

# Reconstructing Cooperation from Civil Conflict: Experimental Evidence from Colombian Development Policy\*

Eric A. Coleman<sup>†</sup>  
Florida State University  
Department of Political Science  
ecoleman@fsu.edu

Maria-Claudia Lopez  
Universidad Javeriana, Colombia  
School of Environmental and Rural Studies  
mlopez@javeriana.edu.co

March 5, 2010

## Abstract

This paper presents an impact assessment of a peacebuilding program in Colombia. The program was started in one region of the country and, through development aid, transmitted to 15 other regions. A common criticism of impact evaluation is that evidence of program impact is localized to the study population. Little evidence exists to confirm if successful programs achieve the same impacts when “scaled-up” to the national level or when transmitted to different regions—fundamentally an issue of external validity. We evaluate program transmission success by comparing outcomes in towns located in the original region to outcomes in towns from five different regions of Colombia where the program has been transmitted. We assess outcomes based on the program goals of increasing cooperation. To measure cooperation we conducted a series of experimental games (a trust game, a public goods game, and a voluntary contribution to local public services) with almost 2000 Colombian citizens from these regions. We show that the metric by which one measures cooperation is important to determining the effects of the program’s expansion. We find that beneficiaries of the peacebuilding program are more likely to exhibit cooperative outcomes in these games. We then investigate the conditional average treatment effects (CAFE) of the program to determine what types of subjects are most likely to benefit from the program in the expansion regions and find that male subjects were the most likely to benefit from the program.

**Keywords:** Community-driven Reconstruction, Colombia, Experiment

---

\*The data used for this paper was collected for the report “Evaluación de Impacto de los programas Paz y Desarrollo y Laboratorios de Paz: Línea de Base e Impactos Preliminares Hacia la consolidación de una propuesta para evaluar el impacto de los Programas Regionales de Desarrollo y Paz” done by the Departamento Nacional de Planeación, Colombia 2008.

<sup>†</sup>543 Bellamy, P.O. Box 3062230, Tallahassee, FL 32303–2230; Tel: 850–644–4540; Fax: 850–644–1367

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Background</b>	<b>5</b>
<b>3</b>	<b>Measuring Cooperation</b>	<b>6</b>
3.1	Experimental protocols . . . . .	10
3.2	The trust game . . . . .	10
3.3	Contribution to a public service . . . . .	11
3.4	The public good game . . . . .	11
<b>4</b>	<b>Estimating Treatment Effects</b>	<b>12</b>
4.1	Average treatment effects . . . . .	13
4.2	Separating treatment effects . . . . .	15
4.3	Conditional average treatment effects . . . . .	15
<b>5</b>	<b>Data Analysis</b>	<b>16</b>
5.1	The trust game . . . . .	16
5.2	Contributions to a public service . . . . .	19
5.3	Public goods contributions . . . . .	21
5.4	Correlations between Measures of Cooperation . . . . .	21
<b>6</b>	<b>Discussion</b>	<b>23</b>
6.1	Comparison to previous research . . . . .	23
6.2	Can the program be successfully transmitted to other regions? . . . . .	24
6.3	Which measure of cooperation? . . . . .	25
<b>7</b>	<b>Conclusion</b>	<b>26</b>
<b>A</b>	<b>Appendix A</b>	<b>30</b>
<b>B</b>	<b>Appendix B</b>	<b>32</b>
<b>C</b>	<b>Appendix C</b>	<b>33</b>

## List of Tables

1	Description of Research Sites . . . . .	7
2	Summary Statistics by Region . . . . .	17
3	Reciprocity Treatment Effects by Groups . . . . .	18
4	Public Service Contribution Treatment Effects by Groups . . . . .	20
5	Public Goods Treatment Effects by Groups . . . . .	22
6	Correlations between Measures of Cooperation . . . . .	23
7	Component means across PCA wealth index quantities . . . . .	32

## List of Figures

1	Geographic location of peace laboratories. . . . .	8
2	Measures of Trust from Various Artefactual Field Experiments. . . . .	23
3	Distribution of Wealth by Region and Beneficiary Status . . . . .	33

# 1 Introduction

The World Bank is beginning to give large amounts of development aid to post-conflict states under the rubric of Community-Driven Reconstruction or CDR (Fearon, Humphreys and Weinstein, 2009). CDR development aid is implemented as block grants to community councils that manage and plan their own reconstruction priorities (Cliffe, Guggenheim and Kostner, 2003). The goal of such aid is to reestablish local institutions and promote reconciliation after civil conflict. The programs are based on the premise that post-conflict areas suffer from “the breakdown of the institutions normally used to prevent localized conflicts from re-igniting, the need for armed populations to see an immediate return from peace, and the opportunity to re-define the social and institutional relationships that led to the conflict in the first place (Cliffe, Guggenheim and Kostner, 2003, p.1).” The program implicitly suggests that if social capital, cooperation, and civil society institutions are reconstructed then attitudes toward political violence will change and future conflict will be less likely.<sup>1</sup>

Recent work in political science suggests that much of the brutality in civil conflict is explained by deeply rooted local polarization which is only manifested in when conflicts emerge (Kalyvas, 2006). It seems a tall order for CDR programs to effectively address this polarization and reconcile this animosity to such a degree that cooperative institutions are successfully reconstructed. Even in non-conflict areas there seems to be a consensus that cooperation depends on the existence of variables that take some time to evolve, such as social norms and social capital (Ostrom, 1990) or are fixed, such as ethnic, religious, and economic heterogeneity (Habyarimana et al., 2006, 2009). Such variables either cannot be changed at all in the short term or at least cannot be easily manipulated. Such problems seem to be even more limiting in post-conflict areas.

Despite the large hurdles that CDR aid must overcome in order to reestablish local cooperation, some evidence now suggests that CDR aid can have immediate effects. Fearon, Humphreys and Weinstein (2009) find that short term CDR aid did, in fact, induce citizens to act cooperatively in post-conflict Liberia. This paper is in much the same spirit as Fearon, Humphreys and Weinstein in that we measure the extent to which internationally funded, yet locally planned, development projects will increase cooperation. However, our primary interest is to investigate to what extent successful CDR projects in one region can be transmitted into other regions of a country which may have very different demographic characteristics (Duflo, 2004) or different experiences with conflict (the intensity and duration of conflict varies across regions). Identifying regions for which such transmission may be successful is beneficial for at least three reasons: first, we can be more confident in the robustness of scientific conclusions that the program has a given impact; second,

---

<sup>1</sup>There are a variety of mechanisms through which this may take place. Such institutions may lower transaction costs and facilitate information sharing (North, 1990) or may facilitate the transmission of participatory norms (Putnam, 1993). See Chapman (2008) for a dissenting view and empirical test of this relationship in relation to opinions on political violence.

it provides important insights to policymakers to target regions where the chance of impact is greatest; and third, it provides estimates about how different groups respond to the program (Duflo, 2004). The fact that results from a policy in one region cannot be generalized to other regions has been a stinging criticism of those opposed to field experiments in international development aid (see, for example many of the responses in Banerjee (2007) and (Duflo, 2004)). As researchers begin to identify the conditions under which programs can be transmitted, policymakers will be able to more effectively target the populations that can benefit most from intervention.

We assess the effectiveness and transmittability of CDR aid from a program in Colombia.<sup>2</sup> Since 1995 a local Colombian NGO, Paz y Desarrollo (Peace and Development), has provided development services in the region of Magdalena Medio—a region hit particularly hard by violent conflict. The World Bank and EU recognized the effectiveness of the program and decided to fund the transmission of the program into 15 other regions of Colombia. Our analysis is to assess the relative degree of cooperation in Magdalena Medio with that in five of the program-transmission regions and a control region. Cooperation is measured by individuals' behavior in a trust game, public goods game, and actual charitable contributions from their earnings.

A number of difficult issues arise from assessing the effectiveness of these programs. First, there are concerns that the types of regions which are selected to participate are somehow fundamentally different from those that are not. We will refer to this problem as the policy-selection problem. This is really a problem of randomization: if peace and development and peace laboratories were randomly assigned to regions than any difference among the regions in their ability to reconstruct cooperative institutions would also be random. Another randomization issue emerges in the types of citizens that participate in the programs within any given region. We will refer to this issue as the self-selection problem. It is also fundamentally an issue that could, in principle, be solved by randomly selecting individuals to participate in the program. Much of the methodological portion of the paper will describe our method for establishing credible comparison groups to provide counterfactuals of behavior for program participants.

In the next section we provide background to the programs in Colombia. In the following section we describe our method for establishing the comparison groups and in section four we explain the statistical model we used to measure program effects. In section five we discuss our estimation results and in section six we conclude.

---

<sup>2</sup>There is disagreement about the present state of civil conflict in Colombia; nonetheless, the EU and World Bank have been giving CDR aid since 2002. Although the conflict has not entirely subsided, reconstruction efforts have been underway since that time.

## 2 Background

Colombia has been embroiled in civil conflict since 1964. While the extent and intensity of the conflict has ebbed and flowed, its effects have been widespread throughout the country. In 1995 state oil workers in the region of Magdalena Medio founded an NGO called the Peace and Development Programme, Magdalena Medio (PDPMM or Peace and Development). The PDPMM sought to establish peace and respect for human rights in the midst of the civil conflict, by financing local development projects. The Peace and Development program funded regional efforts to address the core causes that gave rise to the conflict to begin with: strengthening local democratic institutions, instilling values of human rights, and providing basic economic goods and human services.

Until 2002, 80 percent of the money used to fund PDPMM came from the Colombian government, World Bank Learning Loans administered through the Colombian government, and the state oil company, while the remaining 20 percent came from various international donors. In 2002 the EU designated the region of Magdalena Medio a “Peace Laboratory” and gave substantial assistance to the program. In 2004 the EU funded the transmission of the PDPMM model to be implemented in other regions in Colombia.<sup>3</sup> At this same time the World Bank gave a loan to the Colombian government both to bolster the Magdalena Medio program and to expand it into the other regions. The expansion of program funding took place in waves: the first expansion of the Peace Laboratories was in 2002 to Magdalena Medio; the second expansion in 2004; and the third expansion in 2006.

The Peace and Development and the Peace Laboratories programs start with the premise that development and peace are not only interconnected but need to be achieved on multidimensional levels in order to be sustainable over time. In a macro sense, then, the Peace and Development and the Peace Laboratory programs are meant to attain peaceful resolutions to conflicts by addressing the underlying social, economic, political, and cultural challenges. On a micro level, specific projects are implemented to achieve these objectives. Broadly, programs address respect for human rights, sustainable development, and the establishment and strengthening of responsive democratic political institutions. Public forums (from civic education to art activities) are used to educate individuals on issues regarding human rights and ethnic conflict. Extension services are provided to farmers in an effort to increase crop yields and advise on environmental impacts of different farming techniques. New political institutions have been funded and developed to deal with local public administration and policy and to facilitate political responsiveness to local needs (Henriques, 2007, p.10). The programs work as block-grants to local communities; communities apply for funds to implement

---

<sup>3</sup>Funding from the World Bank loans, as administered through the Colombian government, were targeted to displaced families of the conflict, while the EU funds did not have such an emphasis. Whether funded through the EU or World Bank, both programs use the blueprint of the Magdalena Medio model; that is, to fund locally-driven development projects in the areas of human rights, strengthening local democracy, and sustainable development.

a specific project that tie program objectives.

In 2006, the Colombian government and the donor agencies funded an impact evaluation of the programs operating in six different regions. The primary goal was to assess if beneficiaries of the program would act cooperatively with others in the region. To measure cooperation it was decided to have citizens from program regions participate in a series of experimental games (Habyarimana et al., 2006; Carpenter, Daniere and Takahashi, 2004; Karlan, 2005). Behavior among beneficiaries of the program would then be compared to non-beneficiaries in each region.

Participants played two separate experimental games and were also given an opportunity to donate some of their earnings from those games to a local public service. We report the findings from a series of trust game and public good game experiments conducted in 2006 in 42 rural and urban municipalities in six regions of Colombia with a total of 1993 participants. Among them, 957 were direct beneficiaries of at least one of the projects of the programs and the remaining 1036 were non-beneficiaries but living in a municipality where one of the programs was implemented. The program evaluation took place in the following regions (see Figure 1 and Table 1): Alto Patia y Maciso Colombiano, Magdalena Medio, Montes de Maria, Norte de Santander, Oriente Antioqueo, and Meta (a control region where the program is yet to take effect).

This evaluation was conducted in most regions several years after the laboratories started. In the one region, Meta, the program was just beginning (it started in 2006); program beneficiaries had been identified but had not yet received benefits. While there is regional variation in the types of programs that were implemented, all of the laboratories' basic premise was that programs from local and regional people would be funded in the areas of human rights, political capacity building ("governability"), and sustainable development.

### 3 Measuring Cooperation

Cooperation is a two-edge sword. On the one hand, cooperative behavior is necessary to provide local public goods. Citizen-groups that are able to foster group cooperation may provide a means whereby citizen demands are articulated in the democratic process, and may increase community, regional, and national peace. Fearon and Laitin (1996) argue that ethnic groups regularly exposed to other ethnic groups are likely to develop mechanisms to prevent interethnic violence from spiralling into larger conflict. On the other hand, cooperative individuals can also cooperate to produce public goods for which only they benefit, such as looting and violence (Colletta and Cullen, 2000). Cooperative gangs, for example, are able to coordinate violence and disrupt peace. Such groups may be mobilized in order to destroy peaceful progress (Berman,

Table 1: Description of Research Sites

Region Name	Region Type	Programs	No. of Experiments	No. of Subjects	Description of Implementation
Alto Patía y Macizo Colombiano	Expansion	Peace and Development Second Peace Laboratory	7	328	Part of the second phase of the peace laboratory expansion (2004)
Magdalena Medio	Original	First Peace Laboratory	9	480	The first region of Peace and Development (1995) and the first peace laboratory (2002)
Montes de María	Expansion	Peace and Development Third Peace Laboratory	7	248	Part of the third phase of the peace laboratory expansion (2006)
Norte de Santander	Expansion	Peace and Development Second Peace Laboratory	6	270	Part of the second phase of the peace laboratory expansion (2004)
Oriente Antioqueño	Expansion	Second Peace Laboratory	8	322	Part of the second phase of the peace laboratory expansion (2004)
Meta	Control	Third Piece Laboratory	5	174	Part of the second phase of the peace laboratory expansion (2004)

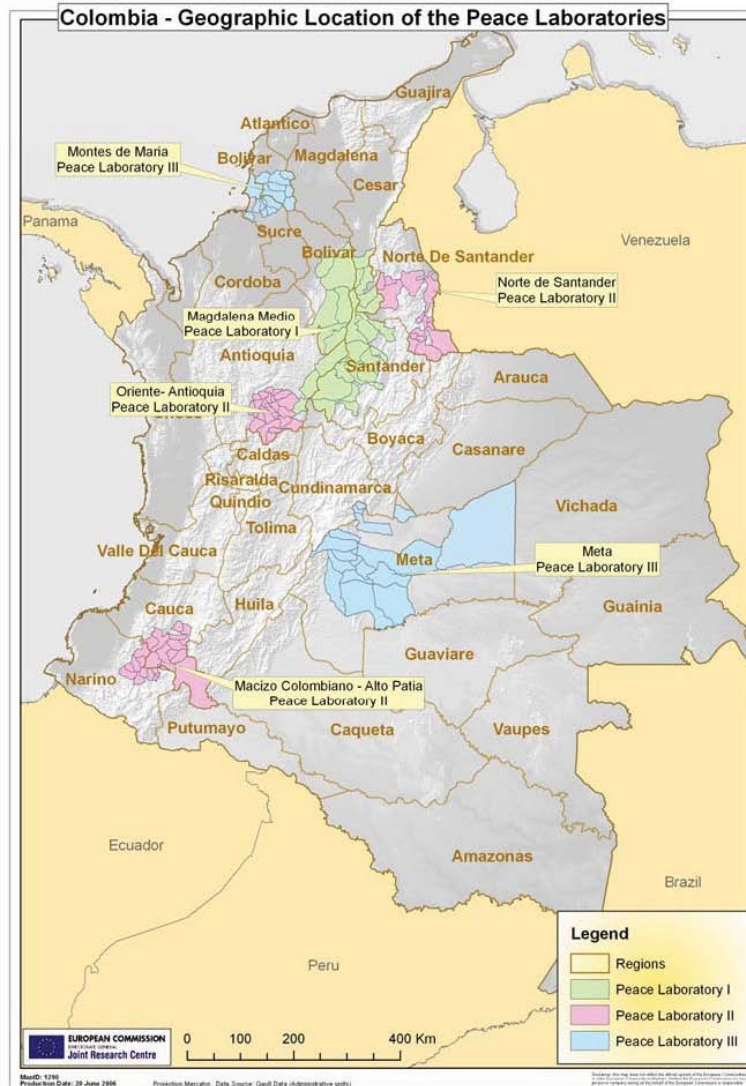


Figure 1: Geographic location of peace laboratories. *Source:* Henriques (2007).

1997).

It is important, then, to differentiate between insular and inclusive cooperation. Insular groups, those willing to act cooperatively with a large group of people, are more likely to engage in activities that benefit society as a whole (Yamagishi, 1998; Glaeser et al., 2000; Rothstein, 2005). Yamagishi's (1998) insular "collectivist societies" that rely only on a small group, are less likely to provide public goods, such as peaceful interactions, for which society as a whole benefits.

In order to distinguish between the insular and inclusive of cooperation, three different experiments were conducted. The first experiment, the trust game, is played when two individuals are paired together. It is our most insular measure of cooperation, behavior in this game affects only two individuals. The second activity,



for which individuals can take their earnings in the trust game and donate some portion to a charitable contribution, is our most inclusive measure of cooperation. In this activity, the society as a whole benefits from the cooperativeness of the subjects. In the third experiment, subjects play a binary, n-person public goods game. This measure falls somewhere in-between the trust game and the charitable contributions in terms of inclusiveness as cooperation benefits the group of subjects participating in the experiment but not society more broadly.<sup>4</sup> The measures of cooperation that are more inclusive are likely to be the types of cooperation most plausibly correlated with the ability to establish peace. Thus, we examine and compare cooperation in all three tasks but put special emphasis on the charitable contributions as the most important measure for cooperation in this program.

Before proceeding it is important to emphasize that none of these measures are intended to be direct measures of peace or conflict resolution. Measuring such variables at the individual level is extremely difficult. Subjects would be highly hesitant to confide peaceful and non-peaceful behaviors to researchers. Another alternative is to measure attitudes towards conflict such as in (Chapman, 2008). These measures, however, may suffer from a bias on the part of respondents to honestly reveal their attitudes. We prefer the behavioral measures because they represent costly decisions; subjects who act cooperatively in the experimental games do so at a cost to themselves. While the willingness to incur such costs has not been firmly empirically related to conflict resolution, there is a plausible causal connection which is implicitly endorsed by the CDR programs. Thus, while these measures are imperfect measures of the potential for conflict resolution we think they provide strong indicators of intermediate program outcomes that are plausibly related to the long-term outcomes of peaceful reconstruction.

With these caveats in mind it is also important to note that behavior in laboratory environments has been shown to be positively correlated with “real life” behavior in a number of realms. Karlan (2005) finds that behavior in trust and public goods games is correlated with the propensity to repay microfinance loans. Falk (2007) finds that behavior in a public goods games was highly indicative of general willingness engage in local public goods provision. Bouma, Bulte and van Soest (2008) find that measures in the trust game are correlated with investments in community projects dealing with natural resource management in India. In this paper we use both strictly laboratory measures from the trust and public goods game, but also supplement these measures with a more general willingness to contribute to public goods in the community.

---

<sup>4</sup>It is important to note that play in each game is anonymous, unconditional on ethnic group, economic background, political persuasion, involvement in the conflict, or any other criteria that subjects might use to exhibit insular cooperation. We strongly feel that conditions of anonymity are essential to field experiments and that despite the limitations in distinguishing the types of individuals targeted for cooperation, the scale effects from cooperation are theoretically relevant in this case. NEED TO FLESH THIS OUT MORE.

### 3.1 Experimental protocols

One laboratory session, wherein a variety of experimental games were played, was conducted in each of 42 experimental locations.<sup>5</sup> Among the six different regions of the study, five have been involved in a program for a sufficient time so that there was a group of individuals that had received benefits. In the remaining region, by the time of the evaluation, the program was just starting and the participants of that region are not considered as current beneficiaries but will receive benefits in the future.<sup>6</sup>

Subjects were told the experiment would last about four hours. There was no show-up payment, but local transportation expenses were covered, and a complimentary snack and modest lunch were provided. Individual earnings ranged between 3,000 and 21,400 pesos with an average of 10,477.36 pesos (about US\$4.55).<sup>7</sup> This amount is about 1 day's wage of those participating in the experiment. Earnings were paid in cash at the end of the experiment. Following the experiment, each person also completed an extended survey with demographic and socioeconomic information, together with inquiries about the life in the community, as well as knowledge perception, and experience with the Peace and Development and Peace Laboratory programs.

Each experimental session was conducted in the same order. Because the order of the experiments was the same in every session, differences in play between regions cannot be because of the ordering of play. Each session was hand-run with native Colombians and was conducted in Spanish; each consisted of a version of the trust game, followed by a charitable contribution to a public service, and finally a public goods game. At the start of each session, the subjects were seated in a semi circle away from each other and were not permitted to communicate among themselves. The experimenter read the instructions aloud and questions were answered in public. Participants with reading and/or writing difficulties received additional assistance, but they were required to make their contribution decisions on their own.

### 3.2 The trust game

The trust game was first reported by Berg, Dickaut and McCabe (1995) and for this experiment it was implemented as follows: Two players A and B were given an endowment of \$4,000 pesos. Player A decided how much of the endowment to send to the anonymously paired player B. Each peso sent from A to B was tripled by the experimenter. Player B then decided how much to send back to Player A from a total value of his initial endowment plus the tripled amount sent by Player A. If subjects maximize own money income then, theoretically, player A would send nothing to player B and then player B would return nothing. However,

---

<sup>5</sup>Experimental locations were randomly selected to be in 42 municipalities out of the possible 127 municipalities in the six regions.

<sup>6</sup>The number of subjects invited to participate in each experiment based on a desire to have a relatively balanced sample across regions. The reader is again referred to Table 1 on the description of the regions, the number of experiments conducted in each region, and the number of subjects who participated in each region.

<sup>7</sup>The exchange rate by the time of the experiment was 2300 pesos for 1 dollar.

the Berg et al. experiment have been replicated by several authors in many places of the world finding that first movers, player A, send between 50%-65% of their initial endowment. This behavior has been identified as trust. Second movers, Players B, send back between 30%-40% of their earnings, behavior explained in the literature as reciprocity, depending upon the experimental conditions (Cardenas and Carpenter, 2008).

### 3.3 Contribution to a public service

At the end of the trust game, each participant calculated their earnings and was then asked to make a voluntary contribution to one of the following organizations: one of the schools of the municipalities, the local health service, the school lunch program, or another organization from the community indicated by the participant.<sup>8</sup> Each person was allowed to donate up to the quantity made in the trust game. Donations were anonymous; nobody in the group knew who was donating how much or to which cause.

### 3.4 The public good game

The public goods game has been designed to elicit information regarding the degree of social-mindedness that participants display. Each participant is given a token and with that token he decides in private either to invest it in a private account or in a public account. If the token is invested in the private account he will make 2,000 pesos in addition to 100 hundred pesos for each token invested in the public account. If the token is invested in the public account, that token will be add with the other tokens invested to the public account and for each token each person will receive 100 pesos.<sup>9</sup>

In this experiment social welfare is maximized when all individuals invest their token to the public account. However, if fewer than 21 subjects invest the token in the public account, subjects will make less by investing in the public account than what they can make by investing in the private account. This game was played twice, in the first round participants were not allowed to communicate in any way with the other subjects in the room. After the first round, and not knowing the outcome of the first round, participants were allowed to engage in a voluntary and open discussion for 10 minutes. The subjects then played the public goods game for a second round.

---

<sup>8</sup>In each one of the municipalities, the experimenter together with the person on charge of the Program determined if there was more than one school, health service, or school lunch program for the municipality. If subjects had a preference for a particular school, health service, or school lunch program they were instructed to indicate it.

<sup>9</sup>Note that this specification of the public goods game allows for a constant marginal per capita return (MPCR) of  $1/20=0.05$  despite the size of the group. This was done because group sizes were uncertain; the number of subjects that show up to any particular experiment was not known beforehand. Isaac and Walker (1988) show that group size effects in the public goods game operate through changes in the MPCR and not through increases in group size itself. Thus, since the MPCR was held constant at different group sizes there should be no differences between sizes of groups and subject behavior in the public goods game.

## 4 Estimating Treatment Effects

We emphasize for this section that the experimental games that were explained in the previous section are used only as a means to measure cooperation and not to test treatments within the laboratory. Our goal, however, is not to merely measure cooperative behaviors, but rather to measure how the Peace and Development and Peace Laboratory programs have altered cooperative behavior. The manner by which we evaluate these programs is not experimental; that is, the programs are not randomly given to some regions and withheld from others and beneficiaries of the programs are not randomly selected to participate in the programs.

Our goal is to assess if the program in Magdalena Medio and then subsequently transmitted to the other expansion regions, was able to engender inclusive cooperation. The very nature of the program, that development projects are financed according to local preferences, prevents us from examining the effects of particular policies—a problem of non-standardized treatment (See Rossi, Lipsey and Freeman, 1989, p.198). Instead, the program is adapted so that it fits a given region's preferences. The treatment, then, should not be considered a set menu of programs, but rather the ability of a community to choose its own menu of programs and implement them with CDR aid.

In an ideal randomized field experiment we would randomly choose individuals to participate in one of the six regional programs or withhold the ability to participate in any regional program (for a review see Humphreys and Weinstein (2009)). We would then measure the cooperation of each individual before and after implementing each program. Any differences in outcomes could be attributed to the programs. The Peace and Development and Peace Laboratory programs depart from the ideal randomized field experiment in two important ways. First, individuals are not randomly placed into a program; this is the self-selection effect. Participation in the program was voluntary. Second, the regions were not chosen to randomly administer the program or not; this is the policy selection effect.

Fortunately, however, the way the policy was implemented allows us to establish credible comparison groups whose behavior can be used to construct counterfactuals of how program beneficiaries would have behaved absent participation in the program. We identify two control groups to take account of the self-selection effect and policy-selection effect.

As identified in Table 1, the region of Meta implemented their program after all of the remaining regions. When the laboratory experiments described in the previous section were implemented, those that will participate in that program in the future were identified, but they had yet to receive any benefits from the program. This feature of the policy implementation allows us to compare those who have selected into the program and have already received treatment to those who have selected into the program but have yet to

receive treatment. Even though we, as the researchers, did not randomize who received treatment, policy-makers unwittingly created a natural experiment by delaying the opportunity for treatment to a subset of individuals. This feature of the program allows us to control for the self-selection effect.<sup>10</sup> For the remainder of the paper we will refer to Meta as the control region.

Despite the fact that program participants in the control region are similar to other beneficiaries in other regions, in that they have similar propensity to self-select into the program, there may be important differences between regions. A different control group is necessary to assess if regional differences are determining the behavior and not simply treatment into the programs. The non-beneficiaries in the Peace and Development and Peace Laboratory regions can be compared to the non-beneficiaries in the control region to control for the policy selection effect.

Thus, we use two control groups to identify the effects of the programs. The first is a nonequivalent control group consisting of those subjects living in the region of a particular program who are not beneficiaries of that program. They are nonequivalent because subjects self-select into treatment; thus they may differ in some fundamental respect from the beneficiaries. We refer to the second control group as the accretion group (Aiken et al., 1998). The accretion group consists of the group of subjects that will be beneficiaries in the new program of the control region. These subjects are plausibly similar to the self-selecting beneficiaries of the remaining regions in that the mechanisms whereby they select into participating in the program are similar. However, subjects in the accretion group are sampled from a different geographical region and therefore may exhibit heterogeneity in some characteristics correlated with spatial differences.

Thus, the first nonequivalent control group is similar geographically to those receiving treatment, but differs on self-selection criteria while the accretion group is similar on selection criteria but different on the policy-selection criteria. By combining both controls into the analysis we are able to control for both policy-selection and self-selection and thus construct valid counterfactuals to enable us to measure the treatment effects of the programs.

## 4.1 Average treatment effects

Consider a behavioral measure of cooperation  $Y_{ij}$  which is a function of a set of control variables,  $X_{ij}$  (such as age, education, etc.), and a program treatment,  $T_{ij}$  for subjects  $i = 1, \dots, n$  in regions  $j = 1, \dots, J$ . The treatment variable is a binary variable indicating whether the subject is a beneficiary of a program. Treatment is composed of two different considerations. First a subject must (geographically) live in an area

---

<sup>10</sup>See McIntosh (2008) for an example of an attempt to identify those who would have taken treatment had they been given the opportunity to do so, but were not given such an opportunity. In that application the author went to villages that were not eligible to participate in the program and had village councils vote on whether they would participate in a hypothetical scenario that if the program were offered. In this paper, the control group is not engaging in hypothetical actions; the identified control group will, in fact, participate in the program at a later date.

where treatment is offered, and second, subjects must voluntarily accept treatment. We first define a binary variable  $L_j$  indicating whether or not the subject lives in a given location. Next define a policy location variable  $A_j$  indicating whether or not the location of the subject is in a region with an active program. We also define a binary voluntary participation variable,  $P_{ij}$  which indicates the subject’s willingness to participate in the program, whether or not the subject lives in a location that offers the program. We next define a multiplicative function  $\tau$  which maps the active policy variable and voluntary participation into treatment, such that  $\tau(A_j, P_{ij}) = A_j P_{ij} \rightarrow T_{ij} \in \{0, 1\}$ . The receipt of the treatment requires both that subjects are in a location where treatment is possible and that they have accepted the treatment. Formally, then, the equation determining outcomes for each subject is:

$$Y_{ij} = \beta X_{ij} + \delta T_{ij} + \theta_j L_j + \gamma P_{ij} + \varepsilon_{ij}, \quad (1)$$

where

$$T_{ij} = A_j P_{ij}.$$

Note that any spatial heterogeneity is accounted for in  $\theta_j$  because there are multiple observations at each location, allowing a dummy variable to be used to identify all spatial effects for the subjects at that location. This is particularly important for this application because there might be large differences in the duration or intensity of civil conflict across locations. However, these differences will be captured by the model.

The parameter  $\delta$  in Equation 1 is a Difference-in-Differences (DID) measure of the average treatment effect (ATE) of the program. Note that this is not the traditional DID that calculates the differences in the change (over time) of outcomes between participating and non-participating individuals.<sup>11</sup> Instead, the two differences come not from the distinction with respect to time and participation, but rather come from the distinction between location and participation. This can be seen by deriving the following equality from Equation 1:

$$\delta = (E[Y_{ij}|A_j=1, P_{ij}=1] - E[Y_{ij}|A_j=1, P_{ij}=0]) - (E[Y_{ij}|A_j=0, P_{ij}=1] - E[Y_{ij}|A_j=0, P_{ij}=0]). \quad (2)$$

That is,  $\delta$  differences the difference between participants and non-participants in the treatment regions with the difference between future participants and non-participants in the control region, the only region where  $A_j = 0$ . For a more formal treatment of the conditions under which the estimates will be unbiased refer to Appendix A.

---

<sup>11</sup>Indeed this is a shortcoming of the research design. If we had pre-treatment data on the behavioral measures of social capital we might be able to triple difference across time, space, and participation as in McIntosh (2008).

There are a variety of ways to measure the DID estimator including performing  $t$ -tests for difference of means from the aforementioned groups. We estimate the DID in a regression framework following Equation 1, where the dependent variables of interest are the various behavioral measures of cooperation: second mover behavior in the Trust Game (Reciprocity), contributions to public services, and behavior in the binary Public Goods games. The regression framework allows us to control for other potentially confounding factors. These estimates will be discussed below.

## 4.2 Separating treatment effects

In this application there are really two types of programs: programs in Magdalena Medio, the region where the Peace and Development program started and the expansion regions. We analyze separate treatment effects for both regions. Some notation is necessary for this task. Let  $A_j^m$  be a binary variable indicating being located in Magdalena Medio region, where there is an active program, and let  $A_j^e$  be a binary variable indicating location in the expansion regions, where there is an active program. With this notation we modify Equation 1 to measure the average treatment effects over locations within a region, such that:

$$Y_{ij} = \beta X_{ij} + \delta^m T_{ij}^m + \delta^e T_{ij}^e + \theta_j L_j + \gamma P_{ij} + \varepsilon_{ij}, \quad (3)$$

where the notation  $\delta^m$  and  $\delta^e$  are now estimates for the ATEs in Magdalena Medio and the Expansion regions, respectively, and  $T_{ij}^e = A_j^e P_{ij}$  as well as  $T_{ij}^m = A_j^m P_{ij}$ .

## 4.3 Conditional average treatment effects

The treatment effects of the peace and development programs may vary according to the covariates in  $X_{ij}$ . For example, there are specific goals in the expansion regions to target the displaced and the poor. Thus, there is theoretical interest in examining if such a program can be especially effective among those subpopulations. In the following empirical analysis, we look at conditional average treatment effects (CATE), where treatment effects are interacted with a subset of covariates of interest: wealth, displacement status, and gender. Separating out a given covariate,  $Z_{ij}$  from  $X_{ij}$  the CATE is estimated by:

$$Y_{ij} = \beta X_{ij} + \beta^z Z_{ij} + \delta^{mz} T_{ij}^m Z_{ij} + \delta^{ez} T_{ij}^e Z_{ij} + \delta^m T_{ij}^m + \delta^e T_{ij}^e + \theta_j L_j + \gamma P_{ij} + \varepsilon_{ij}, \quad (4)$$

where  $\delta^{mz}$  and  $\delta^{ez}$  represent the CATE in each region.

## 5 Data Analysis

The data used to estimate Equations 3 and 4 are described in Table 2. A set of control variables includes the following: male, a binary variable indicating the gender of the subject; age, measured in years; education, measured as an ordinal variable ranging from 0-9, where 0 is no primary school and 9 is completion of a university degree; urban, a binary variable indicating if the subject lives in an urban center; years at residence, indicating how long the subject has lived in the area; displaced, a binary variable indicating if a subject had been displaced as a result of violent conflict; and, the PCA wealth score, which is an index of the assets held by the individual which were recorded in the survey portion of the experiments.<sup>12</sup>

The summary statistics indicate that about half of the subjects are male, half are direct beneficiaries of a Peace and Development program or a Peace Laboratory, half are urban, the mean age is just over 40 years, the mean time lived in the present home is 20 years, while about 20 percent of the subjects have been displaced due to violent conflict. There are some differences in demographic characteristics between non-participants and participants and between Magdalena Medio, the control region, and the expansion regions.<sup>13</sup> For instance, those in the control region, appear more likely to be male, more likely to have more education, and are more wealthy.<sup>14</sup> Overall, beneficiaries of the programs are slightly more likely to be younger, urban, female, and displaced than non-beneficiaries.

### 5.1 The trust game

In Table 3 we report results from second mover behavior in the Trust Game. We do not find first mover behavior in the trust game to be a good indicator of the propensity to cooperate and thus omit the analysis of that outcome.<sup>15</sup> We report separate estimates of the unconditional regional ATEs (Pooled) and the CATEs

---

<sup>12</sup>Data on a variety of assets was collected for each subject. We developed a wealth asset index by performing principal components analysis. The index is formed by using the first principle component scores. Table 7, in Appendix B, reports means for each component used in constructing the index by quintile. For example, the bottom 20 percent of individuals in the wealth index were, on average, older, more likely to be male, had less education, were more likely to be rural, more likely to be displaced by the conflict, have larger households, smaller homes, fewer assets, and have less income than each of the other quintiles of the index. See Montgomery et al. (2000) for an exposition of PCA applied to asset holdings to form wealth indexes.

<sup>13</sup>More detailed summary data on the distribution of each of these variables in the various regions and for participants and non-participants are available upon request.

<sup>14</sup>Subjects in the control regions appear to be, on average, much less wealthy than subjects in the other regions. To ensure that the control groups has observations on low ends of the distribution for wealth we examined the distribution of the wealth score over regions by beneficiary status, which are reported in Figure 3 of Appendix C. This figure shows that the support of control region largely, although not completely, spans the support of the other regions.

<sup>15</sup>First mover actions in the trust game may conflate the propensity to cooperate with risk (Karlan, 2005). Insofar as first mover behavior is not indicative of cooperation but rather the willingness to engage in risk we may not expect programs designed to increase cooperation to modify this behavior. In field setting in Peru Karlan (2005) showed that those sending more in the trust game were less likely to be cooperative, while those more likely to reciprocate trust were more likely to be cooperative, as measured by their propensity to repay group microfinance loans. Schechter (2007) finds large differences in first mover behavior between men and women in the trust game, but then identifies that this difference is due to differences in gender in risk aversion, and that risk acceptance is highly correlated with first mover behavior. We concur with the interpretation by Karlan (2005) that trust may be a measure of riskiness; thus, trust is not linked to cooperation and therefore any program effects on first mover behavior are incidental. Indeed, we found that ATEs and CATEs measured on risk were not significant. These results are available from the authors upon request.



Table 2: Summary Statistics by Region

	Control Region		Magdalena Medio		Expansion Regions		All Regions	
	Beneficiary	Non-Beneficiary	Beneficiary	Non-Beneficiary	Beneficiary	Non-Beneficiary	Beneficiary	Non-Beneficiary
Trust Game (n=868)								
Amt Player 1 Sent	2.872 (0.95)	2.854 (0.87)	2.321 (0.80)	2.566 (0.89)	2.301 (0.91)	2.196 (0.97)	2.355 (0.90)	2.353 (0.97)
Amt Player 2 Return	5.682 (2.32)	5.070 (2.20)	4.014 (1.99)	4.521 (2.23)	3.734 (2.21)	3.819 (2.21)	3.991 (2.22)	4.105 (2.25)
Public Goods Game (n=1739)								
PG Contributions	1.253 (0.76)	1.033 (0.78)	1.139 (0.82)	1.250 (0.71)	1.267 (0.79)	1.174 (0.81)	1.228 (0.79)	1.177 (0.79)
Contributions to Public Services (n=1739)								
Proportion Earnings Contributed	0.424 (0.29)	0.350 (0.27)	0.387 (0.24)	0.441 (0.28)	0.345 (0.25)	0.339 (0.23)	0.364 (0.25)	0.363 (0.25)
Male	0.554 (0.50)	0.571 (0.50)	0.379 (0.49)	0.442 (0.50)	0.479 (0.50)	0.483 (0.50)	0.456 (0.50)	0.483 (0.50)
Age	43.855 (12.57)	40.330 (12.21)	38.296 (13.58)	40.392 (13.21)	42.496 (13.81)	40.271 (13.51)	41.377 (13.78)	40.305 (13.29)
Education	4.157 (2.28)	4.833 (2.21)	3.675 (1.81)	3.920 (2.33)	3.256 (1.65)	3.268 (1.78)	3.458 (1.78)	3.583 (2.03)
Urban	0.663 (0.48)	0.495 (0.50)	0.704 (0.46)	0.598 (0.49)	0.561 (0.50)	0.438 (0.50)	0.612 (0.49)	0.481 (0.50)
Years at Residence	13.219 (15.14)	19.169 (15.83)	17.118 (15.84)	17.879 (16.89)	22.413 (19.05)	21.443 (18.42)	20.049 (18.11)	20.379 (17.87)
Displaced	0.349 (0.48)	0.198 (0.40)	0.204 (0.40)	0.322 (0.47)	0.249 (0.43)	0.308 (0.46)	0.244 (0.43)	0.300 (0.46)
PCA Wealth Score	0.882 (2.77)	1.552 (2.79)	0.348 (2.22)	0.371 (2.67)	-0.302 (1.99)	-0.429 (2.08)	-0.007 (2.17)	-0.038 (2.39)

Means reported in top row, standard deviations in parentheses underneath.

Table 3: Reciprocity Treatment Effects by Groups

	(1) Pooled	(2) Displaced	(3) Wealth	(4) Gender
Magdalena Medio Beneficiary	0.470* (0.24)	0.543* (0.28)	0.435* (0.25)	0.390 (0.25)
Expansion Region Beneficiary	0.528** (0.24)	0.432 (0.28)	0.534** (0.24)	0.270 (0.23)
Displaced Magdalena Medio Beneficiary		-0.155 (0.56)		
Displaced Expansion Region Beneficiary		0.379 (0.36)		
Wealth X Magdalena Medio Beneficiary			0.041 (0.06)	
Wealth X Expansion Region Beneficiary			0.068 (0.05)	
Male X Magdalena Medio Beneficiary				0.185 (0.30)
Male X Expansion Region Beneficiary				0.526** (0.26)
Amt Player 1 Sent	1.173*** (0.07)	1.172*** (0.07)	1.179*** (0.07)	1.175*** (0.07)
Beneficiary	-0.358* (0.19)	-0.364* (0.19)	-0.343* (0.19)	-0.352* (0.18)
Male	0.304** (0.13)	0.296** (0.13)	0.307** (0.13)	0.132 (0.18)
Age	0.021*** (0.01)	0.021*** (0.01)	0.022*** (0.01)	0.021*** (0.01)
Education	0.026 (0.04)	0.023 (0.04)	0.025 (0.04)	0.027 (0.04)
Urban	0.131 (0.14)	0.129 (0.14)	0.122 (0.14)	0.144 (0.14)
Years at Residence	-0.006 (0.00)	-0.006 (0.00)	-0.006 (0.00)	-0.006 (0.00)
Displaced	-0.000 (0.21)	-0.113 (0.22)	-0.005 (0.21)	-0.002 (0.22)
PCA Wealth Score	0.022 (0.04)	0.020 (0.03)	-0.004 (0.04)	0.021 (0.04)
Constant	0.173 (0.33)	0.236 (0.35)	0.161 (0.34)	0.243 (0.33)
R-Squared	0.314	0.316	0.315	0.317
Adj. R-Squared	0.305	0.305	0.305	0.307
F	44.617***	41.881***	47.573***	40.171***
N	874	874	874	874

Fixed effects for each experiment location included but not reported. Cluster-robust standard errors (clustered on experiment location) in parentheses. Two-tailed hypothesis tests: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

of covariates for displacement status (Displaced), the wealth index (Wealth), and gender (Gender).

In order to assess the amount returned by the second mover it is imperative to control for the amount the first mover originally sent. As expected this amount is statistically significant and positive in all models. From the pooled model it appears that on average, and holding all else constant, that for every 1000 pesos sent by the first mover, the second mover returns about 1173 pesos (significant at the 0.01 level). This coefficient is robust across the different model specifications. This is a large rate of return compared to similar studies of the trust game.<sup>16</sup>

<sup>16</sup>For comparisons, see (Cardenas and Carpenter, 2008).

The regional program treatment effects are included under the headings: Magdalena Medio Beneficiary and Expansion Region Beneficiary. In the pooled model the ATE are estimated at 0.47 in Magdalena Medio (significant at the 0.10 level) and an even larger 0.528 in the expansion regions (significant at the 0.01 level). Intuitively, this means that beneficiaries in the expansion regions sent 528 more pesos than they would have had they not participated in the program, holding all else constant. The pooled model also indicates that male subjects returned more than female subjects. To assess the possibility that the program had a disproportionate impact on male subjects, the fourth column of Table 3 reports the CATE of gender. This analysis suggests that male beneficiaries were especially effected by the program, sending 526 more pesos than they otherwise would have, while female beneficiaries of the program sent only 270 more pesos than they otherwise would have. In fact, the treatment effect among female beneficiaries of the program is not significant.

The second and third columns of Table 3 report the CATE of wealth and displacement status. It appears, however, that the treatment effects on second mover behavior in the trust game are not significantly affected by those variables.

## 5.2 Contributions to a public service

Immediately after playing the trust game the subjects were asked if they would care to donate a portion of their earnings to a public service. The subjects could only donate money from their earnings in the trust game (they were not permitted to donate any other money outside of the experiment or from the public goods game). Earnings from the trust game differed across players; to we control for the income effects generated by their previous play we code a dependent variable as the proportion of their earnings which they contribute to a public service. The value of the dependent variable ranges from 0 to 1. We report the results from estimating ATEs for each region on these contributions in the first column of Table 4. The remaining columns report CATEs for displacement status, wealth, and gender.

We find again positive and significant ATEs for both the Magdalena Medio region and the expansion regions. The ATE for Magdalena Medio is almost twice as large as the ATE for the expansion regions. Beneficiaries in the Magdalena Medio program give 11.4 percentage points more of their earnings than if they had not participated in the program (significant at the 0.05 level), while beneficiaries in the expansion regions give about 6.8 percentage points more of their earnings than they otherwise would (significant at the 0.10 level). This may be because of a superior program in Magdalena Medio or it may be because beneficiaries have operated in the program for a longer period of time.

Estimates of the CATEs indicate that the treatment effect do not appear to be conditioned on wealth.

Table 4: Public Service Contribution Treatment Effects by Groups

	(1)	(2)	(3)	(4)
	Pooled	Displaced	Wealth	Gender
Magdalena Medio Beneficiary	0.114** (0.05)	0.133** (0.06)	0.114** (0.05)	0.113* (0.06)
Expansion Region Beneficiary	0.068* (0.03)	0.050 (0.04)	0.065* (0.03)	0.036 (0.04)
Displaced Magdalena Medio Beneficiary		-0.042 (0.04)		
Displaced Expansion Region Beneficiary		0.068** (0.03)		
Wealth X Magdalena Medio Beneficiary			0.004 (0.01)	
Wealth X Expansion Region Beneficiary			-0.011 (0.01)	
Male X Magdalena Medio Beneficiary				0.004 (0.08)
Male X Expansion Region Beneficiary				0.066** (0.03)
Beneficiary	-0.071** (0.03)	-0.073** (0.03)	-0.072** (0.03)	-0.070** (0.03)
Male	0.010 (0.01)	0.009 (0.01)	0.011 (0.01)	-0.011 (0.02)
Age	0.002*** (0.00)	0.002*** (0.00)	0.002*** (0.00)	0.002*** (0.00)
Education	0.010** (0.00)	0.010** (0.00)	0.010** (0.00)	0.010** (0.00)
Urban	0.026* (0.02)	0.026* (0.01)	0.027* (0.02)	0.027* (0.02)
Years at Residence	0.000 (0.00)	0.000 (0.00)	0.000 (0.00)	0.000 (0.00)
Displaced	-0.005 (0.02)	-0.021 (0.02)	-0.003 (0.02)	-0.004 (0.02)
PCA Wealth Score	0.016*** (0.00)	0.016*** (0.00)	0.018*** (0.00)	0.017*** (0.00)
Constant	0.227*** (0.03)	0.235*** (0.03)	0.229*** (0.03)	0.239*** (0.03)
R-Squared	0.055	0.059	0.057	0.058
Adj. R-Squared	0.049	0.053	0.050	0.052
F	8.926***	8.397***	8.159***	9.051***
N	1739	1739	1739	1739

Fixed effects for each experiment location included but not reported. Cluster-robust standard errors (clustered on experiment location) in parentheses. Two-tailed hypothesis tests: \* p<0.10, \*\* p<0.05, \*\*\* p<0.01

However, there are strong CATEs for displaced persons and males in the expansion regions.

Comparing displacement status, it appears that displaced beneficiaries in the expansion regions largely accounted for the treatment effects of the program and that non-displaced beneficiaries were not significantly affected. That is, displaced expansion region beneficiaries gave, on average, 6.8 percentage points more of their earnings than they otherwise would have (significant at the 0.05 level), while non-displaced beneficiaries in the expansion regions actually gave a lower percentage of their earnings to the charitable service than they otherwise would have, although this estimate is small and not statistically distinct from no effect at all.

Comparing genders, it appears that male beneficiaries in the expansion regions largely accounted for the treatment effects of the program and that female beneficiaries were not significantly affected. That is,

male expansion region beneficiaries gave, on average, 6.6 percentage points more of their earnings than they otherwise would have (significant at the 0.05 level), while female beneficiaries in the expansion regions only gave about 0.4 percentage points more than they otherwise would have, although this estimate is small and not statistically distinct from no effect at all.

### 5.3 Public goods contributions

We also assessed program effects in the Public Goods game as described in the previous section. Subjects played two rounds of a public goods game and we code their behavior as 0 meaning no contribution in either round, 1 meaning contribution in one round, and 2 meaning contribution in both rounds of the game. We continue to use the regression framework of Equations 3 and 4 to estimate this model.<sup>17</sup> In Table 5 we report the results from this analysis. We note that these estimates provide a poor fit to the data, so caution should be used in their interpretation.<sup>18</sup> Nonetheless, given the theoretical setup of the problem, the linear estimates are unbiased estimates of the ATEs and CATEs of the program.

Given the caution necessary for inference in these models, we here note that beneficiaries in Magdalena Medio were more likely to contribute to the public good than they otherwise would have been, but this effect is most pronounced among males in the region (both significant at the 0.10 level). There is also some evidence that displaced beneficiaries in the expansion regions are more likely to donate to the public good than they otherwise would have (significant at the 0.10 level).

### 5.4 Correlations between Measures of Cooperation

To what extent are the different measurements of cooperation are correlated? To answer this question we calculate simple correlations between behavior in the public goods game and the trust game. In Table 6 we list these correlations. One should notice that the behavioral measures are only weakly correlated, implying that the type of cooperation measured in the activities is different. We interpret the difference in behaviors

---

<sup>17</sup>While linear regression implies the potential for predicted values that lie outside the support of the dependent variable ( $0 - 2$ ), such a framework nevertheless allows one to consistently estimate the ATE, and does so without the additional distributional assumptions of the data generating process (Angrist and Pischke, 2009, p.94-99). Estimates of the ATEs and CATEs will be unbiased in the linear model because the dependent variable is interval (i.e. not ordinal in the sense that their is uncertain distance between categories) (Long, 1997, p.114). Linear regression also avoids the incidental parameters problem when estimating a fixed-effects ordinal logit or probit model. See Wooldridge (2002, p.484) for a discussion of the incidental parameters problem.

<sup>18</sup>Evidence for poor fit is given by low model significance (F-statistics) and a low R-squared measures of goodness of fit. Maximum likelihood estimation of a fixed effects ordinal model is not possible because of the identification problem that the number of parameters to be estimated increases at the same rate as the number of locations increases. Two different specifications of the data generating process were also estimated: a fixed effects binary logit model where the dependent variable is coded as =1 if the subject donated to the public good in any round; and a a fixed effects binary logit model where the dependent variable is coded as =1 if the subject donated to the public good in all rounds. However, in neither of these models were the variables jointly significant and they did not produce substantively different conclusions from the linear models reported here. These estimates are available upon request from the authors.

Table 5: Public Goods Treatment Effects by Groups

	(1)	(2)	(3)	(4)
	Pooled	Displaced	Wealth	Gender
Magdalena Medio Beneficiary	0.255*	0.273*	0.272*	0.184
	(0.13)	(0.14)	(0.14)	(0.13)
Expansion Region Beneficiary	0.108	0.055	0.107	0.146
	(0.13)	(0.13)	(0.14)	(0.14)
Displaced Magdalena Medio Beneficiary		0.004		
		(0.11)		
Displaced Expansion Region Beneficiary		0.218*		
		(0.12)		
Wealth X Magdalena Medio Beneficiary			-0.026	
			(0.02)	
Wealth X Expansion Region Beneficiary			-0.021	
			(0.02)	
Male X Magdalena Medio Beneficiary				0.163*
				(0.09)
Male X Expansion Region Beneficiary				-0.077
				(0.08)
Beneficiary	-0.194	-0.205*	-0.201	-0.195
	(0.12)	(0.12)	(0.12)	(0.12)
Male	0.020	0.019	0.019	0.025
	(0.04)	(0.04)	(0.04)	(0.06)
Age	0.002	0.002	0.002	0.002
	(0.00)	(0.00)	(0.00)	(0.00)
Education	0.015	0.015	0.016	0.016
	(0.02)	(0.02)	(0.02)	(0.02)
Urban	-0.038	-0.039	-0.034	-0.040
	(0.06)	(0.06)	(0.06)	(0.06)
Years at Residence	-0.002	-0.001	-0.002	-0.002
	(0.00)	(0.00)	(0.00)	(0.00)
Displaced	-0.062	-0.135**	-0.062	-0.062
	(0.06)	(0.07)	(0.07)	(0.06)
PCA Wealth Score	-0.017	-0.019	-0.008	-0.017
	(0.01)	(0.02)	(0.02)	(0.02)
Constant	1.177***	1.199***	1.167***	1.168***
	(0.12)	(0.11)	(0.12)	(0.12)
R-Squared	0.008	0.012	0.009	0.010
Adj. R-Squared	0.002	0.005	0.002	0.003
F	1.082	1.431	1.207	2.377**
N	1742	1742	1742	1742

Fixed effects for each experiment location included but not reported. Cluster-robust standard errors (clustered on experiment location) in parentheses. Two-tailed hypothesis tests: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

as indicative of the degree of insularity of cooperation; second mover behavior in the trust game (reciprocity) measures more insular cooperation, contributions to public services measure a more inclusive tendency to cooperate, while the public goods contributions are somewhere in-between. These results also imply that analysts should think carefully about the measures of cooperation they wish to employ when assessing the effects of a social program. Again, we feel that the measure of inclusive cooperation as measured by charitable contributions is the most relevant measure for evaluating the peace and development program.

Table 6: Correlations between Measures of Cooperation

	Reciprocity/Trust	Contributions to Public Service	PG Contributions
Reciprocity/Trust	1.000		
Contributions to Public Service	-0.022	1.000	
PG Contributions	-0.023	0.062	1.000

Fixed effects for each experiment location included but not reported. Cluster-robust standard errors (clustered on experiment location) in parentheses. Two-tailed hypothesis tests: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

## 6 Discussion

### 6.1 Comparison to previous research

First we discuss our results in light of the literature from previous experimental games conducted in the field (what Harrison and List (2004) term “artefactual field experiments”). We followed standard trust game protocols so our results are most directly comparable to those as measured in the trust game literature.<sup>19</sup> In Figure 2 we plot the average proportion of the endowment sent by the first mover in the trust game, across all individuals from all regions, and compare our results (the light gray bar) to other studies conducted in various locations around the world (the darker bars). While our measures of trust are within the bounds of those previously estimate by other authors, it does appear that trust is relatively high in Colombia compared to other countries.

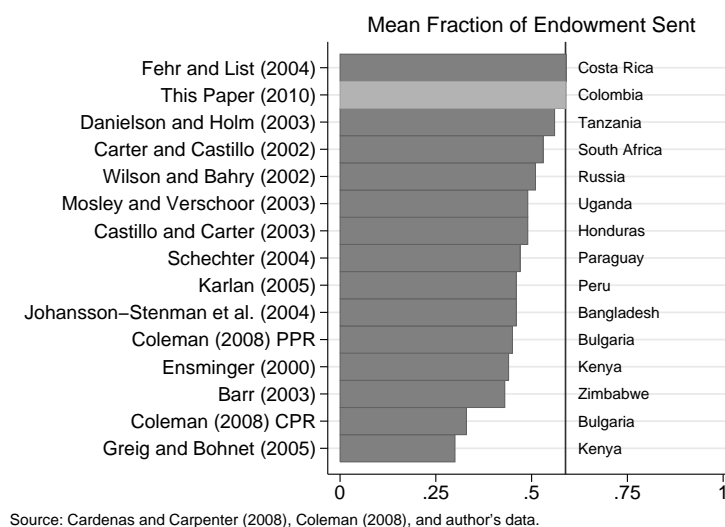


Figure 2: Measures of Trust from Various Artefactual Field Experiments.

This result is counterintuitive if one uses first mover behavior in the trust game as a proxy for cooperation,

<sup>19</sup>There are many variants of the public goods game, but the binary nature of the decision (contribute to the public good or do not contribute) is not as common in that literature. Thus, we rely mainly on our comparisons to the trust game.

but perhaps not counterintuitive if one considers this behavior to be more indicative of risk (Karlan, 2005; Schechter, 2007). Perhaps individuals in a society with civil conflict are more likely to take the moderate risks in the experimental game environment after facing large-stakes risk decisions for many years. That it, their references point for taking risky decisions might change in an environment of perpetual high-stakes risk.

## **6.2 Can the program be successfully transmitted to other regions?**

The main purpose of this paper is to investigate if a successful CDR-driven program can be transmitted to other regions of a country and to try and specify some of the conditions under which successful transmission is most likely. It is apparent from our previous results that the Magdalena Medio program causes its beneficiaries to be more cooperative than they otherwise would be (as measured by second mover behavior in the trust game, contributions to a public service, and behavior in the public goods game). Program beneficiaries in the expansion regions were also likely to exhibit cooperative behavior because of the program for all of the measures except that of the moderately insular measure of the public goods game. However, expansion region beneficiaries were more likely to act cooperatively on issues that were of a broad scale by donating to public services, as well as by cooperating in more insular endeavors as measured by reciprocity in the trust game. In summary, based upon the measures of cooperation which we employed the program has been successfully transmitted from Magdalena Medio to the expansion regions.

Because norms of cooperation may be a part of deeply held beliefs or may be motivated by strong political cleavages one might expect the effects of CDR aid to be minimal. It would seem that CDR aid would be least effective with those who had experienced the negative consequences of violent conflict. However, we found that when explicit program goals were targeted at displaced persons, as they were in the expansion regions, the program was able to effectively target those displaced by violent conflict and the treatment effects among this subgroup were particular strong. Thus, displacement status does not seem to be an insurmountable impediment to the expansion of a CDR program.

The other key impediment to expansion would appear to be the overall level of conflict within any particular location where the experiment was conducted. However, even controlling for such conflict (through the use of fixed effects location-specific dummies) the program still had significant effects on participant behavior. Another potential difficulty might be that only wealthy individuals would benefit from the program. However, we found no evidence in any model that the wealthy disproportionately benefited from being a beneficiary of the program (in terms of their willingness to cooperate).

This is not to say that expansion of a CDR program from a successful region into less successful regions



will be easy or costless, but only that it is possible. There are still some ongoing difficulties with the program as well. In particular, it appears that male subjects disproportionately benefit from the program; male subjects are more likely to increase their cooperativeness because of the program than female subjects. These effects were not measured in the initial program region of Magdalena Medio, but are present in the expansion regions. This suggests that expansion of CDR programs may be difficult to female subjects and that program staff should think carefully about programs that might target that demographic.

### **6.3 Which measure of cooperation?**

Fearon, Humphreys and Weinstein (2009) measure the affects of CDR aid in Liberia and conclude that “Villages exposed to a community driven reconstruction program exhibit higher subsequent levels of social cooperation than those in the control group, as measured through a community-wide public goods game.” In our analysis we measured cooperation through a variety of experimental games to assess if the cooperation that is induced from a CDR program is more insular than inclusive. If we were to confine our analysis to just the trust game we would not have been able to assess if the induced cooperation had the potential to lead to more peace or not. On the other hand, if we were to confine our analysis to the contributions to public services, we might miss some of the more insular types of cooperation that may have other potential benefits at smaller scales. In addition, while beneficiaries in Magdalena Medio had almost two times the treatment effect as those in the expansion regions for the contributions to public services, beneficiaries in the expansion regions actually experience larger treatment effect than those found in Magdalena Medio. Thus, the experimental measure that is employed matters for the policy conclusions.

Furthermore, the correlations between the different measures of cooperation, as reported in Table 6, are very weak. Thus, measurement of cooperation from a public goods game, for example, is a poor proxy for cooperative behavior as measured by reciprocity in a trust game. We do not wish to argue that any particular measure of cooperation is always better than another, but simply that they measure different things. There is no single correct measure of cooperation. Reciprocity, for example, seems to be a measure of cooperation in a small group. Cooperation in the public goods game is a measure of cooperation in a relatively larger group. Cooperation as measured through donations to public services is a broader measure of cooperation, but it is tied to specific projects and may not only reflect an underlying tendency for cooperation as in the other games, but also on the preferences for particular public services and the willingness to provide them cooperatively. Whereas cooperation in the other games is measured by a liquid asset that can be turned into any other good (money) cooperation in public service provision is tied to specific goods. However, donations to public services have the advantage of reflecting “real-world behavior” and simulate actual experiences of

the experimental participants. Thus, analysts should shift from the question, “does the program increase cooperation?” to assess, “what types of cooperation does the program affect?”

## 7 Conclusion

By the measures of cooperation we employ in this paper, it appears that overall the Peace and Development program has been successfully transmitted to other regions of Colombia. This result suggests that it is possible to transmit CDR programs (adapted to local conditions) which are meant to bolster cooperation in one region of a country with civil conflict to other regions of the country with CDR development aid. It also suggests that cooperative behavior can be changed in relatively short time frames (Fearon, Humphreys and Weinstein, 2009). We also show that there is variation in program effects among subgroups in the population. Displaced subjects in expansion regions were most effected by the program. While there does not appear to be any disproportion benefit among the more wealthy individuals, the poor do not particularly benefit from the program either. In addition, it appears that male subjects benefited disproportionately from the program and that future efforts to transmit CDR aid should take special caution to target female subjects and tailor programs to fit their needs.

We hope that our analysis will be useful both for policymakers in the Peace and Development and Peace Laboratory programs, but also for policymakers dealing with CDR aid more generally. Our results suggest that this type of aid can be effective at increasing cooperation. Local policymakers appear to have a good feel for what types of programs can engender cooperation. CDR aid can help fund programs, adapted to local conditions, and increase cooperation.

## References

- Aiken, Leona S., Stephen G. West, David E. Schwalm, James L. Carroll and Shenghwa Hsiung. 1998. “Comparison of a Randomized and Two Quasi-Experimental Designs in a Single Outcome Evaluation: Efficacy of a University-Level Remedial Writing Program.” *Evaluation Review* 22(2):207–244.
- Angrist, Joshua D. and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist’s Companion*. Princeton, NJ: Princeton University Press.
- Banerjee, Abhiji Vinayak. 2007. *Making Aid Work*. Cambridge, MA: MIT Press.
- Berg, Joyce E., John Dickaut and Kevin McCabe. 1995. “Trust, Reciprocity, and Social History.” *Games and Economic Behavior* 10:122–42.

- Berman, Sheri. 1997. "Civil Society and the Collapse of the Weimar Republic." *World Politics* 49(3):401–429.
- Bouma, Jetske, Erwin Bulte and Daan van Soest. 2008. "Trust and Cooperation: Social Capital and Community Resource Management." *Journal of Environmental Economics and Management* 56:155–166.
- Cardenas, Juan Camilo and Jeffrey P. Carpenter. 2008. "Behavioural Development Economics: Lessons from Field Labs in the Developing World." *Journal of Development Studies* 44(3):337–364.
- Carpenter, Jeffrey P., Amrita G. Daniere and Lois M. Takahashi. 2004. "Cooperation, Trust, and Social Capital in Southeast Asian Urban Slums." *Journal of Economic Behavior and Organization* 55:533–551.
- Chapman, Terrance L. 2008. "Unraveling the Ties Between Civic Institutions and Attitudes Toward Political Violence." *International Studies Quarterly* 52:515–532.
- Cliffe, Sarah, Scott Guggenheim and Markus Kostner. 2003. Community-Driven Reconstruction as an Instrument in War-to-Peace Transitions. CPR Working Papers, Paper No. 7, World Bank Social Development Department.
- Colletta, Nat J. and Michelle L. Cullen. 2000. *Violent Conflict and the Transformation of Social Capital: Lesson from Cambodia, Rwanda, Guatemala, and Somalia*. Washington, DC: World Bank.
- Duflo, Esther. 2004. "Scaling Up and Evaluation." Paper presented at the Annual World Bank Conference on Development Economics, Washington, DC.
- Falk, Armin. 2007. "Gift Exchange in the Field." *Econometrica* 75(5):1501–1511.
- Fearon, James D. and David D. Laitin. 1996. "Explaining Interethnic Cooperation." *American Political Science Review* 90(4):715–735.
- Fearon, James D., Macartan Humphreys and Jeremy M. Weinstein. 2009. "Can Development Aid Contribute to Social Cohesion after Civil War? Evidence from a Field Experiment in Post-conflict Liberia." *American Economic Review: Papers & Proceedings* 99(2):287–91.
- Glaeser, Edward L., David Laibson, Jose A. Scheinkman and Christine L. Soutter. 2000. "Measuring Trust in Transition: Preliminary Findings from 26 Transition Economies." *Quarterly Journal of Economics* 65:811–46.
- Habyarimana, James, Macartan Humphreys, Daniel N. Posner and Jeremy Weinstein. 2006. "Why Does Ethnic Diversity Undermine Public Goods Provision? An Experimental Approach." *American Political Science Review* 101:709–725.

- Habyarimana, James P., Macartan Humphreys, Daniel N. Posner and Jeremy M. Weinstein. 2009. *Coethnicity: Diversity and the Dilemmas of Collective Action*. New York: Russell Sage Foundation.
- Harrison, Glenn W. and John A. List. 2004. "Field Experiments." *Journal of Economic Literature* XLII(December):1009–55.
- Henriques, Miguel Barreto. 2007. Peace Laboratory of Magdalena Medio: "A Peace Laboratory"? CERAC Working Paper 6, Centro de Recursos para el Análisis de Conflictos.
- Humphreys, Macartan and Jeremy M. Weinstein. 2009. "Field Experiments and the Political Economy of Development." *Annual Review of Political Science* 12:367–378.
- Isaac, R. Mark and James M. Walker. 1988. "Group Size Effects in Public Goods Provision: The Voluntary Contribution Mechanism." *Quarterly Journal of Economics* 103(1):179–199.
- Kalyvas, Stathis. 2006. *The Logic of Violence in Civil War*. New York: Cambridge University Press.
- Karlan, Dean. 2005. "Using Experimental Economics to Measure Social Capital and Predict Financial Decisions." *American Economic Review* 95(5):1688–99.
- Long, J. Scott. 1997. *Regression Models for Categorical and Limited Dependent Variables*. Advanced Quantitative Techniques in the Social Sciences. Thousand Oaks, CA: Sage Publications.
- McIntosh, Craig. 2008. "Estimating Treatment Effects from Spatial Policy Experiments: An Application to Ugandan Microfinance." *Review of Economics and Statistics* 90(1):15–28.
- Montgomery, M.R., Gragnolati K., Burke A. and Paredes E. 2000. "Measuring Living Standards with Proxy Variables." *Demography* 37:155–74.
- North, Douglass C. 1990. *Institutions, Institutional Change and Economic Performance*. The Political Economy of Institutions and Decisions. Cambridge, UK: Cambridge University Press.
- Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge, UK: Cambridge University Press.
- Putnam, Robert D. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton University Press.
- Rossi, Peter H., Mark W. Lipsey and Howard E. Freeman. 1989. *Evaluation : a systematic approach*. 7th ed. ed. Thousand Oaks, CA: Sage Publications.

- Rothstein, Bo. 2005. *Social Traps and the Problem of Trust*. Theories of Institutional Design. New York: Cambridge University Press.
- Schechter, Laura. 2007. "Traditional Trust Measurement and the Risk Confound: An Experiment in Rural Paraguay." *Journal of Economic Behavior & Organization* 62:272–292.
- Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.
- Yamagishi, Toshio. 1998. *The Structure of Trust: The Evolutionary Game of Mind and Society*. Tokyo: University of Tokyo Press.

## A Conditions for unbiased estimates of the ATE

First, we restate Equation 1 for reference:

$$Y_{ij} = \beta X_{ij} + \delta A_j P_{ij} + \theta_j L_j + \gamma P_{ij} + \varepsilon_{ij}, \quad (5)$$

The general problem is that the treatment effects are biased because of either the self selection effect:

$$E[\varepsilon_{ij}|P_{ij}=1] - E[\varepsilon_{ij}|P_{ij}=0] = \text{Self Selection} \neq 0, \quad (6)$$

or the policy selection effect:

$$E[\varepsilon_{ij}|A_j=1] - E[\varepsilon_{ij}|A_j=0] = \text{Policy Selection} \neq 0. \quad (7)$$

Because the error term is not independent of either participation or the presence of an active program, the parameter estimates will be biased, according to the standard exogeneity assumption (See, for example, Long, 1997, p.12).

Compare the difference in expected outcomes of those who will be willing to participate in the program with those who are unwilling to do so, conditional on being at a location with an active program. The difference between those individuals will be the treatment effects of the program in addition to the self selection effect to entering the program. From Equation 5 and the assumption of Equation 6 we have:

$$E[Y_{ij}|A_j=1, P_{ij}=1] - E[Y_{ij}|A_j=1, P_{ij}=0] = \delta + \gamma + \text{Self Selection}. \quad (8)$$

Now consider the expected difference in outcomes between those who are willing to participate in the program and those unwilling to participate conditional on the fact that they are located in a place without an active program. From Equation 5 and the definition of Equation 6 we have:

$$E[Y_{ij}|A_j=0, P_{ij}=1] - E[Y_{ij}|A_j=0, P_{ij}=0] = \gamma + \text{Self Selection}. \quad (9)$$

This implies that:

$$E[(Y_{ij}|A_j=1, P_{ij}=1] - Y_{ij}|A_j=1, P_{ij}=0]) - (Y_{ij}|A_j=0, P_{ij}=1] - Y_{ij}|A_j=0, P_{ij}=0]) = \delta. \quad (10)$$

Thus, the two control groups solve the self-selection problem.

We can also compare the expected outcomes of those where the program is active with those in areas where the program is not active, conditional on being willing to participate in the program. The difference between those individuals will be the treatment effects of the program in addition to the policy selection effect to entering the program. From Equation 5 and the assumptions of Equation 7 we have:

$$E[Y_{ij}|A_j=1, P_{ij}=1 - Y_{ij}|A_j=0, P_{ij}=1] = \delta + \text{Policy Selection.} \quad (11)$$

Now consider the expected difference in outcomes between those who are located in an area with an active program with those who are not located in such an area, conditional on them begin unwilling to participate in the program. From Equation 5 and the assumptions of Equation 7 we have:

$$E[Y_{ij}|A_j=1, P_{ij}=0 - Y_{ij}|A_j=0, P_{ij}=0] = \text{Policy Selection.} \quad (12)$$

This implies that:

$$E[(Y_{ij}|A_j=1, P_{ij}=1 - Y_{ij}|A_j=0, P_{ij}=1) - (Y_{ij}|A_j=1, P_{ij}=0 - Y_{ij}|A_j=0, P_{ij}=0)] = \delta. \quad (13)$$

Thus, the two control groups solve the policy-selection problem.

In summary, if the assumption of functional form in Equation 5 holds and if either assumption of Equations 6 or 7 hold (underlying the definition in Equation 6 is there is no difference in the expectation of self selection between individuals and regions and no difference in Equation 7 that there is no difference in the expectation of policy selection between individuals within a region), then  $\delta$  is an unbiased estimate of the ATE.

## B PCA Wealth Index

Table 7: Component means across PCA wealth index quantities

	Poorest	Second	Third	Fourth	Richest
PCA Wealth Score	-2.389	-1.415	-0.535	0.655	3.684
Male	0.554	0.472	0.488	0.428	0.428
Age	45.730	41.885	40.675	39.790	36.556
Education	2.100	2.879	3.202	3.861	5.514
Urban	0.315	0.436	0.522	0.656	0.816
Displaced	0.373	0.367	0.299	0.197	0.123
Household Size	4.827	5.283	5.391	5.344	5.052
No. School Age Children	1.433	1.559	1.564	1.559	1.270
Own House?	0.677	0.627	0.667	0.675	0.690
Rooms in House	2.058	2.394	2.748	3.234	4.152
Bedrooms	1.562	1.843	2.045	2.270	2.780
Telephone	0.018	0.079	0.113	0.370	0.667
Mobile Phone	0.354	0.680	0.801	0.848	0.945
Refrigerator	0.034	0.218	0.486	0.740	0.955
Audio System	0.105	0.234	0.381	0.512	0.751
Laundry Machine	0.000	0.003	0.005	0.055	0.454
Oven	0.000	0.003	0.008	0.031	0.228
Sewing Machine	0.010	0.042	0.092	0.160	0.325
Black and White TV	0.239	0.260	0.121	0.071	0.037
Color TV	0.121	0.388	0.646	0.874	0.966
DVD or VHS	0.000	0.008	0.042	0.123	0.396
Computer	0.000	0.000	0.005	0.008	0.239
Bike	0.110	0.249	0.375	0.475	0.638
Motorbike	0.003	0.016	0.037	0.094	0.304
Car	0.000	0.003	0.005	0.029	0.144
Moto-Taxi	0.000	0.000	0.000	0.010	0.021
Annual Earnings from Home (10,000 pesos)	14.005	19.780	28.722	36.616	79.003
Last Week Food Expenditure (10,000 pesos)	3.554	4.775	5.453	6.905	9.949
Last Month Frequency of Food Shortage	1.819	2.226	2.404	2.648	2.990
Last Year Frequency of Savings	3.971	3.911	3.845	3.819	3.423

Mean value of the asset reported.



# C Distribution of Wealth by Region and Beneficiary Status

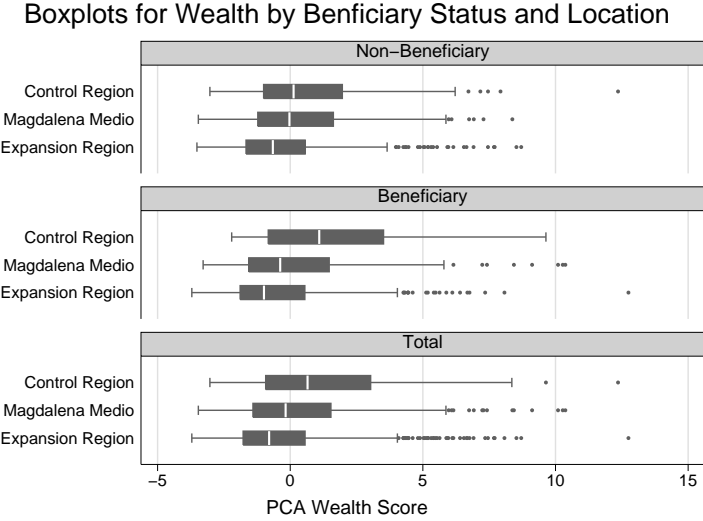


Figure 3: Distribution of Wealth by Region and Beneficiary Status