Extermination as a substitute for assimilation or deportation: an economic approach

Mario Ferrero
EXTERMINATION AS A SUBSTITUTE FOR ASSIMILATION OR DEPORTATION: AN ECONOMIC APPROACH

Mario Ferrero

Institute of Public Policy and Public Choice (POLIS)
University of Eastern Piedmont
Via Cavour 84
15100 Alessandria (ITALY)
Fax: +39.0131.283704
Email: ferrerom@sp.unipmn.it

ACKNOWLEDGMENT. Earlier drafts of this paper were presented at the UCSIA workshop on “Preventing Genocide: Root Causes and Coping Strategies” (University of Antwerp, November 24-25, 2011) and at the 13th Jan Tinbergen European Peace Science Conference (Catholic University of Milan, June 24-26, 2013); participants in both events provided interesting discussion. The author is indebted to Henry Theriault and Ugur Ungor for references and advice on the Armenian genocide and to Petros Sekeris for general comments.

KEYWORDS: extermination; exile; deportation; assimilation; cost minimization.

JEL codes: D74; F51

ABSTRACT. This paper places genocide or mass murder in a continuum of actions that a ruling power can take to remove an unwanted group from a society; that is, it views extermination as a means to an end, and it assumes that perpetrators are rational in the sense that they will choose the combination of means that can achieve the goal at the minimum cost to themselves. The means are assimilation into the general society, physical removal from view (which may involve either deportation within the country or exile from the country), and extermination. The available options and their costs will depend on the type of group, viz.: ethnic/national/racial, religious, income/property class, political. After developing the theoretical framework, the paper surveys a range of historical case studies from different types of group and finds good support for the cost-minimization hypothesis. In particular, it finds that in most cases the choice of means is an interior solution, as the hypothesis would lead one to expect, and that the chosen combination shifts as relative costs change.
EXTERMINATION AS A SUBSTITUTE FOR ASSIMILATION OR DEPORTATION:
AN ECONOMIC APPROACH

1. Introduction

This study places genocide or mass murder in a continuum of actions that a ruling power can take to remove an unwanted group from a society. The economics here has nothing to do with the perpetrators’ motives, as some scholars of genocide have suggested; for it still seems to be incumbent on those scholars to show that to grab the wealth of a group, nothing short of killing them all would do. Rather, the economic approach has to do with the perpetrators’ choice of means to achieve a given goal: it assumes that perpetrators are rational in the sense that they will choose the combination of means that can achieve the goal at the minimum cost to the perpetrators; so genocide or mass murder will be observed to occur only when resort to alternative means turns out to be more costly.

Specifically, suppose a ruler wants to remove a group. There are three ways in which this can be done: assimilation into the general society, physical removal from view (which may involve either deportation to a separate place or exit from the country), and extermination. Because many individuals and groups freely choose to assimilate to the majority in many societies, what we include here among the means of removal at the disposal of a ruler is forced assimilation, over and above the rate that would result from free individual choice. Availability of each option depends upon what dimension defines the target group. Although many groups are defined by several dimensions, four basic dimensions, or types of group, are of interest: ethnic/national/racial, religious, income/property class, political. Clearly not all options are available for all groups: for example, a religious group can assimilate (i.e. convert to the dominant religion) while a strictly racial group cannot. Even when the same set of options is available for two groups, however, the costs to the ruler will be different from case to case.

Suppose rulers are completely cynical: they would not flinch at killing as many people as necessary but would do so only if extermination is the least-cost way of removing the target group. If so, attention must focus on the costs involved by each option for each type of group, including not only the direct costs (such as running the death camps or the conversion procedure) but also, and especially, the indirect, long-term costs such as: having around false converts or disloyal people; losses of human capital; missed opportunities to extract slave labour from prisoners before, or instead of, killing them; the possibility that the exiles may stage a comeback or fuel international opposition to the ruler; the possibility that a day of reckoning may eventually come for perpetrators of mass murder. Seen this way, options become means to an end, or inputs used to produce an output which is the group’s removal, and the ruler is an agent who seeks to minimise the cost of achieving the desired output. As is generally the case in economics, the expectation is that the cost-minimizing input choice will be an ‘interior’ solution combining different means to the same end in some degree; thus in most cases the ruler will not exclusively rely on one means but exploit a variety of options.

This paper takes a first step to estimating repression costs for the ruling power in each type of situation; this will be carried out by surveying a selection of historical cases for each type of group. For example, for class groups, we will test the hypothesis that Pol Pot, who killed huge numbers, may have been no more and no less mean than Fidel Castro, who let them go: the cost difference may account for the different observed choices. Likewise, for ethnic groups, the hypothesis to be tested is that Hitler, who killed huge numbers, may have been just as mean as the Young Turks, who deported their victims while killing many of them in the process, again based on cost differences. If, however, cost differences cannot account for the different outcomes, then we are forced to conclude that preferences must have been different –
for example, some rulers were more restrained than others at killing. Needless to say, beyond this first stab, much further work will be needed to test the usefulness of our approach.

It must be stressed that we will never question the goal itself, i.e. ask why the ruler has decided that the target group is to be removed from society rather than accommodated, bargained with, or kept in a subordinate, exploitative relationship. We take the ruler’s goal as a given, for whatever reason – even though, as many scholars have pointed out, it is usually the case that removal of a group is not an end in itself but is instrumental to a further, overriding objective: the maintenance and enhancement of political power. Even so, as is proper to an economist, we deal with the rationality of the means, not of the ends. The new twist of this paper is to place extermination among the means, not the ends.

To the best of my knowledge, no such project has ever been attempted in the relevant literature; specifically, no one has looked at genocide or extermination as just an alternative to deportation, exile and conversion as means to reach the same goal. However, historians have often found just that in their study of specific cases, so we can do no better than use their work. It must be emphasized that the nutshell historical descriptions that comprise the empirical part of this paper are in no case meant to be comprehensive, or even adequate, accounts of each historical episode; they are selectively focused on those aspects that are relevant to the specific questions we ask and the economic approach we take. We will typically consider a state government’s dealings with a minority group, but will not address fighting between groups such as mob violence or civil war. Also, we will generally use the word ‘extermination’ rather than ‘genocide’ so as to get around the legal limitations of the concept as embodied in the UN genocide convention and be able to address class and political groups.

The remainder of the paper is organized as follows. The next section lays out a simple economic model of choice among alternative means of removal. The following three sections address a sample of cases from each of the three principal types of group, viz. ethnic/racial groups, religious groups, class/political groups, in this order. The final section brings together our findings and concludes.

2. An economic model of group removal

Assume that the ruler cares only about the number of people “removed” from among the target group; removals are his objective, which can be achieved by several means, singly or in combination. A useful way to think about the choice involved is to describe the ruler’s problem as a production problem, in which the output (the removals) can be produced by employing different inputs or combinations thereof: typically, in our setting, killing, exile, deportation, and possibly assimilation (depending on what kind of group we are dealing with). These input combinations that produce the same output do not however, as a rule, have the same cost to the ruler: for example, in a given case, killing 30 and forcing out 70 may be the

---

1 This narrower, and less ambitious, notion of rationality is the main difference between our approach and the “rational choice” theories of ethnic war and genocide discussed and criticized by Kaufman (2001, 2006), whose range of application partly overlaps with the cases examined in Section 3 below. A key implication of this difference is that we do not address the larger question of whether the chosen strategy was ex-post “rational” in the sense of strengthening the ruler’s power, which is one of the main points of Kaufman’s critique – and indeed in one of the cases examined below and used as a key test by Kaufman (2006), that or Rwanda, it is clear that the choice of genocide led directly to the ouster of the perpetrator regime. Arguably, a rational choice of means in the sense developed in this paper not only is compatible with, but should be an integral part of, both rational choice and symbolic politics theories.

2 This section is a straightforward application of introductory-textbook production theory. Readers with some training in microeconomics can skip it entirely.
least-cost way of disposing of 100 people. Then, if the ruler cares only about the removals and not the means per se, we should observe him resorting precisely to that combination of means.

The cost to the ruler of using each (combination of) inputs to do the job will depend on how effective the input is in turning out output – its “productivity” – and how difficult or expensive it is for the ruler to muster henchmen and other necessary resources. So for example the productivity of killing will be different if it is done by pickaxes or machine guns, the productivity of deportation will vary with the difficulty of rounding up the deportees, the productivity of assimilation will depend on how “distant” the minority group is from the majority and how easy it is to fake assimilation and cheat the assimilator. Executors may be squeamish at killing children and women so that carrying out a thorough extermination may require higher compensation or resort to expensive devices to camouflage the act; similarly, the cost of internal deportation will vary with the availability of unpopulated land, the cost of emigration will hinge on how valuable the lost manpower is to the ruler, and the cost of assimilation will depend on how damaging it is thought to be to have faked converts around. External factors are also very important: exiles may or may not be expected to stage a comeback from their host countries, foreign agencies or governments may or may not be expected to punish the perpetrators or otherwise retaliate against them, and so on. The historical cases reviewed in the following sections provide evidence of wide variation along all these dimensions. So in the end the relative costs of different removal strategies are necessarily specific to a given point in time and space.

However, there is one generalization that it is fairly safe to make: the productivity of each additional unit of each input will decrease, and/or its cost will increase, as the number of units used increases – what economists call the principles of diminishing marginal product and increasing marginal cost. This is simply due to the fact that hunting down the last runaway victims, resettling more and more batches of deportees, driving executioners to carry the killing beyond “battle-age” males, converting the most refractory subjects – all of this requires more, or more expensive, resources than dealing with the large initial numbers of victims does. It follows that the total cost of reliance on any one input alone will tend to increase more than proportionally to the output. This by itself tends to drive the cost-sensitive ruler’s choice toward an “interior” solution, i.e. one that combines different inputs in some proportion, rather than a “corner” solution, i.e. one that relies on the exclusive use of one input. The latter is still a possibility consistent with the desire to minimize total cost, but only in rather special situations in which use of all the other inputs turns out to be exceedingly costly from the very first units.

The choice problem facing the ruler can be illustrated by a diagram. In Figure 1, the horizontal axis measures the number of people killed while the vertical axis measures the number removed by other means (exile/deportation or assimilation); alternatively, we can measure percentages of the target population. The downward-sloping straight line (an isoquant) is the locus of all the combinations of number killed and number removed by other means that exactly achieve the target of removing a given total quantity of people (say, 100). As drawn, the line has a slope of –1, reflecting the assumption that the ruler is indifferent to the means per se and cares only about the total numbers removed, so that in his preferences (i.e. on his target isoquant) the dead are “traded” for the otherwise removed at a one-to-one rate. Each of the “C” curves (isocosts) is the locus of all the combinations of number killed and number

---

3 The model can accommodate different ruler preferences, if desired. If for example, contrary to the working hypothesis of this paper, the ruler had a greater liking for killing as opposed to other means of removal, the substitution rate would be higher than one (it would take for example two exiles to compensate the ruler for the “loss” of one killing), the line would be steeper than that in our figures, and the tangency point would shift down and to the right on the isocost curve, resulting in a more killing-intensive optimal combination.
removed by other means that involve the same total cost to the ruler; higher-numbered curves correspond to higher total costs. The slope of a given curve at any point measures the rate at which the two inputs “substitute” for one another at the margin for the same total cost. Therefore the concave shape of the isocosts as drawn in all the figures, which implies that this rate of substitution increases from left to right, embodies the basic “technological” assumption of diminishing marginal product, or increasing marginal cost, discussed above: moving down from left to right the curve becomes steeper, implying that as the use of killing increases it becomes more expensive to obtain as it requires a larger sacrifice of other removals for the total cost to remain the same.

If the ruler seeks to attain his quantitative target at the least feasible cost, as we are assuming, the input combination that allows him to do so is identified by the point at which the target line touches the lowest possible isocost curve. So, in Figure 1, curve C1 lies everywhere below the line: the target cannot be attained at so low a cost. All the other curves do touch the target line at some point: clearly the optimal choice is C2; the other two curves – which exemplify the cost of exclusive killing (C3) and exclusive removal by other means (C4) – involve unnecessary excess cost. If the curves are smooth, as drawn, the contact point between C2 and the line is a tangency point where the two slopes are equal to –1: at that combination of inputs, the cost of an additional killing is equal to the cost of an additional removal by other means, a one-to-one marginal cost ratio exactly equal to the substitution rate between alternative means of removal on the target line, i.e. in the ruler’s preferences. The tangency point in Figure 1 characterizes an interior solution, where it is optimal for the ruler to use both inputs in some proportions. The dashed lines from the tangency point pin down these optimal proportions.

(Figure 1 about here)

While concavity of the isocosts in this setting is a general principle, how “curved” they are – i.e. at what rate relative marginal costs change as the input combination is changed along a given curve – will depend on the specifics of each situation. The minimum-cost input combination that meets the target line will change accordingly The limit cases are the corner solutions, where costs are minimized by exclusive resort to one means of removal. Figure 2 gives an example where it is optimal to kill 100% of the victims: the isocost curves are so flat that the target line touches the lowest isocost (C1) at its starting point on the horizontal axis; any other input mix, such as the one where C2 crosses the line, involves a higher total cost. Here, even at its steepest point on the right, the isocost’s slope is still less than 1, implying that killing is everywhere cheaper than any alternative means of removal.

(Figure 2 about here)

So far we have framed the ruler’s problem as one of minimizing the expenditure of resources under the constraint of achieving a given quantity target. It may be the case, however, that the ruler is resource-constrained or cost-constrained: there is a maximum total cost that he can afford to incur, one that is insufficient to attain the ideal target. He will then try to come as near as possible to the target subject to the constraint of the available resources. The problem then becomes one of maximizing the quantity of removals for the given total expenditure. If the shapes of the isocost and isoquant curves are the same as posited above, the nature of the solution does not change: an interior solution will generally be optimal. Figure 3 depicts one isocost curve, which is now the ruler’s constraint, and several isoquant lines corresponding to different total quantities of removals. All three lines are within reach, but
since the ruler seeks to achieve the maximum number of removals, the input combination that allows him to do so is identified by the tangency point with the 90-90 line. The other two isoquant lines correspond to the two opposite corner solutions, which involve unnecessarily low quantities.

(Figure 3 about here)

A final remark concerns the informational requirements of the cost-minimizing, or quantity-maximizing, solution. As we have seen, curvatures are of the essence: to yield a well-defined solution with the properties we have discussed, isoquants must be linear and isocosts must be concave. By definition, the ruler will know his own preferences – the constant slope of the isoquant line. But will he know the marginal cost of each means of removal at each step? If he lacks sufficient information about the victim group, he may miscalculate those costs. Specifically, the concavity of isocost curves embodies the principles of decreasing marginal products and/or increasing marginal costs; but these in turn depend on arranging the victims in the “right” way from the ruler’s point of view, from those easier or cheaper to dispose of to those more difficult or expensive. If, due to incomplete information, the ordering of victims is wrong so that the more costly ones are brought up first and the cheaper ones last instead of the other way round, the marginal cost of disposal by any one means will be decreasing rather than increasing over the target population, concavity of isocosts will be disrupted, and the point of contact between the curves will not identify the true least-cost combination. This highlights the difficulty for the ruler to pinpoint the optimal input combination on the diagram, as it were. Fortunately for the ruler, however, there is a way to get around this difficulty: self-selection by the victims. If the ruler allows victims to choose from among the acceptable options – such as: “you either convert, leave the country, or die; you choose” – they will sort themselves out by increasing resistance, or aversion, to each option, so as to spontaneously produce an ordering by increasing marginal cost of removal to the ruler. Thus those more willing to assimilate will volunteer first, those less refractory to emigration or deportation will line up first, those less able to fight or escape apprehension will let themselves be killed first (a way of disclosing the marginal cost of killing….), and so on. Self-selection, that is, helps overcoming the informational limitations that constrain the ruler’s choice of means. Therefore we should expect a rational ruler to make use of the self-selection opportunity whenever available.

Summing up, our model yields the following five predictions: in the vast majority of cases a combination of several means of removal will be observed; specifically, whenever (forced) assimilation is a feasible option in addition to exile/deportation and extermination, it will be used too; choice by victims, which triggers self-selection on their part, will be allowed to ease the informational limits that hinder achievement of the minimum-cost combination of means; whether across instances or over time in a single instance, the relative costs of different means of removal will drive the observed outcome; and exclusive resort to one means of removal will only be observed when that means is inordinately cheap to use or all the other means are inordinately difficult or expensive. Armed with these predictions, we are now ready to engage some historical evidence.

3. Ethnic, racial and national groups

Ethnic, racial and national groups hardly lend themselves to assimilation (except through mixed marriages over many generations – typically not an interesting option for the
rulers), so the only alternatives are (total or partial) extermination and deportation or exile.\(^4\) We begin by contrasting the Nazi experience with Jews\(^5\) and Gypsies and then turn to the Ottoman Armenians; moving closer to our day, we then address the Bosnian ethnic cleansing and the Rwandan genocide.

Most scholars today accept that Nazi anti-Jewish policy developed incrementally and that the infamous Final Solution to the ‘Jewish problem’ was not in the back of Hitler’s mind from the beginning but was arrived at through experimentation and response to outside circumstances (Hilberg 1985; Marrus 1989; Aly & Heim 2002; Paxton 2005; Bloxham 2009). The basic facts of the case are well known so a short summary will do. The first stage, starting with Hitler’s accession to power in early 1933, was segregation: stripping Jews of citizenship rights and freedoms, depriving them of property and jobs, and finally marking them with a yellow star. Following the Kristallnacht riots of November 1938, the Nazis’ objective centred on emigration and remained such until the fall of 1941. Of the half million Jews living in Germany in 1933, some 150,000 emigrated before Kristallnacht and another 150,000 thereafter. The Nazis anticipated that the ‘final solution’ would come at the end of the war, in the form of a negotiated mass departure of Jews from Europe. Then the conquest of Poland in September 1939 brought two millions of Polish Jews under German control; these were soon concentrated in urban ghettos as a preparation for a future mass transfer to the East and as a slave workforce for German war production in the meantime. Murder and ‘manslaughter’ (i.e., criminal, deadly negligence)\(^6\) did occur throughout, but there was as yet no general extermination plan. Experimentation and improvisation were the hallmark of Jewish policy for at least a year and a half after the Polish defeat. The initial plan involved the transfer of Jews from the part of Poland incorporated into the Reich to the part ruled by a German governor. This plan was sunk by the governor’s protest at the status of dumping ground for his domain and by the priority given to plans for the resettling of ethnic Germans in the newly acquired German Lebensraum, which competed for transport facilities. In 1940, a plan for the mass deportation of Jews to Madagascar was seriously considered and then dropped because the seas were under the control of the British navy. If the outbreak of war simply provided Hitler with a ‘cover’ for previously decided extermination, ‘why were the millions of Polish Jews in his hands since the fall of 1939 granted a thirty-month “stay of execution”?’ (Marrus 1989: 43)

The turning point came with the attack on the Soviet Union in June 1941. This promised to give the German army vast new territories but also millions more Jews to deal with. The campaign’s planning probably already included a decision to massacre the Soviet Jews that would be captured, along with communists, Gypsies, and other ‘Soviet spies’. This was accomplished by shooting by special military units – the Einsatzgruppen. Then at some point during the Russian campaign, probably in the fall of 1941, in frustration at the stalemate on the Russian front, the decision was taken by Hitler and his circle to move on to European-wide killing: the ‘dumping ground’ to the East was now foreclosed and with it the ‘territorial’ solution had to give way to the Final Solution. Construction started on the first two death camps, which began operation by end of year. In January 1942 the Wannsee Conference coordinated deportation from across Europe to the six extermination camps in Poland – a more

\(^4\) A more finely grained approach would distinguish race, which is defined by blood lines and hence unchangeable, from ethnicity or nationality, which is a bundle of characteristics bound together by tradition some of which can occasionally be unbundled and forcibly changed, allowing for ‘partial’ assimilation – for example, religion for the Armenians or nomadism for the Gypsies. The scope for assimilation remains narrow at best, however, so we will disregard the distinction in this paper.

\(^5\) Of course, in the Nazi view, unlike the Christian view, the Jews were a racial problem, so religious conversion would not avail.

\(^6\) I am borrowing this distinction from Wheatcroft (1996: 1320). Bloxham’s (2011) phrase “death by attrition”, as distinguished from “killing”, conveys much the same meaning.
efficient device for mass killing than firing squads. Now the job had to be finished before the end of the war, not after it, so operations accelerated dramatically. As a consequence, also the Jews of the occupied, allied, or satellite countries of Western and Central Europe, who had so far mostly been left alone and who were more integrated in their societies and therefore more difficult to round up, began to be sought in earnest – an effort sometimes frustrated by local authorities or populations.

In the last months of the war, secret negotiations took place between the allies and some Nazi leaders, notably Himmler, about a possible bartering of the Jews still alive in German hands for some concessions in the upcoming German surrender. Although nothing came of it, this shows the pragmatic cost-benefit approach to the problem by at least some of the Nazis (Marrus 1989: 185-195). Overall, the evolution of Nazi policy toward the Jews seems strongly to support the view that costs were an important consideration affecting choices.

The Nazi policy toward the Gypsies provides an interesting contrast to their Jewish policy (Lewy 2000). In Nazi racial theory they were considered an inferior race, but only because of their widespread admixtures with non-Aryan races: unlike the Jews, the ‘pure’ Gypsies, supposedly Aryan, were good, whereas the Mischlingen (mixed types) were bad and hence in need of cleansing. Some important differences with the ‘Jewish problem’ are apparent. First, the Gypsies were few (some 30,000 in the German Reich, including Austria, on the eve of the war), and the Mischlingen only a subset of these. Second, they had no international connections so that, in the Nazi mind, they never posed anything like the threat of the ‘Jewish conspiracy’. Third, partly as a consequence, they were always a low-priority problem, dealt with on the basis of expediency and improvisation. Fourth, the push to take action often came from below, not from above: the German people, and the local cells of the Nazi party, pressured the authorities to get rid of the Gypsy nuisance. As a result, Gypsy policy alternated between neglect, segregation, and deportation, but there was never at any time a plan for universal extermination of all Gypsies. To be sure, they were dispensable, so murder and manslaughter occurred on a large scale, but this was not unlike the fate that befell the Polish and Soviet prisoners of war.

Actual policy followed a haphazard course. In 1940 a few thousands were deported to eastern Poland, where many died of hardships and abuse in forced labour, while others drifted back home; further deportation, initially envisaged for all German Gypsies, was cancelled because of more urgent priorities in deportation. During the war they were forced to work in labour camps or firms in Germany, and their marriage permits were restricted. Some 5,000 from Austria were deported to the Lodz Jewish ghetto in November 1941 and hence transported to a death camp and gassed shortly thereafter. In Russia they were shot by the Einsatzgruppen as suspected collaborators with the enemy, and the same happened in occupied Serbia. Following Himmler’s so-called ‘Auschwitz decree’ of December 1942, 13,000 German Mischlingen were deported to a special ‘family camp’ in Auschwitz, while the ‘pure breed’ were exempted, and others were also exempted provided they subjected themselves to sterilisation. Czech and other Gypsies followed, so that a total of 23,000 went to Auschwitz. Here, unlike the Jews, they were not screened for the gas chambers upon arrival, and were not forced to work. Uniquely, in this ‘family camp’ families could live together and initially pregnant women and children were issued special food rations, so that other inmates ‘envied’ the Gypsies. Nevertheless, abysmal hygienic conditions, epidemics and starvation, as well as transfer to other work camps, killed thousands. All in all, at least 90% of those deported died one way or another. As this evidence suggests, the underlying idea was unlimited detention or deportation to the East for selected groups, not extermination, as is consistent with a low-priority, and hence low-benefit, objective combined with high opportunity costs for the Nazi leadership.
The Armenian genocide of 1915-16 in Ottoman Turkey\(^7\) is similar to the Jewish genocide in its evolutionary character, shaped by reaction to shifting conditions and opportunities, but differs in one crucial respect: the territorial dimension. Unlike the Jews, who were scattered across Germany and Europe and did not have an ancestral homeland that they could or would even conceivably claim as a national homeland, the Armenians did have such a homeland in eastern Anatolia. While on the eve of World War I only a minority of Armenians were striving for independence, and even fewer were willing to take up arms against the empire, they were being used by Russia and other European powers as a wedge to encroach upon the empire’s affairs – a situation that would become especially threatening with the outbreak of war. At the same time, the Balkan Wars of 1912-13 had not only downsized the total territory and population of the empire and triggered an inflow of millions of Muslim refugees who were desperate for a new home, but also left the Armenians – with the Greeks of Anatolia – as the only important non-Muslim minorities that remained in the shrunken empire. This became important as the Young Turk leadership that controlled the Ottoman government at the time were looking for a new foundation for the state, turning it from a traditional, multinational and multi-religious empire to a nationalist Turkish or Pan-Turkic state; and the ‘heartland’ of the Turkish nation was supposed to be in Anatolia. Thus from this point of view, the Armenians were no longer seen as a minority group to be kept in a subordinate status as a loyal millet, as they had been for centuries, but as an obstacle to the grand nationalist plan – and logically the only ‘final solution’ to this newly constructed Armenian problem could only be ethnic cleansing. Matters came to a head with the war, which provided both an incentive and an opportunity for action. In the first year of the war, many Armenians enlisted with volunteer units in the Russian army and contributed to a serious Turkish defeat and retreat on the Caucasus front, including the loss of the important city of Van which had risen in arms to support the Russian advance. Paramount security and military considerations thus dictated now a thorough removal of the Armenians from the eastern front of the war. After a series of more restricted repression measures, the general deportation decree issued in late May 1915 ordered that all Armenians residing in central and eastern Anatolia be deported and resettled in desert areas of eastern Syria, under the supervision of local authorities and the military; the Armenians of western Anatolia followed shortly afterwards. That forced relocation, not extermination, was thought to be enough for the purpose of ethnically cleansing a specific territory is shown, among other things, by the fact that the sizable Armenian communities of Istanbul and Izmir, as well as those residing in the Arab provinces of the empire, were not deported.

Implementation of the decree, however, rapidly went from manslaughter to murder, that is, genocide. There was no provision for transportation or food supplies; the deportees were marched for hundreds of miles without help or protection and time and again were set upon by Turkish and Kurdish villagers, Circassian tribesmen, and killing squads (the Special Organisation) answerable to the Committee of Union and Progress – the ruling party; those who survived such raids died of disease, hunger, and exhaustion by the thousands. Although detailed outcomes depended to some extent on local conditions and the attitudes of local authorities, the murderous intention of the central government, which tightly supervised the whole operation all along, is shown by the fact that those Armenians – mainly elderly, women and children – who, against all odds, managed to survive to their final destination in the

\(^7\) This discussion relies on Kevorkian (2011) for a full reconstruction of facts and on Akçam (2006) for the issue of responsibility and intent of the Ottoman leadership. Suny (2011) is useful for perspective on the evolution and current state of the scholarship, Bloxham (2011) (which summarizes the essentials of his larger study, Bloxham 2005) for a crisp interpretive framework that stresses the cumulative radicalisation process, and Melson (1996) for comparison with the Holocaust and other genocides.
concentration camps in the Syrian desert were then massively slaughtered in 1916. Thus in this interpretation, the Young Turks had no prior long-term blueprint for genocide, any more than Hitler had for the Jews, but building on pre-existing ill will toward the Armenians, they rapidly escalated from repression to deportation to extermination under the pressure of circumstances and driven by the logic of events that they themselves had set into motion: in the emergency conditions of war and general shortage, death by attrition, and then expediting death by resort to the Special Organisation squads, was the most cost-effective way of ‘solving’ a problem they had created for themselves. The head count is problematic but estimates suggest that anywhere between 600,000 and one million, or between one third and one half of the pre-war Armenian population, died; to these must be added an unknown, but presumably large, number of people, especially women, who were forced to convert to Islam in order to survive. In the end, the Armenians lost their existence as a national community in their original homeland, which suggests that, even though they did not go after every Armenian in the country as the Nazis did with the Jews, the Young Turks did achieve their goal.

The Bosnian war of 1992-95, which claimed over 100,000 lives, displaced more than two million people, and involved the rape of anywhere between 20,000 and 50,000 – mostly Muslim – women, highlights another facet of the territorial dimension of ethnic conflict: the asymmetry of objectives between the parties to the conflict. As we have seen, the Ottomans were intent on cleansing the Armenians out of the Anatolian heartland, while the Armenians – to the extent that an objective can be inferred from their behaviour before and during the genocide – only wanted to live on as an enfranchised minority group in the empire. By contrast, in Bosnia-Herzegovina, the Bosnian Serb leaders, backed by the army of rump Yugoslavia, started a secession war with the goal of achieving an ethnically homogeneous Serb state that included all the Serbs of BH. Since their aim was not control of the whole country but partition and population exchange, they pressed Serbs to leave areas outside of their control en masse. By contrast, Muslims living in Serb-controlled areas were encouraged by the government to stay so as to uphold its claim to the whole, undivided BH; hence they were driven out by terror. The Croats were partly driven out and partly withdrew “voluntarily” (Burg & Shoup 1999: 172). As a result, the Muslims disproportionately bore the brunt of the violence.\(^8\) While the overall death toll was 2.3% of the pre-war population, it was 3.2% for the Muslims, 1.7% for the Serbs, and 1% for the Croats. It bears noting that these figures are minuscule compared to most other extermination episodes discussed in this paper. On the other side of the coin, population displacement was huge. A large part – though by no means all – of it occurred between the two autonomous “entities” established by the Dayton peace agreement which ended the war, the Serb-controlled Republika Srpska and the Federation of Bosnia and Herzegovina, jointly controlled by the Muslims and Croats, while another substantial amount of more or less forced relocation occurred within each entity, and especially between Muslim-dominated and Croat-dominated areas within the Federation. While the intra-entity relocation escapes quantification, we can compute an inter-entity displacement rate for each ethnic group: the ratio of the numbers who by the end of the war had fled to the other entity or abroad to the pre-war population in all BH. This 1991-96 displacement rate amounts to 19.5% for the Muslims, 30.2% for the Serbs, and 28.3% for the Croats – much lower for the Muslims than for the other groups because, due to the asymmetric strategies explained above, the Muslims were mostly cleansed out by force while the Serbs, and in part the Croats, mostly relocated “voluntarily” – meaning in anticipation (right or wrong) of future retribution or persecution. It needs to be stressed that these numbers, large as they are, take no account of displacement within each entity.

\(^8\) All the figures that follow are computed from data in Tables 1, 2, and 4 in Ferrero (2013). Calculations are available upon request.
By this careful mix of murder and population transfer the Bosnian Serb leadership clearly achieved its goals: in the Republika Srpska, whose borders and population composition essentially sanctioned the situation on the ground at the end of the fighting, the degree of ethnic purity (ratio of Serbs to total entity population) and the degree of Serb concentration (ratio of Serbs in the RS to all Serbs in the country) in 1996 both approached 100%, from pre-war levels of 57% and 64% respectively. These achievements were there to stay: refugee returns in the following years were mainly directed to each group’s majority areas, not to one’s pre-war residence, and what “minority returns” there were only marginally diluted, but did not overturn, the ethnic cleansing pursued and accomplished by the Serbs through the war (Bieber 2006). It can also be shown (Ferrero 2013) that, from the Serb leaders’ point of view, the results achieved by the strategy actually chosen far exceed anything that could possibly have been achieved through “peaceful”, negotiated ethnic cleansing, i.e. a strategy of population exchange based on pre-war territorial Serb power without war and without redrawing boundaries. Finally, that the Serbs used the minimum violence, or the minimum-cost choice of means, necessary to achieve their ends is evidenced by the gender-selective character of the killing: consistent with a goal of winning a war of partition and ethnic cleansing, as opposed to wiping out an entire group, only “battle-age” males from 16 to 60 years of age (plus the political and social elites of any age or gender) were targeted for elimination (even in the ominous Srebrenica massacre), leaving women, children, and elderly men alive and deporting them out of the territory of interest — something one scholar describes as “gendercide”, as opposed to “root-and-branch” genocide (Jones 2006).

The genocide in Rwanda, which in the space of three months (April to July 1994) killed anywhere between half a million and one million people and wiped out some 75% of the pre-war Tutsi population of the country, can be seen as a limit case of a cost calculus. Tutsi and Hutu speak the same language, share the same range of Christian denominations, and have lived in the same region of Africa for centuries in various intermixing patterns. So there could be no question of “ethnic” or “national” differences and whatever limited flexibility toward forced assimilation these could in principle allow. The difference between the two groups had previously been mostly a question of class and power, but it was categorized as strictly racial under colonialism and enforced through race labels on ID’s, giving Tutsi privilege in government administration and education. The so-called Hutu revolution of 1959, which led to the country’s independence in 1962, retained the racial discrimination but turned it on its head, bringing Hutu to power and systematically discriminating against Tutsi. The violence of those years killed some Tutsi and drove hundreds of thousands to flee to neighboring countries, giving rise to a festering, massive refugee problem. The Tutsi refugees tried to return in arms in 1963 but failed, and thereafter tried to start a new life and gain citizenship in the host countries. They failed in this endeavour too: for reasons to do with the tangled geopolitics of Central Africa, the Tutsi were denied citizenship in Uganda and other neighboring countries, leaving them with no alternative but try to win back their place in Rwanda. Thus a group of demobilized soldiers who had fought in Uganda’s civil war in the 1980’s – the children of the Tutsi refugees of 1959-63 – formed the Rwanda Patriotic Front and invaded Rwanda in October 1990, starting a war that would drag on until July 1994 when the RPF completed the conquest of the country.

---

It was this war that started the Hutu ruling elite on a genocidal course: an extremist faction known as Hutu Power rapidly gained ascendancy, rejected any plan for negotiated settlement and power-sharing with the RPF, and made plans and preparations for a thorough removal of the Tutsi in the country — now the “enemy within”. The government forces were losing the war and this added urgency and persuasion power to the extremist propaganda to enlist the population’s support and participation in the upcoming “final solution”. By rallying the Hutu masses around a fight against the common enemy, the ruling elite hoped to be able to buttress its fledgling regime and hold on to power. It is significant that the death lists drawn up in advance included, before ordinary Tutsi, all the Hutu politicians and intellectuals opposed to the regime or “complicit” with the Tutsi. The genocidal plan was then implemented with astonishing effectiveness in the three months following the murder of President Habyarimana on April 6, 1994, no doubt on the reckoning that speed and thoroughness could create a fait accompli before the international community would or could meddle in. In this the ruling group turned out to be sadly right, even as they finally fled to Zaire in disarray in July.

From the point of view of the cost calculus that concerns us in this paper, this short account makes clear why, from the point of view of Hutu Power, the “final solution” to the “Tutsi question” in Rwanda could not possibly be, not even in part, a territorial solution: mass deportation and exile produce refugees, which in the Central African context meant more trouble ahead. It had happened in 1959-63, and the refugees of 30 years ago, or their children, were now returning with a vengeance. Furthermore, if the peace settlement with the RPF signed at Arusha in August 1993 were implemented, the provision of a right of return for refugees would come into effect, and this would open up a nightmarish scenario for Hutu Power: on the heels of the RPF fighters, hundreds of thousands of Tutsi returnees would flood the country (Rwandan refugees in neighboring countries numbered some 900,000 by this time). So it is telling that from day one of the genocide, with no mass involvement of ordinary Hutu yet forthcoming and the army, police and militia strained by the multiple tasks of fighting the war and executing the mass killings, lots of effort and manpower were devoted to setting up roadblocks across the country to check ID’s and prevent Tutsi flight. If removal of Tutsi as a group was the goal, and if the prospective costs of future returning refugees are factored in, the least-cost technique to accomplish the goal was one that employed exclusively one method: extermination.

4. Religious groups

Religion is the paramount case that, in addition to death and exit, offers the option of assimilation — that is, conversion. Other things equal, one would expect rulers to take advantage of this extra opportunity because, by so doing, the rulers’ cost of achieving their goal would presumably be lowered. Other things, however, need not be equal: forced conversion may often produce simulation and dual allegiance, which may in turn create further problems, as will be seen below.

The first task is to dig out relevant examples of a three-way alternative. Perhaps surprisingly, a quick survey of the historical record failed to turn out many such examples. Forced conversion on a large scale seems not to have been a popular policy except when rulers were dealing with a majority, such as a newly conquered people, in which case conversion obviously was the only feasible option (like in the medieval crusades against Northeast Europe’s pagan peoples).

For a three-way alternative, when rulers were dealing with a religious minority, we are left with a string of cases from Christian history. The first is the late Roman Empire, which
outlawed the practice of paganism and established orthodox Christianity as the official religion of the state under emperor Theodosius I in 391 AD. At this time, half the empire’s population may still have been pagan (MacMullen 1984: 92-95; Lane Fox 1988: 669-673), but these pagans were not a random sample of the population. They were mainly clustered at the opposite ends of the social spectrum: at one end, the literate aristocracy (and presumably, their slaves), at the other end the peasant countryside. The latter were obviously not mobile, and in any case their conversion was to be a protracted operation through the following centuries. The elite were also not mobile as their social status was tied to civic duties and public office in the imperial cities. So voluntary flight beyond imperial borders, though technically possible, could not be a substantial problem. Now this elite group was badly needed to staff the upper ranks of the state and the church, yet they were most reluctant to conversion. Therefore suppressing the practice of their religion – in particular, overthrowing temples and other cultic shrines, so important to a religion centred on sacrifices like ancient paganism – lowered the opportunity cost of their joining Christianity (Ferrero 2008). This proved sufficient for the purpose – violence to persons and even murder did occur but always started from below, never as an imperial policy. ‘Genuine’ faith was not an important consideration for either civil service or church office at the time – so the main drawback of forced conversions was no hindrance in this case.

The Spanish monarchy’s drive to remove all non-Christians from the kingdom in the aftermath of the reconquista is an interesting case study. Beginning with the Jews, they were ordered, under threat of execution, to either convert or leave the country within four months of the fall of Granada – the last Muslim emirate – in 1492. So they did, although it is not known in what proportions they stayed or left. From the crown’s point of view, this made sense. The Jews were few, and had no Jewish state to go to, so those who chose to leave would scatter throughout Europe and the Orient and never again try to come back. On the other hand those who chose to stay and convert would, by so doing, reveal themselves as the less committed to their tradition and more amenable to give in. It was expected that many of these people would only pretend to convert and stay on as ‘crypto-Jews’ (marranos), but this involved no serious risk as Judaism, under threat of draconian punishment, had ceased to attract Christian converts throughout Europe since the eleventh century (Fletcher 1997, ch. 9). Therefore, extermination would really have been overkill for the crown’s purpose.

The Muslims were different (Wheatcroft 2004, chs. 5, 6). Unlike the Jews, they were a very large number; their labour was needed, so throwing them all out was not practical. Also, unlike the Jews, they had powerful Muslim states in North Africa and elsewhere to go to, which in previous centuries had repeatedly sent forces to bolster the existing Muslim regimes in Spain. The latter consideration was indeed part of the reason for the removal, as the crown feared that a continuing Muslim population in the country could be both an attractor and a fifth column for another invasion. So it seemed that the self-selection strategy that had worked for the Jews would work for them too: the Muslim hotheads would choose to leave while the more lukewarm part would stay and convert or at least pretend to, thus getting rid of the worst troublemakers while retaining some useful workforce. After a bloody mass insurrection of Muslims in the Granada region in 1499-1501, in 1502 all the Muslims of the kingdom of Castille (including Granada) were forced to take baptism or leave; some years later, King Charles V extended the decree to the whole of Spain. Despite the fact that many had already fled to North Africa since 1492, the policy did not work. The nominally Christianised former Muslims, now called moriscos, stubbornly refused to behave and continued their traditional practices. A new, harsher decree by King Philip II in 1567 sparked another bloody rebellion in the Granada region in 1568-70, suppressed at huge human and financial cost. Then internal deportation was tried: in 1570, 80,000 Muslims from Granada were scattered across Castille,
over 20% of whom died on the road – something reminiscent of the Armenian story. As a result, Granada was seriously depopulated and trouble and hostility spread to the receiving areas, as the moriscos still were not behaving. In the closing decades of the sixteenth century, the search intensified for what was in effect a ‘final solution’ to the morisco problem. Several genocidal proposals were seriously discussed, from cramming all moriscos on ships and then sinking them off-shore, to castrating them all and then dumping them off in empty places in the Americas, to working them to death as galley slaves. But the logistic and policing costs were prohibitive and the fleet was badly needed for the war in the Flanders. As extermination was ruled out, the final solution boiled down to expulsion to North Africa in 1609-14 of all moriscos, with some exemptions for mixed marriages and individuals certified by a priest.

Interestingly, since religion had failed as a criterion to identify the victims (everyone was now a Christian), implementation was necessarily based on blood or ethnic background – ‘old Christian’ versus former Muslim convert. Finally, it bears noting that (as with the Holocaust and the Armenian cases) a political or security dimension weighed heavily on the Spanish state’s decision that a final solution was needed despite the staggering opportunity cost in terms of lost manpower: the moriscos were perceived as an ‘internal enemy’ that could conspire and ally itself with the real external enemies of the country – the North African emirates, the Ottomans, even the French Huguenots.

Our final example is the European religion wars sparked by the Protestant Reformation. Since each king or prince granted religious monopoly in his territory to one church, each wanted a one-religion dominion, so each in effect presented his subjects with a choice between conversion, exile or death. Both voluntary conversion and exile did occur on a mass scale, but not nearly enough to prevent the carnage. As the latter was obviously an inefficient solution that inflicted avoidable costs on all participants, the search for a cooperative solution yielded the principle of cuius regio eius religio (one’s religion shall be that of the sovereign ruling the territory of one’s abode), stipulated by the peace of Augsburg (1555) for German Catholics and Lutherans and extended by the treaty of Westphalia (1648) to all denominations throughout Europe. This was an efficient equilibrium solution as it ended the wars and provided a safe passage and a safe haven to those who would emigrate, while achieving the goal of religious uniformity in each territory. The possibility to choose between conversion and safe relocation triggered self-selection and thus minimised the welfare costs of achieving the goal.

5. Class and political groups

Social classes – defined by property ownership, wealth or income – were the special focus of the communist revolutions of the twentieth century. In theory, class groups, like religious groups, can be eliminated through assimilation plus the necessary amount of “re-education”: if a capitalist is expropriated, he/she will take up a factory job; if a large landowner is expropriated, he will be brought down to a small farmer. That is, a class can be wiped out by property redistribution – no need to go further than that; as Stalin emphasised during the Soviet collectivisation campaign of 1929-31, the kulaks (rich peasants) must be liquidated ‘as a class’, not as individuals. However, two problems arise. First, the expropriated group will be resentful and hostile to the regime, so expropriation breeds noncooperation or ‘sabotage’; on the other hand, if they are thrown out of the country they may want to return and claim their former property back. Second, since class boundaries are seldom crisp and clear – how rich must a peasant be to be classified a kulak? – and since several classes may be targeted at the same time, fear and opposition to the regime may spread throughout society and fuel a broad political opposition, whether or not class-based. These problems are compounded if the
country faces an external enemy that may be thought to find support from internal enemies, as happened to the Soviet Union around World War II. We will briefly discuss four cases: the polar opposites on a repression scale – Cambodia under the Khmer Rouge and Cuba under Fidel Castro – and the Stalin era in the Soviet Union and post-unification Vietnam which fall somewhere in between.

In the Soviet Union, the landed aristocracy, big merchants and industrialists were expropriated early in the revolution and mostly fled abroad. By the end of the 1920s, when Stalin jointly launched the collectivisation of agriculture and the crash industrialisation program, the remaining ‘class enemies’ to be dealt with were the kulaks in the countryside and the petty businessmen in small-scale industry and trade – a fairly large, heterogeneous group. Furthermore, the dramatic radicalisation and acceleration of socio-economic change aroused resistance and opposition, or suspicion thereof, in ever-widening groups, from old communist party members to technical specialists and engineers, from army officers to intellectuals to non-Russian nationalities. All these were targeted as enemies of the state, saboteurs, or complicit with fascism, and fell victim to the Great Terror of 1936-38. Finally, with the approach of the war and then the German invasion, whole nationalities were removed and deported en masse to prevent contact and cooperation with the advancing enemy, including Soviet Germans, Poles, Baltic peoples, nations from North Caucasus, and others.11

The Soviets used three methods to deal with their class and political enemies. One was execution by shooting, which happened throughout the period but peaked in the Great Terror. The second was deportation and resettlement to remote, often uninhabitable areas of Siberia, the Far East, and Central Asia, and was applied first to the kulaks and later to the national minorities from the Terror through the war. The relocation was supposed to be temporary but it often took these peoples decades to be allowed to move back. The third method was internment in the labour camp system known as the Gulag, where the goal was never to kill the inmates (unlike the Nazi death camps) but to punish them and extract their labour in the process, although in such harsh conditions that mortality was high. Significantly, the kulaks and the nationalities, which were relatively easily identified, rounded up and moved about, were deported (when not shot), while the sundry collection of other ‘enemies’ that were harvested in successive but overlapping waves, and in which the class factor shaded into the political, were sent to the Gulag (when not shot). Stalinist terror proceeded by quotas, i.e. by setting percentages of the target groups to be picked up and eliminated one way or another. The quotas were knowingly set at levels that included innocents where the boundaries of the target group were fuzzy (i.e. the politicals, but not the nationalities), so as to make sure that the ‘bad guys’ would not escape capture (Gregory 2009; Gregory, Schroeder & Sonin 2011). The head count is controversial. An authoritative assessment (Ellman 2002) of recent research on the ‘victims of repression’ from 1921 to 1953 puts the number deported (starting in 1930) at about 6 million and the number sentenced for political offenses at another 6 million. Of these 12 million, the number of deaths seems to have been 3 – 3.5 million, of which 1 million shootings, 1 – 1.5 million deaths of deportees, and 1 million deaths of prisoners.12

10 Good overviews, including quantitative detail as well as comparison with Nazi German repression, can be found in Wheatcroft (1996) and Ellman (2002).

11 Contrary to allegations by western writers of the Cold War era, the Soviet state never planned genocide but assimilation of selected national minorities, and the high attrition rates among them can be shown to be due to the state’s utter neglect of the human costs of its social engineering projects, worsened by the wartime conditions of dire shortage, rather than to an intention to annihilate them (Statiev 2009).

12 Many categories are left out of these figures, including those killed in the civil war 1918-20, those who died shortly after release from the Gulag as a result of their treatment in it, the victims of famines, the non-Soviet victims, and others. See again Ellman (2002) for discussion. In particular, the famine of 1932-33 is a controversial
In proportional terms, these figures pale in comparison to the record of the Khmer Rouge regime in Cambodia, which in only three and a half years of rule (1975-79) managed to directly kill, or starve or work to death, anywhere between 21% and 31% of the starting population. Here there was scarcely any industry or modern services to speak of, so the ‘real’ class enemies (landlords, merchants and intellectuals) were a small group. Ethnic minorities – a collection of different groups all suspected of being disloyal to the regime or, like the ethnic Chinese, “capitalist elements” – were disproportionately targeted and suffered an overall death rate\(^{13}\) of 44%; but these groups together amounted only to some 10% of the 1975 population, so the repression overwhelmingly fell on the ethnic Khmer. The enormity of the death toll is due to the fact that the egalitarian, agrarian utopia that the Pol Pot group tried to implement was so extreme that it turned virtually everyone into a perceived political enemy, if nothing else because of foot-dragging and insufficient zeal. This was the Soviet radicalisation syndrome pushed to the limit. On the other hand the country is small so internal deportation was not an option, while forcing mass exile to powerful, hostile neighbouring countries (especially Vietnam, which would eventually overthrow the Khmer Rouge regime) would have created an even bigger threat to the regime. So the solution was the Soviet solution magnified many times: extra-judicial sentencing to immediate execution and turning the whole country into a huge forced labour camp where mortality – unlike in the Soviet camp system – was staggering.

At the opposite extreme of the repression range we find the Cuban case. A small, densely populated, ethnically homogeneous country, Cuba had no place for internal deportation. Like all communist regimes, it initially went through a stage of collectivisation and centralisation of the economy but then, despite bouts of ‘volunteer’ mobilisation, it never went to the extremes of revolutionary upheaval and coercion that we find in the Soviet Union or China, to say nothing of Cambodia. Political executions were relatively few and a labour camp system was never developed. The distinctive Cuban solution to the problem of class-enemies-turned-political was emigration: in successive waves those who were unhappy with the regime were allowed to leave, or indeed shown to the door. According to an authoritative estimate\(^ {15}\), in the period 1959-87 the regime killed 0.24% of the mid-period population (not counting the boat people who died at sea while escaping) while an astonishing 12.2% left Cuba for the U.S. and other countries; thereafter, a partly legal, negotiated emigration has continued to this day while executions have all but stopped. This helped to stem the buildup of pressure inside the country that creates the problem of how to deal with those repressed; since exit was voluntary, not coerced, participation in Cuban society for those who remained was also to an extent ‘voluntary’. It remains to be seen if the buildup of pressure outside the country – among those Cuban expatriates who are still looking forward to returning home – will make the cure worse than the disease in the future. So far, however, the strategy of ‘letting out steam’ has

---

\(^{13}\) Estimates of the death toll vary widely. I found in Sharp (2008) the best comparative evaluation of existing research. He estimates total 1975 population at 8 million and excess deaths ranging from a minimum of 1.7 million (Ben Kiernan’s estimate) to a maximum of 2.5 million; these numbers yield the death rates reported in the text. Sharp’s own favored, or “most likely”, estimate is 2.18 million excess deaths, which yields a death rate of 27%.

\(^{14}\) Calculated from data in Kiernan (1996: 458).

\(^{15}\) The basic estimates are by Rummel (1997, Table 15.1B), of which I used the mid-range estimates in the middle columns. Summing 15,000 “executed” (line 788) and 7,000 “prison/concentration camp dead” (line 846) yields 22,000 for “domestic democide” (excluding 51,000 “boat people killed/dead” of line 800). Dividing this number by 9,000,000 for the population in 1973 (line 858) yields a democide rate of 0.24%. The emigration rate is found by dividing 1,100,000 “refugees” (line 856a) by the same population.
spared Cuba the worst excesses of repression and murder that marked other communist regimes.

We may locate the Vietnamese case somewhere near the Cuban example. After the end of the war and the unification of the country in 1975, the newly conquered south, and especially Saigon, had to be “cleansed” of class enemies and former employees and supporters of the American-backed government and of the American occupation troops. Vietnam was densely populated but a degree of internal deportation was possible to previously uncultivated lands; besides, there were the options of “re-education” (a form of assimilation), murder, and (forced or voluntary) emigration, so the government could choose from among the full range of techniques. With respect to the emigration option, Vietnam, unlike Cuba, had a sizable and territorially concentrated ethnic minority: the Hoa, or Sino-Vietnamese, a middle-class group which largely controlled business and commerce in pre-communist South Vietnam and was well connected to its government. In 1975 the Hoa were about 1.3 million, all but 200,000 living in the south, for the most part in the Saigon area (Cima 1987, section “Hoa”). Given their class condition, they disproportionately bore the brunt of the collectivization of the economy and therefore, to escape re-education or labor conscription, they were very much over-represented among the emigrants especially in the late ‘70s, when the socialist transformation was at its height. Their plight was further worsened by the short border war with China in February 1979, which cast many of them – particularly those living in the north – in the light of internal enemies in the government’s eyes. Many fled overland to China from northern Vietnam and many more by boat. Around this time, the government began granting the Hoa legal departure permits for a fee in gold. Unlike in the Cuban case, the sea voyage in unseaworthy craft was long, dangerous and very deadly, especially for the poorer ethnic Vietnamese, but this happened out of the government’s sight.

The existence of a program of extra-judicial executions after 1975 is documented in Desbarats (1990). The same study (p. 196) reports that a total of 2,500,000 people went through the re-education camps in the decade following re-unification, amounting to 9.5% of the 1979 population of South Vietnam. Finally, several million people were resettled in the New Economic Zones, established on virgin lands close to the borders with the ostensible purposes of increasing agricultural output and decongesting the southern cities and the northern countryside. Although the resettlement was partly voluntary, Desbarats convincingly argues that political aims largely trumped economic and demographic aims, and reckons (p. 200) that at least one million people were subjected to forced resettlement, involving in particular “undesirable” elements of the southern cities and the Montagnards, the hill tribes of the Central Highlands who were fighting the new government. Based on this and other studies, in the period 1975-87 the murder rate can be calculated at 0.93% of the South Vietnamese population (not including the boat people who died at sea while escaping) and the emigration rate’s upper-bound estimate at 5.7%. Firm overall figures on the numbers and ethnic composition of emigrants do not seem to be available, but what estimates are available suggest that upward of

---

16 The basic data are in Rummel (1997, Table 6.1B), of which I again used the mid-range estimates in the middle columns. Summing 95,000 “unnatural” deaths in re-education camps (line 673), 48,000 “unnatural” deaths in forced labor in the New Economic Zones (line 687) and 100,000 executions (line 698) yields 243,000 for “domestic democide”. Dividing this number by 26,200,000 for the South Vietnam population in 1979 (line 905: this is the last year for which South Vietnam’s population is given separately) yields a democide rate of 0.93%. This does not include 500,000 boat people “dead at sea” (line 751) out of a total “number fled” – including also those who crossed land borders – of 1,500,000 (line 713); of these deaths at sea, one half can be attributed to government responsibility (line 753). The emigration rate in the text is found by dividing these 1,500,000 “fled” by the same population; this rate is somewhat overestimated as it implicitly assumes that all the emigrants came from the south. Rummel’s executions number coincides with the lower-bound estimate of “extra-judicial executions” by Desbarats (1990: 197).
450,000 Hoa in all fled to China and elsewhere, making up almost a third of a total exodus of perhaps 1,500,000 people. This number implies that the emigration rate for the Hoa was 35% of their 1975 number. Even if starting in the late 1980’s large numbers of both Hoa and ethnic Vietnamese emigrants were allowed to return to Vietnam, the self-selection involved by both the flight and the return – as well as the very substantial death rate of the boat people while at sea, which might have been as high as one in three – worked to the benefit of the regime, which could dispose of potential troublemakers at low cost (and even profit from their departure by selling permits). The regime clearly reckoned that most of the Hoa, arriving in large and thriving overseas Chinese communities in Southeast Asia besides the outlet represented by China herself, would not press for repatriation, and that the United States, weary of the long war and keen to show the world that many Vietnamese were pro-American, would not use force to stop the exodus but take in large numbers of refugees, as indeed it did.

6. Summary and conclusion

This paper has developed the hypothesis that rulers who, for whatever reason, decide to remove an ethnic, religious, class or political group from their society have a choice between extermination, exit or deportation, and assimilation, and that their choice will be driven by relative costs. Thus we innovate with respect to the main thrust of the genocide literature, and at the same time borrow from the communism literature, by putting genocide or extermination among the means, not the ends, of a ruling power. To work out observable empirical implications from this hypothesis, we sketched out a simple economic model in which a ruler chooses the combination of means of removal that allows him to achieve his goal at the least possible cost. A review of selected historical cases seems to provide good initial support to the general hypothesis and to its five specific implications, listed at the end of Section 2, as follows.

First, these means of removal were usually used jointly rather than exclusively, as cost-minimizing behaviour by a ruler would lead one to expect. Even the Nazi policy toward the Jews, which ended in exclusive resort to total extermination, began with emigration and deportation accompanied by manslaughter and evolved over time as changing costs dictated.

Secondly, when assimilation was a feasible option, as was the case with religious and class groups, it was widely resorted to in combination with exile or deportation and killing, as a richer range of tools broadens the rulers’ opportunities for cost reduction.

Thirdly, choice by victims was widely used, apparently because it allows self-selection and hence revelation of individual disposal costs, which in turn makes it easier for the ruler to find and implement the minimum-cost combination of means of removal.

Fourthly, for each of the types of group examined (ethnic/racial, religious, and class/political groups), the actual mix of means of removal used in the different cases was found to follow fairly closely the different relative costs faced (or perceived) by the rulers.

Lastly, in keeping with the previous point, the extreme cases of near-exclusive resort to one means of removal appear to be those in which use of the other means turned out to be prohibitively costly. These include, in different ways, the Rwandan genocide, the forced Christianization of the Roman Empire, the Cambodian killing fields, and the Cuban mass emigration.

\(^{17}\) The total exodus number is from Rummel (see note 16 above). The Hoa number, or a number close to that, is found in Brush (2007), Minorities at Risk Project (2003), and Minority Rights Group International (2008). This number is larger than, but consistent with, a number of 250,000 Hoa who “had taken refuge in China” by the end of 1979, as reported by UNHCR (2000: 82) and confirmed by Cima (1987, section “Hoa”).
These findings seem encouraging enough to warrant further research on the cost approach. A natural extension will be an application of the model to the removal of indigenous peoples in (past or current) colonial empires. If further research does support our hypothesis, then the policy implications for the international community are straightforward, if not necessarily agreeable. Assuming that for the victim loss of life is obviously worse than territorial displacement, however despicable the latter may be, the risk of extermination in a given country can be reduced by offering the target group safe exit and relocation in a way that is acceptable to the perpetrator regime, that is, accompanied by some credible assurance that the exiles will not attempt to retaliate or come back. Such a prescription can be seen as a realist, if pessimistic, effort at minimizing violence, very much in the spirit of Kaufmann’s (1998) argument for population transfer and partition as tools to end ethnic civil wars in cases of extreme hostility between groups.
References


20


Figure 1. Quantity-constrained cost minimization
Figure 2. A corner solution: exclusively killing is the least-cost technique
Figure 3. Cost-constrained quantity maximization
Recent working papers

The complete list of working papers is can be found at 
http://polis.unipmn.it/index.php?cosa=ricerca.polis

*Economics Series

**Political Theory and Law

α ALEx Series

ε Quaderni CIVIS

2013 n.206* Mario Ferrero: *Extermination as a substitute for assimilation or deportation: an economic approach*

2013 n.205* Tiziana Caliman and Alberto Cassone: *The choice to enrol in a small university: A case study of Piemonte Orientale*

2013 n.204* Magnus Carlsson, Luca Fumarco and Dan-Olof Rooth: *Artifactual evidence of discrimination in correspondence studies? A replication of the Neumark method*

2013 n.203** Daniel Bosioc et. al. (DRASD): *OPAL – Osservatorio per le autonomie locali N.2/2013*

2013 n.202* Davide Ticchi, Thierry Verdier and Andrea Vindigni: *Democracy, Dictatorship and the Cultural Transmission of Political Values*

2013 n.201** Giovanni Boggero et. al. (DRASD): *OPAL – Osservatorio per le autonomie locali N.1/2013*

2013 n.200* Giovanna Garrone and Guido Ortona: *The determinants of perceived overall security*

2012 n.199* Gilles Saint-Paul, Davide Ticchi, Andrea Vindigni: *A theory of political entrenchment*

2012 n.198* Ugo Panizza and Andrea F. Presbitero: *Public debt and economic growth: Is there a causal effect?*

2012 n.197c Matteo Migheli, Guido Ortona and Ferruccio Ponzano: *Competition among parties and power: An empirical analysis*

2012 n.196* Roberto Bombana and Carla Marchese: *Designing Fees for Music Copyright Holders in Radio Services*

2012 n.195* Roberto Ippoliti and Greta Falavigna: *Pharmaceutical clinical research and regulation: an impact evaluation of public policy*

2011 n.194* Elisa Rebessi: *Diffusione dei luoghi di culto islamici e gestione delle conflittualità. La moschea di via Urbino a Torino come studio di caso*

2011 n.193* Laura Priore: *Il consumo di carne halal nei paesi europei: caratteristiche e trasformazioni in atto*
2011 n.192** Maurilio Guasco: L’emergere di una coscienza civile e sociale negli anni dell’Unita’ d’Italia

2011 n.191* Melania Verde and Magali Fia: Le risorse finanziarie e cognitive del sistema universitario italiano. Uno sguardo d’insieme

2011 n.190* Gianna Lotito, Matteo Miglieli and Guido Ortona: Is cooperation instinctive? Evidence from the response times in a Public Goods Game


2011 n.188* Gianna Lotito, Matteo Miglieli and Guido Ortona: An experimental inquiry into the nature of relational goods

2011 n.187* Greta Falavigna and Roberto Ippoliti: Data Envelopment Analysis e sistemi sanitari regionali italiani

2011 n.186* Angela Fraschini: Saracco e i problemi finanziari del Regno d’Italia

2011 n.185* Davide La Torre, Simone Marsiglio, Fabio Privileggi: Fractals and self-similarity in economics: the case of a stochastic two-sector growth model

2011 n.184* Kristine Forslund, Lycia Lima and Ugo Panizza: The determinants of the composition of public debt in developing and emerging market countries

2011 n.183* Franco Amisano, Alberto Cassone and Carla Marchese: Trasporto pubblico locale e aree a domanda di mobilità debole in Provincia di Alessandria