV, Coombs, and Bidirectional Elimination.

The success ratio is rather poor for Instant Runoff Voting. The column with eight candidates already drops below 90% (when M=5).

The success ratio is much improved with the Coombs Method, never once dropping below 90% throughout the entire 10x10 matrix. It is more difficult to arrange preferences in such a way that A remains the Condorcet winner, but is eliminated using the Coombs Method.

The success ratio for Bidirectional Elimination is even stronger than the former, not once dropping below 99.4%. Throughout the 10x10 matrix, Bidirectional Elimination comes extremely close to Condorcet efficiency.

7. Conclusion

The system of Bidirectional Elimination, an algorithm using IRV and the Coombs Method in various stages, comes very close to meeting the Condorcet criterion and may in fact be a more just alternative to Condorcet methods—depending on the designer's emphasis on Condorcet efficiency versus strategy proofness. While there are some possible arrangements of voter preferences such that the mechanism will fail to elect the Condorcet winner, these cases comprise less than .6% of the total in a 10x10 matrix of candidates and voters.

Bidirectional Elimination is probably less susceptible to strategic voting, but this should be proven rigorously. I say "probably" because IRV and the Coombs Method are each individually less susceptible than most Condorcet methods; it seems logical Bidirectional Elimination would be as well. This question leaves room for future study and experimentation: Given most Condorcet methods' vulnerability to tactical voting, Bidirectional Elimination may ironically be more successful at electing the Condorcet winner. I encourage future economists to rigorously examine the susceptibility of Bidirectional Elimination to strategic voting, and compare with that of Condorcet methods, as well as IRV and the Coombs Method individually.

Experimentation for Strategy-Proofness

This study could be facilitated with a program similar to Jaehyun Park's, that would work as follows. Given inputs for N candidates and M voters, the program would generate $(N!)^M$ permutations of true preferences for voters V_1 through V_M. A "case counter" would start at zero. For each possible permutation of true preferences, the following would happen (a nested *for* loop): (1) The voting algorithm in question would be performed on true preferences, and a winner W would be elected. (2) Holding all other voters' preferences constant, the program would then generate all remaining (N!-1) *false* permutations of voter V_1's preferences, and perform the algorithm on each of these. If at least one set of *false* preferences existed for V_1 such that reporting these false preferences would elect a candidate whom V_1 preferred over W, one would be added to a "case counter," and the program would move onto the next set of true preferences, skipping steps 3 and 4. (3) If nothing had been added to the case counter after trying all permutations of *false* preferences for V_1, the program would repeat step 2 for $V_{2\ldots M}$. If at least one set of *false* preferences existed for any V_i such that reporting these false preferences would elect a candidate whom V_i preferred over W, one would be added to a "case counter," and the program would move onto the next permutation of true preferences. (4) If after step 3 no set of *false* preferences were found for any voter V_i that could improve his outcome, the "case counter" would remain at its current value, and the next permutation of true preferences would be tested similarly, until all permutation of true preferences had been tested. (5) Dividing the final "case counter" value by $(N!)^M$ would output the percentage of cases, given N and M as inputs, for which the voting algorithm in question is susceptible to tactical voting.

The program would need to be designed, as Park designed his, to be capable of random sampling in addition to testing all cases. As N and M grow, $(N!)^M$ grows astronomically—and given the number of operations required to conduct this simulation, testing all cases could require vast amounts of time. A Monte Carlo method will be necessary.

Bidirectional Elimination's robustness should come as promising news to economists. The algorithm holds potential to significantly improve democratic processes. Moreover, this new simulation method for testing criteria can be applied to other voting methods, enabling further experimentation and eventual comprehensive, data-driven comparisons of voting methods.

Works Cited

Arrow, Kenneth (1963). Social Choice and Individual Values. New York: John Wiley & Sons Inc., Second Edition.

Caplin, Andrew and Nalebuff, Barry (1988). "On 64% Majority Rule." Econometrica, Vol. 56, No. 4. 787-814.

Dennis, Gregory (2008). "Why I Prefer IRV to Condorcet." Fairvote: The Center for Voting and Democracy.

Gehrlein, William (2006). Condorcet's Paradox. Theory and Decision Library. Series C: Game Theory, Mathematical Programming, and Operations Research. Springer.

Gibbard, Allan (1973). "Manipulation of Voting Schemes: A General Result." Econometrica, Vol. 41, 587-601.

An Econometric Inquiry into the Effect of Religious Profiles on National Wealth

Brian Joel Feldman
Northeastern University

Introduction

With the ever expanding debate over economic wealth raging on around the world, many countries (especially developing ones) have gone under-microscope to see what stimulates growth and what hinders it, but are social scientists and economists looking at the right things? So much of their time is spent talking about the technologies, workforces, and governments of these countries that often the religious profiles and cultures of developing countries get shoved aside, tossed into footnotes, or are all together omitted from the conversation. With this in mind, the emerging field of economics of religion is under immense pressure to come forward and possibly answer the question on many peoples' minds, "Why are these developing countries not growing?"

Review of the Literature

Numerous studies have been produced during the brief history of economics of religion, all of which have asked important questions about the impact of religion on the surrounding world. The most notable topics and papers that arose in my research for this original study include: Robin Grier's "The Effect of Religion on Economic Development: A Cross National Study of 63 Former Colonies"[1], Rachel M. McCleary's "Religion and Economic Development: The advantage of moderation"[2], and Marcus Noland's "Religion, Culture, And Economic Performance"[3]. It was these articles that laid the framework for this study by setting forth a precedent on which to build.

In her study on "a pooled sample of 63 ex-colonial states" Grier contributed significantly to the argument of religion as a significant factor in real GDP growth. In her argument, Grier opposes the hypotheses of Weber, most notably the 'Protestant Ethic' (which states that the industrial revolution boom and subsequent rise of capitalism in Europe can be attributed to the strong Protestant roots of the area) , by saying "my results show that religion is not the sole determinant of differential development and growth." Although Grier aims to refute the points of Weber's 'Protestant Ethic,' she clearly suggests the effect of religion on national wealth to be a significant one4. This is later reflected in her econometric analysis of her collected data. "The results support my hypothesis that the growth rate of Prot-

1 Robin Grier, "The Effect of Religion on Economic Development: A Cross National Study of 63 Former Colonies" ***Kyklos: International Review for Social Sciences***, 50, (1997). Accessed Oct 2011. http://faculty-staff.ou.edu/G/Robin.M.Grier-1/religion.pdf

2 Rachel M. McCleary, "Religion and Economic Development: The advantage of moderation", ***Hoover Institution, Stanford University,*** 147 (March 28, 2008) Accessed Oct. 2011 http://www.hoover.org/publications/policy-review/article/5729

3 Marcus Noland, "RELIGION, CULTURE, AND ECONOMIC PERFORMANCE" (Working Paper***), The Peterson Institute for International Economics*** Accessed Oct. 2011 http://www.iie.com/publications/wp/03-8.pdf

4 Robin Grier, "The Effect of Religion on Economic Development: A Cross National Study of 63 Former Colonies" ***Kyklos: International Review for Social Sciences***, 50, (1997). Accessed Oct 2011. http://faculty-staff.ou.edu/G/Robin.M.Grier-1/religion.pdf

estantism is significantly correlated with real GDP growth, and that Protestantism is one of many determinants of development."[4] As she has implied with her arguments, and stated explicitly with her analysis, religion clearly has a significant impact on national economies. The information in Grier's study has been adapted into this study in the form of the inclusion of specific religions, which Grier had proved relevant in her conversation specifically about Protestantism.

The second study to be analyzed is McCleary's study on religious moderation. One of the stronger topics the article discusses is the 'two-way causation'[5] of religion and development in which McCleary states "The more educated a person is, the more likely he is to turn to science for explanations of natural phenomena… [a]ccording to this view, the higher the levels of educational attainment, the less religious people will be… On the other hand, an increase in education will also spur participation in religious activities, because educated people tend to appreciate social networks and other forms of social capital."[5] This two-way causation, from the viewpoint of the study at hand, suggests that variable literacy would share in the effect of religion on per capita GDP possibly creating a bias. But more importantly what McCleary suggests is the possibility that the full effect of religion very well may not be felt until the population is educated, and although questions of causality arise, this increasing returns in wealth presents a situation of exponential returns to religion, possibly observing a sort of economies of scale in which wealth increases faster as more people capture their human capital potentials.

Another important point that McCleary makes is on religious terrorism, where she says "Religious terrorists feel a sense of alienation from the larger society... [Krueger] found that the deterioration in economic conditions over time is associated with the likelihood of educated men becoming terrorist attackers."[5] This spurs an intriguing argument into the effects of religious diversity and tolerance. Although the two are not perfectly correlated, it is presumed that the more diverse religious bodies are the less likely they are to breed religious extremism because of the increases in tolerance and decrease of the sense of 'alienation'. This, according to McCleary's article and Krueger's quoted research would mean that less religiously diverse populations would hinder economic stability and growth. Ultimately, McCleary's findings on religious diversity within populations suggest that my four religion ratio should have both a large and statistically significant impact on per capita GDP, as well as fitting into a larger context of religious variables that have significance.

Finally, in the article most similar and relevant to this one, Noland states that there is "[a]bundant evidence affirm[ing] that religious belief affects a wide range of behavioral outcomes (Iannaccone 1998), and religious activity can affect economic performance at the level of the individual, group, or nation." 6 As Noland goes through the observations of Weber, Greif, and Solow though, he goes into extensive detail about historical examples only to conclude by saying "What does this have to do with economic performance? … [T]hese cultural measures [can't] directly explain economic performance, and indeed, they do not appear to be correlated with economic growth

5 Rachel M. McCleary, "Religion and Economic Development: The advantage of moderation", *Hoover Institution, Stanford University,* 147 (March 28, 2008) Accessed Oct. 2011 http://www.hoover.org/publications/policy-review/article/5729

6 Marcus Noland, "RELIGION, CULTURE, AND ECONOMIC PERFORMANCE" (Working Paper*), The Peterson Institute for International Economics* Accessed Oct. 2011 http://www.iie.com/publications/wp/03-8.pdf

rates."[6] This allusion to his results suggests that religion's impact is insignificant on his dependent variables (total factor productivity)[6], and ultimately this suggests that religious variables might not be significant in studies of religious variables impact on national wealth. This of course is not reason enough to halt since there are a few major distinctions to be made between this study's model and Noland's, most notably the difference in dependent variables.

It is clear that although economics of religion is still emerging, a great amount of research has been brought forth, but even more clear is the room for expansion and improvement on previously existing models. This is made obvious in the contradictory points of many articles in the field, such as McCleary's and Roland's articles that have been analyzed here.

Explanation of the Model

For this model pertaining to per capita gross domestic product (GDP) numerous independent variables were chosen for several reasons. The variables nominated have been consider either because of research done by prominent economists in the field of economics of religion, or because they have been suggested as relevant through economic theory. These bases exclude the explanatory variables on religion, which are the focus of this study, because the correlations of these variables are currently only speculative; meaning these were not chosen from either previous studies or from theory since they are believed to be novel and therefore have not been explored directly through any previously existing works.

The variables chosen to explain the variation in per capita GDP included the real GDP growth rate, the inflation rate, the unemployment rate, urbanization, the literacy rate, and finally the industrial makeup of the country. These variables have been included in the model because economic theory suggests a strong correlation between these variables and per capita GDP or because their omission may have led to a bias in the coefficients of other independent variables.

Since there are various major religions in the world, an observed country's religious identity has been captured with variables in a few different forms.

The first form that the observed country's religious makeup, or religious profile, was included as, was a direct measure of what religions made up the population of the observed country. Since there was little to no unbiased literature on the specific economic effect of one religion over another, numerous specific religions have been included in the model independently of one another as to test their specific effects on per capita GDP. These specified religions include: Christianity, Buddhism, Islam, Hinduism, Sikhism, Judaism, and Irreligion.

The second form chosen to incorporate a country's religious profile was a four religion ratio within an observed country. This measurement, adapted from the 'four firm concentration ratio' measurement common in industrial organization, serves as a quantification of the religious diversity within an observed country. Since it was observed that more oppressive governments often reduce religious tolerance, as well as the same corrupted regimes also hindering wealth, it was expected that as this measure of religious diversity will be negatively correlated with GDP per capita.

Finally, the last form that was chosen to incorporate religious profiles in to the model was the measurement of years of independence that the observed country had experienced. This measurement was include because it was theorized that a country that has been occupied by a foreign power most likely would have had its religious identity

Variable	Formula
Industrial Makeup (e.g. Service)	$Industrial\ Makeup = \frac{GDP_{Service}}{GDP_{total}}$
Urbanization	$Urbanization\ Rate = \frac{Population\ in\ Urban\ Areas}{Total\ Population}$
Literacy Rates	$Literacy\ Rate = \frac{Population_{older\ than\ 15, can\ read}}{Population_{total}}$
Percent of Population (by religion) (e.g. Islam)	$Percent\ of\ Population_{Islam} = \frac{Population_{Islamic}}{Population_{total}}$
Four religion ratio (FRR)	$Four\ Religion\ Ratio = \sum 1st, 2nd, 3rd, 4th\ Largest\ Religions\ \ (percentages)$
Years Independent	$Years\ of\ Independence = 2012 - Year\ of\ Declared\ Independe$

Table 1. *Formulae of calculated variables.*

stripped or suppressed in that time of occupation, such as the 15 countries formerly part of the Soviet Union in which religion was objectively abolished and replaced with atheism.[7]

Description of the Data

The data collected for this study consists of a cross-sectional collection from 188 of the 196 countries in the world. The omitted eight were left out due to either unreliable gross domestic product (GDP) figures, missing religious profile estimates, or because of other anomalies in the data. All of the data, religious or economic, was derived from the most recent publishing and estimates when available.

As noted above, some of the variables specified in my model required calculation. Below is a table of formulas explaining each variable's calculation (see table).

Since much of the data collected is directly related to the definition of the collector, an important note is the sources of the data. The industrial makeup, GDP real growth rate, inflation rate, unemployment rate, urbanization, literacy, and years of independence were all calculated using data collected from the CIA World Factbook. The remainder of the data (religious profiles and FRR) has been based of data collected by the World Christian Encyclopedia8. This data was deemed "best" since it was the most thorough and large-scale of any data source in the field since it had estimates for

7 "Revelations from the Russian Archives: ANTI-RELIGIOUS CAMPAIGNS". *Library of Congress*. July 22, 2010. Web accessed October 2011. http://www.loc.gov/exhibits/archives/anti.html

8 David B. Barrett, et.al, "World Christian Encyclopedia". *Oxford University Press*. 2001. Web Accessed 2011.

GDP/capita ($)		*Unemployment (%)*		*Inflation rate (%)*	
Mean	**15431.15**	Mean	14.00	Mean	4.88
Standard Error	**1572.54**	Standard Error	1.20	Standard Error	0.34
Median	**8400.00**	Median	8.50	Median	3.70
Mode	**2500.00**	Mode	6.70	Mode	2.30
Standard Deviation	**21272.98**	Standard Deviation	15.21	Standard Deviation	4.55
Minimum	**300.00**	Minimum	0.50	Minimum	-2.40
Maximum	**179000.00**	Maximum	95.00	Maximum	28.20
% of GDP Services		***% of GDP Industry***		***% of GDP Agriculture***	
Mean	56.43	Mean	29.63	Mean	13.85
Standard Error	1.13	Standard Error	1.04	Standard Error	1.03
Median	57.90	Median	26.60	Median	9.50
Mode	58.20	Mode	24.40	Mode	2.40
Standard Deviation	15.46	Standard Deviation	14.25	Standard Deviation	13.99
Minimum	3.80	Minimum	5.40	Minimum	0.00
Maximum	92.50	Maximum	93.90	Maximum	76.90
Literacy		*Urbanization (% in cities)*		*GDP real growth rate (%)*	
Mean	82.28	Mean	56.52	Mean	3.80
Standard Error	1.49	Standard Error	1.72	Standard Error	0.27
Median	91.80	Median	58.00	Median	3.85
Mode	99.00	Mode	52.00	Mode	4.20
Standard Deviation	20.08	Standard Deviation	23.17	Standard Deviation	3.70
Minimum	21.80	Minimum	11.00	Minimum	-13.00
Maximum	100.00	Maximum	100.00	Maximum	16.30
Buddhism%		*Muslim %*		*Christian %*	
Mean	3.25	Mean	24.12	Mean	56.16
Standard Error	0.95	Standard Error	2.60	Standard Error	2.81
Median	0.02	Median	3.47	Median	70.14
Mode	0.00	Mode	0.00	Mode	None
Standard Deviation	12.89	Standard Deviation	35.14	Standard Deviation	38.07
Minimum	0.00	Minimum	0.00	Minimum	0.10
Maximum	85.27	Maximum	99.09	Maximum	98.07
Judaism %		*Sikhism %*		*Hinduism %*	
Mean	0.50	Mean	0.04	Mean	2.26
Standard Error	0.39	Standard Error	0.01	Standard Error	0.68
Median	0.01	Median	0.00	Median	0.01
Mode	0.00	Mode	0.00	Mode	0.00
Standard Deviation	5.28	Standard Deviation	0.20	Standard Deviation	9.26
Minimum	0.00	Minimum	0.00	Minimum	0.00
Maximum	71.39	Maximum	2.19	Maximum	72.82
Years of independence		*Four-firm concentration (of religions)*		*Irreligion %*	
Mean	127.82	Mean	92.00	Mean	5.88
Standard Error	18.81	Standard Error	0.995	Standard Error	0.66
Median	52.00	Median	97.96	Median	1.99
Mode	21.00	Mode	100.00	Mode	None
Standard Deviation	253.75	Standard Deviation	13.464	Standard Deviation	8.94
Minimum	0.00	Minimum	32.01	Minimum	0.00
Maximum	2000.00	Maximum	100.00	Maximum	49.76

Table 2. *Summary Statistics of the Data.*

every country in the world, which, even if estimated, makes for an even playing field on which to compare all religions, even in countries without proper census data. This was also important because the definition of irreligion differed from source to source particularly since different sources defined irreligion differently, so using data from multiple sources would have been out of the question.

Table 2 contains a summary of the data that has been collected and entered for all of the independent variables from this model testing the effect of religious profiles on per capita GDP.

There are a few major clarifications that need to be made about the data specifically. The first difference is described by the discrepancy in average populations by religion from their counterparts that reflect the world's population. According to adhearents.com, 33% of the world's population is Christian9, whereas the data for this model shows the average country at 56.16%. This difference arises because the data collected was based on each country producing an unweighted percent of the population that does not take into account the relative size of the country, so for the purposes of this study, the Cook Islands have the same weight as United States or India.

The second clarification that needs to be made refers to the interpretation of means from the above data. Since smaller countries have larger impacts on this data set, a variable with a mean above what is separately published should be interpreted differently. This means that the $15,431 GDP per capita mean should not be compared to the $9,216 published by the World Bank. This is because the World Bank's publications are based on a per person average and not a per country average like suggested in this study's data. Similar observations can be made about the 56.16% Christian that was mentioned before.

Some numbers from the above table have some very interesting implications. The variation in the dependent variable (GDP per capita) was expectedly great from prior knowledge of poor wealth distribution of society and was reflected as such in the standard deviation of $21,272, which compared to the mean of $15,431 was very large. The median of $8,400 per capita GDP has the most interesting implications though. The median suggests that the majority of countries around the world are struggling (portrayed by a median well less than the mean, suggesting that more that 50% of countries in the world live below the world average) and the mean statistic was so high likely because of some outliers skewing the data upwardly, suggesting that a few countries have abnormally high levels of wealth.

Another set of data worth discussing is the data retaining to the variable describing Sikhism in a population. This variable's data has relatively low variation and therefore is expected to have little to no explanatory power over GDP per capita. This is important to keep in mind since it will like lead to the variables omission. The same can be said for Judaism since the mean is only 0.5, meaning a sum of Jewish percentages is only 94 (since there are 188 observations) and 71.39 of that comes from one country hinting at relatively low variance, particularly if you were to exclude the outlying observation.

Estimation of the Model

When initially estimating the model for religion's impact on Gross Domestic Product (GDP) per capita, seventeen variables were incorporated including eight variables that economic theory considers to be strong factors that account for GDP. The remaining nine variables have been included to account for the suspected impact that religion would have on the model. With all seventeen variables included, initial estimates of the model did not live up to expectations and immediately raised major questions about possible error.

With all seventeen variables incorporated, the model was only able to account for 26.62% of the variation found in GDP per capita. This relatively low adjusted R2 did not meet expectations since variables were included that were expected to be strong indicators of GDP per capita. This hinted towards the possible existence of a specification error and more specifically that the wrong form function may have been chosen. This was supported by a graphical presentation of the data collected for GDP per capita, which portrays the dependent variable as exponentially growing from observation to observation, suggesting that the log-level form should be used.

Once the model was specified in the log-level form, suggested to be the correct form from the graphical depiction of GDP per capita, some of the expectations that were previously unmet became fulfilled. The new model accounted for 77.07% of the variation of the dependent variable (now the natural log of GDP per capita and therefore not comparable to the original model in which GDP per capita was the dependent variable) but still showed signs of some major econometric issues including multicollinearity (as suggested through variance inflation factors), irrelevant variables, and heteroskedasticity (as suggested through White's test).

The first of the issues addressed, multicollinearity, was identified by creating variance inflation factors (VIFs) which can be compared against the rule of thumb of 5. This means that if more than 80% of the variance of one independent variable can be accounted for by the other independent variables then we assume multicollinearity exists. This creates issues with the VIFs associated with the industrial makeup variables (the measurement of GDP from the agricultural, service, and industrial sectors of an observed countries economy), and the variables representing the Christian percentage of the population of a country and a Muslim percentage of the population of a country. The VIFs for these problem variables were calculated as 105.9, 150.73, 123.79, 20.1, and 15.97 respectively; all of which are substantially over the accepted rule of thumb amount.

To correct for multicollinearity in the model, the culprit variables were examined to determine the root causes of the multicollinearity and to establish which variables should be removed to correct the model. Since the VIFs on the industrial makeup of the observed country were so high, and the three make up an identity in which the three must equal 100% since the entirety of GDP is generated through either agriculture, service, or industry and therefore represent a textbook example of multicollinearity. To solve for this, the variables representing the percent of GDP from service and percent of GDP from industry were removed. This left just the variable measuring the percent of GDP from agriculture which, with the other two variables excluded, had a VIF of just 2.24; a great improvement from 105.9.

The correction for multicollinearity associated with the religious variables (specifically the variables measuring the percentage of the population that is Christian in the observed country and the percentage of the population that is Muslim in the observed country) was achieved in a slightly different method than that of the industrial makeup solution since the multicollinearity was not as obvious as the identity observed in the industrial makeup variables.

To investigate which variables might be causing the high VIFs for the Christian

9 "Major Religions of the World Ranked by Number of Adherents".*Adherents.com*. August 9, 2007. Web. Accessed Oct. 2011. http://www.adherents.com/Religions_By_Adherents.html

and Muslim variables, separate regressions were run in which the Christian and Muslim variables were set as the dependent variables. The variables that most likely caused the multicollinearity, which would be represented by high t-statistics in both models, included the other specific religions measured (Sikhism, Judaism, Hinduism, Buddhism, and Irreligion) and the variable measuring the number of years independence the observed country had experienced. These variables, particularly Judaism, Sikhism, and Irreligion, were not statistically significant in measuring GDP per capita and did not contain a great amount of variation in the data collected, leading me to believe they were weak indicators of the observed GDPs and could be readily dropped from the model. Similarly, the t-statistic associated with the years of independence, like the t-statistics of Judaism and Sikhism, were relatively low meaning that we (as researchers) could not accept the variable to be statistically different from zero, leading me to believe that the best course of action would be to drop the variable.

With the above corrections made, the VIFs for the variables Christian and Muslim both dropped below 10, but were unable to drop below 5 without omitting the other variable. In the best interest of the investigatory power of the model both variables were left in since the effect of specific religions would be best observed in the two largest religions in the world. This of course doesn't completely rule out the possibility of multicollinearity, but because the VIFs were low enough, suggested that the effect that the multicollinearity has on the variable's standard errors is low enough to proceed with the regression analysis and more specifically hypothesis testing.

The next issue that arose in the model accounting for variations in the natural log of GDP per capita is in reference to heteroskedasticity. The model, which now incorporated eleven variables, was put through the White test to test for heteroskedasticity, to test for the problem, and indeed the results suggested that heteroskedasticity did exist.

To fix this issue, the model was estimated again with robust standard errors. This new model using robust standard errors saw all of the standard errors increase, except for the standard errors of the four religion ratio variable and the Hinduism variable. Although the Hinduism variable remained at a level of insignificant impact on the dependent variable, the four religion ratio variable went from insignificant to significant with the inclusion of robust standard errors. Oppositely, the variable accounting for urbanization became insignificant with the increase in its standard error. This ultimately led me to the realization that urbanization was the variable creating heteroskedasticity in the model, and led me to omit the variable. Omitting the variable resulted in a correction of the results from the White test, as well as the same desired result that the robust standard errors generated with regard to four religion ratio.

Now down to nine independent variables, the model accounting for variation in the natural log of GDP per capita was significantly stronger as it was cleared of multicollinearity and heteroskedasticity following the above fixes. The last specification test that was run was the Ramsey's RESET test to test for omitted variables. The results of the RESET test suggested that there was in fact an omitted variable, which is believed to be the omitted measure of religiosity for each country. This variable would affect each variable in the model differently, by giving weight to the religions on a country by country basis instead suggesting that all members of a religion practice with the same devotedness from country to country. Unfortunately this data is not available, and admittedly leaves something to be desired for future expansion of the model as more countries conduct extensive censuses.

Ultimately, the remaining variables (percent of GDP from agriculture, inflation rate, unemployment rate, and literacy rate) were all tested in multiple functional forms to determine the positive or negative impact that the change in form would have on the model. After testing each variable in their level form as well as the natural log form the model's R2, RESET test f-statistic, VIF, and White test results were compared to determine the best form function on a variable by variable basis. The results of these comparisons suggested that the variables associated with inflation rate, unemployment rate and percent of GDP from agriculture would all be better served in the log-log form instead of the semi-log form. It is important to note here though that when minimizing the F-statistic of the RESET test, versions of the model exist that were able to get the F-statistic as low as 1.50 (compared to the 4.11 produced by the final model) but this required dropping the R2 down as low as .66 compared to the .82 that the final model produced, and accepted heteroskedasticity, which can be corrected by using Robust standard errors. The results for this alternative model which reduces the RESET f-statistic are in appendix 1.

With the model finally correctly specified, the correlation coefficients, robust standard errors, and t-statistics were calculated for the remaining variables as well as the F-statistic for the model as a whole. This final model accounted for a large amount of variation in the dependent variable as represented by the R2 of 0.82; a very high R2 for cross sectional data. This coupled with the F-statistic of 99.37 (significant even at the 0.0001 level) gave good indication to the strong explanatory power of the model.

Some other highlights of the regression results include the statistically significant effect that the four religion ratio has in a negative direction, which verifies the assumptions that diversity would have a positive effect of GDP per capita. Similarly, literacy (hypothesized to have a positive impact on GDP per capita) and GDP from agriculture (hypothesized to have a negative impact on GDP per capita) were both significant, while other key variables such as inflation rate and unemployment were not significantly different from zero.

To focus on the main point of the research though, it was necessary to test the significance of the religious variables as a whole. This entailed running an F-test on the five variables describing religious statistics. This produced an F-statistic of only .7639 which at the degrees of freedom provided by the data produced a P-value of 0.42 on the f-statistic10. This is interpreted as meaning that the inclusion religious indicators in a model explaining GDP per capita has an insignificant impact on the model's statistical significance.

Model Results

See Table 3.

Conclusions

This study and subsequent regression analysis shone a light on a real world issue that haunts all political systems, not just that of developing nations. Although the regression proved that the religious profile of a country does not have a significant impact on the country's observed Gross Domestic Product, the statistical significance of the four religion ratio and Christian population of an observed country could be interpreted as evidence for the

10 Calculated by using the "p-Value Calculator for an F-Test" on Dr. Daniel Soper's website.

-Dr. Daniel Soper "Statistics Calculators" *Statistics Calculators Index*, Web. 2006. http://www.danielsoper.com/statcalc3/calc.aspx?id=7

	Semi-Log Model		Semi-Log Model w/ (Robust Std Err)	
Percent of GDP from Agriculture	-.7113~	(.0469)*	-.7113~	(.0584)*
Unemployment Rate	-.0475~	(.0504)	-.0475~	(.0716)
Inflation Rate	-.0017~	(.0573)	-.0017~	(.0593)
Literacy	.0183	(.0035)*	.0183	(.0037)*
Percent of Population Christian	.0074	(.0029)*	.0074	(.0030)*
Percent of Population Muslim	.0049	(.003)	.0049	(.0031)
Percent of Population Buddhist	.0081	(.0046)*	.0081	(.0049)
Percent of Population Hindi	.0038	(.0056)	.0038	(.0049)
Four Religion Ratio	-.0738	(.0494)	-.0738	(.0342)*
R Square	0.8231		0.8231	
Adjusted R Square	0.8120		--	
F-statisticistic (9, 143)	73.94*		99.37*	
*-Significant at Alpha of .05	~ - Log-Log Est.			

Table 3. *Regression results.*

argument supporting government policy to control religious profiles in its country. It is observed that some other coefficients, specifically the percent of GDP from the agricultural sector, have much stronger impacts on the model and would make for more productive policy topics with respect to many of the developing nations around the world.

This study is admittedly limited due to the lack of data for so many countries around the world and in the future, studies within economics of religion could be better served by a model similar to this one if panel data could be collected. This would prove substantially stronger since it would be implied (since it is one observation over time) that similar methodologies of data collection would be used in panel data, whereas the consistency of religious data available today, like that used in this model, is not nearly as accurate as it could be.

Appendix 1 (opposite top). Linear Regression.

Appendix 2 (opposite bottom). Final Regression.


```
Linear regression                                      Number of obs =     152
                                                       F(  9,   142) =   44.80
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.6598
                                                       Root MSE      =  .72672

------------------------------------------------------------------------------
             |             Robust HC3
       lngdp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lninflatio~e |  -.1595233   .0680949    -2.34   0.021     -.294134   -.0249126
lnunemploy~t |  -.0275235   .0911726    -0.30   0.763    -.2077545    .1527074
urbanizati~s |   .0224103   .0052269     4.29   0.000     .0120778    .0327428
    literacy |   .0349639   .0056383     6.20   0.000     .0238179    .0461098
   christian |   .0126922   .0045552     2.79   0.006     .0036875     .021697
      muslim |   .0141609   .0045273     3.13   0.002     .0052112    .0231106
    buddhism |   .0148563   .0068734     2.16   0.032     .0012688    .0284437
    hinduism |   .0158676   .0083307     1.90   0.059    -.0006006    .0323358
fourfirmco~s |  -.1535181   .0876867    -1.75   0.082    -.3268581    .0198219
       _cons |   19.15472   8.758101     2.19   0.030     1.841615    36.46783
------------------------------------------------------------------------------

. estat ovtest

Ramsey RESET test using powers of the fitted values of lngdp
       Ho:  model has no omitted variables
                 F(3, 139) =      1.50
                  Prob > F =      0.2180
```

```
      Source |       SS       df       MS              Number of obs =     153
-------------+------------------------------           F(  9,   143) =   73.94
       Model |  182.347048     9  20.2607831           Prob > F      =  0.0000
    Residual |  39.1831359   143  .274007943           R-squared     =  0.8231
-------------+------------------------------           Adj R-squared =  0.8120
       Total |  221.530184   152  1.45743542           Root MSE      =  .52346

------------------------------------------------------------------------------
       lngdp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
lnofgdpagr~e |  -.7112694   .0469424   -15.15   0.000    -.8040601   -.6184786
lninflatio~e |  -.0017442   .0572585    -0.03   0.976    -.1149267    .1114382
lnunemploy~t |  -.0475402   .0503664    -0.94   0.347    -.1470991    .0520186
    literacy |   .0183415   .0034515     5.31   0.000     .0115189     .025164
      muslim |   .0048916   .0030124     1.62   0.107    -.0010631    .0108463
   christian |   .0074016    .002925     2.53   0.012     .0016197    .0131834
    buddhism |   .0080783   .0046208     1.75   0.083    -.0010557    .0172123
    hinduism |   .0037645   .0056215     0.67   0.504    -.0073475    .0148765
fourfirmco~s |  -.0737745   .0494227    -1.49   0.138     -.171468     .023919
       _cons |   15.73076   4.916268     3.20   0.002     6.012808     25.4487
------------------------------------------------------------------------------

. estat ovtest

Ramsey RESET test using powers of the fitted values of lngdp
       Ho:  model has no omitted variables
                 F(3, 140) =      3.53
                  Prob > F =      0.0166

. estat imtest, white

White's test for Ho: homoskedasticity
         against Ha: unrestricted heteroskedasticity

         chi2(54)      =     50.66
         Prob > chi2   =    0.6038
```

Works Cited

Brown, Timothy T. "A Monetary Valuation of Individual Religious Behavior: The Case of Prayer" (Working Paper), The Association of Religion Data Archives, Accessed Nov. 2011 http://www.thearda.com/workingpapers/download/Valuing%20Religious%20Behavior.pdf

Central Intelligence Agency, "The World Factbook", cia.gov, Accessed Oct. 2011 https://www.cia.gov/library/publications/the-world-factbook/

Fitzgerald, Michael , "Satan, the great motivator" Boston.com (Boston, MA) Nov.15.2009 http://www.boston.com/bostonglobe/ideas/articles/2009/11/15/the_curious_economic_effects_of_religion/?page=full

Franck, Raphaël and Iannaccone Laurence R., "Why did religiosity decrease in the Western World during the twentieth century?" From the April 1, 2010 Seminar, Bar-Ilan University, April 1, 2010. Accessed Oct 2010 http://www.biu.ac.il/soc/ec/seminar/data/4_01_2010/franck%20iannaccone%20religiosity%20biu.pdf

Grier, Robin , "The Effect of Religion on Economic Development: A Cross National Study of 63 Former Colonies" Kyklos: International Review for Social Sciences, 50, (1997). Accessed Oct 2011. http://faculty-staff.ou.edu/G/Robin.M.Grier-1/religion.pdf

Iannaccone, Laurence R. , "Introduction to the Economics of Religion" Journal of Economic Literature, XXXVI (September 1998). Accessed Oct 2011. http://www.scribd.com/doc/62863336/Iannaccone-Introduction-to-the-Economics-of-Religion

McCleary, Rachel M. , "Religion and Economic Development: The advantage of moderation", Hoover Institution, Stanford University, 147 (March 28, 2008) Accessed Oct. 2011 http://www.hoover.org/publications/policy-review/article/5729

Noland, Marcus , "RELIGION, CULTURE, AND ECONOMIC PERFORMANCE" (Working Paper), The Peterson Institute for International Economics Accessed Oct. 2011 http://www.iie.com/publications/wp/03-8.pdf

World Christian Database Vol.2 , "Excel Datafile" worldmapper.org, http://www.worldmapper.org/data/nomap/religion_data.xls

Equity Volatility across Industries from IPO Day

Andrew Nielsen Lewis
Princeton University

Abstract

While the volatility of any given stock changes over time, certain identifiable factors consistently contribute to increased stock price volatility. For example, stocks are typically more volatile immediately following their IPOs than they are after years of active trading. In addition, industries that are characterized by increased uncertainty typically include companies with more volatile stock prices. This study examines the trends of intraday stock volatility by industry over time from IPO date and determines which industries have the most volatile IPOs, which industries' stocks take the longest amount of time to stabilize after IPO day, and which industries' stocks stabilize the most over time. It uses a nonlinear model to estimate the trends of stock price volatility in various industries as well as the overall market.[1]

Section I: Introduction

While the volatility of any given stock changes over time, certain identifiable factors consistently contribute to increased volatility. For example, stocks are typically more volatile immediately following their IPOs than they are after years of active trading. In addition, industries that are characterized by increased uncertainty typically include companies with more volatile stock prices. This study examines the trends of intraday stock volatility by industry over time from IPO date and determines which industries have the most volatile IPOs, which industries' stocks take the longest amount of time to stabilize after IPO day, and which industries' stocks stabilize the most over time.

Although many studies have investigated returns following IPO date or general trends in equity volatility, they have not focused on modeling trends in volatility relative to IPO dates. The study of de facto IPO returns is more appealing in that it has quick and obvious applications – after discovering a way to predict abnormal stock returns around an IPO, traders or hedge fund managers can quickly apply their new knowledge to realize excess gains. The findings of this study regarding volatility trends from IPO, however, can also be used by investors to profit. The ability to predict volatility can lend insight into forecasting option prices by using the Black-Scholes options pricing model; being able to predict with some certainty the change in a stock's volatility from its IPO day will also allow a trader to predict with more accuracy the motion of related option prices. With information about volatility trends in an industry, an investor can more accurately value a stock's options.

Additionally, the findings of this study can help companies to make better choices regarding their own IPOs – having an understanding of typical industry volatility will help to guide decisions on whether to go public and how to price IPOs. This will not only benefit the few who stand to make extraordinary profits from such business deals but the entire stock market and economy as the IPO market becomes more efficient. The findings can also be used for risk management purposes, in behavioral

1 Acknowledgements: I greatly appreciate the guidance provided to me by Professor Yuliy Sannikov (Princeton University Department of Economics), Wei Cui (Princeton University Department of Economics), and Oscar Torres-Reyna (Princeton University Data and Statistical Services) during this study.

finance studies, and other industry-based stock studies.

Section II: Literature Review

Existing literature has not focused on trends in volatility from IPO date or industry-specific volatility trends.

Differences between volatility levels and IPO behavior across industries are recognized phenomena. In "Stock volatility in the new millennium: how wacky is Nasdaq?," G. William Schwert notes that technology stocks and recent IPO stocks exhibit much higher trading volatility than other stocks (2002, p. 3-4). A model constructed by Michelle Lowry, Micah Officer, and Schwert shows that technology stocks see returns that are different from other stocks on a statistically significant level both including and excluding the technology bubble of the late 1990s and early 2000s (2010, p. 438).

Existing literature is more concerned with IPO returns or IPO return volatility – which can help to predict immediate returns from IPO investments – than with absolute stock price volatility near IPO. Investigations of volatility trends focus more on period-specific trends or other general trends. For example, the Schwert 2002 study focuses on volatility trends in the NASDAQ during the dot-com period. In another example, Schwert's "Why Does Stock Market Volatility Change Over Time?" deals only with trends in overall market volatility throughout modern market history (1989, p. 1120). David Simon investigates market volatility trends around the dot-com bubble period in "The Nasdaq Volatility Index During and After the Bubble." Just as with IPO inquiries and industry-based investigations, much volatility trend research relates to predicting security returns. For example, Baillie and DeGennaro search for (but do not find) correlation between volatility and equity returns in "Stock Returns and Volatility" (1990, p. 211). This study differs in that it only analyzes and models trends in price volatility instead of returns.

Section III: Data

Data for all stocks used in this study were retrieved from the Center for Research in Security Prices (CRSP) daily stock database via Wharton Research Data Services. For each stock, all available trading data was retrieved during the period ranging from January 1, 1925 through December 31, 2009; data after December 31, 2009 was not comprehensive or fully accurate in the CRSP database. The decision to use trading data from before 2010 should not have a significant impact on results.

First, the stocks that were components of the S&P 500 on December 31, 2009 were studied in aggregate to create a broad market benchmark. Then a series of Dow Jones sector indexes (which were not limited to S&P 500 stocks) were used to represent the ten sectors examined in this study. The components of these indexes were provided by MarketWatch.2 An inaccurate intraday trading range assigned to one observation (IdeaRC on December 31, 2009) was removed from the Consumer Services industry data set prior to regression.

For each stock, the data retrieved from CRSP to complete the study included company name, CRSP-assigned PERMNO, trading date, dividend amount, daily low price, daily high price, and daily closing price. All price values were automatically adjusted by CRSP to reflect the impact of stock splits on real changes to equity value. A database error caused some daily closing prices to be returned as negative numbers; as a result, all negative closing prices were

2 Components used for each index can be found at http://bigcharts.marketwatch.com/industry/bigcharts-com/default.asp (originally accessed March 22, 2011)

replaced with their absolute values before continuing. Since CRSP cannot automatically adjust prices to reflect the impact of dividend payments, a cumulative dividend payment variable was generated and added to all price variables; this adjustment was meant to reflect the real effects of dividend payments on shareholder value. For example, Microsoft issued a one-time special dividend of $3.00 on November 15, 2004; because the price drops associated with dividend payouts do not affect shareholder value, 3 was added to the closing prices, low prices, and high prices for all Microsoft observations on or after the day the dividend was issued. This does not lead to a perfect adjustment for dividend payouts. Since no stocks paid out dividends within their first thirty days of trading, however, no observations used in regression were affected; as a result, this inaccuracy did not affect results.

The data is subject to some level of selection bias. All of the companies chosen were still trading as of December 31, 2009; as a result, companies taken private for reasons including bankruptcy and acquisition may be underrepresented. Furthermore, the accuracy of the S&P 500 as a true market benchmark is not perfect because the S&P 500 includes only large companies that have maintained consistent profitability. If stock resilience or company sizes are correlated with IPO period volatility trends, the results of this study may be skewed as a result.

Section IV: Methodology

The methodology followed throughout this study consists of two processes. First is the calculation of $\bar{\nu}$ values that represent average intraday price volatility for each industry. Second is the nonlinear regression of $\bar{\nu}$ values against time from IPO in order to estimate volatility trends following IPO day in each industry.

Calculating $\bar{\nu}$

The $\bar{\nu}$ calculation process was run individually using data sets that include the trading history for the S&P 500 and each of the ten selected industry groups provided by MarketWatch. For all examined data sets, a day counter variable was generated for each observation to mark the number of days that had passed from IPO. On a stock's IPO day, the variable equals 1. On the day immediately following IPO, the variable equals 2. For each additional day from IPO, the variable is incremented by one. The creation of this variable makes comparison of the chronologically scattered IPO periods for different stocks in each industry possible.

Actual volatility (ν) was calculated for every observation using the stock's intraday trading range squared relative to closing price. Each observation's high price (P_H), low price (P_L), and closing price (P_C) were used to calculate volatility ν for each trading day:

$$\nu = \frac{(P_H - P_L)^2}{P_C} \tag{1}$$

Microsoft Corporation had its initial public offering on March 13, 1986. On its first day of trading, its adjusted low price was $25.50, its adjusted high price was $29.25, and its adjusted closing price was $28.00.

$$\nu_{MSFT:IPO} = \frac{(29.25 - 25.50)^2}{28.00} = 0.5022321 \tag{2}$$

As a result, Microsoft's volatility rating ν for IPO day was 0.5022321. A higher calculated ν represents increased volatility.

This calculation is based on "the standard definition of volatility" as defined by Euan Sinclair in *Volatility Trading* (2008, p. 16). Sinclair's definition, where N defines the number of periods sampled, is:

$$\nu = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2} \quad (3)$$

For the purposes of this study, N is equal to 1 because volatility is measured on a day-by-day basis:

$$\nu = \sqrt{\frac{1}{1}\sum_{i=1}^{1}(x_i - \bar{x})^2} = \sqrt{(x_i - \bar{x})^2} \quad (4)$$

Since the $(x_i - \bar{x})^2$ term merely represents deviation with respect to an average, intraday trading range squared as a percentage of closing price was used instead to normalize differently priced stocks.

$$\nu = \sqrt{(x_i - \bar{x})^2} = \sqrt{\frac{(P_H - P_L)^2}{P_C}} \quad (5)$$

The square root of the calculated percentage was not taken. The final ν calculation measures variance instead of standard deviation:

$$\nu = \frac{(P_H - P_L)^2}{P_C} \quad (6)$$

Volatility ν was calculated for stocks in the ten Dow Jones Sector Indexes provided by MarketWatch in addition to the S&P 500 which serves as a market benchmark. Volatility ratings ν for stocks in each industry group were then divided by the number of stocks in the index (N) to calculate an average composite measure of industry volatility $\bar{\nu}$ for each day relative to IPO. Where N is the number of stocks in an index with an available ν value:

$$\bar{\nu} = \frac{1}{N}\sum_{i=1}^{N}\nu_i \quad (7)$$

For example, on Day 1 the S&P 500 index contained 500 stocks with available data. The sum of all ν values divided by 500 was 0.2221825 on Day 1 and 0.114652 on Day 2. This suggests that an average large-cap stock's volatility decreases by about 50% after its first day of trading.

Regression Model

After calculating $\bar{\nu}$ values for every day from IPO in each industry (represented using a variable), a nonlinear regression was used to model $\bar{\nu}$ as a function of day from IPO. No values besides one $\bar{\nu}$ value for each $counter$ value were required to run the regression. To investigate the IPO period in isolation, observations for any day past 30 were dropped from the regression in order to minimize the impact of significant non-IPO events on the regression results; the 30-day period was chosen to capture the full effect of the IPO event while minimizing the effects of non-IPO events on the data. The regressions run to estimate all three φ values for each index used the following model given known values for $\bar{\nu}$ and $counter$:

$$\bar{\nu} = \varphi_1 + \frac{\varphi_2}{1+\varphi_3*(counter-1)} + \varepsilon_t \quad (8)$$

The regression estimates average industry volatility ($\bar{\nu}$) as a function of the number of days from IPO ($counter$). On IPO day, the ($counter$ - 1) term is equal to 0; this means that average IPO day volatility is estimated by $\varphi_1 + \varphi_2$ since the denominator of the second term becomes equal to one. For each day following IPO, volatility decays as the ($counter$ - 1) term grows and the second regression term approaches 0. The model effectively represents the significant impact of an IPO event on stock volatility.

This model was chosen for its ability to break down volatility into two types: IPO-related volatility and baseline volatility. Since the model's second term rapidly decays as the counter variable increases, its value represents the effect of IPO on stock volatility; φ_2 captures the expected additional volatility on an IPO day when ($counter$ - 1) is 0 and the denominator of the model's second term becomes equal to 1. The φ_1 term is the value of the horizontal asymptote above which the regression settles as increases and the second regression term approaches zero; φ_1 is an indicator of the average level of long-run volatility

Sector	φ_1 (Baseline Volatility)	φ_2 (IPO-Driven Volatility)	φ_3 (Stabilization Rate)	R^2
Basic Materials (548)	0.0303853 (47.53)	0.0632364 (27.65)	1.496436 (7.88)	.9675
Consumer Goods (653)	0.0390784 (31.01)	0.1378587 (30.74)	1.463953 (8.81)	.9929
Consumer Services (798)	0.0598604 (24.60)	0.4793099 (47.82)	3.806189 (8.87)	.9884
Financials (2058)	0.0395426 (69.25)	0.0907731 (42.87)	1.769266 (11.59)	.9860
Healthcare (821)	0.0575546 (25.57)	0.2345721 (25.97)	2.959527 (5.58)	.9619
Industrials (1141)	0.0412802 (47.95)	0.1061883 (31.54)	2.424831 (7.50)	.9740
Market (S&P 500)	*0.0470464 (20.04)*	*0.1760342 (19.88)*	*1.936521 (5.20)*	.9379
Oil & Gas (578)	0.049487 (17.32)	0.1118282 (10.66)	1.686358 (2.93)	.8142
Technology (890)	0.1173751 (26.78)	0.7072644 (40.15)	2.997637 (8.57)	.9837
Telecommunications (143)	0.0297135 (12.20)	0.1283577 (16.27)	1.045903 (4.98)	.9137
Utilities (206)	0.0276752 (33.32)	0.0460092 (13.93)	2.732502 (3.12)	.8794

Table 1. *Regression results.*

that stocks will exhibit after the effects of the IPO period have ebbed. The φ_3 variable represents stabilization rate – the larger the value of φ_3, the quicker stock volatility converges to its long-term volatility level φ_1. For all industries, this model approximates $\bar{\nu}$ values very well with all regression R^2 values above 0.8100 and seven of eleven above 0.950. In addition, all estimated coefficients were significant at the 1% level.

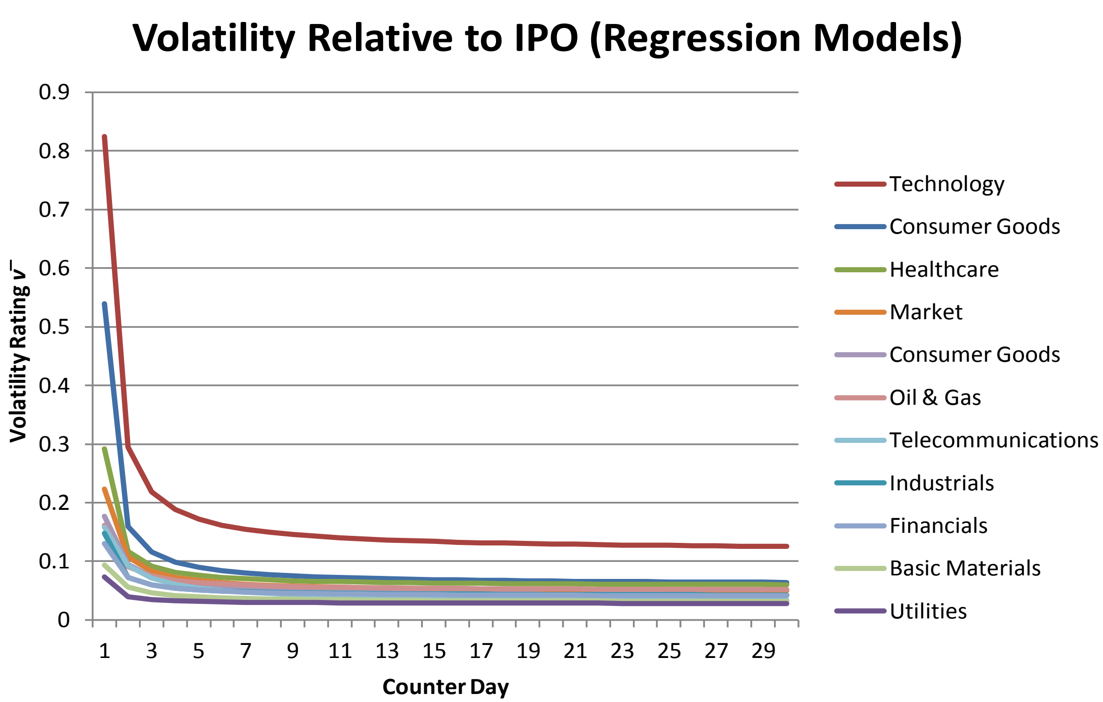
All calculations and regressions were run using Stata version 11.1. For a discussion of Stata-specific methodology, see **Appendix I**.

Section V: Results

The results of the regressions for each industry and the market benchmark are in **Table 1** with corresponding sample sizes and t-statistics in parentheses.

For tables ordered by regression outcomes, see **Appendix II**. For graphs of the regression models with calculated values, see **Appendix V**.

In every regression, all three φ variables were statistically significant at the 1% level. Since all stock prices change regularly, baseline volatility φ_1 was expected to be positive and statistically significant from zero. The consistency of increased volatility on IPO dates means that IPO-driven volatility φ_2 was expected to be positive and nonzero, as well. Finally, since the levels of volatility observed on IPO dates are typically expected to decline over time, stabilization rates φ_3 were also expected to be positive nonzero values. The regression results confirmed these expectations on all counts. Furthermore, the high R^2 values associated with the regressions indicate that the model described by Equation (8) captures the typical trends in price volatility exhibited by stocks in each sector very accurately. At the 10% level, sample size had an insignificant effect on fit quality R^2 (see **Appendix III**); this implies that variations in sample size be-



tween the industries studied are no reason for concern.

The sectors with the highest baseline volatility φ_1 are technology, consumer services, and healthcare. Technology, the most volatile industry, exhibits volatility levels that are equal to nearly double those seen in the runner-up Consumer Services sector. For the Technology sector, this can be explained by the unpredictability of earnings linked to the dynamic, rapidly-evolving nature of the industry and the uncertainty surrounding the future profitability of technological innovations. For similar reasons, the healthcare industry ranks in the top three most volatile industries. For pharmaceutical companies, profitability is determined by the approval (or rejection) of drugs by the FDA; the future cash flows generated by a drug awaiting approval can be very large if the drug receives approval or can be zero if the drug is ultimately rejected. Consumer Services is a very broad sector and includes companies ranging from 4 Kids Entertainment, a children's film and television production company, to Boyd Gaming Corp, a casino operator. On some level, however, their performances are all correlated with both the availability of consumers' disposable income in addition to consumers' constantly-shifting preferences; the unpredictability of these factors adds to uncertainty for all Service companies.

The three stocks exhibiting the lowest amount of baseline volatility – basic materials, telecommunications, and utilities – are all characterized by relatively stable revenue streams and predictable earnings. Responsible for functions earlier in the production line, companies in the Basic Materials industry deal with a customer pool with less elastic demand than those in other industries. For Telecommunications businesses, client contract volume is not typified by sharp drops or big spikes. Utilities, the sector with the lowest baseline volatility, is perhaps characterized by the most predictable future cash flows; profits seen in this sector are usually stabilized by both steady consumer demand and heavy regulation in cases where monopolies are allowed.

A significant positive correlation exists between φ_1 and φ_2. A linear regression showed that for every 0.01 increase in φ_1, the value of φ_2 is expected to increase by about 0.0757; this indicates that, as a general rule, industries with higher IPO-time volatility will also exhibit relatively high long-run volatility (See **Appendix IV**). True to this observation, the top and bottom three sectors by φ_2 are the same as when ordered by φ_1 except for the Financials sector replacing Telecommunications.

The values represent an estimation of how quickly the effects of an IPO on volatility fade; the higher the value of φ_3, the faster an industry regresses towards its baseline volatility level φ_1. Although the three industries with the quickest drops in volatility are the same as the three with the highest φ_1 and φ_2 values, the value of φ_3 does not appear to be significantly correlated with either φ_1 or φ_2; for example, the Utilities sector has the fourth highest φ_3 value but the lowest φ_1 and φ_2 values. As a result, we can conclude that IPO period stabilization rate and volatility magnitude are not closely linked when observing industry averages.

Section VI: Conclusion

The models used in this study's regressions provide both the tools to predict trends in stock price volatility around the 30-day IPO period as well as an effective breakdown of IPO period volatility into its two main components: IPO-fueled volatility and baseline volatility. The results indicate which industries have the most volatile IPOs, which stabilize the fastest, and about where each industry stabilizes; they provide both ordinal and cardinal measures of volatility for all sectors examined. While no two stocks can be expected to behave exactly the same regardless of their commonalities, the industry volatility forecasts found here can help to lend insight into the impact that a stock's industry has on the way it behaves in the market.

Estimating Annualized Volatility and an Options Arbitrage Opportunity

The results of the study have a number of applications. Most practical is the potential to compare industry volatility trends with related option prices; since option prices are directly affected by the volatility of their underlying stocks, the regression models generated in this study can be modified to calculate expected option prices using the Black-Scholes model. This could be used to help predict option price movements for new IPOs. For example, when N() is the cumulative distribution function of the standard normal distribution, T–t is the time to maturity, S is the spot price of the underlying security, K is the strike price, r is the risk-free rate, and σ is the volatility of the underlying asset, the standard Black-Scholes model for the pricing of call options follows the equations:

$$C(S,t) = N(d_1)S - N(d_2)Ke^{-r(T-t)} \quad (9)$$

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}} \quad (10)$$

$$d_2 = \frac{\ln\left(\frac{S}{K}\right) + \left(r - \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}} \quad (11)$$

Having greater insight into the motion of σ in this model given a company's industry will allow for more accurate forecast of d_1 and d_2 and, as a result, a recent IPO's call options. This information can also be applied to put option pricing using the Black-Scholes model:

$$P(S,t) = \left(1 - N(d_2)\right)Ke^{-r(T-t)} - \left(1 - N(d_1)\right)S \quad (12)$$

Once again, if d_1 and d_2 can be predicted with more accuracy across industry, put options can be better valued and their price trends predicted more accurately. Most important, however, is simply the use of the regressions to identify typical volatility trends by industry and, as a result, predicted directional options price movement.

Determining the precise effects on option price movement provided by the regression results exceeds the scope of this study. A starting point, however, would be to estimate annualized volatility using the regression results – the Black-Scholes model uses volatility on an annualized basis. Assuming that the $\bar{\nu}$ ratings represent variance, that any given year includes 252 trading days, and that volatility levels remain close to φ_1 after IPO, estimated annualized volatility as variance σ^2 for any industry could be calculated using:

$$\sigma^2 = \sum_{i=1}^{30} \left(\varphi_1 + \frac{\varphi_2}{1+\varphi_3 * (i-1)}\right) + (222 * \varphi_1) \tag{13}$$

The corresponding annualized standard deviation σ, which is also a component of the Black-Scholes model, can be calculated using:

$$\sigma = \sqrt{\sum_{i=1}^{30} \left(\varphi_1 + \frac{\varphi_2}{1+\varphi_3 * (i-1)}\right) + (222 * \varphi_1)} \tag{14}$$

The resulting σ^2 value represents the sum of all daily expected daily volatility ratings – the first term represents the 30-day IPO period as modeled by the regression results while the second term represents the remaining 222 trading days in the 252-day trading year. Additional adjustments could be made to this estimation; for example, the long-run volatility term could be adjusted to account for unforeseen volatility spikes over the course of the trading year. Using some percentage ρ to represent the effect of abnormal volatility spikes on long-run volatility, we can calculate:

$$\sigma^2 = \sum_{i=1}^{30} \left(\varphi_1 + \frac{\varphi_2}{1+\varphi_3 * (i-1)}\right) + (222 * \varphi_1) * (1 + \rho) \tag{15}$$

$$\sigma = \sqrt{\sum_{i=1}^{30} \left(\varphi_1 + \frac{\varphi_2}{1+\varphi_3 * (i-1)}\right) + (222 * \varphi_1) * (1 + \rho)} \tag{16}$$

After being further scaled by price for compatibility with the Black-Scholes model, σ^2 and σ values can be used to estimate the fair value of options for IPO stocks; estimated prices can be compared with real prices to lend insight into pricing efficiency and, when studied using historical discrepancies, can be used to develop an arbitrage-based options trading strategy for IPO stocks.

Although such further investigation exceeds the scope of this study, some high-level hypotheses can be fairly formulated without further analysis. While options traders are probably forward-thinking enough to recognize that volatility will decline following the IPO period, the drop in volatility they expect may be under- or overestimated. For example, the study's results suggest that stocks in the Technology sector typically exhibit about 7.5 times their long-term volatility levels on IPO day. In the case of an underestimated drop, the market may forecast that IPO volatility for Technology stocks is typically 4 times long-run volatility and that volatility will ultimately drop by 75% by the end of the IPO period; using our results, we forecast that volatility will actually be reduced by about 85%. As a result, associated put or call options could be overpriced because of inaccurate volatility forecasting. Given market nonrecognition of industry-based volatility trends, industry-based arbitrage opportunities exist and can be identified with further study.

Regardless of whether significant arbitrage opportunities do exist, it is important to note that the final values and associated regression coefficients reflect only trends in average industry volatility; no regressions were run using values for individual stocks. Although the regression models do accurately predict the motion of $\bar{\nu}$ values from IPO day, it cannot be assumed that these $\bar{\nu}$ trends can be used to effectively predict trends in individual stocks without further analysis.

Other Applications

Stock price volatility on and following IPO date is also relevant for companies trying to launch their own IPOs and the investment banks helping them to do so. Having access to a model that can help to predict a company's volatility after IPO can help to guide decisions on whether a company should sell equity in the public market in addition to how its shares should be priced. These findings on volatility are most relevant when value preservation is a key objective for the IPO candidate; discovering that its share prices may be more volatile than expected in the public market may lead a more conservative company to seek alternative sources of capital. For example, companies in the Technology and Consumer Services sectors may decide not to issue common stock to raise capital because their industries can be characterized by above-market volatility during both the IPO period and the long run.

When attempting to manage any portfolio, minimizing risk is of utmost importance. In "How Relevant is Volatility Forecasting for Financial Risk Management?" Christoffersen and Diebold observe that the ability to forecast volatility of a security is incredibly useful for risk management purposes (2000, p. 12). The new insight that this study provides into industry-based IPO volatility forecasting can advance risk management proficiency. For example, the values of the φ_3 variables lend insight into the rate of regression from IPO volatility levels to baseline volatility. When deciding whether to invest in the stock of a recent IPO, the regression values can help portfolio managers decide when prices have stabilized enough to buy the stock safely. For example, stocks in the Telecommunication sector exhibit total IPO-time volatility of about 0.16 while their baseline volatility levels are close to 0.03. If a portfolio manager wants to wait until the stock has stabilized and volatility has dropped below 0.05 to evaluate it, this study's regression suggests that the manager should wait until after day 7.

It is surprising that the effects of industries and the IPO period on stock volatility have not been studied previously. Although volatility trends were successfully modeled during this study, much work is left to be done to fully understand the impact of these factors on volatility and the applications of a deeper understanding; further study into these relationships will certainly provide valuable insight for both academics and profit-seeking arbitrageurs.

Variable Name	Value
`counter`	*Day from IPO*
`total`	*Sum total of all values for a given* `counter` *value*
`permno`	*CRSP-assigned permanent identification number for company*
`denominator`	*Number of companies for a given* `counter` *value*
`vbar`	*Calculated value of industry volatility rating for a given* `counter` *value*

Table 2. *STATA variables.*

Appendix I: Description Of Nonlinear Regression Process For Stata

To calculate the values for a given industry data set pulled from CRSP, the value of all volatility ratings were added into one total sum using:

```
by counter:egen total=sum(v)
```

The number of individual companies in the industry was then determined using:

```
by permno counter:generate
denominator=_N
```

To calculate the average volatility ratings for the data, set, a simple average of these values was taken:

```
generate vbar = total/denominator
```

After using these commands, all stocks have the same "vbar" rating on any given counter day so duplicates can be dropped by counter:

```
duplicates drop counter, force
```

The leftover observations represent industry volatility with vbar and use counter to show how much time has passed from IPO. Before running the nonlinear regression, all other variables can be dropped.

Nonlinear regression models can be specified using Stata's nl command. Representing using C1, using C2, and using C3, the regression is generally run using:

```
nl (vbar = {C1} + {C2}/
(1+{C3}*(counter-1)))
```

Because the regression requires the estimation of three coefficients with only two given variables, however, initial value estimates for all three variables must be passed in order generate the desired results. After experimentation, it was discovered that 0.05, 0.2, and 1.75 can be used as initial values for C1, C2, and C3 respectively in every industry to produce the regression results with the best fit:

```
     nl (vbar = {C1} + {C2}/
(1+{C3}*(counter-1))), initial
(C1 0.05 C2 0.2 C3 1.75)
```

Appendix II: Ordered Regression Results

For all regression tables, relevant sample sizes and t-statistics are listed parenthetically.

Sector	φ_1 (Baseline Volatility)	φ_2 (IPO-Driven Volatility)	φ_3 (Stabilization Rate)	R^2
Technology (890)	0.1173751 (26.78)	0.7072644 (40.15)	2.997637 (8.57)	.9837
Consumer Services (798)	0.0598604 (24.60)	0.4793099 (47.82)	3.806189 (8.87)	.9884
Healthcare (821)	0.0575546 (25.57)	0.2345721 (25.97)	2.959527 (5.58)	.9619
Oil & Gas (578)	0.049487 (17.32)	0.1118282 (10.66)	1.686358 (2.93)	.8142
Market (S&P 500)	*0.0470464 (20.04)*	*0.1760342 (19.88)*	*1.936521 (5.20)*	*.9379*
Industrials (1141)	0.0412802 (47.95)	0.1061883 (31.54)	2.424831 (7.50)	.9740
Financials (2058)	0.0395426 (69.25)	0.0907731 (42.87)	1.769266 (11.59)	.9860
Consumer Goods (653)	0.0390784 (31.01)	0.1378587 (30.74)	1.463953 (8.81)	.9929
Basic Materials (548)	0.0303853 (47.53)	0.0632364 (27.65)	1.496436 (7.88)	.9675
Telecommunications (143)	0.0297135 (12.20)	0.1283577 (16.27)	1.045903 (4.98)	.9137
Utilities (206)	0.0276752 (33.32)	0.0460092 (13.93)	2.732502 (3.12)	.8794

Table 3. *Regression results ordered by* φ_1 *values.*

Sector	φ_1 (Baseline Volatility)	φ_2 (IPO-Driven Volatility)	φ_3 (Stabilization Rate)	R^2
Technology (890)	0.1173751 (26.78)	0.7072644 (40.15)	2.997637 (8.57)	.9837
Consumer Services (798)	0.0598604 (24.60)	0.4793099 (47.82)	3.806189 (8.87)	.9884
Healthcare (821)	0.0575546 (25.57)	0.2345721 (25.97)	2.959527 (5.58)	.9619
Market (S&P 500)	*0.0470464 (20.04)*	*0.1760342 (19.88)*	*1.936521 (5.20)*	*.9379*
Consumer Goods (653)	0.0390784 (31.01)	0.1378587 (30.74)	1.463953 (8.81)	.9929
Telecommunications (143)	0.0297135 (12.20)	0.1283577 (16.27)	1.045903 (4.98)	.9137
Oil & Gas (578)	0.049487 (17.32)	0.1118282 (10.66)	1.686358 (2.93)	.8142
Industrials (1141)	0.0412802 (47.95)	0.1061883 (31.54)	2.424831 (7.50)	.9740
Financials (2058)	0.0395426 (69.25)	0.0907731 (42.87)	1.769266 (11.59)	.9860
Basic Materials (548)	0.0303853 (47.53)	0.0632364 (27.65)	1.496436 (7.88)	.9675
Utilities (206)	0.0276752 (33.32)	0.0460092 (13.93)	2.732502 (3.12)	.8794

Table 4. *Regression results ordered by* φ_2 *values.*

Sector	φ_1 (Baseline Volatility)	φ_2 (IPO-Driven Volatility)	φ_3 (Stabilization Rate)	R^2
Consumer Services (798)	0.0598604 (24.60)	0.4793099 (47.82)	3.806189 (8.87)	.9884
Technology (890)	0.1173751 (26.78)	0.7072644 (40.15)	2.997637 (8.57)	.9837
Healthcare (821)	0.0575546 (25.57)	0.2345721 (25.97)	2.959527 (5.58)	.9619
Utilities (206)	0.0276752 (33.32)	0.0460092 (13.93)	2.732502 (3.12)	.8794
Industrials (1141)	0.0412802 (47.95)	0.1061883 (31.54)	2.424831 (7.50)	.9740
Market (S&P 500)	*0.0470464 (20.04)*	*0.1760342 (19.88)*	*1.936521 (5.20)*	*.9379*
Financials (2058)	0.0395426 (69.25)	0.0907731 (42.87)	1.769266 (11.59)	.9860
Oil & Gas (578)	0.049487 (17.32)	0.1118282 (10.66)	1.686358 (2.93)	.8142
Basic Materials (548)	0.0303853 (47.53)	0.0632364 (27.65)	1.496436 (7.88)	.9675
Consumer Goods (653)	0.0390784 (31.01)	0.1378587 (30.74)	1.463953 (8.81)	.9929
Telecommunications (143)	0.0297135 (12.20)	0.1283577 (16.27)	1.045903 (4.98)	.9137

Table 5. Regression results ordered by φ_3 values.

Sector	φ_1 (Baseline Volatility)	φ_2 (IPO-Driven Volatility)	φ_3 (Stabilization Rate)	R^2
Consumer Services (798)	0.0598604 (24.60)	0.4793099 (47.82)	3.806189 (8.87)	.9884
Financials (2058)	0.0395426 (69.25)	0.0907731 (42.87)	1.769266 (11.59)	.9860
Technology (890)	0.1173751 (26.78)	0.7072644 (40.15)	2.997637 (8.57)	.9837
Industrials (1141)	0.0412802 (47.95)	0.1061883 (31.54)	2.424831 (7.50)	.9740
Consumer Goods (653)	0.0390784 (31.01)	0.1378587 (30.74)	1.463953 (8.81)	.9929
Basic Materials (548)	0.0303853 (47.53)	0.0632364 (27.65)	1.496436 (7.88)	.9675
Healthcare (821)	0.0575546 (25.57)	0.2345721 (25.97)	2.959527 (5.58)	.9619
Market (S&P 500)	*0.0470464 (20.04)*	*0.1760342 (19.88)*	*1.936521 (5.20)*	*.9379*
Telecommunications (143)	0.0297135 (12.20)	0.1283577 (16.27)	1.045903 (4.98)	.9137
Utilities (206)	0.0276752 (33.32)	0.0460092 (13.93)	2.732502 (3.12)	.8794
Oil & Gas (578)	0.049487 (17.32)	0.1118282 (10.66)	1.686358 (2.93)	.8142

Table 5. Regression results ordered by R^2 values.

Appendix III: Correlation Between R2 And Industry Sample Size

```
      Source |       SS       df       MS              Number of obs =      11
-------------+------------------------------           F(  1,     9) =    3.19
       Model |  .007842007     1  .007842007           Prob > F      =  0.1075
    Residual |  .022093514     9  .002454835           R-squared     =  0.2620
-------------+------------------------------           Adj R-squared =  0.1800
       Total |  .029935521    10  .002993552           Root MSE      =  .04955

------------------------------------------------------------------------------
          r2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        size |   .0000539   .0000302     1.79   0.108    -.0000143    .0001222
       _cons |   .9027976    .027312    33.05   0.000     .8410135    .9645817
------------------------------------------------------------------------------
```

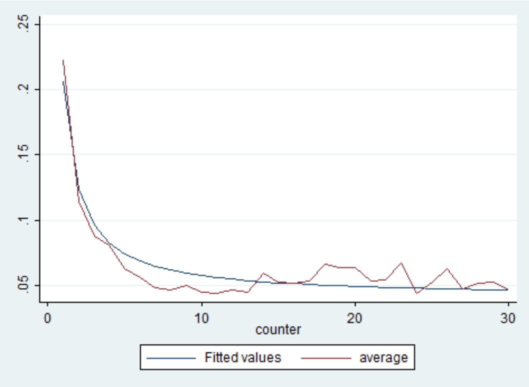
The results of a linear regression of sample sizes on R^2 values show that sample size and quality of regression fit are not significantly correlated.

Appendix Iv: Correlation Between Long-Run Volatility And IPO Volatility

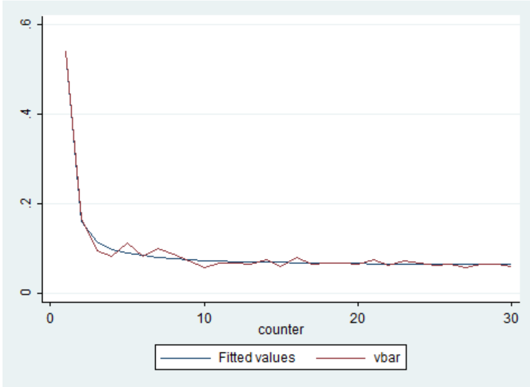
```
      Source |       SS       df       MS              Number of obs =      11
-------------+------------------------------           F(  1,     9) =   57.66
       Model |  .360193636     1  .360193636           Prob > F      =  0.0000
    Residual |  .056221351     9  .006246817           R-squared     =  0.8650
-------------+------------------------------           Adj R-squared =  0.8500
       Total |  .416414988    10  .041641499           Root MSE      =  .07904

------------------------------------------------------------------------------
          c2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          c1 |   7.566707   .9964793     7.59   0.000     5.312514      9.8209
       _cons |  -.1633648   .0543324    -3.01   0.015    -.2862731   -.0404565
------------------------------------------------------------------------------
```

Appendix V: Calculated And Regression Results As A Function Of Day From IPO

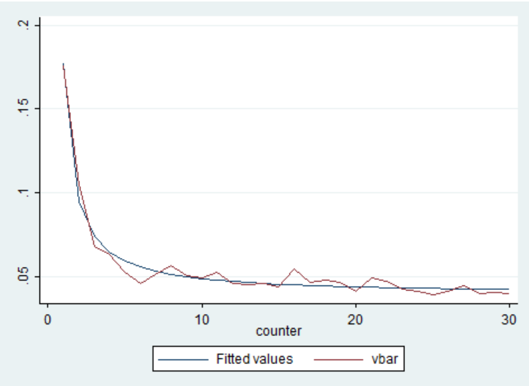
The following graphs were generated by Stata and show both observed values and estimated regression models as a function of day from IPO . Observed values are represented in red and are labeled "vbar" for each industry graph and "average" for the S&P 500 graph. Regression estimates are represented in blue and are labeled "Fitted values" for all graphs. All graphs were generated using Stata version 11.1. For all graphs, day relative to IPO is measured on the x-axis and average industry volatility is measured on the y-axis.



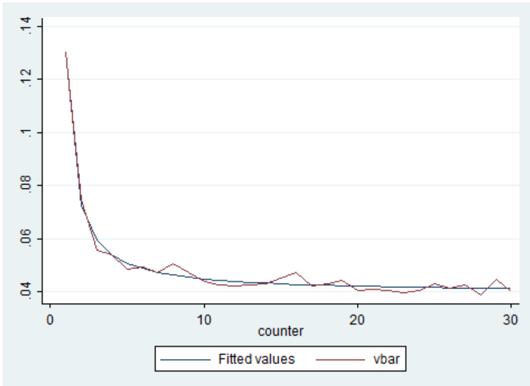
S&P 500



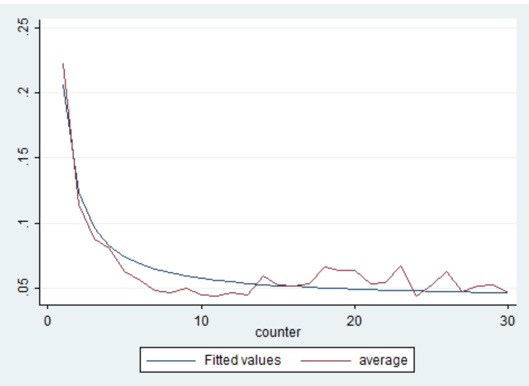
Consumer Services



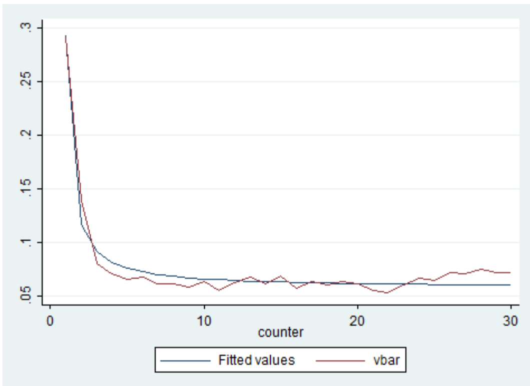
Base Materials



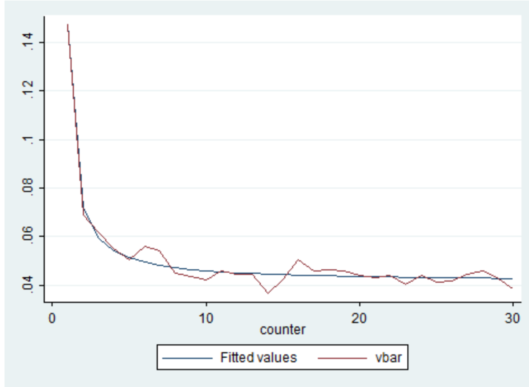
Financial Services



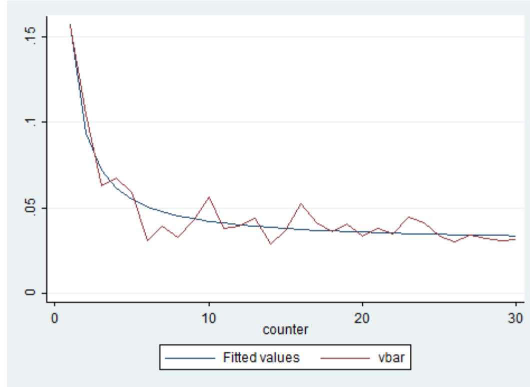
Consumer Goods



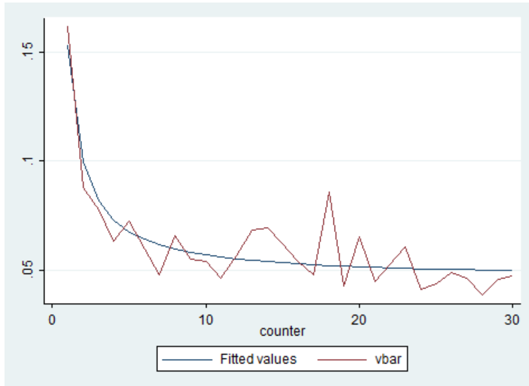
Healthcare



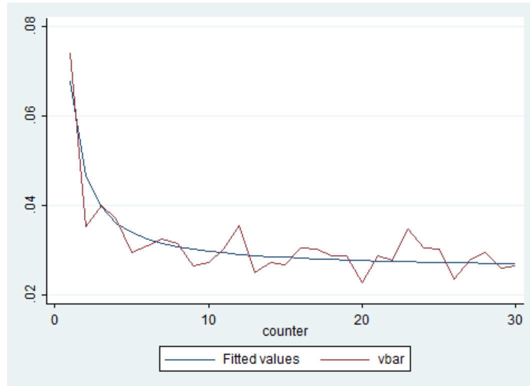
Industrials



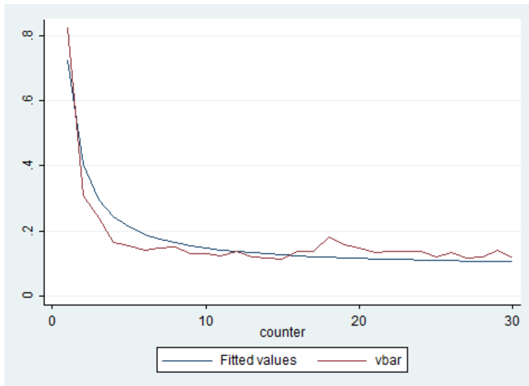
Telecommunications



Oil & Gas



Utilities



Technology

Works Cited

Baillie, Richard T. and Ramon P. DeGennaro. "Stock Returns and Volatility." The Journal of Finance and Quantitative Analysis, Volume 25, Number 2, pages 203-214.

Christoffersen, Peter F. and Francis X. Diebold. "How Relevant is Volatility Forecasting for Financial Risk Management?" The Review of Economics and Statistics, Volume 82, Number 1, February 2000, pages 12-22.

Draho, Jason. The IPO Decision: Why and How Companies Go Public. Edward Elgar Publishing, Inc.: Northampton, MA (2004).

Lowry, Michelle, Micah S. Officer, and G. William Schwert. "The Variability of Initial IPO Returns." The Journal of Finance, Volume 65, Issue 2, April 2010, pages 425-465.

Schwert, G. William. "Stock volatility in the new millennium: how wacky is Nasdaq?" Journal of Monetary Economics, Volume 49, Issue 1, January 2002, pages 3-26.

Schwert, G. William. "Why Does Stock Market Volatility Change Over Time?" The Journal of Finance, Volume 44, Number 5, December 1989, pages 1115-1153.

Simon, David P. "The Nasdaq Volatility Index During and After the Bubble." The Journal of Derivatives, Volume 11, Number 2, Winter 2003, pages 9-24.

Sinclair, Euan. Volatility Trading. John Wiley & Sons, Inc: Hoboken, NJ (2008).

Should Punitive Damages be Scaled to Wealth?

Harrison A Seitz
Northeastern University

Introduction

In cases in which society, and by proxy, courts, want to deter future problematic behavior, punitive damages serve a fundamental role. As an additional source of damages that can be levied upon an injuring party, punitive damages can excise gains that were obtained through illegitimate or harmful means and also prevent future similar actions from being taken. However, there are many questions that must be resolved in order for punitive damages to have any merit in our justice system. How large do punitive damages need to be in order to provide a socially optimal level of deterrence? How large ought punitive damages be in relation to compensatory damages awarded in the same case? Should the injurer be liable for harms that may have occurred to a third party?

The discussion of punitive damages, especially in their relation to tort reform, has been a popular topic in the public discourse in the past couple of years. There are many scholars, supported by legislation and decisions at the state and federal levels, who advocate strongly for restricting punitive damages or eliminating them entirely. On the other hand, there are those who would suggest that the extent of punitive damages ought to be modified or increased. This paper will discuss one proposed modification of punitive damages, which is the inclusion of the wealth of the injurer as a factor in considering the magnitude of the damages.

Literature Review

There have been several opinions published in the literature regarding this issue. The first piece of relevant literature is a response paper by A. Mitchell Polinsky, titled, "Are Punitive Damages Really Insignificant, Predictable, and Rational? A Comment on Eisenberg et al." (1997). In this paper, Polinsky rebuffs three assertions made by separate study of punitive damages. First, Polinsky demonstrates that punitive damages, although seemingly insignificant on a trial level because the amount awarded over a large number of cases is relatively small, still have significance in the pretrial settlement process and are therefore still a part of an injurer's decision calculus. Second, Polinsky shows that punitive damages are typically unpredictable not only in when they are awarded, but also in determining their magnitude. For example, he notes that the injurer's conduct as well as the type of misconduct may have varying impacts on whether or not punitive damages are awarded. Finally, Polinsky addresses the argument that there is a positive correlation between punitive and compensatory damages. He argues that there are cases in which there may be positive or negative correlations between the two values, thus assumptions about how punitive damages ought to be constrained based on proportionality are not rational. For example, he addresses cases in which the injurer may have zero chance of escaping liability. In accordance with theory that punitive damages are meant as a deterrent, in these cases there can be no meaningful deterrence from punitive damages, thus they ought to be lower even if the compensatory damages are high.

The second piece of relevant literature on this topic is a paper titled, "A Theory of Wealth and Punitive Damages," by Keith Hylton (2008). In this paper, Hylton argues that wealth should be a consideration in the calculation of punitive damages for some cases. In general, the conclusion that he reaches is, "Where the parameter of interest in the determination of an optimal punitive award, external cost or internal gain, is unobservable and correlated with the defendant's wealth, the optimal punitive award will be a function of the defendant's wealth". Put simply, this tells us that in some cases, an injurer's valuation of committing a harmful act may increase according to their wealth, much like how a normal good's value increases with a consumer's wealth. In order to achieve an optimal amount of punishment, Hylton would argue that punitive damages should be scaled to wealth.

The third piece of related literature that will be discussed is a paper from A. Mitchell Polinsky and Steven Shavell titled, "Punitive Damages: An Economic Analysis" (1998). This article is an extensive look at punitive damages with attempts to explain a formula-based approach to determining optimal awards. Additionally, this piece covers the issue of wealth in the determination of punitive damages. Polinsky and Shavell argue that wealth should never be a consideration in determining the scale of punitive damages for several reasons. First, they cite that it discourages corporations from expanding and taking any sort of risks because greater wealth actually increases their liability. Under this point, they also argue that the deciding factors in determining the amount of punitive awards ought to be limited to solely the amount of harm caused and the chance of escaping liability. In contention with these claims, they note some circumstances in which wealth could be a useful determinant of the magnitude of punitive damages. They first demonstrate that injurers who are poorer are more likely to be risk-averse since they cannot afford liability insurance and will take proper prevention measures. In this case, they ought to receive lower levels of punitive damages. The next argument they advance is that wealthier individuals tend to value money less, so in order to offset the utility that they gain from causing harm, they must be punished proportional to their wealth.

The next piece of relevant literature is a paper by Robert Cooter titled, "Punitive Damages, Social Norms, and Economic Analysis" (1997). In this paper, Cooter describes the economic reasoning behind punitive damages and their purpose of either internalizing costs or deterring future behavior. In the discussion of how punitive damages ought to be determined, Cooter argues that the criteria for punishment to achieve deterrence should be determined on the basis of the actual harm caused and the expected liability. In cases in which the harm caused may be unclear, such as when the illicit gain may be satisfaction, Cooter argues that even though there is no clear standard for how to value that satisfaction, courts ought to use aggregate statistical data to form the amounts of punitive damages, much like fines for criminal acts.

The last piece of relevant literature that will be discussed is a study from Theodore Eisenberg, et. al, titled, "The Decision to Award Punitive Damages: An Empirical Study" (2009). Using cross-sectional data from 2005, this study determined a model to explain the relationship between how often punitive damages are requested in civil cases, how often they are awarded, and how proportional they are to compensatory damages. They found that punitive damages are very rarely pursued, were likely to be awarded if they were, and were often proportional with regards to the compensatory damages awarded in the case.

Analysis

Before analyzing the effects of considering the injurer's wealth, it would be useful to first define what situations should deserve punitive damages. According to Cooter, punitive damages are awarded when the injurer's behavior is determined "bad enough", or malicious enough to warrant additional punishment (1997). Punitive damages are considered to be any damages that exceed the combined amount of the harm caused to the victim and gains that the injurer may have made from committing a malicious action. This suggests that the goal of punitive damages is not to internalize costs (as compensatory damages do this already, assuming they are set at the correct amount), but rather to deter future malicious behavior.

A common theory of how punitive damages ought to be determined in order to achieve optimal deterrence is the simple calculation of the cost of the harm multiplied by "punitive damages multiplier", or essentially the chances that the injuring party will actually be held liable for their action (Polinsky and Shavell, 1998). This method of deterrence seeks to reach an optimal level based on a valuation of the social costs of harmful behavior. Proponents of this calculation would argue that because this is the only way to achieve a specific, non-arbitrary level of punitive damages, wealth should not be a consideration. It can be said that any rational actor will only consider taking preventative action based on calculations of expected costs. In this case, only the direct value of the harm that an injuring party causes can be expected. For example, in a case of an automobile accident, the wealth of the driver is irrelevant since they will both presumably take the same measures to prevent an accident from happening and cause the same amount of harm. If the damages were instead scaled to wealth, it would actually cause inefficient over-deterrence since the punishment is based off a more arbitrary and far less contributory factor.

On the other hand, there are situations in which the value of the harmful action committed by the injuring party may be relevant. As a person's wealth increases, their valuation of nominal amounts of money decrease due to diminishing marginal returns. However, we can consider the utility that an injuring party gains from causing harm to be similar to a normal good (Hylton, 2008). As their wealth increases, they would be more willing to pay for that amount of utility. This would tell us, then, that the amount of punitive damages shouldn't be based on the exact monetary harm caused by the injurer, but rather on how much they would value that harm. The justification behind this is because in order to achieve deterrence, the punitive measures must have some sort of impact in meaningfully changing the injurer's behavior. If punitive damages were restricted to proportionality to the harm caused, then wealthier individuals would be able to internalize the punishments in many cases and would not be deterred from committing harmful actions in the future.

Conclusion

Based on the goal of deterring future harmful behavior, it would seem as if considering wealth in punitive damages fulfills the philosophical intent as well as providing an optimal level of deterrence based on the injurer. However, this is problematic for determining legal consistency or a bright line in calculating the magnitude of punitive damages. In addition to muddling the standard for acceptable punishment, scaling punitive damages to the wealth of the injurer may discourage individual and firm behavior in ways that are unjustified. For example, increased liability because of wealth may disincentivize corporations from conducting research and development. In another case, wealthier people may be subject to a disproportionate amount of risk from punitive damages that they could not meaningfully be able to prevent in some cases, causing over-deterrence. Based on the context, it may be useful for wealth to be considered in determining the amounts of punishment, but it should be restricted to some reasonable standard in order to have solvency and to check abuse.

Works Cited

Cooter, R. (1997). Punitive damages, social norms, and economic analysis. Law and Contemporary Problems, 60(3), 73-91.

Eisenberg, T., Heise, M., Waters, N., & Wells, M. (2009). The decision to award punitive damages: an empirical study. Cornell Law School Legal Studies Research Paper Series. Retrieved from http://ssrn.com/abstract=1397587.

Hylton, K. (2008). A theory of wealth and punitive damages. Widener Law Journal, 17. Retrieved from http://www.bu.edu/law/faculty/scholarship/workingpapers/documents/HyltonKWealthPunitives2REV.pdf

Polinsky, A.M. (1997). Are punitive damages really insignificant, predictable, and rational? a comment on eisenberg et al. The Journal of Legal Studies, 26(S2), 663-677.

Polinsky, A.M., & Shavell, S. (1998). Punitive damages: an economic analysis. Harvard Law Review, 111(4), 869-962.

International Collective Action on Climate Change

Evan Richards Sankey
Northeastern University

Climate change is a negative externality which afflicts humanity on a global scale. If it is left unchecked, it is reasonable to expect that warming will negatively affect the diversity of our biosphere, our economy and our way of life. Any remedy for global climate change – one which slows or halts further increases in the global temperature by curbing greenhouse gas emissions – would classify as a pure public good. That is, it would be both non-rival and non-excludable. On a global level, this has huge implications. A welfare improvement brought to the citizens of Bangladesh by climate change mitigation could be equally enjoyed by Americans and Europeans. Additionally, no body anywhere could be prevented from enjoying these benefits. These characteristics, however, seem to present policy-makers with a problem. The guaranteed benefits of mitigation to all may generate a destructive incentive for some to free-ride on its provision by others. This paper will begin with a brief introduction to the scientific community's assessment of climate change. It will follow with an overview of the literature on collective action and its application to the problem at hand. Then there will be a critical examination of current and past international efforts to address climate change. It will conclude by suggesting possible alternatives to these efforts.

Climate Change And Uncertainty

Over the last hundred years, the mean global temperature has risen by approximately .6 degrees Celsius. This change has coincided with a drastic increase in the concentration of greenhouse gases (primarily carbon dioxide) in the Earth's atmosphere. Though there is a vocal minority of climatologists who deny it, the vast mainstream opinion favors the proposition that the industrial revolution caused the increase in greenhouse gases (concentrations are thought to have risen by about 30% since 1750) which in turn is heating the Earth up (Barrett, 2007). The basic warming mechanism is that an excessive level of greenhouse gases in the upper atmosphere absorbs excessive levels of infrared radiation from the sun and increases the overall temperature of the Earth.

Unfortunately, climatologists are far less certain of the future path of warming and the consequences it will have. The most reliable climate models predict that over the next century, the average global temperature will increase by a further 1.4 to 5.8 degrees Celsius (Barrett, 2007). It is widely believed in the scientific community that at some point, these temperature changes could cause catastrophic changes in climate patterns. The shrinking of the ice sheets in Greenland and Antarctica could lead to changes in the undersea currents in the Atlantic Ocean which might lead to drastic cooling in Northern Europe. Melting ice may also make life very difficult for countries at low elevations like the Maldives and Bangladesh. Average sea levels rose by about two tenths of a meter in the last century. Some expect them to rise by half a meter in this century. Another possibility is that changes in rainfall patterns could render large parts of the southwestern United States uninhabitable. Events such as these could entail huge costs to the global economy. William Nordhaus, an economist at Yale University, has made a widely publicized calculation that unabated climate change could potentially lower the gross domestic product of the world by 2% by 2100. An alternate estimate made the British Government that global GDP will take a hit of between 5.3% and 13.8% by 2200 (Barrett, 2007). Complicating this picture even further is the notion that the costs of climate change will not be evenly distributed. Countries at specific latitudes are expected to benefit from increased agricultural opportunities which warming will present. The key point here is that the benefits of climate change mitigation – the benefits of attempting to avoid a variety of disastrous possibilities – are highly uncertain.

A similar pattern holds for the cost side of the equation. Policies which aim to limit greenhouse gas emissions will necessarily target the sources of those emissions: transportation, power generation and industrial machines. An otherwise unnecessary investment in a massive expansion of nuclear power and low emission vehicles seems in the cards. This will undoubtedly be costly, but again it is unclear exactly *how* costly. Drastic action on climate change such as a commitment to cut carbon dioxide emissions by 50% has been calculated to lower global output by between 1% and 3% with poorer countries incurring less cost than wealthy countries (because they use energy more inefficiently). Other estimates are much higher. The models used to generate these figures necessarily involve certain assumptions which may not hold in reality. There are always unintended consequences. A further consideration is that the costs of action will likely be front-loaded. That is, the pain will come early, and the benefits of mitigation will take possibly hundreds of years to become evident. This makes a direct comparison of costs and benefits inherently difficult. What should be clear from this overview is that there is room for reasonable international policymakers to disagree about the net payoffs and worthwhile goals associated climate change mitigation. Awareness of this capacity for disagreement is critical for understanding the prospects for effective collective action.

Olson And The Collective Action Problem

The modern conception of collective action in the context of political economy is based on the work of Mancur Olson. In *The Logic of Collective Action* published in 1965 Olson mounted a massive critique of the prevailing wisdom at the time on group behavior. Groups had been conceived as perfect embodiments of the common wills of their individual members. Effective collective action was taken as given (Sandler, 2004). Olson's model imagines groups as existing to provide their members with intra-group public goods – that is, goods which no group member can be prevented from benefitting equally from. A classic example is a labor union which wins a wage increase from management by striking. Obviously all union members, *even* those who did not strike will enjoy the wage increase. This tension is the crux of Olson's contribution. There emerges within groups a gap between the level of a public good (in this case a strike) which is optimal for the group as a whole to provide, and the level which is optimal for an individual member of that group to provide. Since all union members benefit from a strike-induced wage hike, each member will try to "free-ride" on every other member's walkout. As a result, this particular public good is undersupplied. Though Olson considers this problem to be endemic to group behavior in general, he highlights some group characteristics which can variously aggravate or remedy the free-rider problem.

Small groups are most likely to supply an optimal level of public goods. Their most important advantage stems from the fact that their members' contributions to the provision of their goal are clear and large. If a member is contemplating free-riding, he will see that the venture will fall apart without his contribution. Contribution becomes a dominant strategy. Smaller groups also find it easier to coordinate solutions to the the problem of verification and enforcement and arbitrate conflicts of interest between its members. Finally, members of small groups will tend to engage with one another repeatedly over time. The prospect of repetition can be a powerful incentive to cooperate with your peers (Shepsle, 2010). Small groups' tendency to achieve their stated goals makes them what Olson calls "privileged" groups.

Large groups, on the other hand, are seen as being much less effective at providing an optimum level of public goods to their members. In general, as groups get larger each individual contribution will count for a smaller percentage of the collective outcome. Therefore, holding costs constant, the probability that the individual will find contribution worth his while declines. Then too, a large group will find it difficult to verify and enforce compliance and reconcile more frequent conflicts of interest between its members. Free riders in a large group are anonymous and probably only identified at a high cost. Furthermore, it is difficult to prevent them from using a good which is by definition non-excludable. Overall, large groups' size makes them unlikely to be able to surmount these various coordination problems and therefore unlikely to supply the optimal level of public goods to their members. As a result, Olson labels them as "latent" groups.

Another important determinant of the effectiveness of groups at supplying public goods is the nature of their membership. Some groups are highly homogenous. Their members are more likely to look alike, talk alike and think alike. They are less likely to get caught up in conflicts of interest because their members' preferences are more homogenous. Unsurprisingly, small groups tend to be more homogenous (that is often why they form in the first place). This status further augments their capacity to provide public goods. Large groups, however, are usually highly heterogeneous. Their members have widely divergent priorities, beliefs and levels of knowledge. This variety massively increases the probability that the aforementioned problem of anonymity and enforcement as well as the scope for deep conflicts of interest will make cooperation impossible. Surprisingly, the heterogeneity of large groups may sometimes be a blessing in disguise. As heterogeneous groups grow, the probability that a small group of large and highly capable group members will choose to supply the public good grows. The reason for this is that large and capable members usually have more to gain overall from the provision of a public good than the small, less capable members. They may decide that their potential gains are so large that they will supply the entire public good themselves (thereby making the group privileged), despite the inevitability that smaller members will free-ride (Sandler, 2004). This notion will prove invaluable for understanding the role of rich countries in climate change mitigation.

In the absence of a government which can force cooperation and public good provision through outright coercion, some groups (especially large ones) will remain perpetually latent with their members worse off as a result. Olson refuses to accept this sorry state of affairs and proposes a remedy. In large groups, the core impediment to optimal public good provision is, as already noted, the fact that individual members notice that their individual contributions seem to count for less while the costs they incur to contribute remain fixed. Therefore, as the group grows, members' willingness to work for the group declines. Olson's solution is to attach *private* benefits to contribution towards the public good. If those who contribute (and only those who contribute) are guaranteed a wholly private benefit like a cash payment, members' overall willingness to contribute would be greatly enhanced (Shepsle, 2010). It would be a challenge, of course, to determine the optimal level of private payments to confer on contributors. Too low and the public good will remain undersupplied, but too high and the group may exhaust its collective resources making excessive payments to its members. The payments must come from somewhere!

International Collective Action On Climate Change

The first, most obvious observation to make about the international community is that it is mostly comprised of sovereign nation states which are fiercely protective of their rights to self-determination. There is no world government with unlimited coercive powers which can be expected to save the day simply by mandating global cuts in carbon emissions. There are (depending on the definition) about 193 sovereign states in the world. On their own, many of them have succeeded in providing public good solutions to various regional externalities like excessive water and air pollution. However, at a global level, the community of nations is not so different from a group of self-interested individuals. Since there is no known other civilization to compare it to, it is unclear exactly how large and heterogeneous we can consider our global community to be. However, in order to proceed, this analysis will consider it to be large and heterogeneous. This seems reasonable.

There are many more countries now than there were at the close of World War II and their interests seem to be as divergent as ever. The weaknesses of large heterogeneous groups correspondingly apply. The countries of the Earth will attempt to free-ride on the provision of climate change mitigation by others and there will be little to keep them from doing so because global verification and enforcement is quite difficult (as evidenced by our nuclear proliferation problems). Furthermore, as already noted there is scope (even if just based on the uncertainty in the data) for reasonable countries to come to quite different conclusions about the costs and benefits of climate change mitigation and therefore about the emissions targets which will play a huge part in any mitigation agreement. At a glance then, it would seem that cooperation on supplying the public good of climate change mitigation is not forthcoming and that the world is not a privileged group.

On closer examination, however, our prospects are not so terrible. The heterogeneity of the world is also a major strength. There are severe inequities in wealth, population size and carbon emissions amongst those 193 countries. The United States alone comprises approximately 20% of global output and 16% of global carbon dioxide emissions. Similarly, a relatively small number of countries have most of the Earth's population and wealth and generate most of its carbon emissions. In fact, a formal organization of states which fits this description already exists. It is called the Group of Twenty or G20. It includes (among others) China, India, the United States, the United Kingdom, Russia, Germany, and Japan – countries which are all major players on the world stage. Precisely because of this, however, these nations have significantly more in common with each other than with the majority of small and relatively less influential states in the world. Each of these countries would sustain big loses to their economies (if only because they *have* more to lose) were some or all of the disastrous climate scenarios predicted by climate models to come true. None would gain from climate change, as some small agriculturally dependent countries would. This is not to say, however, that their interests are perfectly aligned. As we will see, there is considerable disagreement between the developed and developing wings within this group of major players over what to do about climate change. As a result, it is unclear whether these commonalities are enough to encourage them to provide climate change mitigation for the world.

There are two other factors working in favor for these major countries (like those in the G20) possibly supplying the public good of climate change mitigation. The first concerns the nature of mitigation itself. This small, but weighty group of countries might be thought of as a "k-group" – a grouping of actors whose full participation is necessary to supply the public good (Shepsle, 2010). The smaller countries' free-ridership can be tolerated because they do not matter very much to mitigation. However, if even one large country free-rides, the goal (the emissions target) cannot be reached (McGinnis & Somin). This fact serves as an incentive for each nation to contribute lest the public good not be provided at all. However, as noted by Kenneth Shepsle in *Analyzing Politics*, this will be no consolation if even some of the countries involved believe that the costs of contribution outweigh the benefits. The dominant strategy will turn to defection. Secondly, these major players, the G20 members, tend to interact frequently. This is not a one-shot game. Life will go on and there will be future issues (like this one) which many of these countries will have to cooperate with each other on in order to solve. In contemplating torpedoing the provision of climate change mitigation, a large country may find the prospect of future reprisals by its peers too costly to be engaging in intransigence. Given all of these favorable factors, it seems odd that we don't already have an effective mitigation agreement.

Montreal

Before considering the results of the most important climate change mitigation treaty, it is instructive to examine the Montreal Protocol of 1987 which imposed limits on the consumption and production of chlorofluorocarbons (CFCs) internationally. It is often held up as the prime example of an international environmental agreement which achieved its stated intentions. It is therefore important to understand why it worked so well.

In the 1970s and 1980s a number of scientific studies were published which made it abundantly clear that emissions of CFCs from industrial processes and aerosol cans were drifting into the upper atmosphere, interacting with sunlight and causing serious damage to the earth's ozone layer. The ozone layer is important because it protects the surface of the Earth from harmful ultraviolet rays from the sun. Ultraviolet rays are known to cause damage to plant life and sea life and well as skin cancer in humans (Barrett, 2007). Like carbon dioxide, CFCs drift around in the atmosphere far from where they are emitted. As a result CFC emissions by multiple emitters combine to make everyone worse off. Ozone damage mitigation can therefore be considered a public good, suffering from the same ostensible problems of international collective action as climate change mitigation. Yet, the treaty was successful.

The Montreal Protocol as well as the dozen amendments which followed and strengthened it has led to a sharp decline in the concentration of CFCs in the atmosphere and a recovery in the strength of the ozone layer which is expected to continue. How was the Montreal Protocol so fantastically successful at solving a problem which seems so similar to climate change? There is some disagreement on this question, but it seems the single biggest reason is that there was far less uncertainty in the cost-benefit analyses provided by the scientific community at the time. In addition to being more certain, most studies confirmed that the benefits of acting to limit CFC emissions outweighed the costs by very large margins (Sandler, 2004). Unlike with climate change, no countries benefited from the deterioration of the ozone layer. As a result, more countries' preferences were aligned. There were fewer conflicts of interest. Many countries (like the United States) had started limiting CFC emissions even before the treaty was signed. Another important factor which helped the Montreal Protocol succeed was the extreme concentration of CFC emissions amongst a relatively small number of countries. In 1987, it was estimated that 12 countries accounted for more than 78% of emissions (Sandler, 2004). This made the relevant "k-group" of major countries smaller and easier to coordinate. It also meant that those countries which had the most to lose from CFC emissions were also the ones who produced the most CFCs. This made it relatively easy for the major players to make the world a privileged group. Though most small nations did sign the treaty, their participation was not strictly necessary to achieve the favorable outcome we have today.

Finally, some observers have drawn attention to the role that external incentives (not unlike Olson's private benefit remedy idea) played in compliance with the Montreal Protocol. The treaty provided two types of incentives – one positive, one negative. First, developed countries agreed to make payments to developing countries to help defray the cost of compliance with the treaty. Second, punishments were imposed on countries which judged the optimal level of emissions reductions to be below what the treaty demanded. These punishments came in the form of trade restrictions. Compliant countries were permitted to impose severe restrictions on imports of products containing CFCs from noncompliant countries. Since participation in the treaty was so high, these restrictions proved to be highly costly to noncompliant countries (Barrett).

Kyoto And Copenhagen

The Kyoto Protocol is an international environmental agreement which was signed and ratified by many United Nations member states (under the auspices of the UN) in 1997. It entered force in 2005. It was intended to force binding cuts on carbon dioxide emissions sufficient to keep global temperature increases at or below 2 degrees Celsius. Though Kyoto only entered force recently, it has not yet had any noticeable effect on carbon dioxide emissions and it is not expected to. Earlier, it was noted that understandable disagreement over the costs and benefits of climate change mitigation may be enough to scuttle the whole enterprise (even within a smaller k-group). This is precisely what has happened (Sandler, 2004).

In order to secure an agreement, the drafters of the treaty at Kyoto had to exempt a number of major developing countries (like China and India) from binding emissions cuts. These countries effectively concluded that the costs of the requisite massive cuts exceeded the benefits (especially in the context of rapid economic development). The problem is that those countries are massive and necessary members of any k-group coalition of major powers which might decide to provide climate change mitigation. China now generates more emissions overall than the United States. Since major developing polluter nations whose contributions are necessary to provide the public good of mitigation effectively defected, it became a dominant strategy for other otherwise cooperative nations to defect (McGinnis and Somin, 2007). As a result, the United States never ratified the treaty. Despite, the obvious disincentives, Europe did adopt it – perhaps because of some vague unifying notion of "European values." A further problem is that compliance by countries which had agreed to binding cuts is not enforceable. As an example, Canada has missed its targets by considerable margins and will not face any sanction for doing so. What incentive does it now have to comply anyway?

Similar problems plagued the more recent 2009 Copenhagen Summit on Climate Change. At Copenhagen, imposition of binding targets wasn't even attempted. Most developed countries made incredible promises of further emissions cuts and affirmed their general commitment to some kind of climate change mitigation. Developing nations, meanwhile, skated by with promises to reduce the "carbon intensity" of their economies – to lower the carbon emissions per dollar of output they produce. Even if they kept these promises, the speed of their economic growth means that their emissions would continue to rise. Most observers have concluded that Copenhagen was a bigger failure than Kyoto. What is to be done?

One possibility would be an incentive mechanism analogous to the one proposed by Olson and similar in form to the one implemented in the Montreal Protocol. However, given the wide divergence of opinion on the costs and benefits of climate change mitigation (as opposed to in the Montreal Protocol) and already weak participation levels, it seems unlikely that rich countries would be willing to pay *enough* compensation to developing countries and that potentially noncompliant countries would find noncompliance too costly due to trade restrictions. Trade restrictions tend to be more effective when there is already high compliance (Barrett, 2007).

An alternative solution would be to provide a decision making environment which excluded the interests of small countries. One problem with the United Nations decision-making apparatus is that it demands absolute unanimity for an agreement to enter force (Barrett, 2007). Given the large scope for conflicts of interest over climate change mitigation (especially amongst the smaller countries), the unanimity requirement tends to unnecessarily water-down treaties and prevent the big countries from cooperating more effectively with each other. The appropriate k-group of nations could assemble under the Group of Twenty (a ready-made k-group) framework. As already noted, small groups suffer from fewer coordination problems and so are more likely to be privileged. More substantial cooperation from developing nations could potentially be secured in a small group and enforcement and verification would be easier.

Ultimately, though, the cold reality boils down to the fact that the net benefits of climate change mitigation are not nearly as high as the net benefits of ozone layer protection were (Sandler, 2004). Additionally, there are more big carbon emitters than there were big CFC emitters at the time the Montreal Protocol entered effect (the relevant k-group is bigger). These differences may make the climate change compliance gap unbridgeable – at least until there is more certainty about its costs and benefits and the net benefits are shown to be unambiguously positive.

Works Cited

Barrett, Scott. Why Cooperate? The Incentive to Supply Global Public Goods. New York, New York: Oxford University Press, 2007. Print.

McGinnis, John O. and Ilya Somin. "Should International Law be Part of our Law?" Stanford Law Review. 59.5 (2007): 1175-1247. Print.

Olson, Mancur. The Logic of Collective Action: Public Goods and the Theory of Groups. 2nd ed. Cambridge, Massachusetts: Harvard University Press, 1971. Print.

Sandler, Todd. Global Collective Action. Cambridge, the United Kingdom: Cambridge University Press, 2004. Print.

Shepsle, Kenneth A. Analyzing Politics: Rationality, Behavior and Institutions. 2nd ed. New York, New York: W. W. Norton & Company Inc, 2010. Print

Submissions Process and Guidelines

ECONPress is a peer reviewed journal where anonymized papers are reviewed multiple times and may undergo several rounds of revisions before the paper is accepted for publication. All potential authors should visit our website to submit their papers(s).

Peer Review Process

When an Author submits a paper to ECONPress, it is initially assigned to two or more Reviewers. These Reviewers make a recommendation to the Section Editor who then communicates a decision to the author:

- *Accept Submission*
- *Revisions Required*
- *Resubmit for Review*
- *Decline Submission*

Accept Submission - The paper will be reviewed by a Copy Editor and after the small errors are corrected, the paper will undergo a final proofread before it is put in a queue for publication.

Revisions Required - The paper is suitable for the journal but some revisions are necessary to become ready for publication. The author will receive feedback related to the paper and will be required to upload a revised copy of their paper. The paper will again be reviewed by at least two reviewers, one having read the paper in the previous round. This process may continue for several rounds and if it is ultimately deemed fit for publication, it will be marked *Accept Submission*.

Resubmit for Review - The paper is suitable for the journal but requires significant revisions. The author is suggested to make extensive changes to the paper and complete a new submission.

Decline Submission - This paper is not suitable for this journal. This may be for many reasons, including the paper not relating to economics, the paper is discovered to have been plagiarized, etc.

Submission Information

Authors should submit their papers through our web site, www.ECONPress.org

We welcome all papers related to the study of Economics from students enrolled as an undergraduate at an accredited university. This includes (but is not limited to) interviews with economists, research papers and opinion papers.

Submissions are accepted on an ongoing processes, with the selection process occurring twice per year: once in the Spring and once in the Fall.