Discrimination in the Warplace:
Evidence from a Civil War in Peru

by

Marco Castillo*   Ragan Petrie
Georgia Institute of Technology   Georgia State University

November 2007

Abstract: Few events give the opportunity to observe the full range of human behavior as wars do. In the case of civil wars in ethnically-mixed societies, the distribution of violence across various segments of the population can provide evidence on the extent and nature of discrimination. As in the case of markets, identifying discrimination in the warplace is challenging. There is uncertainty on the reconstruction of events as well as the rationale behind the violence. We use a unique data set collected by the Peruvian Truth and Reconciliation Commission on war crimes during the 1980's to show that there is evidence of taste-based discrimination by agents of the state towards ethnic minorities and women. The evidence is robust to different assumptions on the logic of repression and missing data problems.

1. Introduction

Colonial rule and slavery has left a legacy of discrimination and exclusion. They are thought to be at the root of persistent inequality in the developing world. Sadly, the press is filled with reports of ethnic conflict and ethnic cleansing and with evidence that social status is stubbornly correlated with ethnic background. This paper discusses the identification of discrimination from patterns of human right violations committed by agents of the state trying to quelch social unrest. Identifying the source of discrimination from observational data on political violence is not any easier than identifying the source of discrimination in labor markets. Civil wars, however, allow us to observe interactions among populations that do not normally interact with one another, and this helps us to more clearly identify discrimination. The recent work of Truth Commissions trying to identify the extent of human right violations in countries with civil wars shows how unequal is the distribution of violence towards poor and indigenous populations. This paper is the first, to our knowledge, to look at the nature of that unequal treatment.

There is a large literature addressing the issue of discrimination in labor markets. Riach and Rich (2002) review the literature on audit studies conducted in many places. The general finding is that across a variety of markets there is evidence of discrimination based on race and gender.[1] However, as discussed by Heckman (1998), a problem with audit studies is that it is difficult to control for the distribution of unobservables. It is possible that differences in the dispersion of unobserved factors related to productivity vary across race and gender. This might mask the presence of taste-based discrimination in some cases and generate the appearance of discrimination among subjects based on observables.

More closely related to this paper is the literature on discrimination in police searches. Knowles, Persico and Todd (2001) discusses identification strategies using data on police searches of drivers. They show that taste-based discrimination can be detected if the probability of being guilty given search varies with race. The authors exploit the fact that if the police officers are homogeneous and risk neutral, they should invest their manpower to the point where the returns of searching a white driver or black driver are equal. Antonovics and Brown (2007) show that taste-based discrimination might be hidden if returns to searches and crime are different across the race of the police officer. They show that the probability of

---

[1]In a more recent study, Bertrand and Mullainathan (2004) find that those with black-sounding names tend to be discriminated against in a study using fake resumes.

search given the driver's race should be constant conditional on the race of the police officer. The key identifying assumption of their model is that drivers are randomly matched to a police officer. In the absence of that, it is not possible to identify discrimination based on searches conditional on the race of the officer. Anwar and Fang (2006) present an alternative test of discrimination based on the ordering of search intensity and the probability of finding a person guilty given search. They show that in the absence of statistical discrimination the patterns of search and guilt given search must not depend on the race of the officer. Dominitz and Knowles (2006) argue that tests of discrimination are not robust to the rationale behind the actions of the police. They show that if the motive of the police is to minimize crime rather than maximize successful searches then tests of discrimination must rely on non-testable hypotheses on the distribution of propensities to commit crime.

This paper discusses how these theoretical results can be used to test for the nature of discrimination during political violence. Our basic theoretical result is that under similar conditions and facing a similar population, it must not be the case that if an agent of the state searches a group more intensely they find more guilty people per search. The number of guilty people per search must be decreasing in search. Evidence counter to this cannot be reconciled with theories of statistical discrimination. We also exploit periods of martial law, during which the cost of search decreases for all agents, to further test for taste-based discrimination.

The Truth and Reconciliation data on the 1980's Peruvian civil uprising provides an excellent opportunity to test these ideas. Peru is an ethnically diverse country with close to 20% of its population claiming Indigeneous languages as their mother tongue. Due to several waves of migration, the population also claims European, Asian and African heritage. In Peru, being poor and being Indian are thought to be synonymous (see de la Cadena, 2000), so we can use ethnicity to proxy for income. Social divisions in the country are manifested in the composition of the different Armed forces. The Navy and Air Force have always been associated with upper and upper-middle classes, while the Army and the Police are primarily composed of people with middle and lower-middle class backgrounds. Most importantly, the civil conflict placed these different groups in the poorest and more distant areas of the country. This provides an ideal situation to observe the nature of discrimination because groups that might not normally interact with each other did. During the conflict, certain areas of the country were under almost exclusive control by a select few of the Armed Forces.

We show that the distribution of violence in Peru during the period of civil

unrest cannot be solely reconciled with statistical theories of discrimination. While violence on the part of the Armed Forces was widespread, in areas where violence was most intense, the behavior of the Army, Navy and Police towards non-Spanish speakers, especially women, is not consistent with statistical discrimination. These results are shown to hold even after we account for the problem of missing covariate data, which is particularly severe in this data set.

The next section presents the theoretical framework. A brief description of the data is given later. The subsequent sections present the results and test their robustness to missing covariate data problem. We then conclude.

2. Theory and Hypotheses

This section starts with a version of Knowles, Persico and Todd's (2001) model of discrimination to derive testable hypotheses of taste-based discrimination. We introduce an alternative specification that might be relevant for the case of repression of political violence. Also, we discuss alternative models (versions of Anwar and Fang, 2005, and Dominitz and Knowles, 2006, models) and their hypotheses on taste-based discrimination.

A person in group i compares the benefit of joining a violent group, v_i, against the cost of being punished, θs_i. A person joins a violent group (and perform tasks for it) if

$$v_i \geq \theta s_i$$

where v_i is private information drawn from a known distribution F_i. Agents of the state do not observe whether any particular person in group i is guilty and therefore searches anyone with equal probability, s_i. Agents of the state discover guilt after search with certainty, but are only able to punish with probability $\theta \in (0, 1]$. θ represents an institutional parameter determining how much freedom agents of the state have to punish. For example, during periods of martial law, the rules of law are relaxed and agents have greater freedom to punish, $\theta = 1$, however, during normal times $\theta < 1$. The existence of this parameter will be important to our identification strategy. Given these incentives, the likelikood that a person in group i is guilty is $\Pr(G|i) = 1 - F_i(\theta s_i)$.

An agent of the state maximizes the number of guilty people punished. That is, an agent will search a person in group i if the cost of searching that person, τ_i, is below the expected gain,

$$\theta \Pr(G|i) \geq \tau_i.$$

The expression shows that incentives to perform searches increase with the likelihood that a guilty person is punished, θ. An agent of the state will choose a probability an individual in group i is searched, s_i, different from 0 or 1 if $\theta \Pr(G|i) = \tau_i$. That is, in equilibrium, we will have that,

$$\theta(1 - F(\theta s_i)) = \tau_i$$

This expression implies that the probability of a person being guilty increases with the cost of search of group i, τ_i. This also implies that if the cost of search does not vary across groups ($\tau_i = \tau_j$, all i, j), the proportion of people punished given search should be constant across groups. This expression also implies that even if institutional changes allow θ to increase (e.g., agents of the state are allowed to punish more freely), the proportion of people punished given search should remain constant. Following previous literature, we will consider that an agent taste-base discriminates if $\tau_i \neq \tau_j$.

Since repression is expected to be carried out by different agents of the state (say the Army might have better training in certain areas than the Navy), one might consider the possibility that search costs vary across agents of the state but not across groups. If this is the case, we will have that $\theta(1 - F_i(\theta s_i^{Army})) = \tau_i^{Army}$ and $\theta(1 - F_i(\theta s_i^{Navy})) = \tau_i^{Navy}$ for all groups i. That is, the proportion of people being punished given search is expected to be constant for each agent of the state, but may be different across agents of the state. This is consistent with statistical discrimination. Note also that because both agents of the state face the same institutional constraints, θ, it must be the case that those agents searching more intensely are the ones who punish guilty people less frequently.

Anwar and Fang (2005) present an alternative model of discrimination that does not reject statistical discrimination even if variation in search and guilt given search varies across groups for each agent of the state. The authors, however, show that a simple pattern of behavior must be observed, for a given θ, if agents of the state do not taste-base discriminate. Agents with larger cost of search should search less and punish/convict more given search. This test require comparisons across agents of the state for a given institutional setting θ, or across institutional settings, θ and θ', for a given agent of the state.

Recently Dominitz and Knowles (2006) have called into question the premise behind the models of discrimination based on Knowles, Persico and Todd (2001).

The authors consider the case in which an agent of the state's objective is to minimize crime and not to maximize succesful searches. This point is particularly relevant for the case of political uprising where one might think that agents of the state try not only to capture guilty people but to discourage participation in violent activities. An adaptation of Dominitz and Knowles' (2006) model to our context suggests that agents of the state solve the following problem:

$$\max_{s_i} \sum_i -(1 - F_i(\widetilde{s}_i))\eta_i\alpha_i$$
$$\text{subject to}$$
$$\sum_i \widetilde{s}_i\eta_i p_i \leq \theta Y$$

where $\widetilde{s}_i = \theta s_i$, η_i is the size of the population in group i, p_i is the cost of searching group i, and Y is a resource constraint. The parameter α_i denotes a taste for searching group i, or a source of taste-based discrimination. However, the model shows that differential costs of searching a group, p_i, might also represent discrimination. Since $\widetilde{s}_i$ is increasing in θ, increasing the probability of punishing guilty people works as an income effect on search. Dominitz and Knowles (2006) show that even if $p_i = p_j$ for all i, j, it is in general not possible to derive non-parametric tests of discrimination without imposing non-testable hypotheses on the distribution F. While this is true, it is clear that if different agents of the state face the same group i for a given institutional constraint, θ, it must be that those agents searching a group more intensely must also find fewer people guilty and therefore must punish that group less frequently. In addition, it can be shown that the model does not have testable hypothesis across institutional constraints, θ's. The term $\theta(1 - F_i(\widetilde{s}_i(\theta)))$, with $\widetilde{s}_i(\theta)$ being the optimal solution to the crime minimization problem, can increase or decrease with θ even when $\widetilde{s}_i(\theta)$ is increasing in θ.[2]

In summary, the models in this section yield the following predictions. Knowles, Persico and Todd (2001) predict that the probability of being found guilty given search should be constant for a given agent across groups and across institutional constraints, θ and θ'. Anwar and Fang (2005) predict that if agents differ only in their search costs, then the probability of search must be inversely related to that cost and the probability of being found guilty given search must be directly related, holding the group constant. This result also holds across institutional constraints. Finally, Dominitz and Knowles' (2006) model of crime minimization implies that

---

[2]This is possible because the reduction in crime must be smaller than an increase in the likelihood of punishment. We have not found general restrictions on F, say log-concavity, that would imply that this term always be increasing or decreasing in θ.

holding an institutional constraint, θ, constant, agents of the state who search a group more intensely must also find fewer guilty people per search. This result does not necessarily holds across institutional constraints. The following sections will discuss assumptions necessary to make these tests implementable.

3. Methodology

The models above suggest that the ideal data to test for discrimination must include the proportion of people found guilty of terrorism given that they were searched by an agent of the state. Unfortunately, no such data exist. The next best measure of guilt is the number of people killed extrajudicially by agents of the state and the number of people currently missing and last seen in the custody of the agent. From now on, we will refer to these two cases as killing.[3] Our measure of search is if a citizen reported any act perpetrated by any agent of the state. This measure include reports on imprisonment and interrogation. The data do not tell us if a citizen was briefly stopped and let go.

The previous section suggests some tests of discrimination. For a person of group i, let y_i be equal to 1 if a person is guilty of committing acts of terrorism and 0 otherwise, and let j, k denote the agent to whom an intervention has been attributed. θ is the probability that a guilty person is punished, and let s be equal to 1 if a person is searched and 0 otherwise.

The first test of discrimination is that:

$$E[\theta y_i | j, \theta, s = 1] = E[\theta' y_i | k, \theta', s = 1] \text{ for all } i, j, k, \theta, \theta'$$

That is, if all agents of the states have the same cost of search, they must equalize the return per search if they are maximizing successful searches. If different agents of the state have different search costs, the test requires that:

$$E[\theta y_i | j, \theta, s = 1] = E[\theta' y_i | j, \theta', s = 1] \text{ for all } i, k, \theta, \theta'$$

Under weaker assumptions, as in Anwar and Fang (2005), the test can be modified as follows:

---

[3]Rarely would any agent of the state ever say that killings were made without a good reason (on average). This paper is therefore not about whether human rights violation occured, which is clear from the data, but whether human rights violations reveal distate for a particular segment of the population.

Either $E[s_i|j,\theta] \leq E[s_i|k,\theta']$ and $E[\theta y_i|j,\theta,s=1] \geq E[\theta' y_i|k,\theta',s=1]$ for all i,j,k,θ,θ'

or $E[s_i|j,\theta] \geq E[s_i|k,\theta']$ and $E[\theta y_i|j,\theta,s=1] \leq E[\theta' y_i|k,\theta',s=1]$ for all i,j,k,θ,θ'

This test says that agents with lower search cost must search more intensely and therefore find fewer people guilty. This ranking must hold across institutional constraints. That is, holding a group constant, more intense search must be associated with less successful search.

Finally, assuming that agents of the state minimize crime rather than successful search, as in Dominitz and Knowles (2006), further relaxes the conditions for detecting taste-based discrimination:

Either $E[s_i|j,\theta] \leq E[s_i|k,\theta]$ and $E[y_i|j,\theta,s=1] \geq E[y_i|k,\theta,s=1]$ for all i,j,k,θ

or $E[s_i|j,\theta] \geq E[s_i|k,\theta]$ and $E[y_i|j,\theta,s=1] \leq E[y_i|k,\theta,s=1]$ for all i,j,k,θ

This test says that the ranking of searches and successful searches must hold only for a given group and a given institutional constraint, θ, across agents of the state.

The covariate data we use is sex and mother language (i.e. Spanish or non-Spanish). Mother language is highly correlated with social status and ethnicity. The mother language of indigenous people is not Spanish.

In our analysis, we look at the extent and nature of discrimination. To look at the extent of discrimination, we look at differential treatment in the probability of being killed, $E[y|i,j,\theta]$, and the probability of being searched, $E[s|i,j,\theta]$, during periods of martial law and during normal times. The probability of being killed is calculated by taking the number of observations of people killed with covariates i by agent j in time period t (i.e. martial law or not) in a particular district and weighing it by the population in 1981 of people with covariates i in that district.[4] The probability of being searched is calculated in the same fashion, but we use the number of people reporting an act, instead of the number of people killed. The probability calculation only includes the districts in which a particular agent of the state was active.[5] Agents were often only active in certain districts, so this gives more accurate comparisons across agents.

---

[4] Peru is made up of 24 departments, and departments in Peru are divided up into provinces. Provinces are further divided up into districts.

[5] We do not include in the denominator population counts of districts in which the agent was not active.

To look at the nature of discrimination, we test whether behavior is consistent with the models of statistical discrimination presented above. If our tests do not hold, this suggests taste-based discrimination. To calculate the probability of being killed given search, $E[\theta y | i, j, \theta, s = 1]$, we count the number of observations of people killed with covariates i by agent j in time period t (i.e. martial law or not) in a particular district and weight it by the number of observations of acts reported by people with covariates i by agent j in time period t (i.e. martial law or not) in that district. In our empirical analysis, we look across regions, so probabilities by district are averaged across all districts in that region. This applies to all calculated probabilities.

The following sections present the data and results.

4. Overview of the Data

The data we use are from the Peruvian Truth and Reconciliation Commission and from the 1981 Peruvian Census. The former data were gathered as the Commission travelled around the country during 2001-2003 and invited victims, survivors, family members and neighbors to report any acts of violence during the period of 1980-2000. Acts of violence include killing, disappearance, torture, detention, and rape. The majority of violent acts fall under the category of killing and disappearance, and it is those reported acts that we focus our analysis on. Since the data are only reported acts, the data are considered a lower bound measure on the number of violent acts that occurred during 1980-2000. Each act includes subrecords detailing if a person was detained, tortured, disappeared, etc. Our measure of search is the total number of separate acts and our measure of "success" is the proportion of those acts that ended up in extrajudicial execution or disappearance. Note that a large portion of the data has missing covariate data (e.g. on level of education and age), but the sex of the individual is typically always known.[6] While the collected data span from 1980-2000, we restrict the data to the period 1980-1993 when the terrorist movement was at its strongest.[7]

We use district-level census data from 1981 broken down by sex and language

---

[6]The data set of acts we use in this paper is a subsample of all registered acts. We do this because this subsample has been cross-validated and therefore is less likely to suffer from measurement problems.

[7]The nature of the terrorist movement changed dramatically in 1993 with the capture of the leader of Shining Path, Abigail Guzman, the previous year and the suspension of the democratic regime.

(i.e. Spanish or non-Spanish speaker) to establish differential treatment across groups. Census data for language are available for all departments in Peru but two (Apurimac and San Martin). These departments are excluded from our estimations of killings and search weighted by population data. While we do not have population data from these departments, the Truth and Reconciliation data do contain observations from these departments. The results of our tests for discrimination using the probability of being killed given search do not change by the inclusion of these departments.

Another important element in our analysis is the identification of periods of martial law, or as they are called in Peru, states of emergency. A state of emergency means that the state cannot guarantee a citizen's rights and agents of the state are given more political power. During these periods, we would expect an increase in the number of human rights violations. Through the course of the political violence in Peru, there were various states of emergency, both at the national and regional level. We have classified each event from the T&R data as having occurred during a state of emergency or not based on where and when the event took place. For instance, according to the compilation of states of emergency made by Landa (1995), the department of Ayacucho and Huancavelica were under state of emergency 38% of the time. In our data, 74% of the incidents of violence attributed to the Army, Navy and Police occurred during states of emergency. Similarly, the departments of Huanuco, Loreto and Ucayali were under state of emergency 28% of the time, and 58% of the registered incidents occurred during states of emergency.

The Truth and Reconciliation Commission roughly divides agents of the state into the three main Armed Forces (Army, Navy, and Air Force) and the Police. We concentrate our analysis on the Army, Navy and Police since these were the groups most involved in the political violence.

The Final Report (Peruvian Truth and Reconciliation Commission, 2004) states that the total number of people killed during the eighties and early nineties might have reached 70,000. The majority of the cases are attributable to Shining Path, a terrorist group behind the uprising. Approximately 30,000 are believed to be the responsibility of agents of the state.[8]

---

[8]The Truth and Reconciliation Commission arrived at this conclusion using much-debated statistical methods. The actual number of casualties recorded is about 30,000. This paper only uses reported cases, i.e., cases with a paper trail.

Table 2

Descriptive Statistics

Types of Crimes Committed and Demographics by Perpetrators

	State	Terrorists
Killed	29.1%	69.9%
Killed and Missing	49.0%	75.4%
Tortured	38.4%	12.7%
Men	77.1%	78.1%
Primary Education*	16.0%	19.9%
Non-Spanish Speakers*	68.6%	73.6%
* as percentage of observed		

Table 2 shows a brief summary of the types of crimes committed and the demographic composition of the victims. Terrorist groups were more prone to killing than agents of the state. Two thirds of those coming across terrorists were killed, more than twice the probability corresponding to agents of the state. There are a large number of cases of missing people, and while these people are not classified as killed, this is most likely because there was no body found to verify that the person was killed.[9] To be more inclusive with the number of people most likely killed during the terrorist movement, we combine killed and missing together to have a broader measure.[10] All the analysis in the paper uses this combined classification. Also, as Table 2 shows, most of the victims were men and non-Spanish speakers.

---

[9]For example, many people were killed during the violence and their bodies were thrown into rivers. It was impossible to find the bodies or identify them. So, while the person may have actually been killed, he or she would be classifed as missing in our data.

[10]No person was classified as both killed and missing in the Truth and Reconciliation data.

Table 3
Percentage of the Population in 1981
Reporting a case by Agent of the State

Department	Army	Navy	Police
Amazonas	0.03%		0.04%
Ancash	0.02%	0.05%	0.12%
Arequipa	0.02%		
Ayacucho	0.67%	0.15%	0.32%
Cajamarca			0.14%
Cusco		0.03%	0.08%
Huancavelica	0.10%		0.13%
Huanuco	0.32%		0.10%
Ica	0.30%	0.03%	0.07%
Junin	0.14%	0.01%	0.07%
La Libertad	0.11%		0.08%
Lambayeque	0.01%		0.17%
Lima	0.01%		0.01%
Loreto	0.01%	0.01%	0.01%
Pasco	0.08%	0.03%	0.35%
Piura	0.16%		0.05%
Puno	0.13%		0.28%
Ucayali	0.09%	0.10%	0.03%

Table 3 shows that the intensity of involvement of different agents was not the same during the civil war in Peru.[11] Table 3 shows a sample of the departments that suffered from terrorism and the variance in the data. For example, on average, a person was four times more likely of having an encounter with the Army than having an encounter with the Navy in Ayacucho. And, a person was twice as likely of having an encounter with the Army than with the Police in the same department. However, in Ucayali, the Army and the Navy were involved in similar number of cases. While this is informative, it is an imperfect measure of searches by different agents of the state because we do not have data on the number of agents of the state in each region during the conflict.

---

[11] The numerator is the number of cases attributed to an agent of the state and the denominator is the population in 1981 in the areas where the agent was active.

5. Testing for Discrimination

We look first at differential treatment of various groups by agents of the state. Second, we examine if differential treatment is consistent with statistical discrimination.

5.1. Extent of Discrimination

We start first by looking at differential treatment in the probability of being killed and the probability of search given covariates by various agents of the state. Table 4 shows the probability of being killed for the two regions in which terrorist activity was the most active. The probabilities are estimated by taking into account the population in the area in 1981 prior to the civil war.[12] Probabilities are listed during a state of emergency and during a non-state of emergency. Recall that we assume data were missing at random, so these are point estimates.

---

[12] We only consider the districts in which there are reported acts and aggregate up to the regional level. So, for example, if there are three districts, in a region of 100 districts, in which there are reported acts, we weigh those reports by the population in those districts and aggregate up to the region. This means we are not counting population in districts in which there was no political violence.

Table 4
Probability of Being Killed
South Central Region: Ayacucho/Huancavelica

	States of Emergency			Non-States of Emergency		
	Army	Navy	Police	Army	Navy	Police
Women, Spanish	0.004	0.009	0.005	0.010	0.000	0.012
	(0.004)	(0.009)	(0.005)	(0.007)	(0.000)	(0.008)
Women, Non-Spanish	0.062	0.034	0.038	0.033	0.002	0.020
	(0.005)	(0.006)	(0.005)	(0.004)	(0.002)	(0.004)
Men, Spanish	0.074	0.007	0.042	0.012	0.021	0.014
	(0.016)	(0.007)	(0.013)	(0.007)	(0.021)	(0.008)
Men, Non-Spanish	0.269	0.070	0.177	0.135	0.052	0.067
	(0.012)	(0.009)	(0.011)	(0.009)	(0.011)	(0.007)

Northeast Region: Ucayali/Loreto/Huanuco

	States of Emergency			Non-States of Emergency		
	Army	Navy	Police	Army	Navy	Police
Women, Spanish	0.016	0.003	0.001	0.011	0.000	0.000
	(0.004)	(0.002)	(0.001)	(0.208)	(0.000)	(0.000)
Women, Non-Spanish	0.052	0.043	0.000	0.044	0.000	0.000
	(0.014)	(0.030)	(0.000)	(0.482)	(0.000)	(0.000)
Men, Spanish	0.091	0.033	0.013	0.081	0.028	0.009
	(0.008)	(0.007)	(0.004)	(0.243)	(0.295)	(0.202)
Men, Non-Spanish	0.277	0.326	0.058	0.185	0.137	0.019
	(0.032)	(0.073)	(0.024)	(0.533)	(1.114)	(0.616)

Note: standard errors in parentheses

There are notable patterns in the data. First, men are more likely to be killed than women and non-Spanish speakers are more likely to be killed than Spanish speakers. These differences are significant for all agents in states of emergency in both regions.[13] In non-states of emergency, this holds for all agents in the South Central region only. Second, the Army is significantly more likely to kill any group than the Navy in both states of emergency and non-states of emergency and in both regions. The Army kills more non-Spanish speakers than the Police in states of emergency in both regions.

---

[13]We use a t-test to test for difference in means for specific groups and a chi-test to do a joint test of difference in means across several groups. Significance is measured by a p-value strictly less than 0.10. All results noted in the text of the paper are significant.

Comparing across periods of states of emergency and non-states of emergency, the Army is significantly more likely to kill men and non-Spanish speaking women in states of emergency in the South Central region. The Police also kill more men and non-Spanish speaking women during states of emergency in the South Central region. The Navy is significantly more likely to kill women in states of emergency in both regions.

Table 5 shows the probability of being searched. Since we do not know the number of searches, we measure the probability of search by any act committed against a certain group weighted by the population in 1981.

Table 5
Probability of Being Searched
South Central Region: Ayacucho/Huancavelica

| | States of Emergency | | | Non-States of Emergency | | |
	Army	Navy	Police	Army	Navy	Police
Women, Spanish	0.018	0.009	0.011	0.010	0.031	0.024
	(0.009)	(0.009)	(0.007)	(0.007)	(0.031)	(0.012)
Women, Non-Spanish	0.110	0.037	0.058	0.054	0.009	0.034
	(0.007)	(0.007)	(0.006)	(0.006)	(0.005)	(0.005)
Men, Spanish	0.107	0.021	0.059	0.038	0.021	0.036
	(0.019)	(0.012)	(0.016)	(0.012)	(0.021)	(0.013)
Men, Non-Spanish	0.363	0.087	0.240	0.196	0.055	0.103
	(0.014)	(0.010)	(0.012)	(0.011)	(0.011)	(0.009)

Northeast Region: Ucayali/Loreto/Huanuco

| | States of Emergency | | | Non-States of Emergency | | |
	Army	Navy	Police	Army	Navy	Police
Women, Spanish	0.021	0.007	0.007	0.021	0.004	0.003
	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.002)
Women, Non-Spanish	0.082	0.065	0.022	0.084	0.049	0.000
	(0.018)	(0.037)	(0.015)	(0.019)	(0.035)	(0.000)
Men, Spanish	0.122	0.052	0.028	0.104	0.052	0.020
	(0.010)	(0.009)	(0.006)	(0.009)	(0.009)	(0.004)
Men, Non-Spanish	0.402	0.392	0.175	0.273	0.215	0.047
	(0.038)	(0.080)	(0.041)	(0.034)	(0.065)	(0.021)

Note: standard errors in parentheses

14

Similar patterns emerge in the probability of search. First, men are significantly more likely than women to be searched by all agents in any period of time in both regions. Second, in states of emergency, all agents are significantly more likely to search non-Spanish speakers than Spanish speakers in both regions. This also holds in non-states of emergency in the Northeast region. The patterns are less clear in the South Central region for women, but non-Spanish speaking men are more likely to be searched than Spanish-speaking men in non-states of emergency.

Overall, the Army is significantly more likely to search than the Police, and the Police are more likely to search than the Navy during states of emergency and non-states of emergency in the South Central region. These patterns also hold in the Northeast region for both states of emergency and non-states of emergency. Regarding levels of search across states-of-emergency and non-states-of-emergency, all but the Navy in the Northeast region show a significant increase in the level of activity.

Clearly, there is differential behavior towards different groups by agents of the state. This discrimination in behavior might be statistical or taste-based. We turn to that next.

5.2. Nature of Discrimination

We have seen that killing and search vary by agents and characteristics of the victims. Is this differentiation due to taste-based discrimination? The theory section makes several predictions of behavior consistent with statistical discrimination for agents of the state. We examine this next.

Table 6 shows the probability of being killed given search for various groups across agents of the state. First, we examine if the proportion of guilty people given search is constant by group and state-of-emergency for a given agent of the state in a given region. We hold the region constant for the reasons that geography and resources change across geographical areas. The joint test that the probability of being killed given search is the same for each group across states of emergency is rejected for all agents of the state in the South Central region. In the Northeast region, the Navy is the only group for which we can reject the hypothesis that the probability of being killed given search is constant. This variation in behavior is inconsistent with a model of discrimination where agents maximize "successful" searches (such as Knowles, Persico and Todd, 2001). Also, it is not consistent with a more general version of the model either (such as Anwar and Fang, 2005).

15

Table 6
Probability of Being Killed Given Search
South Central Region: Ayacucho/Huancavelica

	States of Emergency			Non-States of Emergency		
	Army	Navy	Police	Army	Navy	Police
Women, Spanish	25.00	100.00	50.00	100.00	0.00	50.00
	(21.65)	(0.00)	(35.36)	(0.00)	(0.00)	(25.00)
Women, Non-Spanish	55.79	90.63	66.33	60.64	25.00	60.42
	(3.25)	(5.15)	(4.77)	(5.04)	(21.65)	(7.06)
Men, Spanish	68.75	33.33	71.43	30.00	100.00	37.50
	(8.20)	(27.22)	(12.07)	(14.49)	(0.00)	(17.12)
Men, Non-Spanish	74.05	80.28	73.78	68.63	95.65	64.44
	(1.67)	(4.72)	(2.29)	(2.65)	(4.25)	(4.12)
Total	63.73	75.32	62.20	61.51	84.51	54.49
	(1.06)	(2.81)	(1.69)	(1.80)	(4.29)	(2.82)

Northeast Region: Ucayali/Loreto/Huanuco

	States of Emergency			Non-States of Emergency		
	Army	Navy	Police	Army	Navy	Police
Women, Spanish	76.00	50.00	20.00	52.00	0.00	0.00
	(8.54)	(25.00)	(17.89)	(9.99)	(0.00)	(0.00)
Women, Non-Spanish	63.64	66.67	0.00	52.63	0.00	0.00
	(10.26)	(27.22)	(0.00)	(11.45)	(0.00)	(0.00)
Men, Spanish	74.53	64.71	47.83	77.21	53.13	42.86
	(3.43)	(8.20)	(10.42)	(3.60)	(8.82)	(10.80)
Men, Non-Spanish	68.81	83.33	33.33	67.69	63.64	40.00
	(4.44)	(7.61)	(11.11)	(5.80)	(14.50)	(21.91)
Total	68.72	72.97	28.36	67.08	45.10	32.69
	(2.22)	(5.16)	(5.51)	(2.62)	(6.97)	(6.51)

Note: standard errors in parentheses

Second, we look at search patterns and the probability of guilt given search. As shown in Table 5, the intensity of search across most groups increased during the states of emergency, implying that fewer successful cases should be encountered.[14] Let's consider the South Central region first. In this region, all groups

---

[14]Recall that agents of the state take into consideration the probability of punishing a guilty person while searching. This would affect the overall level of search only.

significantly increased search of non-Spanish speakers during states of emergency. Our models predict that all groups should decrease killing given search during states of emergency. Looking at Table 6, we see that the probability of being killed by the Army for non-Spanish speaking women is 60.64% in non-states of emergency and 55.79% in states of emergency. So, the probability of killing has gone down, as predicted. However, this does not hold for the Police and the Navy. The probability of being killed by the Police goes up, but not significantly, and the probability of being killed by the Navy is 25.00% in non-states of emergency and increases significantly to 90.63% in states of emergency. This behavior towards non-Spanish speaking women is not consistent with our models of statistical discrimination. By contrast, the behavior of the Army, Navy and Police towards non-Spanish speaking men is consistent.

In the Northeast region, there is less behavior that is inconsistent with statistical discrimination. For example, search does not increase during states of emergency for any group by any agent, and killing given search does not change. The only exception to this is for non-Spanish speaking men. In this case, the Army, Navy and Police all increase search of non-Spanish speaking men during states of emergency, however there is no significant decrease in killing given search. Statistical discrimination would predict that killing given search should decline.

Third, we look at search and probability of guilt given search within a state of emergency or a non-state of emergency (θ fixed). Looking at the South Central region, Table 5 shows that there is an ordering of search. Given a θ, the Army searches more than the Police, and the Police search more than the Navy.[15] We would, therefore, expect monotonicity in killing given search across agents of the state for a fixed θ. For example, looking at states of emergency, the Army kills significantly fewer women than the Navy, but it kills siginficantly more Spanish-speaking men. Also, the Navy kills significantly more women than the Police, but there is no significant different for men. The Police kill significantly more non-Spanish speaking women than the Army, but not for other groups. These switches across groups for different agents cannot be explained by our models of statistical discrimination. Indeed, this result is not even consistent with a model of crime minimization (as in Dominitz and Knowles, 2006).

Looking now at the Northeast region, we also see some patterns that are inconsistent with our model. In states of emergency, the Army searches more Spanish speakers than the Navy, but there is no significant difference in killing given search. In non-states of emergency, the Army also searches Spanish speakers

---

[15] The only exception is for Spanish-speaking women in non-states of emergency.

significantly more than the Navy, and the Army kills significantly more Spanish speakers than the Navy. None of this behavior is consistent with our models.

In sum, we observe some behavior by the Army, Navy and Police that is inconsistent with statistical discrimination. This occurs in both regions and is slightly stronger in the South Central region where the civil war was the most intense and earlier. In this area, most of the discrimination is aimed towards non-Spanish speakers, especially women.

5.3. Robustness

A concern with the results of the previous section is that they are based on the subsample of acts for which covariate data is available. This section restricts the analysis only to gender and region since this data is available for almost all observations. Table 7 presents the results for the intensity of search and the probability of being killed given search. We confirm some of the general trends. Search increases during states of emergency but is more pronounced in the South Central region than in the Northeast. All these differences are significant with the exception of the Navy in the Northeast region.

Table 7

Intensity of Search and Probability of Being Killed Given Search

South Central Region: Ayacucho/Huancavelica

	States of Emergency			Non-States of Emergency		
	Army	Navy	Police	Army	Navy	Police
	Probability of Search					
Women	0.26	0.08	0.09	0.09	0.03	0.05
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Men	0.66	0.16	0.37	0.31	0.13	0.16
	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
	Probability of Being Killed Given Search					
Women	56.80	72.97	57.95	50.56	58.33	5.39
	(1.99)	(5.16)	(3.72)	(3.73)	(14.23)	(3.24)
Men	66.23	75.00	63.13	64.73	89.83	54.43
	(1.25)	(3.51)	(1.88)	(2.01)	(3.93)	(3.24)

Northeast Region: Ucayali/Loreto/Huanuco

	States of Emergency			Non-States of Emergency		
	Army	Navy	Police	Army	Navy	Police
	Probability of Search					
Women	0.05	0.01	0.02	0.04	0.01	0.01
	(0.00)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)
Men	0.24	0.10	0.06	0.18	0.08	0.04
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
	Probability of Being Killed given Search					
Women	65.82	55.56	7.69	51.61	0.00	20.00
	(5.34)	(16.56)	(7.39)	(6.35)	(0.00)	(12.65)
Men	67.81	75.36	31.67	71.43	56.86	35.56
	(2.40)	(5.19)	(6.00)	(2.73)	(6.94)	(7.14)

Note: standard errors in parentheses

As before, we find patterns that are difficult to reconcile with statistical discrimination. In the South Central region, we see that while the Police search in the same intensity as the Navy during non-states of emergency, the proportion of women being killed given search is significantly higher for the Navy. This pattern is also at odds with the fact that the Army searches women more intensely than the Police in these periods but kills more. We also observe that during states of emergency, the Police search women about three times less frequently than the

Army, but the cases of people killed given search is no different. Notice also the positive correlation between the intensity of search of women and the probability of being killed as we pass from non-states of emergency to states of emergency.

Compared to the disaggregated results for the Northeast region, the evidence for this region from Table 7 gives stronger evidence at odds with statistical discrimination against women. The Army searchs women more intensely than the Navy and the Police both in and out of states of emergency. However, the Army kills a larger proportion of women given search. This evidence is not consistent with either maximization of succesful searches nor minimization of crime.

6. Conclusions

Differences by race and class permeate social relations in the market and in the political scene. One expects to see these crevasses most clearly whenever the rule of law vanishes. Civil wars, conflicts and catastrophes therefore provide an excellent opportunity to assess the nature and extent of discrimination. Using a unique data set on human rights violations collected by the Peruvian Truth and Reconciliation Commission, we find that the pattern of violence cannot be reconciled with current models of statistical discrimination. This is true even if we consider different rationales behind the exercise of violence.

We find several key patterns in the data. There was clearly differential treatment across groups by agents of the state. Men and non-Spanish speakers were more likely to be killed and searched by any agent of the state. This result holds in periods of martial law and periods without martial law. We exploit changes in institutional environment to then test the nature of discrimination. In particular, we find that behavior towards non-Spanish speakers, especially women, cannot be reconciled with theories of statistical discrimination and seems to be driven by taste-based discrimination. This was particularly pronounced in the South Central region of the country where the terrorist activity was most pronounced and earlier.

The behavior we observe in this extreme setting is revealing. Wars are periods of time in which emotions are raw and people respond on instinct. That we observe patterns of behavior inconsistent with statistical discrimination suggests that there are underlying sentiments of animus towards certain segments of the population. We expect this would be difficult to observe outside of these extreme periods of time, especially since many of these groups may not normally interact with one another. While perhaps not blatantly obvious, if these underlying sentiments

exist, they may still manifest themselves in differential hiring practices, segregated neighborhoods, or limited educational opportunities.

Designing policies to address these sentiments is more challenging. If differential treatment were due to incorrect beliefs on behavior, accurate, credible information on behavior might be all that is needed to mitigate this. However, if differential tratment is due to animus, then this should be taken into account. Understanding the source of differential treatment is the first step towards designing successful policies aimed at giving equal opportunities across groups.

References

Antonovics, K. and B. G. Knight, "A New Look at Racial Profiling: Evidence from the Boston Police Department," forthcoming, *Review of Economics and Statistics*, 2007.

Anwar, S. and Fang, H., "An Alternative Test of Racial Prejudice in Moter Vehicle Searches: Theory and Evidence," *American Economic Review*, 96(1), 127-151, 2006.

Bertrand, M. and S. Mullainathan, "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *American Economic Review*, 94, 4, pp. 991-1013, 2004.

de la Cadena, Marisol, *Indigenous Mestizos: The Politics of Race and Culture in Cuzco, Peru.* Duke University Press, 2000.

Dominitz, J. and J. Knowles, "Crime Minimisation and Racial Bias: What Can We Learn from Police Search Data?," *Economic Journal*, 116, F368-F384, 2006.

Heckman, J. "Detecting Discrimination," *Journal of Economic Perspectives*, 12, pp. 101-16, 1998.

Horowitz, J. L. and C. F. Manski, "Censoring of Outcomes and Regressors Due to Survey Nonresponse: Identification and Estimation Using Weights and Imputations," *Journal of Econometrics*, 84, pp. 37-58, 1998.

Knowles, J., N. Persico, P. Todd, "Racial Bias in Motor Vehicle Searches: Theory and Evidence," *Journal of Political Economy*, 109, pp. 203-29, 2001.

Landa Arroyo, César, "Base de Datos Sobre Estados de Emergencia en el Perú : documento de trabajo," Pie de Imprenta: Lima : PUCP, 1995

Riach, P. A., and J. Rich, "Field Experiments of Discrimination in the Market Place," *Economic Journal*, 112, pp. 480-518, 2002.

Peruvian Truth and Reconciliation Commission, *Final Report*, 2004.