Strategic dynamics of social mimicry

James D. Fearon
Department of Political Science
Stanford University
Stanford, CA 94305-6044

June 25, 2013

Abstract

A social mimic is a member of one social category who pretends to be a member of different category in hopes of obtaining some benefit. Social mimicry is a strategic problem, since the extent of mimicry depends on detection efforts, but detection efforts likewise depend on the amount of mimicry. A simple game theoretic model of the problem yields a rich set of comparative statics results concerning equilibrium amounts of mimicry, rates of passing, and how these vary with the stakes at issue, the size of the group of potential mimics, and other factors. One interesting implication is that if detection efforts are moderately accurate, then the equilibrium amount of mimicry may be quite low even in a high stakes case like genocide. The equilibrium rate of mimicry is determined not by the mimic's preferences, but by how much the targets care about catching them.

1 Introduction

Gambetta (2006) describes situations of "social mimicry" in terms of a "mimic" who chooses from a set of signals in hopes of deceiving a "dupe" into thinking that he or she possesses some unobservable property k. In Gambetta's paradigmatic cases, the property k is membership in a social category (an "identity"), but in some applications it might not be. Examples might include:

- a minor using a fake driver's license to buy alcohol;

- a Tutsi pretending to be a Hutu in an attempt to escape the 1994 genocide in Rwanda;
- an IRA operative pretending to be a non-combatant in encounters with the authorities in Northern Ireland.
- a would-be terrorist attempting to board an airplane in order to hijack or blow it up.
- a tax payer misrepresenting her annual income on tax forms.

I describe below a simple game-theoretic model of social mimicry, which proves to have the underlying structure of betting in a card game like poker. Mimicry is analogous to bluffing. It may be useful to develop the analogy formally in order to make precise the nature of the strategic interaction in social mimicry and to produce comparative statics results interpretable for this class of situations.

Social mimicry is a strategic problem in the classic sense that optimal behavior for the mimic depends on what the target (Gambetta's "dupe") is expected to do, and vice versa. If targets always investigate and have a high enough probability of detecting mimics, then it may not be worthwhile for mimics to make the attempt. But if attempted mimicry is quite rare, then targets have little incentive to make the effort to try to detect it.

These observations already suggest that in an equilibrium state of affairs, the expected payoff from attempting to mimic must be about the same as the expected payoff of not mimicking. If mimicry is better than non-mimicry, more mimicry will be attempted and detection efforts will increase. If mimicry is worse on average than non-mimicry, then fewer will try it and detection efforts will decrease, restoring the incentive for more mimicry. Likewise, in equilibrium the rate of mimicry must be such that the expected payoff to a target of attempting to detect must be about the same as the payoff for not bothering. If detection efforts fall below this rate, then more mimicry is attempted, raising the payoff to trying to detect, and vice versa if detection efforts are too intense.

As shown formally below, an interesting implication is that if detection efforts are made sometimes but not always in equilibrium, then the rates at which mimicry is attempted and succeeds depend on how much the *target* (or "dupe") cares about stopping it versus the costs of doing so, and not at all on how much mimics want to pass or their costs for getting caught trying. When the value of detecting mimics (or equivalently the costs of failing to detect) is high, targets will attempt detection for sure unless the rate of attempts is low. Thus in equilibrium attempts will be rare.

This observation might help explain the possible fact that attempts to pass in genocides are surprisingly infrequent (Laitin 2009). More generally, the interesting result is that there are conditions such that the higher the *stakes* at issue – it is reasonable to think that in many cases the value of passing for the mimics is related to the value of detection for the targets – the *lower* the equilibrium rate of attempted mimicry, the greater the equilibrium detection efforts, and the lower the rate of successful passing.

The analysis produces a large set of comparative statics, which could be useful for empirical assessments in particular areas. There are quite a few potential variables of interest here, including

- the percentage of non-k types in the population
- the percentage of non-k types who try to pass
- the percentage of non-k types who successfully pass
- the percentage of non-k types who are caught trying to pass.
- the percentage of people in the population exhibiting k signals who are mimics
- the frequency of detection efforts
- the probability that a mimic passes.

Depending on the specific empirical example, some of these may be known or observable, while others are certainly not. For example, we might know the percentage of 15-to-20 year olds in the population and the number caught trying to buy alcohol with a fake ID card, but not the number with fake ID cards. The formal results can establish expectations about how these various quantities might be expected to be related to each other, and how they would vary with factors like the stakes, the probability of detection conditional on inspection, the cost for a mimic of being detected, the share of non-k types in the population, and so on.

2 The model

We consider the interaction between a person who may or may not want to pass herself off as possessing the valuable property A (which could be membership in some social category), and a potential target or "dupe," which I will call the government, or G for short. The game begins with Nature drawing the person who will interact with G from society, represented as a continuum of agents. A share $1 - \beta$ of society possess the desirable characteristic A, and β do not. Those who possess the characteristic A will be called A's, and those who do not will be called B's. A and B could thus refer to two social groups like Hutus and Tutsis in Rwanda before 1995. A B who tries to pass as an A will be called a "mimic."

B's differ among themselves in how costly they would find it to try to pass as an A. An individual B's cost of trying to pass is $e \geq 0$, where in the population of B's e is distributed by a smooth cumulative distribution function F. These costs could be material, if mimicry requires acquiring certain goods, skills, or changing one's appearance, or psychological, if associated with costs for abandoning one's identity as a B or losing social connections.[1]

After Nature draws the person to interact with the potential dupe, the person chooses whether to try to pass. A types always try to pass, at zero cost. B's who decide not to try to pass get their value for the status quo, which is normalized to be zero.

The government (or government agent) then observes whether the person is emitting A signals or not, and chooses whether to *investigate* if the person is a mimic. This would involve any of the various ways that people have of detecting deception, as appropriate for the specific situation. For instance, they might ask for and inspect an ID card in the underage drinking example, or interrogate the suspect in the IRA example. Assume that investigating has a fixed cost $d > 0$ (d for "detect").

If the government chooses not to investigate, then G realizes a payoff of zero if the person was in fact an A, and suffers a loss $l > 0$ if she was a mimic. For example, l might be the expected cost of fines and loss of business if the police discover that the liquor store sold to an underage buyer. In the IRA example, the cost of failing to

[1] An alternative formulation that is more natural for some examples is to make the cost of mimicking fixed, but assume that B's vary in their value for passing as an A. That is, we would take e as fixed and let the benefit of passing (v below) be distributed according to F. The results in this variant are essentially identical.

detect a mimic is the political and material damage that follow (in expectation) from not jailing an IRA operative.

If the government chooses to investigate, then I assume for convenience that it correctly identifies an A as an A type for sure. However, it correctly identifies a mimic as a mimic only with probability $1 - p$, so that $p \geq 0$ is the probability that a mimic succeeds in passing as an A if investigated.

A mimic who passes gets the benefit $v > 0$. A mimic who is detected pays a cost $c > 0$. Thus if the government chooses to investigate and the person is in fact a mimic, the mimic's expected payoff is $pv - (1 - p)c - e$. If G chooses not to investigate, the mimic gets $v - e$. Likewise, the government gets an expected payoff of $-pl - d$ for investigating when it in fact faces a mimic. If the government in fact faces an A, then investigation just wastes the effort cost d (yielding payoff for G of $-d$).

Table 1 summarizes the payoffs in different scenarios and outcomes.

Table 1. Payoffs from an encounter.

type of person	G's action	mimic's payoff	G's payoff
mimic	G investigates	$pv - (1 - p)c - e$	$-pl - d$
	no investigation	$v - e$	$-l$
A type	G investigates		$-d$
	no investigation		0

p	=	Pr(mimic passes)
v	=	value of passing for a mimic
c	=	mimic's cost if detected
e	=	effort or psychological cost of attempting to pass (varies across individuals)
l	=	G's loss for failing to detect mimic
d	=	G's effort cost for investigating

For many cases, it is natural to suppose that the mimic's value for passing is related to the government's cost of failing to detect a mimic. For example, if the A's are allowed access to some desirable goods that are denied to B's – such as government jobs, more positions in the university, social distinction – then more for the B's can mean less for A's. In the case of a pure taste for discrimination, it may be that the more the A types hate the B's, the greater the value of passing as an A. So in the comparative statics below we will sometimes consider what happens as v and l are varied together (e.g., $v = l$), and in this case we will refer to the *stakes* at issue in mimicry.

3 Analysis

The game is easily solved for a perfect Bayesian equilibrium in which G and every type of B are doing the best they can given the other players' strategies, and in which the government's beliefs about the likelihood of facing a mimic are consistent with the actual frequency with which mimics try to pass. A strategy for a B is a rule specifying whether to try to pass as a function of his or her effort cost e. A strategy for G is a probability with which to investigate a person emitting A-type signals.

Let β' be the government's belief that it faces a mimic conditional on encountering a person claiming to be an A. Given the government's costs and benefits for failing to detect a mimic, the government prefers to pay the costs of investigation when

$$\begin{aligned} \beta'(-pl-d) + (1-\beta')(-d) &> \beta'(-l) \\ \beta' > \beta^* &\equiv d/l(1-p). \end{aligned}$$

G prefers not to investigate when it believes the chance it faces a mimic is less than this threshold value β^*, and G is indifferent when $\beta' = \beta^*$. Appropriately, the government agent will investigate for lower beliefs that it faces a mimic (i.e., when β^* is smaller) when the costs of investigation d are low, the costs of passing a mimic l are high, and when the ability to detect mimics is high (low p).

Let $\beta'(z)$ be the government's belief that it faces a mimic if all B types with costs of mimicry $e < z$ try to mimic. This is the share of mimics in the population of all people sending A-type signals (Bayes' rule):

$$\beta'(z) = \frac{\beta F(z)}{\beta F(z) + 1 - \beta}. \quad (1)$$

Finally, note that $\hat{e} = pv - (1-p)c$ is the effort cost such that a B with cost $\hat{e}$ for mimicking would be just willing to do it even if she was sure to be investigated. Of course, it can be that $\hat{e} \leq 0$, in which case no B would want to try to pass if a detection effort were certain. However, if there is a high enough chance that you could pass even if investigated (large enough p) or if the benefits of passing are very large relative to the status quo and the cost of getting caught as a mimic, then there can be types of B who would want to try to pass even if the government is investigating everyone.

For example, consider a potential victim of a genocide in progress. If the person believes that he is certain to be killed if he tries to pass and fails, but is *nearly* certain to be killed even if he does not try to pass, then trying to pass even if there is only a small chance it will work may be worthwhile even if he knows that detection will be attempted. Formally, this is a case where c is very close to zero, or equivalently, v is vary large. (Recall that we normalized the value of not trying to pass to be zero, so that if the status quo as a B is very bad, then v is large relative to c.)

We are now in a position to describe the game's equilibrium, which is unique for any given set of parameters. Depending on the parameters, patterns of action in equilibrium fall into one of three cases, distinguished by the behavior of the government. In case (1) below, the government does not bother to investigate *anyone*, and all types of B with $e < v$ mimic and get away with it. In case (3), by contrast, the government investigates *everyone*, and only those B's with effort costs for mimicry less than $\hat{e}$ try to pass. In case (2), the government investigates some but not others.

Proposition 1. The following strategies form a PBE of the mimicry game for three mutually exclusive cases, which depend on the relationship of β^* to two cutpoint values, $\beta'(\hat{e})$ and $\beta'(v)$.[2]

Case 1: If $\beta^* > \beta'(v)$ then G never investigates, and all types $e < v$ become mimics.

Case 2: If $\beta'(v) > \beta^* > \beta'(\hat{e})$ then G investigates any person showing signs of being an A with probability

$$g^* = \frac{v - e^*}{(1-p)(v+c)}$$

[2]Since $\beta'(z)$ is increasing and $\hat{e} < v$ when $p < 1$, $\beta'(\hat{e}) < \beta'(v)$. In this description of the equilibrium, I will ignore non-generic boundary cases such as $\beta^* = \beta'(v)$.

where e^* is implicitly defined by

$$F(e^*) = \frac{1-\beta}{\beta}\frac{\beta^*}{1-\beta^*}.$$

B types with $e < e^*$ become mimics and the rest do not.

Case 3: If $\beta^* < \beta'(\hat{e})$ then G investigates all people showing signs of being an A, and only B's with $e < \hat{e}$ try to mimic.

In the first case, either the government finds investigation too costly relative to its expected benefit to make it worthwhile, or there are too few potential mimics in the population to make it worthwhile, or both. The threshold value $\beta'(v)$ is the share of mimics in the population of "A"'s if all Bs who would want to mimic if they could get away with it in fact mimic. If this share is small enough relative to the government's threshold belief for being willing to investigate, then G just permits mimicry. Of course, within the case, the probability that a B succeeds as a mimic is 1. The number of mimics is increasing in the share of B's in the population and in the value of passing (v). It is decreasing in the average cost for appearing as an A.[3]

As argued above, for many empirical examples it can make sense to speak of "the stakes" for mimics and targets. That is, the greater the mimic's benefit for passing, the greater the target's cost for failing to detect; this will be the case, for example, when there is essentially a distributional conflict between mimics and targets. In the model, this means that l is positively related to v. When this is so, increasing the stakes means that more mimicry is attempted while we remain in the bounds of Case 1, but Case 1 becomes less likely to obtain because increasing l means that β^* gets smaller – G is more inclined to investigate when the stakes are larger, of course. And as we will see, increasing the stakes within Case 2 has a very different impact on the rate of attempted mimicry.

In Case 2, government agents investigate some but not all people claiming to be A's.[4] B's whose costs of mimicry are less than a threshold level $e^* < v$ give it a try. In this case the equilibrium probability that an "A" is in fact a B is completely determined by the ratio of a government agent's costs of investigating to the expected benefits of

[3] Say that increasing the costs of appearing as an A means rightward shift in F.

[4] In a slightly more realistic formulation, mimics could differ randomly in the degree of suspicion they raise when trying to pass, or government agents could differ in their degree of perceptiveness. Thus the agents would not need to adopt a consciously mixed strategy, but it would be effectively random from the perspective of B's.

investigating a mimic. If there are "too many" B's pretending to be A's, investigation rates in increase. If there are too few, investigation decreases. The share of B's in the population of apparent A's is "just right" (at β^*) when government agents get the same expected payoff from investigating as they get from not investigating. Note well that this share is completely determined by G's preferences, not B's.

The share of "A"'s who are mimics (β^*) is not the same thing as the share of B's who try mimicry ($F(e^*)$), but it turns out that they act the same way in this case. Since

$$F(e^*) = \frac{1-\beta}{\beta}\frac{\beta^*}{1-\beta^*} = \frac{1-\beta}{\beta}\frac{d}{l(1-p)-d},$$

the share of B's who mimic is increasing in the costs of detection, and decreasing in the danger to G of missing mimics and the probability of getting caught. This all seems intuitive. A non-obvious implication, however, is that within Case 2, *increasing the stakes lowers the rate of mimicry*, since increasing l lowers the share of B's who mimic (and the share of "A"'s who are actually B's), while increasing the importance of passing for B's, v, has no effect at all.

At the same time, however, because increasing the stakes makes G's threshold belief for investigation, β^*, smaller, it becomes more likely that Case 3 obtains rather than Case 2. As discussed below, in Case 3 G investigates everyone and more B's try mimicry as the stakes increase. So the effect of increasing the stakes on the rate of mimicry is complicated – increasing when the government doesn't care enough to investigate potential mimics, decreasing when the government wants to investigate some but not all, and then increasing again when the government cares enough to want to check all "A"s.

Besides the stakes, the share of B's who try mimicry is also determined by the size of the population of B's. In Case 2, the larger the group of potential mimics, the *smaller* the share of them who try mimicry. So if we begin with an extremely small group (β close to zero), G initially doesn't care about a few mimics (Case 1). But eventually, for a large enough group of B's, investigations begin and absolutely fewer B's will mimic as the size of the group grows.

In Case 2, the government's rate of investigation of "A"s (g^*) has the following comparative statics:

- increasing in the stakes,
- increasing in the probability that a mimic passes if investigated,

- decreasing in a mimic's costs for getting caught,
- decreasing in the costs of investigation, and
- decreasing in the share of the population that are Bs.

Because not every "A" is investigated in Case 2, the probability of successful passing is greater than p, the chance of passing when investigated. A mimic's unconditional probability of passing in this case is $1 - g^* + g^*p$ which equals $(e^* + c)/(v + c)$, and has the following comparative statics. Successful passing is more likely when

- the costs of investigation are higher,
- the probability of passing if investigated (p) is higher,
- the costs of getting caught (relative to the benefit of passing) are *higher*,
- and the stakes are lower.[5]

Finally, we come to Case 3. Here, even if the government is going to investigate every single person claiming to be an A, there are enough B's who want to try to pass to make comprehensive investigation worthwhile for the government. This is more likely if the expected savings for the government by investigating are large relative to the costs of detection – that is, $l(1 - p)$ is large relative to d – while the expected benefits of attempting to pass are large for a mimic ($pv - (1 - p)c)$). For this situation to obtain, investigation must not be too accurate – if it were then not enough B's would want to try mimicry for comprehensive investigation to be worthwhile – nor too inaccurate – if it were then it would not be efficient for the government to investigate everyone.

So, provided that investigation is of at best medium accuracy, raising the stakes high enough pushes the interaction towards case 3 by increasing the share and number of B's attempting to pass even if they are sure they will be investigated.

Figures 1 and 2 may help to summarize the most important of the comparative statics results described above. Figure 1 shows how several measures of the extent of mimicry vary with the stakes at issue. Figure 2 shows how the equilibrium shares of mimics from the Bs, among the "A"s, and in the total population vary with the

[5] e^* is decreasing in v when l is positively related to v.

size of the B group. The three cases are visible in each figure by the breakpoints where the slopes of the lines change abruptly. For example, in Figure 2, when β's are fewer than about 20% of the population (given the parameters used to generate these figures), all B's mimic – this is Case 1. The intermediate range is case 2, and the curves after stakes of 1.5 in Figure 1 and β of about .66 in Figure 2 are Case 3.

Figure 1 illustrates how mimicry increases with the stakes until the government is motivated to start investigating, at which point increasing the stakes actually *reduces* mimicry rates, up to the point where the government wants to investigate all claiming to possess the desirable characteristic A. Then greater stakes increase mimicry again as more B's give it a try even though they are sure to face investigation.

An interesting if perhaps unfortunate implication of this result is that it suggests that it may be hard to identify any robust empirical patterns concerning the size of the stakes and the scale of mimicry. Apart from the observation that mimicry will be common and unpoliced if there are few potential mimics and they can do little damage from the target's perspective, mimicry efforts may be rampant or quite uncommon when the stakes at issue are moderate or high. Much would depend on how accurate investigations are.

Figure 2 suggests a simpler pattern: the share of B's who choose to try mimicry is weakly decreasing in the size of the B population. For some empirical settings, this could be a testable prediction.[6] By contrast, the share of mimic's among the As is increasing at first, then stable, and then increasing again as the size of the B group grows.

4 Some evolutionary dynamics

In the equilibrium of the generic social mimicry game analyzed above, government officials do not want to investigate either more or less given the true number of mimics, and B's don't want to mimic at a higher or lower rate given the true probability of being investigated and passing. I gestured at an argument that decentralized decision-making would actually yield such an equilibrium: if there are too many mimics, gov-

[6]For example, did Nazi-occupied European countries, or districts within a country, with more Jews have lower rates of mimicry? There would of course be an "other things equal" problem here in that stakes might also vary with population shares.

ernment officials would increase detection efforts, reducing mimicry; if "too much" detection effort, fewer mimics, and thus a switch to less detection.

But is that right? If we make reasonable assumptions about what G and B's can observe, and about what information they use in adjusting their strategies, would the implied dynamical system actually converge to the equilibrium identified above?

Suppose then that the action plays out over a series of periods $t = 0, 1, 2, \ldots$, which could be days, weeks, months, etc. In $t = 0$, the government agent G randomly chooses $g_0 \in (0, 1]$, a share of people presenting themselves as A's to investigate, while B's with effort cost less than a randomly determined level $\hat{e}_0$ attempt to pass. G then learns the percentage of people she checked whom she determined to be B's trying to pass. If she knows (or can estimate relatively well) the probability that she can correctly spot a true B who is trying to mimic ($1 - p$ in the model), then she can "back out" an estimate of the percentage of people who presented as A's who actually are B's simply by dividing the number of Bs she identified by $1 - p$. This is β' in the model above, so call the first-period estimate β'_0 in this dynamic version.

The government agent has a notion of whether "too many" B's are trying to pass. In the model as described above, if $\beta' > \beta^*$, then the agent finds it worthwhile to investigate more "A"'s, because the expected costs they impose, or the expected penalties for the agent from letting too many B's through, are greater than the costs of investigating another "A" (those costs are d per individual). So it seems reasonable to suppose that in the next period, G will increase the frequency of investigations if $\beta'_0 > \beta^*$, and reduce the frequency if $\beta'_0 < \beta^*$. A simple adjustment rule for periods $t > 0$ is

$$g_t = g_{t-1} + \epsilon(\beta'_{t-1} - \beta^*), \tag{2}$$

where $\epsilon > 0$ is a parameter that reflects the amount by which G changes investigation frequencies. This proposal makes the plausible assumption that G makes smaller changes when the rate of B's trying to pass is closer to the rate at which G is just indifferent between doing and not doing another investigation.[7]

What about the Bs? For many contexts it would be reasonable to suppose that they can get information about the rate at which Bs who try to pass are investigated and the rate at which they succeed. This would be the case, for example, if Bs are a relatively small ethnic or religious minority faced with persecution, so that word

[7] Since g_t is a proportion, I need to add that we should take the median of 0, 1, and the value given by (2).

spreads in the community about who tried to pass, who got away with it, and who got caught. Teenagers in a large high school get some information about how many peers have fake IDs and how often they work. The IRS advertizes its audit rate and it is probably reasonable to assume that the probability of "passing" in that case is quite low.

To the extent this is the case, a simple and natural adjustment rule for the B's is that they "best respond" to the detection frequency of the previous period. A type with cost for mimicing e (recall that this is relative to the person's benefits for passing and cost for being caught) thus estimates an expected payoff for trying to pass of $v(1-g_{t-1}+g_{t-1}p)-g_{t-1}(1-p)c-e$, which implies that in period $t>0$, types less than the threshold value

$$\hat{e}(g_{t-1}) = v(1-g_{t-1}+g_{t-1}p) - g_{t-1}(1-p)c \qquad (3)$$

will mimic.[8]

To close the model, note that β'_t is given by (1) using $\hat{e}(g_{t-1})$ for z. This system is easily simulated or considered analytically. If the government agent's adjustment rate ϵ is not so large she (pathologically) bounces between investigating everyone and investigating no one, then the system converges rapidly to the equilibrium levels of g^* and e^* found in section 3. Figures 3a and 3b illustrate two cases, where in the second case G's adjust the frequency of investigation in large increments, "overshooting" and producing some oscillation. In this example, in both cases G starts out investigating at a rate that deters all B's from attempting to mimic, which leads G to investigate less because there is no one to catch, so A's are just being inconvenienced. Soon enough it begins to be worthwhile for the lowest cost B's to start trying to pass.[9]

Possibly the strongest assumption made above is that the government agent can reliably estimate her own probability of correctly detecting a mimic. In most contexts, a successful mimic leaves no trace. If the goverment, or "dupes," are unable to observe the size of the population of successful mimics, then it is not clear how they can estimate how good they are at spotting them. If they overestimate their ability to detect mimics – which seems psychologically plausible to me on the grounds that we have a tendency not to think about "dogs that didn't bark" – then Bs will be able

[8]Here we need add the technical stipulation that if $\hat{e}(g_{t-1})$ is greater than the upper bound on B's possible costs for trying to pass (if there is one), we set $\hat{e}$ in t to that upper bound, and likewise to zero if the value in (3) is less than zero.

[9]Parameter values are $v=1$, $c=5$, $p=.5$, $\beta=.4$, $\beta^*=.2$, and e is uniformly distributed on $[0,1]$. In Figure 3a the adjustment rate ϵ is .2 while in 3b it is 1.5.

to take advantage of this and will mimic at higher rates than e^*, while G investigates too little. This is most likely if the Bs are a cohesive group with strong social networks, which would make for effective information flows about the probability of being investigated and the ability to pass.

On the other hand, if the Bs are not a cohesive group and if the penalties for being caught trying to mimic are very large, then the reverse might well obtain. That is, very few attempt to pass and little is known about success rates, which means that Bs have little to go on and so may not try to mimic despite the fact that if they did, they would be quite likely to pass. For example, there may be a great many opportunities to successfully and profitably mimic that arise in every day life, but which "too few" take due to underestimates of their ability to get away with it.

5 Institutional commitments to investigate potential mimics

In the game analyzed above, the target chooses whether to investigate for mimicry on a case-by-case basis. It does not have the option of committing ex ante to investigate all people claiming to be As. In some empirical settings – particularly those where the targets are government agents – it is natural to suppose that the government can make institutional commitments on the frequency of investigation. For instance, Congress can fund the Transportation Security Administration and pass laws committing the TSA to screen all airline passengers. The IRS can commit, for the most part, to implement a particular rate of auditing tax returns. Customs systems are institutional commitments to investigate the citizenship claims of people entering a country. In other settings, however, institutional commitment is more difficult. The government cannot commit clerks in liquor stores to check all customers' IDs, for instance. Sometimes in situations of official efforts to prevent mimicry, the government faces a principal-agent problem in figuring out how to motivate and monitor its agents' efforts to detect mimics.

Consider then a modified version of the signaling game in which the target moves first, choosing whether to commit to investigate all people claiming to be an A, or to investigate on a case-by-case basis as above.[10] Suppose for starters that the institu-

[10]A slightly more general version would allow the government to commit to any given probability of investigation g, but I will be consider the full investigation commitments versus no commitment for sake of clarity.

tional commitment to investigate all is itself costless (although the larger number of investigations pursued will be costly).

Proposition 2. If G can commit in advance to investigate all persons claiming to be As, then no commitment is made if the conditions for Case 1 or Case 3 in Proposition 1 obtain, while commitment is made if the conditions for Case 2 obtain. Equilibrium strategies are then the same for Cases 1 and 3, while in Case 2 the government investigates everyone, and B's attempt to mimic only if their costs of doing so are $e < \hat{e}$ (as in case 3).

Proof. Without commitment, in Case 2 G gets an expected payoff per encounter of $-pl\beta^* - d$. By committing ex ante, G induces all types of B with $e > \hat{e}$ to not mimic, yielding an expected payoff per encounter of $-pl\beta'(\hat{e}) - d$. Since in Case 2 $\beta^* > \beta'(\hat{e})$, this means that commitment is always better for G in this case.

In Case 1, G's expected payoff per encounter is $-l\beta'(v)$, and by committing it could get $-pl\beta'(\hat{e}) - d$. Rearranging, commitment is worse for G than no commitment in this case when $\beta'(v) - p\beta'(\hat{e}) < d/l$. This is certainly true when $p = 0$, since then $\beta'(v) < d/l = \beta^*$ is a condition for Case 1 when $p = 0$. Since $\beta'(v) - p\beta'(\hat{e})$ is decreasing in p, commitment is worse than no commitment for G for any p when the condition for Case 1 holds.

For Case 3 parameters, commitment by G is superfluous since investigation is comprehensive in this case anyway. $\square$

So if it were costless to set up, the government would always want to construct ex ante institutional commitments to investigate all if it has an incentive to investigate some. The reason is that there is a gain in terms of deterrence of mimicry. G investigates more often than without commitment, but this is more than compensated by the reduction in mimics who get through. Why then is commitment necessary? Because after setting up the institutions to force investigation of all presenting as A's, the rate of mimicry drops to a point where G would like to stop investigating so much – this is why commitment is needed. If the TSA and other airport precautions are highly effective, then there is a lot of wasted cost and effort, ex post.

Proposition 2 assumed for clarity that there are no costs to setting up an institutional system of commitment to investigate apparent As. If we introduce costs of building a monitoring system, then whether G wants to do so in Case 2 will depend on a comparison of the costs of institutional commitment versus the gains from deterring more mimics. Higher stakes, or an increase in the number of potential mimics will tend to favor moves towards institutional commitment to investigate.

6 Illustrations

6.1 Fake IDs

The costs of procuring a fake ID are presumably small compared to the value of having one if it is successful (v), but the cost of getting caught (c), a minor felony, would be large for many or most kids. Then in a Case 2 equilibrium (some checking), young-looking customers would be checked at a rate of about $v/(v + c)$, which one would think would be less than one half.[11]

The costs of checking ID for a liquor store clerk are small but not nothing – possible loss of sale, possible irritation of of-age buyers, discomfort at implying dishonesty – while at least for the store owner, the expected cost of selling to minors might be quite large (risk of losing licence). If so, then in a Case 2 equilibrium we would expect a low rate of fake ID use among late teens. Further, if the costs of selling to underage buyers are large enough for the store owners that checking IDs is optimal, then they may want to have a policy of "We Check Everyone" to try to commit clerks ex ante, as discussed above. Psychologically, this could also lower the costs of checking for the clerk, by mitigating the specific implication of dishonesty.

6.2 Taxpayers

Falsifying a tax return is practically costless in and of itself. The gain to getting away with it, v, is the tax rate times the amount of income underreported. The cost of getting caught is some sort of large financial penalty and maybe worse (most U.S. citizens, I think, are unclear on this). Many also probably factor in reputational damage as a cost of being found guilty of cheating on tax returns.

[11] Who are the Bs and who are the As here? For a good part of the population, visual inspection can determine with almost perfect certainty that the person is over 21, or under, say, 12. The As would be be everyone else who is over 20, and the B those who are less than 21. A model tailored to this specific problem would incorporate a person-specific signal of age.

This would again imply a low rate of equilibrium auditing by the IRS (that is, $v/(v+c)$ would be small). But the true rate of auditing is in fact *extremely* small – on the order of one in 130 returns in the U.S. – so most people would have to view the costs of getting caught as 130 times the value of cheating to rationalize this audit rate. Further, I believe I have read that the net financial yield on additional IRS audits in terms of penalties and delinquent taxes is very high, which would not be the case if the IRS was simply trying to maximize its (legal) take. Most likely their principal, Congress, does not want too many audits because these are unpopular with constituents.

Since we have good evidence that this should be a Case 2 equilibrium, without institutional commitment the rate of cheating on tax returns (mimicry) would in equilibrium be determined by the ratio of the IRS's political and financial costs for an audit to the expected political and financial returns from an audit. It is not obvious to me if this would be a small number (less than 20%?) or something larger. However, this is also case where it seems plausible that the IRS can commit in advance to audit a given percentage of all tax returns.

6.3 IRA operatives

In this example (and possibly the fake ID example above), a more natural formulation of the model would have Catholics in Northern Ireland varying across individuals in how much they would value being an IRA operative if they could get away with it, and pay no social or other costs for becoming one. That is, the individual variation would be on the benefit of passing v rather than on the cost of becoming an operative e. While some of the expressions in the formal analysis change slightly, there are no substantive differences in results. Types of B with a value for being an IRA operative greater than a threshold value v^* opt for the organization, while the rest do not.[12]

IRA operatives would probably be hard to detect by interrogation of random individuals, and the financial and political costs of individual interrogations are presumably at least moderate relative to the expected damage foregone by catching an operative (high p, moderate d relative to l). This would suggest Case 1: not enough Catholics who want to pay the cost e to be an operative to make individual interrogations of

[12]In case 2 with this formulation, types of B with $v > v^*$ mimic, where v^* is implicitly defined by $1 - F(v^*) = \frac{1-\beta}{\beta}\frac{\beta^*}{1-\beta^*}$. G investigates with probability $g^* = (v^* - e)/(v^* + c)$.

random Catholics worthwhile for the government. This could differ in some neighborhood or bounded social set where many operatives are suspected (higher β).

6.4 Potential victims of genocide

In the midst of an active genocide, the options for a B may reduce to (1) wait for neighbors and government agents to identify and kill you, with high probability, or (2) preemptively seek to escape the country by passing as an A, though facing certain death if you fail to pass. Given these options, it is surprising that, as far as we know, more European Jews did not preemptively seek to escape the Nazis by passing, and more Tutsis did not try to pass as Hutus in Rwanda. It should be stressed that this is a stylized fact at best. I don't think we know much about actual rates of attempted passing and their success in either case, and especially for the latter. Also, since we know what ultimately happened, it is easy for us to overestimate how confident Jews and Tutsis should have been at the time that not mimicking would lead to death. Nonetheless, as Laitin (2009) argues, a fair bit of evidence suggests the rates of successful and attempted mimicry cannot have been large, which seems surprising.

What does the model imply for plausible parameter values in this example? The stakes are clearly very large, which would put us in Case 3 *if* B's chances of passing are not negligible. In Case 3 all are investigated and most B's try to pass. However, if the probability of passing, p, is quite small, so that other options like hiding, bribery, or seeking protection from friends would be more attractive if you were certain to be investigated for mimicry ($pv-(1-p)c < 0$), then we return to Case 2. Here, because the stakes are large for the government relative to the costs of investigating (it wants to kill these people after all) the equilibrium rate of mimicry is predicted to be very small, even if not everyone is thoroughly investigated.

The evidence Laitin presents suggests that at least for the Rwandan case, the probability of successful passing if investigated was in fact extremely small. The main problem seems to have been that in a rural, village-based society, everybody knows your name, appearance, and history, so there were abundant sources of information for genocidaires.[13] Exiting the village and trying to pass at multiple roadblocks appears also to have been highly unlikely to work, in large part because the government had

[13]This raises a more general point: Social mimicry simply cannot be much of a issue except in cities and mass societies.

for many years required individuals to carry official ID cards that specified tribe.[14] The prediction from the model would be high but not comprehensive investigation rates, and low levels of attempted mimicry.

Passing also appears to have been extremely difficult for the most part in Nazi-occupied Europe. For men and boys, circumcision was nearly impossible to disguise.[15] Women and girls did not face this handicap, and Laitin cites a study by Weitzman that claims that "thousands" of Jews, "mostly women," escaped the Holocaust by mimicry. The difficulty of acting independently of a family unit (with males) may help explain why this number is not even higher, but it may still be surprising that it is not larger.

7 Conclusion

In Michael Spence's (1973) classic signaling model, there are good workers and bad workers, the former having higher productivity than the latter. Good types have a lower effort cost for years in school, and, in order to starkly illustrate the logic of signaling, Spence assumes that education has no effect on productivity. In a separating equilibrium, good types choose a level of education that is high enough that bad types would not want to mimic it even though doing so would get them a higher wage from employers. This outcome is inefficient. The education costs are a pure loss relative to the first best outcome that would occur if employers could observe ability directly. The information asymmetry drives good types to incur education costs to distinguish themselves from potential mimics.

In this example, then, a whole system of institutions – higher education – is explained in part by a problem arising from the potential for mimicry (or, in economic terms, an information asymmetry in a situation where there is an incentive to misrepresent). What colleges and universities do, in this argument, is not so much to add value in

[14]Fake IDs were hard to procure on short notice. As Laitin observes, a big puzzle in this case is why more Tutsis hadn't prepared fake IDs long before in anticipation of such dangers.

[15]From the website "Virtual Jewish Library": "Because non-Jews in continental Europe generally were not circumcised, German and collaborationist police commonly checked males apprehended in raids. For boys attempting to hide their Jewish identity, using a public restroom or participating in sports could lead to their discovery. More rarely, they underwent painful procedures to disguise the mark of circumcision or even dressed as girls." See also Gambetta (2006, 232) on how cultural knowledge made it extremely difficult for Polish Jews to pass as Christians.

terms of ability but to screen good from worse types in the admissions process.

Now consider the contrast between social mimicry as described by Gambetta and as modeled above, and the classic signaling model. In the former, the focus is on strategic behavior by the potential mimics and dupes: Do they mimic or not? Do they investigate or not? In the latter, the focus in on strategic behavior by the people the mimics want to mimic, or "models" in Gambetta's language: how can they credibly distinguish themselves from other types? This is not particular to Spence's example. Whenever there is a benefit that many desire but that politics, ethics, or economic efficiency wants to allocate to some subset, then "models" want to be able to distinguish themselves from mimics.[16]

In the approach suggested by the standard signaling game, we would study social institutions with an eye to how they are constructed to screen out potential mimics, to be "incentive compatible," in the language of economics. Social mimicry would then be understood as an instance of, or another way of describing, the problem posed by private information and incentives to misrepresent.

That is, a very large number of social, political, and economic interactions have the following aspect. What should be done to produce a good outcome for a group or for an individual should ideally depend on private information held by different people. But because of distributional conflicts, people might misrepresent their private information if they were simply asked. The deep structure of such problems has been mapped and explored by economic theorists under the rubric of "mechanism design."[17]

A mechanism can be thought of as a formal or informal institution that takes as inputs some form of "announcements" of individual's private information (often called "types") and maps these into social decisions or outcomes. Tax systems, voting systems, legislatures, juries, markets, auctions, pricing policies by businesses, political

[16]It is worthwhile to spell out exactly how the classic Spence model differs from the mimicry model analyzed here. In the Spence model, the employer (or G in the model here) is completely incapable of figuring out if the job applicant is a high or low productivity types except by actually putting them to work; investigation is useless ($p = 1$). However, in Spence's case, job applicants can verifiably report their effort level e, and the marginal cost of effort is higher for low productivity types than for high productivity types. Thus this is a class of situations where the "models" (as Gambetta terms them) are able and have an incentive to take actions that verifiably distinguish themselves from potential mimics.

[17]For an introduction see the Nobel lectures of Maskin (2008) and Myerson (2008).

constitutions, all manner of contracts, police departments, counterinsurgency methods – all these and much more are mechanisms. From the perspective suggested by the mechanism design approach, institutions elicit signals and screen among different types, producing various collective decisions as results. Institutions may be informal, as with customs concerning how buyer-seller bargaining proceeds, or highly formal, as with how a political constitution structures bargaining among branches of government.

Societies that have managed to get the "mass" level with cities and lots of impersonal exchange have done so because they have developed institutions that make social mimicry appear to be a marginal problem at best, in that almost all the time we can trust the representations of others we deal with on a daily basis. But this is precisely because we have developed formal and informal institutions that deter mimicry in anonymous and "one shot" interactions. Seen in this light, solving the problems posed by potential mimicry is a central challenge for low-income countries trying to develop economic and political institutions that will support complex economic exchange among larger-than-traditional groups of people and organizations, and political institutions wherein voters can trust politicians that pretend to be "good types" to actually be good types. High income countries have developed institutions and cultures that resolve this problem to a greater degree, although new technologies like the internet may pose new threats by creating new ways to mimic.

References

Gambetta, Diego. 2006. Deceptive Mimicry in Humans. In *Perspectives on Imitation: From Neuroscience to Social Science*, ed. Susan Hurley and Nick Chater. chapter 10, pp. 221–241.

Laitin, David D. 2009. "Cultural Distance: A Constraint on Mimicry?" Stanford University, unpublished paper.

Maskin, Eric. 2008. "Mechanism Design: How to Implement Social Goals." *American Economic Review* 98(3):567576.

Myerson, Roger. 2008. "Perspectives on Mechanism Design." *American Economic Review* 98(3):586603.

Spence, A. Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics* 87(3):355–74.

Figure 1. Share of B's who mimic as function of stakes

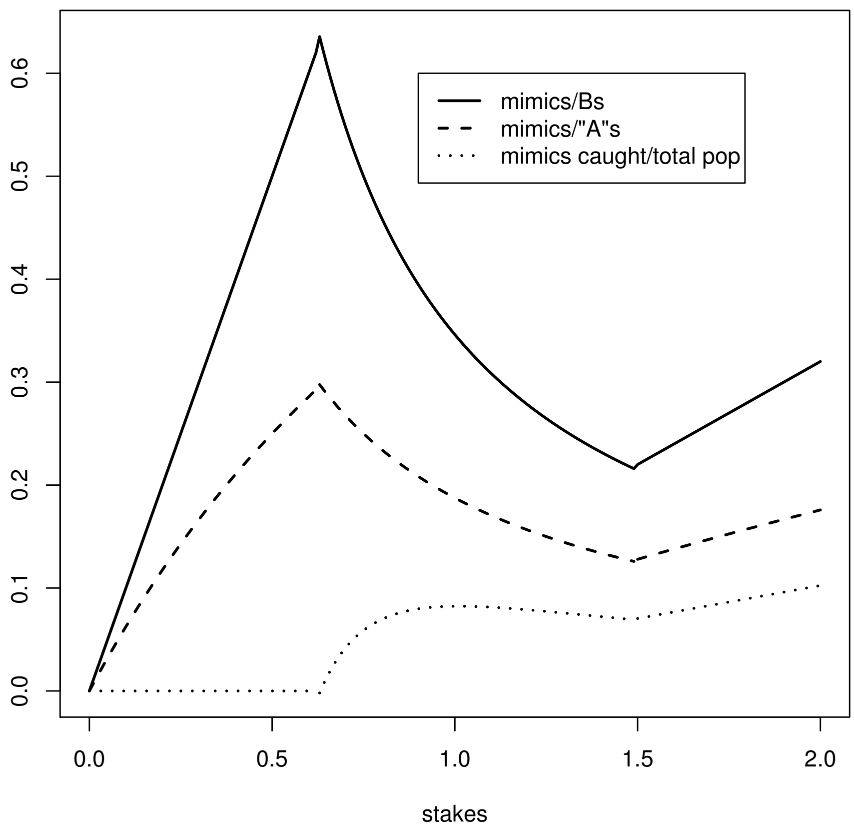

Figure 2. How mimicry varies with pop. share of possible mimics

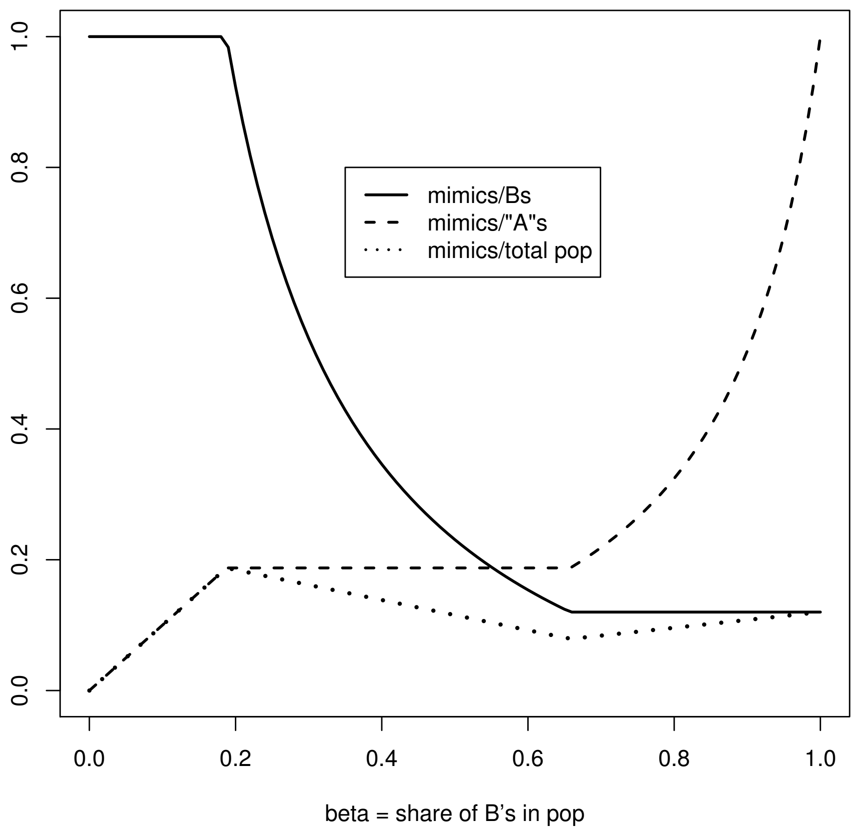

Figure 3a. 'Slow' adjustment

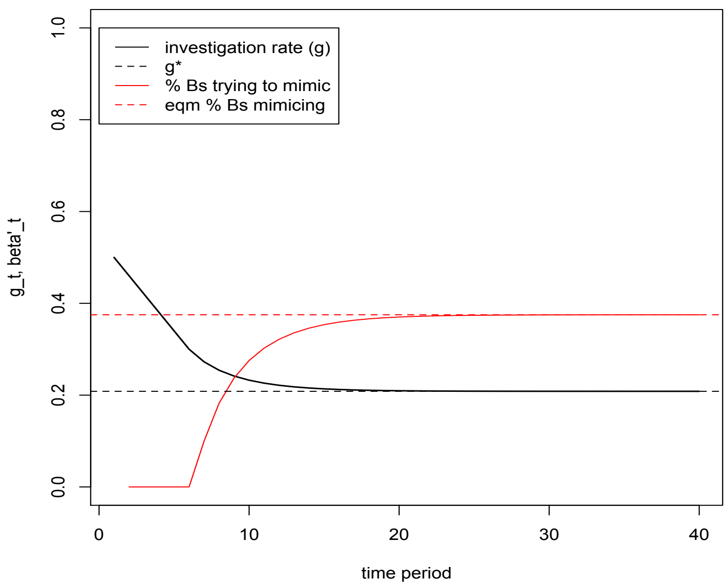

Figure 3b. 'Fast' adjustment