Identity change and economic mobility: experimental evidence

Manuel Muñoz-Herrera[*] [†]

Abstract

I study the impact identity change (assimilation) has on economic mobility. I experimentally assign people to different group identities, majority or minority, before they interact in a social coordination setting. In equilibrium, minority assimilators achieve economic mobility by integrating with the majority. In the experiment, assimilators are discriminated against and cannot integrate, if majority members encounter conformists (non-assimilators) in the minority. Thus, assimilators fail to attain economic mobility because those who maintain the status quo impose negative externalities on those who risk changing it.

Keywords: identity choice, assimilation, networks, coordination, inter-group contact
JEL Classification: C92, D91, D85

[*]Center for Behavioral Institutional Design, New York University Abu Dhabi (e-mail: manumunoz@nyu.edu).

[†]I am grateful to Kanchan Chandra, Aleksei Chernulich, Sanjeev Goyal, Eric Kimbrough, Frederic Moisan, Nikos Nikiforakis, Ernesto Reuben, Pedro Rey-Biel, Roberto Weber, Marc Witte, Yves Zenou and participants at a number of seminars for helpful comments. I am grateful for financial support from Tamkeen under the NYUAD Research Institute award for Project CG005. Ethics approval for this research was granted by the IRB office at NYU Abu Dhabi.

1 Introduction

"Choice of identity may be the most important 'economic' decision people make" (Akerlof and Kranton 2000, p. 717).

Since the seminal work of Akerlof and Kranton (2000), the study of identities in economics has been widespread.[1] Researchers have found that sharing an identity with others promotes cooperation (Chen and Li 2009), trust (Alesina and La Ferrara 2002), and better economic outcomes (Chen and Chen 2011). However, the core assumption that individuals have agency to change their group identities has not received as much attention, even though identity choice can be central to properly understanding inter-group processes and social change (some exceptions include Shayo 2009; Bernard et al. 2016; Desmet et al. 2017; Dasgupta and Goyal 2019). The mechanism is simple, if someone from the out-group assimilates the identity of the dominant group, the sense of a shared identity would drive those in that dominant to perceive the assimilator as an in-group member and change the way and consequences of their interactions (Gaertner and Dovidio 2000; Akerlof 2016). In other words, the stranger becomes a fellow (Eckel and Grossman 2005; Charness and Chen 2022).

Assimilation can, therefore, be a powerful strategy for individuals whose group identities put them at a disadvantage. For example, assimilation has been found to reduce economic sanctions (Saleh and Tirole 2021), increase chances of finding a job (Battu and Zenou 2010), finding a spouse (Bicchieri 2016), receiving help (Choi et al. 2019), and even leading to economic mobility (Boyd and Richerson 1992, 2002; Henrich and Boyd 2001). Aside from the potential benefits for members of the disadvantaged group to successfully assimilate, we know little about what makes assimilation fail. In this paper, I experimentally study what conditions drive assimilation and under which of them assimilation is more likely to successfully lead to economic mobility.

To address this objective, I create a novel experimental design that allows me to induce group identities that are meaningful but can be changed by individual choices. To induce identities, I use a combination of different tools. First, every individual in the experiment is assigned to one of two groups: majority or minority. These labels are not explicitly used in the experiment, but differences in group size are key to the experimental design (see Section 2.) Instead, a unique symbol reflects group membership: a circle for the majority and a triangle for the minority. Members in each group are then tasked to choose a name for their group by voting for their preferred one from a list of 5 options. To do this, in-group members can communicate freely through a chat box. A group name is

[1] For recent surveys of the literature see Charness and Chen (2022) and Li (2020).

chosen when all group members vote for the same option. The options are animal names, different for each group: felines for the majority and canines for the minority.

Therefore, I make use of differences in visual markers (symbols) and names between groups, and a joint voting task plus free-form communication within groups, as combined tools to induce in-group identification and out-group differentiation.

After the group assignment stage, members of the minority and the majority participate together in a social coordination game (Goyal et al. 2021). Individuals decide with whom to form network connections and then play a coordination game with those in their network, choosing between two actions. Individuals benefit from coordinating with others on one same action, and the gains are additive on the number of connections one coordinates with. As such, it is better to be well integrated with others than to be isolated.

Akerlof and Kranton (2000) suggest that identities are tied to prescriptions regarding appropriate behavior.[2] Identity-wise, therefore, I induce group prescriptions on which action is preferred for each identity through differences in payoffs. Simply put, any player earns more when she coordinates on the action that has been prescribed to her identity, and those in the minority prefer the opposite action to those in the majority. This social coordination game allows inter-group differentiation to be enhanced, as the differences in prescribed actions can drive segregation between the two groups. Economically speaking, segregation is problematic because it leads to lower payoffs for everyone, especially for the minority, as gains increase based on network size.

After a series of interactions in the social coordination game with fixed identities, individuals are given the option to either keep their group identity or change it for that of the other group. If a minority member changes her identity and assimilates that of the majority, she earns more by coordinating on the majority's action than on the minority's action (which was not the case prior to assimilating). This is meant to symbolize that by assimilating, the prescription of the new identity is also internalized.[3]

Consequently, individuals can strategically assimilate to integrate with the dominant group and improve their economic standing. As assimilators and the majority have aligned incentives over which action to play, it is in their interest to form connections between them, so that they can all benefit from their shared behavior. The optimal outcome is

[2] Akerlof and Kranton's work take a preference-based approach to social identity, emphasizing that a central difference between identities is that they are associated to prescriptions or norms on how to behave. There are also group-contingent social-preference models, see for example Chen and Li (2009) and Chen and Chen (2011), which predict how group identity affects behavior. I make use of the latter approach at the end of my experiment in the Other-Other allocation decision.

[3] Although strategic lying is not part of this project, in some settings individuals can strategically disguise and signal to someone in power that they share the same identity, without having actually assimilated. See for example Casoria et al. (2022). This is also related to the notion of social free-riding in (Bernard et al. 2016).

one where the entire minority assimilates, as it leads to a completely integrated society where all individuals coordinate on the same action. This reduces inequalities across social groups and allows the minority to attain economic mobility, as all minority members earn more than in the case of no assimilation.

However, if for some reason those in the majority have doubts as to whether the assimilators will actually act in alignment with their newly adopted identity (i.e., choose the majority's action), they may discriminate against them by not reciprocating connections (Lees and Cikara 2020). This leaves assimilators in a worse position than if they had conformed to their original group identity. Such is a case where assimilation fails.

To test for the conditions that drive assimilation choices and the likelihood that it successfully leads to economic mobility for the minority, I look at a series of experimental variations. Specifically, I explore different barriers associated with identity change. In one dimension, I look the effect of having (or not having) an assimilation cost (Austen-Smith and Fryer 2005; Fryer and Torelli 2010).[4] Assimilation costs are usually a result of social sanctions that arise when conformists in a social group punish those who abandon the group. To reflect this, I model the cost of assimilation as increasing on the number of conformists in the minority, so that it is smallest when all minority members assimilate.

In another dimension, I vary whether the assimilation choice is identifiable. If a person who assimilates a new group identity has limited ways to display it, then others may only see in her those external markers of the group she used to belong to (e.g., skin color, accent, height, tattoos, etc). Limited identifiability of one's identity can lead to mistrust from the dominant group and can reduce the chances of assimilators succeeding. I implement this in the experiment by varying whether identities are publicly observable or private.

At the end of the experiment, every individual participates in an other-other allocation decision where they divide a set of resources between a receiver from her in-group and one from the out-group (see e.g., Chen and Li 2009; Lane 2016). I use this decision to empirically test the effectiveness of sharing a common identity, emerging from the assimilation choice and from repeatedly solving a coordination problem together, to enhance social preferences among participants originally assigned to different social groups (see Chen et al. 2014).

The main result of the paper is that assimilation does not necessarily lead to economic mobility. Assimilation fails because the majority discriminates against assimilators when they also encounter conformists in the minority. This despite it being in the best interest of

[4] A complementary line of research looks at the costs individuals pay to avoid assimilating or acting in misalignment with their identities, see for example Bursztyn et al. (2020).

the majority to integrate with those who assimilate. That is, two individuals of the same type, *assimilators*, are judged and treated differently depending on what others around them do. This result may be driven by the ambiguity and tension between individual and collective choices. An assimilator makes an individual decision to change her identity, but that does not necessarily coincide with the choices of her fellow minority members. This can make members of the majority pool assimilators and conformists into one single category (for a model of optimal categorization see Fryer and Jackson (2008)). Thus, the meaning of individual decisions can be tainted by the decisions of others in the same social group.

There are three possible outcomes in terms of identity choices. First, an *all-conform outcome*, where no one in the minority assimilates. There is no ambiguity here, as everyone in the minority is choosing to conform to their original group identity. On the other extreme, there is the *all-assimilate outcome*, where everyone in the minority changes her identity for that of the majority. This is also a collective outcome in which the minority group acts in unity and can provide a clear meaning of their assimilation choices. However, there is the *fractured outcome*, in which there are assimilators but also conformists in the same minority group. This means that when a majority member looks at the minority, she may doubt that those who assimilate will act in alignment with their chosen identity, as the presence of conformist can be a force pulling them back. Thus, the majority may pool assimilators and conformists into a single category and treat them similarly.

I find evidence suggesting that in the case of a fractured minority group, conformist impose negative externalities on assimilators, which hurts the latter's chances to integrate and benefit from assimilating. But, if the minority assimilates collectively (no conformists), then the majority reciprocates to the assimilators' intentions to integrate and economic mobility is attained. This result, that two assimilators are treated differently depending on what others in their social group do, is consistent across experimental conditions. Regardless of whether the minority faces assimilation costs, has no way to publicly signal the assimilated identity, or is not exposed at all to any of these barriers, the majority discriminates against assimilators when the minority is fractured by the presence of conformists.

This paper is a contribution to the research on cultural assimilation. One of the main goals in this area of research has been to identify conditions that prevent assimilation, especially when the choice of not assimilating puts minority individuals in a position of economic disadvantage (see e.g., Bisin et al. 2016). A common finding in the literature suggests that being in a group that imposes sanctions, as a mechanism to police its boundaries, is one of the most frequent drivers of conformism (Battu et al. 2007; Battu and Zenou 2010). My work echoes this result by identifying the case with assimilation costs (derived from social sanctions) as the most frequent driver of a *fractured outcome*, where assimilation is not outright prevented but moderated, fracturing the social group into assimilators and conformists.

My work also extends this line of research by looking at what causes assimilation to fail. That is, conditional on assimilating, what could prevent an assimilator from reaping the benefits of her identity choice. I find that not all assimilators are treated equally, but their success is context-dependent. Assimilators in groups where other minority members around them choose to conform end up being discriminated against by the majority, and as such assimilation fails in procuring economic mobility. But, assimilators from a minority that is not fractured are welcomed by the majority (see Charness et al. 2007; Fryer and Jackson 2008).

The results of this paper also contribute to the literature on group identity, by looking at ways that can help reduce inter-group conflict (Cikara and Van Bavel 2014; Gaertner and Dovidio 2000). As it has been found that individuals are more favorable towards those they share a group identity with and less so to those in other social groups, researchers have been exploring the common sharing of an identity between members of different social groups as a way to increase social cohesion (Mousa 2020; Lowe 2021).

Assimilation of identities is a way in which individuals from different groups can end up sharing a common identity. I find that the power of a shared identity is context-dependent. If a person assimilates while the rest of her group members maintain their original identity, those in the other group may be unwilling to see her as an in-group member (a new comer) because of the ambiguity at the group level. This speaks to the tension that can arise due to categorization, where the identity an individual chooses for herself and the identity others impose on her may be at odds. For example, the minority assimilators choose to identify with the majority, but those in the majority pool assimilators with conformists and assign to them an identity of out-group, leading to discrimination.

Moreover, I find that minority members in groups that collectively assimilate make more equitable distributions of resources between the two receivers in the other-other allocation choices. But, the effect is asymmetric between social groups as there is no differential effect on the distributive preferences of the passive majority, who consistently acts biased against the minority receivers. This result complements the literature on identity change and the effects of shared identities, as it indicates that sharing a common identity may not be enough to reduce inter-group biases, unless members of both groups are active in creating this commonality. Otherwise, members from the passive group will benefit from those who assimilate without adjusting their views or attitudes towards them.

The results of my work suggest that policies targeting assimilation by focusing on publicizing the benefits of assimilation, may not succeed in helping minorities and members of disadvantaged groups attain economic mobility. Even when the benefits from assimilation are clear and stark, those who assimilate would not reap those benefits when their social group is fractured by conformist, because those in the dominant group may be unwelcoming. As such, policies should not only target those who can assimilate but also those in the dominant group.

A possible way to convey this message effectively is to inform majority members that by welcoming those who want to assimilate, not only social processes may be more fruitful, but also social preferences can transform, which in the end will benefit majority members and not just those in the minority. Consequently, efforts to promote integration and to reduce out-group tensions should not be placed only on the identity changers but also on the dominant group.

The rest of the paper is organized as follows. In Section 2, I present the experimental game of social coordination. Section 3 contains the experimental design and hypotheses. I report the results of the experiment in Section 4. In Section 5, I explore the effect of choosing between social sanctions or unobservable identities. Finally, in Section 6, I conclude with a discussion.[5]

2 Game

In the experiment, I use a social coordination game where individuals are assigned to either a majority or a minority group. Minority and majority members identify with different social categories and, as such, have prescriptions to behave in opposing ways. In general, each individual benefits from coordinating her behavior with those around her, and the benefit is larger if she behaves in alignment with what is prescribed by her group identity. As such, the size of the group generates economic differences and inequality. That is, the identity of the majority group is more efficient because more people would want to play its prescribed action.[6] Minority players can, however, abandon their group's identity and assimilate that of the majority. Identity change has the potential to foster network connections across groups, resulting in economic mobility for the minority.

[5] In addition, in Appendix A I include the experimental instructions. Appendix B contains all regression tables. Appendix C provides a detailed description of the sample and some additional analysis of the results. Finally, in Appendix D I include a formal model of the social coordination game.

[6] For related models see Hernandez et al. (2013); Ellwardt et al. (2016); Goyal et al. (2021).

2.1 The experimental coordination game

In the experimental game, the population is composed of a majority group of size 4 and a minority group of size 3. The two groups can be *ex-ante* differentiated by appearance markers and identity markers. Appearance markers symbolize traits, such as skin color, hair, size, etc., which are fixed, exogenous to the individual and initially correlate with group identities (see e.g., Efferson et al. 2008).[7] This is displayed by an external symbol: *empty circle* for the majority, ○, and *empty triangle* for the minority, △. On the other hand, group identities are expressed as prescriptions on how to behave, e.g. preferences, views and values of the group, which can be altered by individual decisions.[8] The group identity is displayed by an internal symbol: *filled circle* for a player with the identity of the majority, ●, and *filled triangle* for one with the identity of the minority, ▲. Therefore, at the beginning of the game each player in the majority displays the pair of symbols (○,●) and one in the minority displays (△,▲).

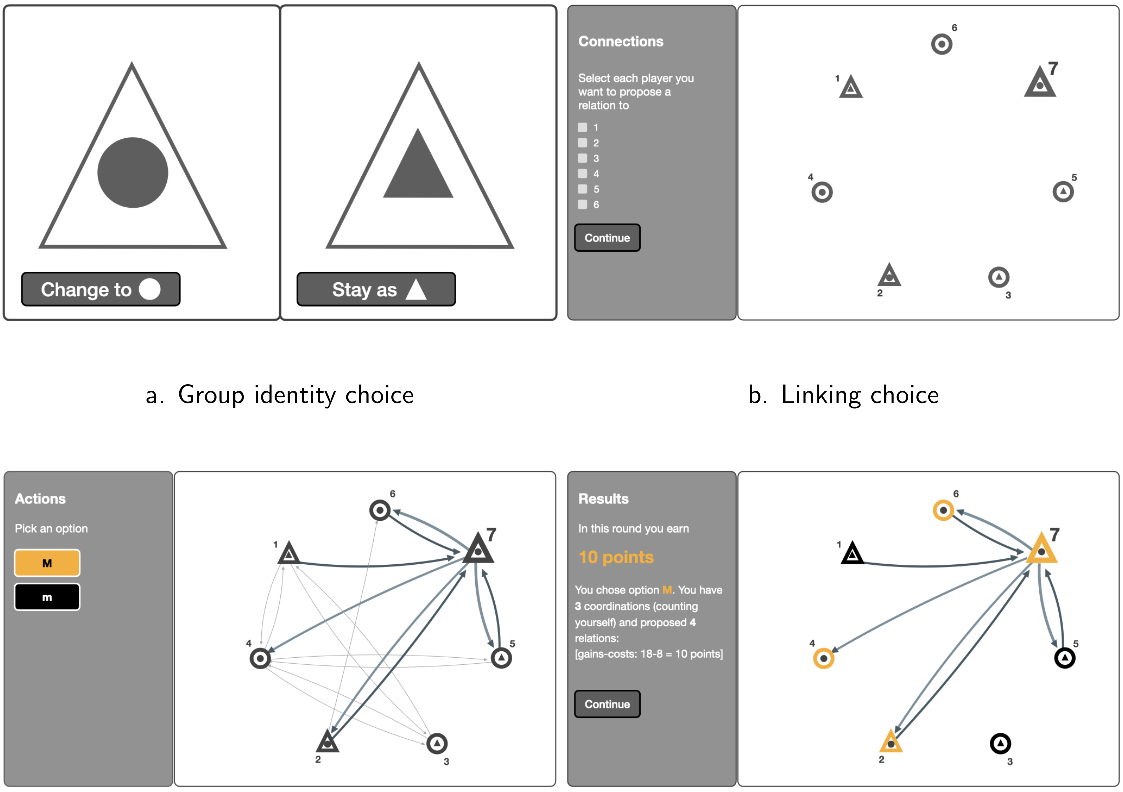
The game has three stages. In Stage 1, every player chooses whether to change their group identity, *assimilate*, or stick with it and maintain the status quo, *conform* (see Figure 1a). While the external symbol is kept fixed, the internal symbol may change depending on the decisions individuals make. This results in four combinations of symbols that visually display external markers and the chosen group identity. Figure 1b shows an example where player 7 is assigned to the minority and assimilates the identity of the majority (△, ●), while 1, also from the minority, is a conformist who sticks to her group identity (△, ▲). Similarly, 6 is a majority conformist who keeps her identity (○, ●), while 5 changes it to assimilate to the minority (○, ▲).

In Stage 2, players observe the external markers and chosen identities of everyone in the population and simultaneously make link proposals to any of the other 6 players. Each proposal, whether reciprocated or not, costs $c = 2$ points. Figure 1b presents the screen players see in the linking stage. Every player is assigned a numeric label to facilitate identification in the current period, and also a node displaying her corresponding pair of symbols (larger for the decision-maker).

In Stage 3, players observe the emerging network. As illustrated in Figure 1c, players are

[7] Arguably, even markers such as skin color could change over time or across generations. Therefore, fixed markers include features that are permanent as well as those that are not, but cannot be altered in the short term. For a detailed discussion on the endogenous construction of identities and the mutability of traits see Chandra (2012).

[8] Group identities and, consequently, prescriptive norms can arise through multiple channels. Some examples include them being exogenously imposed by a process of inter-generational transmission of traits (see e.g., Bisin and Verdier 2001). Also, there are exogenous alterations of social and group identities through publicity and marketing (Akerlof and Kranton 2000), or via organizations, such as universities or the army, which commonly induce identification within its members (see also Andreoni et al. 2021).



a. Group identity choice

b. Linking choice

c. Action choice

d. Results

Figure 1 Screens in the experiment.

1a: Players choose whether to maintain their group identity or assimilate the other group's identity by clicking on the corresponding button. The image above each button illustrates how they will be displayed on the screen given their choice. The internal symbol, ● or ▲, represents the chosen identity, while the external symbol, ○ or △, displays appearance markers that keep track of the original group a player was assigned to. 1b: Players see their own and others' numeric labels and symbols, and choose which links to propose by ticking on the boxes on the left. 1c: Then, they observe the proposals made (outgoing arrows) and the proposals received (incoming arrows) by everyone. Own link proposals are displayed with thicker lines. Players choose an action, M or m, by clicking on the corresponding button on the left. 1d: Players see the action chosen by everyone else, illustrated as the border-color of each node, and a summary of their payoffs in the current period. For simplicity, only the decision-maker's reference network is displayed in the final screen of each period.

informed of the proposals made and received, displayed as incoming and outgoing arrows respectively (thicker for the decision-maker). For example, player 7 has links with $2, 5$ and 6. She does not reciprocate the proposal from 1, while she makes an unreciprocated proposal to 4. Reciprocated proposals result in the creation of links, which lead to the relation of being *neighbors*.

Then, every player decides between two possible actions M or m, the same for all of her neighbors. Building on Akerlof and Kranton (2000), a group identity specifies prescriptions on how to act. I model this through differences in the gains from choosing alike with others. A player who chooses the identity of the majority, ●, prefers action M to action m, and one who chooses the identity of the minority, ▲, prefers action m to M; in both cases irrespective of her external markers. Players earn $v = 6$ points for each coordination if they choose their prescribed action and $v - r = 4$ points for coordinating on their non-prescribed action, where $r = 2$ is the utility loss for acting in misalignment with their group identity.[9] There are no gains from a connection between two players who do not coordinate their actions.

The total payoff a player gets depends on the number of her neighbors choosing the same action she has chosen, denoted by χ, and on the number of link proposals she makes, denoted by n. If a player behaves in alignment with her chosen identity she earns $v(1 + \chi) - cn = 6(1 + \chi) - 2n$. Otherwise, she earns $(v - r)(1 + \chi) - cn = 4(1 + \chi) - 2n$. The 1 in the parenthesis accounts for the choice of the focal player, so that payoffs differ on the alignment of action and identity even when a player does not form any connection in the network and stays isolated, $n = 0$, or when she does not coordinate her actions with anyone else, $\chi = 0$.

At the end of every period, everyone observes the outcome of the game and their net payoffs on the screen, as in Figure 1d. The figure shows that player 7 coordinates successfully on her *prescribed* action with neighbors 2 and 6, and fails to coordinate with 5. She makes link proposals to $2, 4, 5$ and 6. Thus, her net payoff is $(6 \times 3) - (2 \times 4) = 10$ points. At the beginning of any subsequent period, every player is assigned a new numeric label and a new position in the network, to preclude identification across periods.[10]

There is a benchmark version of the game with *fixed identities*, as in Goyal et al. (2021). I will use this game as a within-subject variation in the experiment, where identities cannot be changed. Therefore, unlike the case with identity choice, markers stay fixed, so that

[9] For consistency and comparability with related studies, I use the same parameters as in Goyal et al. (2021). See also Ellwardt et al. (2016).

[10] By randomly assigning new numeric labels and new positions on the screen, direct identification of individual players across periods is prevented, despite interactions being fixed within each group of 7. This is a crucial feature, necessary for treatment IGNORANCE, as will be explained in Section 3.1

all players are displayed with matching internal and external symbols: (○, ●) for the majority and (△, ▲) for the minority. Given identities are fixed, the first stage in the game is the linking choice (see Figure 1b). Subsequently, after making their linking proposals and observing the resulting network, players choose one of the two actions: M or m (see Figure 1c). Payoffs are calculated as in the game with flexible identities described above.

2.2 Equilibrium outcomes

In this section, I provide a brief illustration of equilibrium outcomes. The interested reader can find a formal model of the social coordination game and a more general equilibrium characterization in Appendix D

In a Nash equilibrium of the social coordination game, a player only forms links with others choosing the same action. Otherwise, she receives no benefit from the relation but has to pay the cost of establishing it. This holds independently of the player's identity and chosen action. Now, for a given action profile, a player is better off when she forms a link with everyone else choosing the same action and not only with a subset of them, as every coordination increases her payoffs by 6 points (4 points in misalignment). As this is true for all players, it is easy to see that one can Pareto rank all equilibrium outcomes for a given action profile. The case in which every player is connected to all others choosing the same action and to no one else, Pareto dominates any other outcome where at least one pair of players choosing the same action is not connected.

There are equilibrium outcomes where players who choose a group identity also play an action that aligns with it and others where identity and action are misaligned. Here, I focus on a fraction of the equilibrium set in which the former is true. Specifically, I focus on symmetric equilibrium outcomes in which players choose actions in alignment with their identities[11] and in which any pair of players choosing the same action are connected. Among those outcomes, I narrow my attention to cases in which only minority players change their identity, as these are the relevant outcomes for my research question.[12]

This points to three types of equilibrium outcomes, which vary depending on the identity choices of the minority, as illustrated in Figure 2. The sequence of outcomes displays the

[11] There are other equilibria where players choose an identity in Stage 1 and a behavior in Stage 3 that goes against it. These outcomes are included in the equilbrium characterization, but for simplicity of the description I relegate them to the appendix. Moreover, previewing the findings of the experiment, behavior in the experiment shows strong correlations between the *chosen* group identity in Stage 1 and the action chosen in Stage 3, for both majority and minority players, so that less than 1% of the cases display misalignment.

[12] Outcomes in which the majority players also change their group display the same properties, see Appendix D.

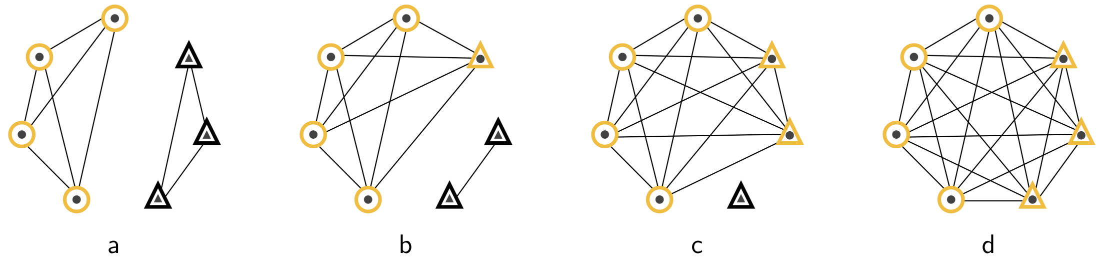


Figure 2 Equilibrium outcomes.

Majority players are illustrated as circles and minority players as triangles. The border color illustrates the action chosen, lighter color mustard when the majority's action M was chosen, and darker color black when the action m was chosen. A line connecting two nodes represents an undirected link. Network 2a portrays the *all-conform outcome*, where all minority players stick to their group identity and segregate. Networks 2b and 2c each illustrate a case of the *fractured outcome*, where the minority is divided into a fraction that assimilates and a remaining fraction of conformists. Network 2d illustrates the *all-assimilate outcome*, where all minority players change their identity and integrate with the majority.

progression from the case where everyone in the minority is a conformist, so no one assimilates the majority's identity, to one where all minority players assimilate. Figure 2a illustrates the *all-conform outcome*, where all players are conformists and maintain their original group identity. Thus, the entire minority segregates from the majority. This is followed by the *fractured outcome*, where either one of the minority players assimilates the majority's identity while the other two conformists stick to their original identity (Figure 2b) or where there is only one minority conformists maintaining the status quo, while the other two assimilate (Figure 2c). Finally, Figure 2d illustrates the *all-assimilate outcome*, where all minority players abandon their group identity to assimilate that of the majority. Thus, the entire minority integrates with the majority.

As every player earns a net of 4 points for each connection in equilibrium (6 points from coordinating minus 2 points for establishing the connection), it is clear that the more integrated a player is the higher her payoffs will be (see Table 1). Consequently, a fraction of minority conformists sticking to their original identity is socially inefficient in the game. This is so, because while assimilators integrate with the majority, conformists fracture the minority and form a separate community from those who assimilate. As such, conformists are connected with fewer others and earn less than assimilators, generating inequalities not only between the minority and the majority, but also within the minority group.

The extreme case where all minority players are conformist leaves the minority players earning 14 points and the majority players earning 18 points. The inequality arises due to differences in size between the majority and the minority. Once a minority player assimilates, her payoffs are the same as those of any majority player. Naturally, the payoffs of a conformist decreases linearly on the number of assimilators, given her network shrinks. On the other extreme, where all minority players assimilate, payoffs are 30 points for everyone, maximizing efficiency and eliminating inequality within and between groups. This is the case where the entire minority attains economic mobility.

3 Experimental design

3.1 Treatments

An experimental session consists of a sequence of four parts that build on each other to test the effect of assimilation on economic mobility: (1) Group assignment, (2) Social coordination game with *fixed* identities, (3) Social coordination game with *flexible* identities, and (4) Other-Other allocation. Individuals know the number of parts in the experiment, but only see the instructions for a specific part once the previous one has finished. I introduce experimental variations in the game with flexible identities (PART 3), while PART 1, PART 2 and PART 4 are identical across treatments. I describe each part below, and explain treatments in the description for PART 3.

PART 1. Group assignment: To induce group identities at the beginning of the experiment, I make use of a combination of different tools. Individuals are randomly assigned to groups and informed they will be represented by an internal and an external symbol (see Figure 1 for an illustration). They are also informed of their group size (4 for the majority and 3 for the minority) and that the symbols identifying those in the other group are different. Subsequently, individuals in each group are connected via a chat box through which they can communicate. The goal is to choose a name for their group out of a list of 5 names. For a name to be chosen, all players in the group must vote for the same option. Those in the majority chose a name between a list of five felines (*Cats, Tigers, Lions, Leopards, Jaguars*) and those in the minority chose between five canines (*Dogs, Jackals, Coyotes, Foxes, Wolves*). I use different symbols and names to prime group identity. In addition, I use chat communication and a group voting task to further induce within-group identification and between-group differentiation (Eckel and Grossman 2005; Chen and Li 2009).[13]

PART 2. *Fixed* group identities: Once the group assignment stage concludes, the 7 players play together the social coordination game without identity choice for 10 periods. This is the 2-stage game where players choose links and then actions. The aim of PART 2 is two-fold. First, it is to further enhance in-group identification and out-group differentia-

[13] The group chat takes place at the beginning of the experiment but before the players receive any information about the coordination game they will play. Because of this, there is no room for players to talk about potential strategies for the game. During the experiment, I do not use *majority* and *minority* as labels, and instead use the symbols and chosen names to refer to each group.

tion. This because preferences over outcomes are induced through differences in payoffs between groups, but players cannot abandon their group identity to assimilate that of the other group. Second, this part serves as a benchmark to compare outcomes with the case of flexible identities (PART 3), which allows me to empirically evaluate whether the possibility to assimilate can be used by minority players to attain economic mobility.

PART 3. *Flexible* group identities: After the 10 periods of PART 2, players play together the social coordination game with identity choice for 10 periods. This is the 3-stage game where players choose identities, links and actions. In PART 3, it is possible for minority players to assimilate the identity of the majority. Identity change can maximize efficiency and reduce inequality when compared to the case in which identities are fixed.[14] I assign players into one of three main treatments: BASELINE, SANCTIONS, or IGNORANCE, which I use to asses how different settings impact the assimilation choices of the minority and its chances of attaining economic mobility (see Table 1 for details). I explain each treatment as follows:

- BASELINE: In the baseline treatment there are no barriers to assimilation, so that identity change is free and visible. That is, any player can abandon her assigned identity and assimilate that of the other group at no cost, independently of the group identity others choose. Also, assimilators can be clearly identified by everyone else before the linking proposals are made, because the internal symbol is set to display the chosen identity of every player (see Figure 1b).

- SANCTIONS: This treatment explores the effect of social sanctions. Anyone who changes her group identity can be distinguished from those who maintain it, as in the BASELINE, but assimilation is costly. The cost of changing identities for a minority assimilator is $\delta = 6 + 2 \cdot (n_m - \tau)$, where $n_m = 3$ is the size of the minority group and τ is the number of minority members assimilating, including herself.[15] This means that anyone who changes her group identity pays a fixed cost of 6 points, plus a variable cost that increases by 2 points for every conformist in her group who sticks to the original identity. In this way, I exogenously impose social sanctions from conformist to assimilators.

- IGNORANCE: This treatment explores the effect of identities not being observable. There is no cost to assimilate, as in the BASELINE, but the choice is not visible to

[14] Note that the aim of a sequence going from fixed to flexible identities is to evaluate the benefits of identity change in solving economic disparities that arise in an environment where assimilation is not possible. Because the interest of this sequence is contingent on its direction, I do not test for order effects by running treatments where the game with flexible identities is played before that with fixed identities.

[15] The assimilation cost for a majority player is $\delta = 6 + 2 \cdot (n_M - \tau)$, where $n_M = 4$.

Table 1 Experimental treatments and payoffs

Description of Parts and Treatments, as well as summary of payoffs (in points). Payoffs are displayed separately for minority players who assimilate or conform, as well as for the majority. For the case of fixed identities, a minority player is said to assimilate when choosing the action of the majority and conforms when choosing the action of the minority, as identities cannot be changed.

	Fixed identities	**Flexible identities**		
		BASELINE	SANCTIONS	IGNORANCE
Identity change				
Costly	N/A	×	✓	×
Visible identity	✓	✓	✓	×
Payoffs if All conform (Figure 2a)				
Min. Assimilate	.	.	.	.
Min. Conform	14	14	14	14
Majority	18	18	18	18
Payoffs if Fractured #1 (Figure 2b)				
Min. Assimilate	12	22	12	22
Min. Conform	10	10	10	10
Majority	22	22	22	22
Payoffs if Fractured #2 (Figure 2c)				
Min. Assimilate	14	26	18	26
Min. Conform	6	6	6	6
Majority	26	26	26	26
Payoffs if All assimilate (Figure 2d)				
Min. Assimilate	16	30	24	30
Min. Conform	.	.	.	.
Majority	30	30	30	30

others. As such, assimilators cannot be differentiated from conformist who stick to their group identity. Players can see each others' external markers, which signal the group they were initially assigned to. However, no one can identify each others' chosen identity. I achieve this by eliminating the internal symbol from all nodes on the screen, irrespective of their identity choice.

PART 4, Other-Other allocation: After the end of the social coordination game with flexible identities (PART 3), every player is asked to allocate 10 points between two anonymous receivers. There is one receiver from the minority and another from the majority, and both are drawn from the same group of 7 as the allocator. No one was allowed to allocate points to herself. The novelty of including this task after the game with flexible identities is to explore if assimilation leads to positive inter-group attitudes, beyond the setting for which the assimilation of an identity was relevant. Specifically, I explore if by sharing a common identity players from the majority and the minority make more equitable allocations between the two receivers.

3.2 Hypotheses

I now present the hypotheses I test in the experiment for PART 2, PART 3, and PART 4.

In PART 2, players play the social coordination game with fixed identities, which replicates Goyal et al. (2021). The main results of their work suggest that when identities are fixed and linking is costly, players only connect with others in their group: minority with minority and majority with majority. Consequently, everyone chooses actions in line with their group identity. Thus, I postulate the following hypothesis for PART 1:

Hypothesis 1 *In a setting with fixed identities, minority and majority players avoid interactions across groups and choose actions in alignment with their assigned group identity.*

Payoffs in PART 2 are the benchmark to which I compare what assimilators get in the game with flexible identities. The minority is said to achieve economic mobility when their payoffs are higher than those attained in the setting with fixed identities.

In PART 3, players play the social coordination game with flexible identities. As mentioned above, groups are assigned into one of three treatments: BASELINE, SANCTIONS or IGNORANCE. In all three treatments, minority players can assimilate to integrate with the majority and coordinate with a larger number of neighbors than in an outcome with fixed identities. Among all outcomes, the *all-assimilate outcome* is socially efficient, leads to economic mobility and reduces inequality, within and between groups. This is true across all treatments. Thus, I postulate the following *null* hypotheses for PART 3.

Hypothesis 2 *In a setting with flexible identities, all minority players assimilate the majority's identity and all together form a complete network where everyone chooses the same action* M. *This leads to economic mobility for the entire minority.*

Alternatively, it may be that not all individuals in the minority assimilate, but instead some conformist stick to their original identity, when they are exposed to social SANCTIONS or IGNORANCE. Exposure to such barriers makes identity change riskier than in the BASELINE. If only one minority player assimilates, the entire minority group could end up worse-off than if there was no assimilation at all. Specifically, for the BASELINE treatment, as shown in Table 1, any number of minority players can assimilate and benefit from this choice.[16] For the SANCTIONS treatment, a minority player is worse-off assimilating if she is the only one. But, when two minority players assimilate, each can get more than in the *all-conform outcome*.[17] In the IGNORANCE treatment, if a single minority player assimilates, it is not possible for her to be individually identified as the one changing her group identity, even if everyone has correct beliefs on the number of assimilators. This is also true in the case of two assimilators, even though the probability of being identified increases. If all three minority players assimilate, and beliefs are correct, it is as if there was complete information, and all players can maximize their payoffs. As such, when there are barriers to assimilation, the *all-assimilate outcome* is the first-best and the *all-conform outcome* is the second best.

In summary, assimilation is riskier in SANCTIONS and in IGNORANCE than in the BASELINE, because both assimilators and conformists can end-up worse-off in a *fractured outcome* than in an *all-conform outcome*. I postulate the following *alternative* hypothesis for treatments SANCTIONS and IGNORANCE in PART 3:

Hypothesis 3 *If the minority is exposed to social sanctions or their identities cannot be observed, not all players assimilate. Players segregate into two communities, one choosing the majority's identity and another conforming to the minority, resulting in either the all-conform outcome or a fractured outcome. In these cases, not all minority members attain economic mobility.*

Finally, in PART 4 I explore how changes in group identities impact social preferences, specifically distribution preferences between the minority and the majority. This is a way to explore if the benefits from assimilating extend beyond the context in which it occurs. The existing experimental evidence on similar types of allocation choices consistently shows that individuals are likely to bias their allocation in favor of the receiver from their same group (Chen and Li 2009; Cikara and Van Bavel 2014).

In my experiment, however, after being exogenously assigned an identity, players have the option to endogenously change it (and consequently change their reference group) in the game with flexible identities. I use this choice to evaluate if assimilation leads

[16] As displayed in Table 1, if a single minority player assimilates, she gets 22 points per period, compared to 14 points she gets in the *all-conform outcome*. Naturally, if all three assimilate, each minority player gets 30 points, which is more than twice the payoff of conforming.

[17] As displayed in Table 1, if only one minority player assimilates, she gets 12 points because she has to pay an assimilation cost of 10 points. Thus, it is only when two assimilate that each would get 18 points and improve compared conforming, given the cost is reduced to 8 points. If all three minority players assimilate, each gets 24 points and the entire minority achieves economic mobility, because the *all-assimilate outcome* is socially efficiency and Pareto dominant. In SANCTIONS, unless two out of three assimilate, payoffs from a *fractured outcome* are not higher than those in the *all-conform outcome*.

majority and minority individuals to see each other as sharing a common identity, despite being originally assigned to opposing out-groups. If this were the case, the more the minority collectively assimilates (the fewer conformists there are) the more equitable the allocations between the two receivers will be. Thus, reducing in-group bias. Consequently, I postulate the following hypothesis:

Hypothesis 4 *The more the minority collective assimilates the more likely majority and minority players are to make equitable allocations between the two receivers.*

3.3 Experimental procedures

The experiment was conducted at the experimental laboratory of the University of Rosario (REBEL). A total of 336 individuals took part in the study.[18] Subjects interacted through computer terminals and the experiment was programmed using oTree (Chen et al. 2016). Each session lasted on average 120 minutes, including the time used to read the instructions and to anonymously pay.

From the beginning, subjects were informed that the study consisted of a sequence of four parts played within the same group of 7. Instructions were administered on the screen at the beginning of each part, so subjects did not know anything specific about a subsequent part before it started. After reading the instructions, subjects were presented with a set of comprehension questions and could not advance until all answers were correct. At all times, they could click on a button on the screen and a summary of the instructions for the corresponding part would be displayed (see Appendix A).

With respect to earnings, one of the 10 periods of play was randomly selected for payment from PART 2 (game with fixed identities) and one from PART 3 (game with flexible identities). For PART 4 (the allocation choice) the decision from one of the 7 subjects in a group was randomly selected for payment. In the instructions for each part, subjects received information about the way payment for the current part would be calculated. However, they were only informed of their actual earnings at the end of the experiment. Subjects were paid using the exchange rate of 2 points = 800 Colombian Pesos (COP). On average, players earned 46,000 COP (Approx. 15.5 USD), including a 3.5 USD show-up fee. The standard conditions of anonymity and non-deception were implemented in the experiment, and no one participated in more than one session.

[18] I conducted 3 sessions with 4 groups of 7 subjects for each treatments, including those in the treatment REVEAL discussed in Section 5. All subjects were undergraduate students, out of which 58% where female. See Appendix C for more details on the sample.

4 Results

In this section, I summarize the main findings of the experiment. The data from the coordination games consist of decisions made in two Parts of 10 periods. There are 12 groups of 7 players in each of the main treatments: BASELINE, SANCTIONS and IGNORANCE (I explore the results from treatment REVEAL in section 5), resulting in 360 observations at the group level for each part. To control for any interdependence across periods, I run random effects GLS regressions, clustering standard errors on groups. I report two-sided p-values in the text and provide all regression estimates in Appendix B.

Table 2 Summary statistics

Average fractions (percentages) for each of the main variables. Standard deviations in parenthesis. Observations are disaggregated by majority and minority; pooled for the game with *fixed* identities and separated by treatment for the game with *flexible* identities.

	Fixed identities		**Flexible identities**					
			BASELINE		IGNORANCE		SANCTIONS	
	Major	*Minor*	*Major*	*Minor*	*Major*	Minor	*Major*	*Minor*
Assimilation	n/a	n/a	0.01	0.95	0.01	0.65	0.01	0.59
			(0.09)	(0.21)	(0.08)	(0.48)	(0.08)	(0.49)
Inter-group connectivity	0.04		0.77		0.42		0.51	
	(0.13)		(0.33)		(0.44)		(0.43)	
Action alignment	0.98	0.89	0.99	0.98	0.99	0.96	0.99	0.91
	(0.14)	(0.31)	(0.06)	(0.16)	(0.09)	(0.20)	(0.06)	(0.28)
Efficiency (mobility)	0.88	0.75	1.42	1.68	1.18	1.15	1.28	0.91
	(0.25)	(0.36)	(0.33)	(0.57)	(0.39)	(0.71)	(0.35)	(0.67)
In-group bias	n/a	n/a	0.36	0.22	0.39	0.37	0.31	0.48
			(0.42)	(0.37)	(0.43)	(0.38)	(0.40)	(0.45)

Table 2 reports summary statistics of the main variables, separate for the majority and the minority. *Identity change* is the fraction of players in either the minority or the majority assimilating the identity of the other group (not applicable when identities are fixed). *Inter-group connectivity* reports the fraction of links formed across the majority and the minority.[19] *Action alignment* is the share of players who play an action in alignment with their (chosen) identity. *Efficiency* is the fraction of points an individual attains per period, normalized by the expected earnings in the *all-conform outcome*.[20] This is an indication of economic mobility; where anything above 1 means earnings are higher than predicted

[19] See Appendix C.1 for a detailed summary of other connectivity variables, such as within-group connectivity and link proposals.

[20] The benefit of normalizing efficiency by the payoffs predicted for the case in which everyone conforms is that the value is the same for the benchmark game with fixed identities and across treatments in the game with flexible identities (see Table 1). Note that the *all-assimilate outcome* is the first-best outcome, for which efficiency is 2.1 for the minority in the BASELINE and IGNORANCE treatments and 1.7 in the SANCTIONS treatment.

for the second-best outcome, where no one assimilates. Finally, *in-group bias* is a measure from the Other-Other allocation choice. It refers to the difference between the allocation to the in-group receiver and the out-group receiver, normalized by the total endowment of 10 points.[21]

The first part of Table 2 summarizes the results for the social coordination game with *fixed* identities. The emerging outcome becomes clear by looking at the sequence of variables described. The first choice is inter-group connectivity, which shows that networks were practically segregated between minority and majority. On average, 4% of the 12 links that could exist between the two groups were formed; this is practically less than 1 connection across groups. In such segregated networks, both majority and minority players are very consistent in playing the action that aligns to their group identities, 98% for the majority and 89% for the minority. Consequently, they both achieve a level of efficiency below but somewhat close to 1, which is the maximum attainable in the segregated outcome. Thus, providing support to Hypothesis 1.

The second part of Table 2, summarizes the results of the game with *flexible* identities, by treatment. It is evident that when given the option to change identities, majority members chose to keep theirs 99% of the times and there are no differences between treatments. In terms of inter-group connectivity, it dramatically increases from fixed to flexible identities. Although, there are still some challenges for the minority and the majority to properly integrate, especially when there are barriers to assimilation. As with fixed identities, action alignment is high for both majority and minority. Specifically, for the minority players, who are the ones assimilating, they play the action that matches their identity 98% in the BASELINE, 96% in IGNORANCE, or 91% in SANCTIONS. Thus, resulting in levels of efficiency above 1 in almost all cases (except for the minority in SANCTIONS), which suggests that, on average, the possibility to assimilate allows the minority to attain economic mobility. I explore this in detail in the next section.

4.1 Assimilation and economic mobility

In this section, I look at how identity choices affect inter-group connectivity and economic mobility. I explore this effect in settings where players are exposed to either social SANCTIONS or IGNORANCE, and compare it to the BASELINE without barriers to assimilation.

With respect to the minority, Figure 3A shows that in the BASELINE the average level of assimilation is 95%, which drops to 65% in IGNORANCE ($p = 0.007$) and to 60% in

[21] For ease in the exposition, I report the allocation bias under the treatment columns. That is why there is n/a for the fixed identity columns.

SANCTIONS ($p < 0.001$), where the effect of SANCTIONS and IGNORANCE is not statistically distinguishable ($p = 0.684$). Thus, at the individual level, both social sanctions and unobservable identities reduce but do not outright prevent assimilation from the minority.

I take a deeper look at the data, evaluating the extent to which assimilators and conformists co-exist in the same group. I also explore which barrier to assimilation is more conducive of fracturing the minority in that way. Figure 3B shows that minority groups exposed to social sanctions are less likely to avert the *fractured outcome* when compared to groups in which identities could not be observed or to groups for which assimilation was cost-less (see the black fraction of the bars): about 50% of the groups in SANCTIONS are fractured, which is twice more than the 25% in IGNORANCE ($p = 0.051$) and four times more than the 12% in the BASELINE ($p = 0.001$).

On the other hand, the setting without barriers is most fertile to promote the *all-assimilate outcome* (88%). This is also the most frequent outcome with unobservable identities (55%) and the second most frequent with social sanctions (35%). This indicates that when a minority player assimilates, she is most likely to be surrounded by other assimilators in the BASELINE, and most likely to be in a fractured minority, surrounded by conformists in SANCTIONS.

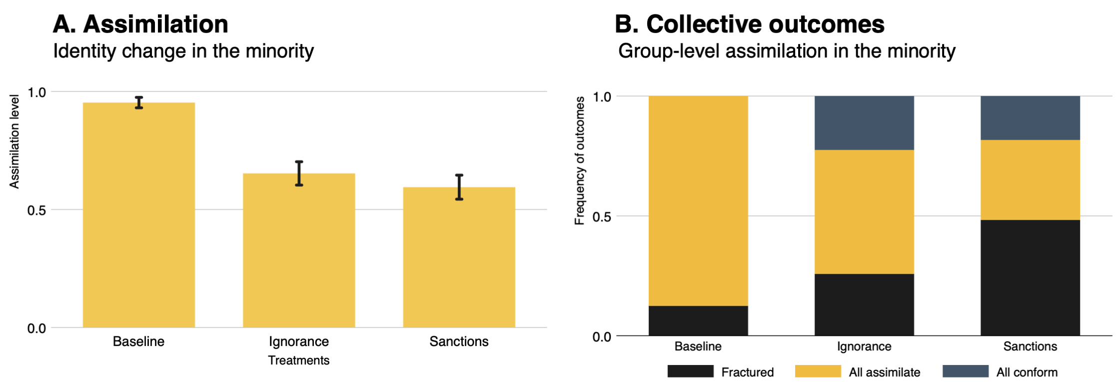


Figure 3 Individual and collective assimilation

Panel A illustrates the average level of assimilation of the minority. Lines represent 95% confidence intervals. Panel B illustrates the frequency of each outcome (*fractured outcome*, *all-assimilate outcome*, and *all-conform outcome*). Results in both panels are disaggregated by treatment.

The natural question to follow is how the difference in assimilation patterns and collective outcomes impact the chances of the minority to integrate with the majority. Figure 4A illustrates inter-group connectivity. For the fixed-identity stage, as shown before, connectivity is close to null, such that players create incomplete networks where they segregate by majority and minority: 96% of the links across groups are missing. Compared to fixed identities, inter-group connectivity significantly increases to 77% ($p < 0.001$) in the BASELINE. The increment is evident but not as large in SANCTIONS ($44\%, p < 0.001$) or in

IGNORANCE ($51\%, p = 0.004$).

Next, I focus exclusively on the minority members that assimilate. It is efficient that any assimilator proposes a connection to each of the four majority members and all proposals should be reciprocated. This because assimilators and majority players share the same identity and thus, have aligned incentives over which action they would prefer to coordinate on. In this case, the four majority players and the assimilator benefit the most if they connect, independently of whether the other minority members are conforming or assimilating. The results show, however, that reciprocity is significantly lower for assimilators when the minority is fractured by conformists, even if the assimilation choice were observable (i.e., the effect is not a consequence of unobservable identities in IGNORANCE).

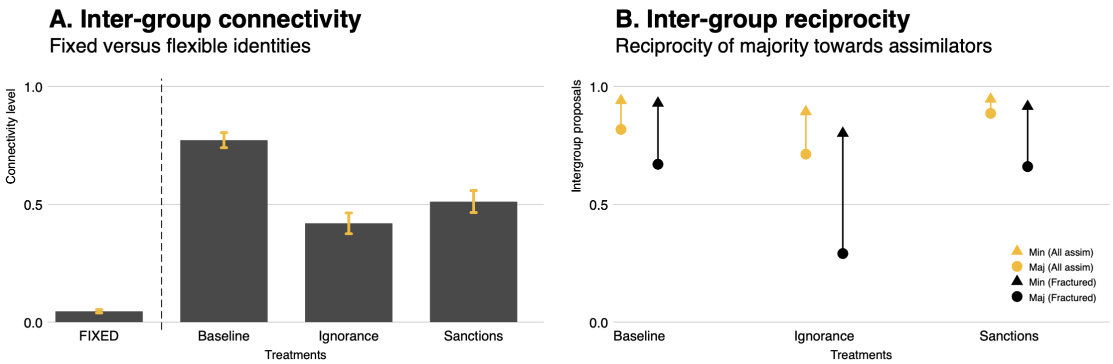


Figure 4 Inter-group connectivity and reciprocity

Panel A illustrates the average level of connectivity between the minority and the majority for the game with *fixed* identities, as well as for the main treatments with *flexible* identities. Panel B illustrates the level of reciprocity from the majority towards assimilators. The top of each line (triangle) indicates the fraction of link proposals made by the assimilator to the four players in the majority. The bottom of each line (circle) is the fraction of proposals from the majority to the assimialtor. The length of the line indicates the fraction of unreciprocated proposals by the majority, separate for the *all-assimilate outcome* (light color) and the *fractured outcome* (dark color). Results in both panels are disaggregated by treatment.

Figure 4B shows the level of inter-group connectivity of minority assimilators, separate for the *all-assimilate outcome* (light color) and the *fractured outcome* (dark color). The top of each connected line (triangle symbol) shows the level of proposals made by an assimilator to those in the majority (1 means she made a proposal to all 4 majority members), the bottom is the level of proposals made by the majority members to the assimilators (circle symbol), and a line's height is the fraction of unreciprocated proposals from the majority to the assimilators.

When the minority acts as a collective unit and all members assimilate, individuals form on average 82% of the links in the BASELINE, 71% in IGNORANCE ($p = 0.194$), and 88% in SANCTIONS ($p = 0.608$). The fact that treatment differences disappear indicates that unity facilitates individual connectivity. However, when the minority is fractured, average inter-group connectivity is 67% in BASELINE, 29% in IGNORANCE ($p < 0.001$) and 66% in SANCTIONS ($p = 0.590$). This suggests that integration is harder for assimilators when others in the minority are conformists, not because they make fewer proposals to the majority, but because the majority proposes back fewer times. That is, assimilators behave in the same way towards the majority regardless of what others do. But, the majority reciprocates differently to the same number of proposals received from those who assimilate, depending on what they see others choose.

The next step is to look at the effect of assimilation on payoffs, and consequently on economic mobility. Figure 5A shows the payoffs minority players got when they chose to conform. If the entire minority conformed in unity (lighter bars) their payoffs were close to the maximum level of efficiency of 1, while if the minority was fractured payoffs for conformists where reduced to almost half. This is the natural effect of sticking to their group identity when others in the minority assimilate, as their networks shrink.

Figure 5B shows the case for assimilators. Naturally, assimilating when others also assimilate (lighter bars) gives the highest payoffs, as minority assimilators integrate much better if there are no conformists fracturing their group. In contrast, assimilating within a fractured minority gives lower payoffs than those of assimilators in a united minority (compare dark versus light bars in Panel B). Furthermore, if one compares the light bars on Panel B with the dark bars on Panel A, for each treatment, it is notable that an assimilator in a fractured minority would have been better off staying a conformist, and persisting on the inefficient group identity, in an all-conformist minority. For instance, a conformist in SANCTIONS, would get an average payoff of 0.79 points if the minority is united (all conform), while she would receive a payoff of 0.63 if she assimilates but some minority players conform. The same effect is observed when comparing the payoffs for the IGNORANCE treatment ($0.93 > 0.48$).

In summary, the main findings of the experiment indicate that minority players can strategically assimilate the group identity of the majority to achieve economic mobility. However, the success of their choice is dependent on whether their social group (the minority) acts in unity, so that all others assimilate as well, or is fractured by conformists.

The presence of conformists imposes negative externalities on assimilators. These are observed in the way the majority reciprocates to their intentions to integrate. Majority players significantly reduce reciprocity towards assimilators when they also encounter conformists in the minority group. Lowering reciprocity reduces inter-group connectivity and directly impacts payoffs. Assimilators in a fractured minority not only get a lower payoff than assimilators in a united minority, but also get less than if they had not assimilated and stayed in a united minority of only conformists. However, economic mobility is

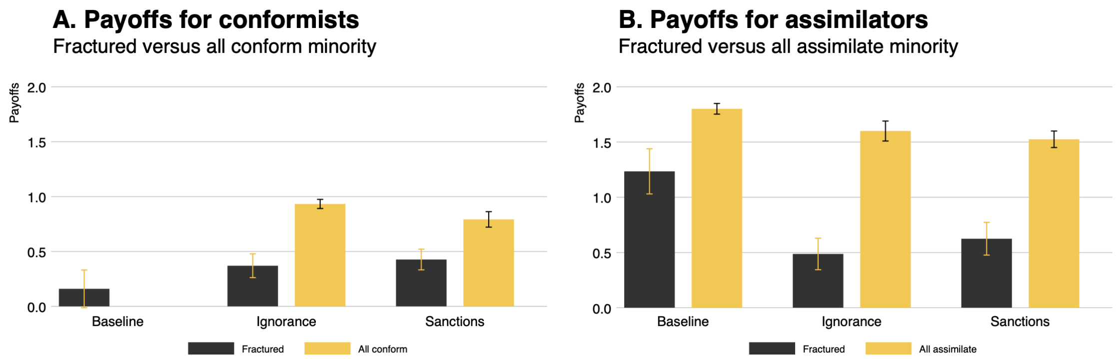


Figure 5 Payoffs (efficiency) and economic mobility.

Panel A illustrates the level of payoffs minority players got when they conformed and the rest of the minority acted in unity (*all-conform outcome*) or was fractured. Panel B illustrates payoffs for minority assimilators, also when the rest assimilated as well or when it was fractured. Results in both panels are disaggregated by treatment.

achieved if the minority assimilates in unity, irrespective of the barriers faced.

The results give support to Hypothesis 3, which predicts lower levels of assimilation and more frequent fractured minorities when there are barriers to assimilate. However, an additional insight from the findings suggests that social sanctions are more likely to produce fractures in the minority. I explore this further in Section 5, by looking at the REVEAL treatment.

In what follows, and to close the results section, I look at whether assimilation and collective unity within the minority has positive spillovers on social (distributive) preferences between groups, in a setting that goes beyond the social coordination game.

4.2 Identity change and social preferences

In this section, I measure differences in the allocation players make to a receiver from their in-group and to a receiver from their out-group, at the end of the experiment. The difference between the allocations represents the magnitude of the *in-group bias*. Following Hypothesis 4, equitable allocations (no bias) are expected to occur the more a player perceives the out-group receiver as someone she identifies with. Figure 6A reports the magnitude of in-group bias for the minority.

The findings support the predictions in Hypothesis 4, that allocations to the out-group are expected to be positively related to the minority collectively assimilating. Allocations in the BASELINE are biased by 22%, the bias in IGNORANCE is higher but not significantly different to the BASELINE (37%, p = 0.101), while the bias in SANCTIONS is significantly larger than in the BASELINE (48%, p = 0.012). That is, the more minority players assimilate collectively as a unit, the more equitable are their preferences towards the out-group.

Figure 6B illustrates the magnitude of in-group bias for the majority, and shows that there are no differences in bias in the BASELINE (36%) with respect to IGNORANCE (39%, p = 0.728) or to SANCTIONS (31%, p = 0.528). So that even when the minority assimilates and acts the way the passive majority prefers, the majority seems unaffected by it. Thus, collective assimilation can transform social (distributive) preferences for the active group (the minority), but not for the passive group that maintains the status quo (the majority).

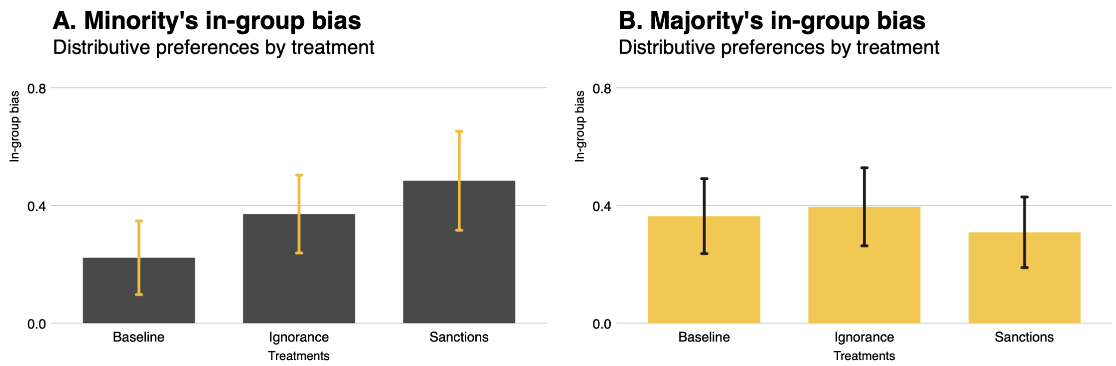


Figure 6 In-group bias in Other-Other allocation, by treatment.

The graph shows the magnitude of allocation bias towards in-group members for the minority (Panel A) and the majority (Panel B). Results in both panels are disaggregated by treatment.

5 Revealing identity choices

In this section, I report results on treatment REVEAL, where I allow individual players to endogenously select whether to conceal or to reveal their chosen identities. I do not conduct a treatment combining social sanctions and limited observability of identities, as argued by (Fryer and Torelli 2010; Bisin et al. 2011), these two barriers are mutually exclusive. An individual can either signal her assimilation of a new group identity or conceal it. Making the assimilation choice visible guarantees integrating with the out-group but exposes the assimilator to social sanctions from her in-group. Keeping it hidden impedes sanctions but puts the assimilator in a setting of limited observability, where her choice is not identified by either the out-group or other assimilators in her in-group. I impose each barrier exogenously and separately in the main treatments (SANCTIONS and IGNORANCE) to better understand the way social groups confront them. To complement this choice in design, I conduct a the REVEAL treatment where I make barriers endogenous.

In the REVEAL treatment, players can assimilate at no cost but their chosen identity will not be observable by others (as in IGNORANCE). Then, players choose whether to *reveal* their identity, which then exposes them to sanctions that increase on the number of conformists (as in SANCTIONS). This is implemented for the minority through a *revelation cost* that has the same form as the assimilation cost in SANCTIONS: $\delta_{\text{REVEAL}} = 6 + 2(n_m - \tau)$, where $n_m = 3$ and τ is the number of in-group members who assimilate.[22]

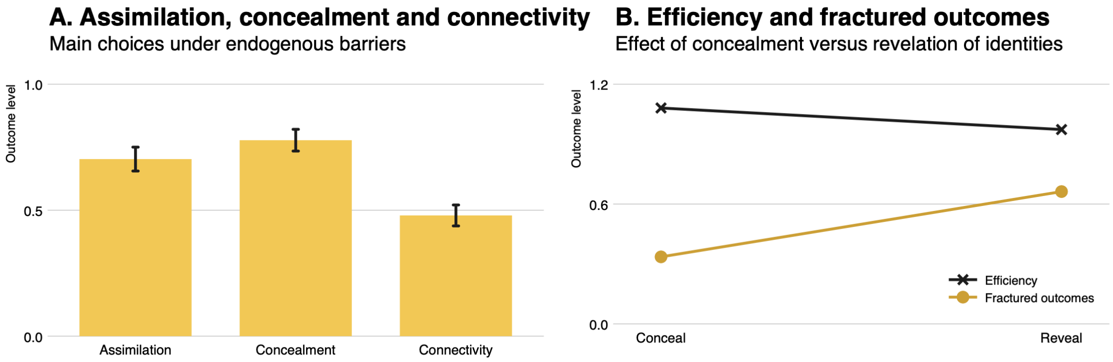


Figure 7 Main choices in treatment Reveal (RV).

Panel A illustrates the average level of assimilation, identity concealment (choosing not to reveal), and inter-group connectivity. Lines represent 95% confidence intervals. Panel B illustrates the level of efficiency (line with X) and the frequency of *fractured outcome* players are in, conditional on revealing or concealing their chosen identities.

The main question with endogenous barriers is whether the choice individuals make of which barrier to expose themselves to leads to outcomes that reflect similar patterns to those observed with exogenous barriers. Figure 7A shows that minority players assimilate 70% of the times. This is not distinguishable from either IGNORANCE ($p = 0.726$) or SANCTIONS ($p = 0.393$). Once players choose their identities, less than a quarter (22%) revealed them. Instead, most players (78%) kept their identities concealed. The consequent level of inter-group connectivity is 48%, which is significantly lower for players concealing than revealing their group identity ($39\% < 77\%, p < 0.001$).

Figure 7B shows that the minority attains a slightly higher, although not significantly different, level of efficiency when identities are concealed than revealed (1.1 versus $0.97, p = 0.196$), in line with the results above presented. But, the most notable outcome is that individuals revealing their chosen identities end up in a *fractured outcome* significantly more often than those who conceal them.[23] In fact, when all three minority players assimilate, there is significantly less revelation of chosen identities than when some assimilate and others conform ($p < 0.001$), which is indicative that being exposed to social sanctions is likely to be related to a fractured minority. Thus suggesting that by endogenously avoiding social sanctions, the minority is able to act cohesively and attain economic mobility, even

[22] For the majority the revelation cost is $\delta_{\text{REVEAL}} = 6 + 2(n_M - \tau)$ given $n_M = 4$.

[23] The frequency of the *fractured outcome* is 40%, which is lower but neither significantly different from SANCTIONS ($p = 0.606$) nor from IGNORANCE ($p = 0.209$).

though their chosen identities are concealed.

6 Conclusions

In this paper, I show that two minority assimilator will be treated differently, either welcomed or discriminated against by those in the majority they want to join, depending on the choices of those other minority players around them. This means that even in a setting where the benefits of assimilation are stark and sizeable, assimilators may fail to benefit from their choices when the minority group is fractured into assimilators and conformists. Those who persist on maintaining the status quo impose negative externalities on those who risk to change it.

The different experimental variations of my design allow me to identify the robustness of this result. No matter how hard it was for a minority player to assimilate, whether she had to overcome social sanctions or risk assimilating in a setting where her identity choices would not be observable, the outcomes are consistent and assimilation fails when the minority fractures. This is a strong indication that the success of assimilation cannot be placed on the individual alone, but must involve the awareness and participation of those in the dominant group.

A limitation of my work is that individuals may be less attached to their induced identities compared to natural or real-world identities (Tajfel and Turner 1979; Eckel and Grossman 2005; Charness et al. 2007). If the induced identities make it easier for individuals to abandon them and assimilate, then the observed levels of assimilation in the real world would be more modest than those in my study. However, if anything, this provides stronger support to the main finding that assimilation fails when the minority group is fractured. If in my setting, where interactions may be thought of a purely transactional, the majority discriminates against assimilators because of what other minority members do, it is likely this will become harsher when individuals are more attached to their group identities.

Arguably, the other-other allocation decision (Chen and Li 2009) at the end of the experiment is a complementary dimension that shows how individuals (to a certain extent) identified and attach meaning to the group they were assigned to. The findings show that across treatments there is an in-group bias in allocations, suggesting thay by making more equitable allocations between majority and minority receivers, assimilators internalize the other group's identity as well. Thus, providing support to the idea, already persistently observed in multiple studies, that experimentally induced identities can be a strong and useful approach to understand group processes and social change (Chen and Chen 2011).

A challenging result from the other-other allocation decision is, however, that those in the majority do not change their social preferences towards the minority, irrespective of how much effort assimilators put into changing their group identities. Arguably, the passivity of the majority prevents the creation of a common identity for them, although it is not precluded for the minority. This inability to see others as part of one's group because of not having to do anything to make it happen, can be an important source of discrimination in real world settings.

My work suggests two avenues for further research. One is to directly evaluate how likely are individuals to change identities in settings where group identities were formed outside the laboratory. This can help calibrate the impact of identity attachment on assimilation choices and outcomes. A second direction is to evaluate the alternative setting of a third social category, so that instead of the minority assimilating the identity of the majority, both minority and majority can create a new group identity together. Possibly, by eliminating the focality of the majority, this would prevent them from being passive. If members from the majority and the minority work together to create a new social category, they could perceive each other as sharing a common identity, which would reduce discrimination and benefit both groups.

References

Akerlof, G. (1997). Social distance and social decisions. Econometrica, 65:1005–1027.

Akerlof, G. (2016). "we thinking" and its consequences. American Economic Review: Papers & Proceedings, 106:415–419.

Akerlof, G. and Kranton, R. (2000). Identity and economics. Quarterly Journal of Economics, 115:1–39.

Alesina, A. and La Ferrara, E. (2002). Who trusts others? Journal of Public Economics, 85:207–234.

Andreoni, J., Nikiforakis, N., and Siegenthaler, S. (2021). Predicting social tipping and norm change in controlled experiments. Proceedings of the National Academy of Sciences, x:1–1.

Austen-Smith, D. and Fryer, R. G. (2005). An economic analysis of "acting white". Quarterly Journal of Economics, 120:551–583.

Battu, H., Mwale, M., and Zenou, Y. (2007). Oppositional identities and the labor market. Journal of Population Economics, 20:643–667.

Battu, H. and Zenou, Y. (2010). Oppositional identities and employment for ethnic minorities: Evidence from england. The Economic Journal, 120:F52–F71.

Bernard, M., Hett, F., and Metchel, M. (2016). Social identity social free-riding. European Economic Review, 90:4–17.

Bernheim, B. (1994). A theory of conformity. Journal of Political Economy, 102:841–877.

Bicchieri, C. (2006). The Grammar of Society: The Nature and Dynamics of Social Norms. Cambridge University Press.

Bicchieri, C. (2016). Norms in the wild. How to diagnose, measure and change social norms. Oxford University Press.

Bisin, A., Patacchini, E., Verdier, T., and Zenou, Y. (2011). Formation and persistence of oppositional identities. European Economic Review, 55:1046–1071.

Bisin, A., Patacchini, E., Verdier, T., and Zenou, Y. (2016). Bend it like beckham: Ethnic identity and integration. European Economic Review, 90:146–164.

Bisin, A. and Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. Journal of Economic Theory, 97:298–319.

Boyd, R. and Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. Ethology and Sociobiology, 13:171–195.

Boyd, R. and Richerson, P. J. (2002). Group beneficial norms can spread rapidly in a structured population. Journal of Theoretical Biology, 215:287–296.

Bursztyn, L., Callen, M., Ferman, B., Gulzar, S., Hasanain, A., and Yuchtman, N. (2020). Political identity: Experimental evidence on anti-americanism in pakistan. Journal of the European Economic Association, 18:2532–2560.

Casoria, F., Reuben, E., and Rott, C. (2022). The effect of group identity on hiring decisions with incomplete information. Management Science, in press:1–10.

Chandra, K. (2012). Constructivist theories of ethnic politics. Oxford University Press.

Charness, G. and Chen, Y. (2022). Social identity, group behavior and teams. Annual Review of Economics, 12:691–713.

Charness, G., Rigotti, L., and Rustichini, A. (2007). Individual behavior and group membership. American Economic Review, 97:1340–1352.

Chen, D., Schonger, M., and Wickens, C. (2016). otree-an open-source platform for laboratory, online, and field experiments. Journal of Behavioral and Experimental Finance, 9:88–97.

Chen, R. and Chen, Y. (2011). The potential of social identity for equilibrium selection. American Economic Review, 101:2562–2589.

Chen, Y. and Li, S. X. (2009). Group identity and social preferences. American Economic Review, 99:431–457.

Chen, Y., Li, S. X., Liu, T. X., and Shih, M. (2014). Which hat to wear? impact of natural identities on coordination and cooperation. Games and Economic Behavior, 84:58–86.

Choi, D. D., Poertner, M., and Sambanis, N. (2019). Parochialism, social norms, and discrimination against immigrants. Proceedings of the National Academy of Science, 116:16274–16279.

Cikara, M. and Van Bavel, J. J. (2014). The neuroscience of intergroup relations: An integrative review. Perspectives on Psychological Science, 9:245–274.

Dasgupta, P. and Goyal, S. (2019). Narrow identities. Journal of Institutional and Theoretical Economics, 175:395–419.

Desmet, K., Ortuño Ortin, I., and Wacziarg, R. (2017). Culture, ethnicity, and diversity. American Economic Review, 107:2479–2513.

Eckel, C. C. and Grossman, P. J. (2005). Managing diversity by creating team identity. Journal of Economic Behavior & Organization, 58:371–392.

Efferson, C., Lalive, R., and Fehr, E. (2008). The coevolution of cultural groups and ingroup favoritism. Science, 321:1844–1849.

Ellwardt, L., Hernández, P., Martínez-Canovas, G., and Muñoz-Herrera, M. (2016). Conflict and segregation in networks: An experiment on the interplay between individual preferences and social influence. Dynamic and Games, 3:191–216.

Fehr, E. and Schurtenberger, I. (2018). Normative foundations of human cooperation. Nature Human Behaviour, 2:458–468.

Fryer, R. G. and Jackson, M. O. (2008). A categorical model of cognition and biased decision making. The B.E. Journal of Theoretical Economics, 8:Article 6.

Fryer, R. G. and Torelli, P. (2010). An empirical analysis of "acting white". Journal of Public Economics, 94:380–396.

Gaertner, S. and Dovidio, J. (2000). Reducing intergroup bias: The common ingroup identity model. Psychology Press.

Goyal, S., Hernandez, P., Moisan, F., Martinez, G., Muñoz-Herrera, M., and Sanchez, A. (2021). Integration and diversity. Experimental Economics, 24:387–413.

Henrich, J. and Boyd, R. (2001). Why people punish defectors: Weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. Journal of Theoretical Biology, 208:79–89.

Hernandez, P., Muñoz-Herrera, M., and Sanchez, A. (2013). Heterogeneous network games: Conflicting preferences. Games and Economic Behavior, 79:56–66.

Jackson, M. O. and Wolinsky, A. (1996). A strategic model of social and economic networks. Journal of Economic theory, 71:44–74.

Lane, T. (2016). Discrimination in the laboratory: A meta-analysis of economics experiments. European Economic Review, 90:375–402.

Lees, J. and Cikara, M. (2020). Inaccurate group meta-perceptions drive negative outgroup attributions in competitive contexts. Nature Human Behavior, 4:279–286.

Li, S. X. (2020). Group identity, ingroup favoritism and discrimination. In Zimmermann, K. F., editor, Handbook of Labor, Human Resources and Population Economics, pages 1–28, Springer Nature Switzerland AG.

Lowe, M. (2021). Types of contact: A field experiment on collaborative and adversarial caste integration. American Economic Review, x:1–32.

Mousa, S. (2020). Building social cohesion between christians and muslims through soccer in post-isis iraq. Science, 369:866–870.

Saleh, M. and Tirole, J. (2021). Taxing identity: Theory and evidence from early islam. Econometrica, 13705:1–80.

Schelling, T. (1960). The Strategy of Conflict. Harvard University Press.

Schram, A. and Charness, G. (2015). Inducing social norms in laboratory allocation choices. Management Science, 61:1531–1546.

Shayo, M. (2009). A model of social identity with an application to political economy: nation, class and redistribution. American Political Science Review, 103:147–174.

Tajfel, H. and Turner, J. C. (1979). An integrative theory of intergroup conflict. In Austin, W. G. and Worchel, S., editors, The Social Psychology of Intergroup Relations, pages 33–47, Monterey, CA: Brooks/Cole.

Online Appendix:

Identity change and economic mobility: experimental evidence

Manuel Muñoz-Herrera[24]

[24] Social Science Division, New York University Abu Dhabi (e-mail: manumunoz@nyu.edu).

A Instructions

Below I include a sample of the instructions participants read for each part in treatment BASELINE. Each set of instructions has comprehension questions that participants were required to answer correctly before continuing. I also specify differences with instructions for the other treatments when necessary.

Instructions were originally in Spanish.

Instructions Part 1

Welcome to this study

You will receive a minimum pay of **10,000 pesos** for your participation in this study. Please read these instructions carefully to find out how you can earn **additional money**.

All interactions take place through the computers. Please do not talk or communicate with the other participants in any other way.

Please raise your hand if you have any question, and an experimentalist will come and answer your question privately.

This study is **anonymous**. Therefore, your identity will not be revealed to the other participants nor theirs to you.

This study has **4 parts**. You will now only read the instructions for Part 1. Once Part 1 is over, you will read the instructions for Part 2, and so on. In all four parts you will interact with the same participants.

You will participate with other 6 participants in the different parts of the study. 4 participants will be assigned to group ● and 3 participants will be assigned to group ▲. This is explained in the next screens.

In this experiment you can earn points depending on your choices and the choices made by the other participants. The amount of points you can earn is explained in the instructions for each part. At the end of the study, we will convert the total number of points you have earned into pesos using the following exchange rate: **1 point = 800 pesos**. You will receive your earnings in cash.

[*page break*]

Part 1: Your Group

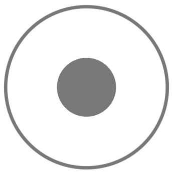
Here we explain the visual information of your group (the way it will be displayed on the screen). This information will be useful for the subsequent parts of the experiment. In this part you will also make a decision together with the other participants assigned to your group: you will choose a name for your group. This is explained as follows:

1. Group image

You have been assigned to **group** ● (internal symbol) and your **appearance** is ○ (external symbol).

There are 4 players in your group ●, you included. There are 3 players in group ▲.

Each player in your group will be displayed on the screen using the following image:



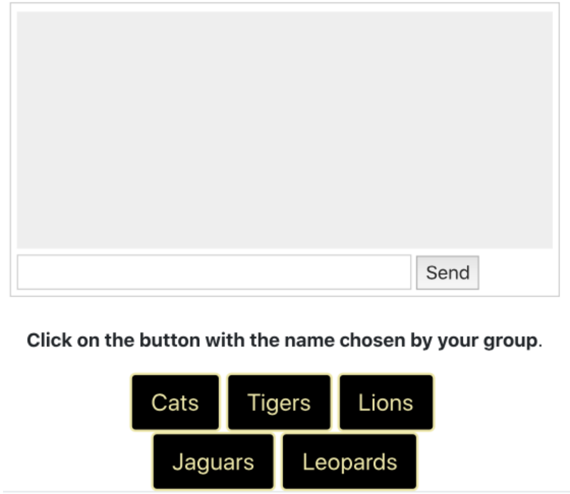
In the chat box on the right hand side you can communicate with the other participants in group ● and choose a name for your group, *by clicking on the button with the name you prefer as a group*. **To choose a name, all 4 players in your group must click on the same button**.

2. Group name

In the chat box below, you will be labeled as **Player 1**. It is not allowed to use any offensive language. And to maintain anonymity, it is not allowed to send any information that can be used by the other participants to identify who you are.

Write on the chat box to communicate with the other players in your group.

[*page break*]



Instructions Part 2

This is Part 2

In the following screens you will see the instructions of Part 2. At the end of the instructions you will see a summary of the most important information. This **summary** will be available in each screen during Part 2.

Part 2 has **10 rounds**. As in Part 1, once Part 2 has ended, you will read the instructions for Part 3. In Part 2 you can interact with the same 6 other players from Part 1, those in group ● (Cats) and those in group ▲ (Dogs).

Each of the 10 rounds in Part 2 have three stages: **Connections** (stage 1), **Actions** (stage 2), and **Results** (stage 3). The choices you make in each stage are explained in the following screens.

You will receive payment for **one of the 10 rounds** of Part 2. At the end of the study, the computer will randomly choose which round will be used for payment.

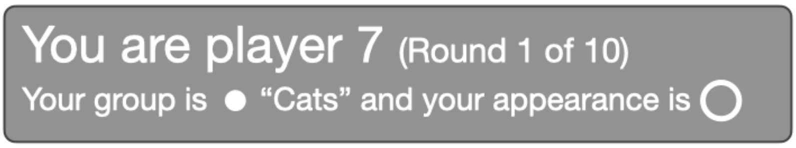
[*page break*]

Decisions in Part 2

Groups, Appearances & Labels

In Part 2, you will participate with 6 other players (the same as in Part 1). Each player is randomly assigned to a group ● “Cats” or ▲ “Dogs” (internal symbol), and will also have an appearance ○ or △ (external symbol). Each player will also have a numeric label from 1 to 7. You will see the group, appearance and label of each player on the screen (see Figure 1).

On the header of each screen you will see a summary of this information as in the example below:



In each round there will be 4 players in group ● “Cats” and 3 in group ▲ “Dogs”. A player’s group and appearance **will not change** between rounds during Part 2, while his numeric label and position in the screen will be **randomly changed in each round**.

For instance, the player in the header above is in group ● “Cats” and has the numeric label 7 in round 1. In round 2, the same player will be in the same group as before but his label may be any number from 1 to 7, and his position may be anywhere on the screen. The same is the case for all other players.

Stage 1: Connections

In the **Connections stage** you can propose a connection to those players you want to participate with in Stage 2. You can propose a connection to any player no matter their group or appearance. If you propose a connection to a player who **also** proposes to you, there is an **active connection** between the two of you. In Stage 2 you will **only participate** with the players with whom you have an active connection.

You will see your connections on the screen (see Figure 2). There is a *light arrow* leaving from you for each connection you proposed, and a ***dark arrow*** heading towards you for each connection proposed to you. **If both** arrows are present, you have an **active connection** with that other player. If only one of you proposes a connection but the other does not, there is no active connection between the two of you. **Your connections are illustrated with thicker lines and those of others are thinner.*

Stage 2: Actions

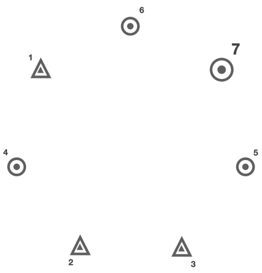


Fig. 1: Groups

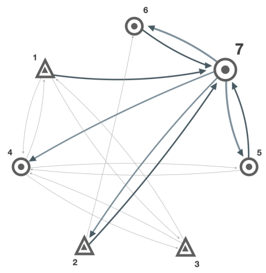


Fig. 2: Connections

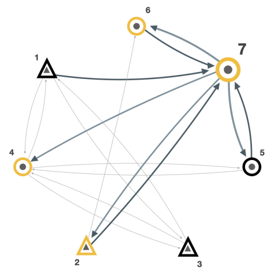


Fig. 3: Actions

In the **Actions stage** you will pick one of two actions: mustard or black. You only pick one action for the stage, not one for each connection. You will get points for each of your active connections choosing **the same action as you**. You will **not** get any points for an active connection choosing **a different action** than you, nor from a player choosing the same action as you, if he **is not** an active connection.

The action chosen by each player is illustrated by the border color of his external symbol (see Figure 3).

[*page break*]

Comprehension Questions

Before you continue, please answer the following comprehension questions:

1. You have been assigned to a group: ● or ▲ (inner symbol). How often does your group change?
    - It is fixed and does not change
    - The computer changes it in each round
    - I can change it in each round
2. You have been given an appearance: ◯ or △ (external symbol). How often does your appearance change?
    - It is fixed and does not change
    - The computer changes it in each round
    - I can change it in each round by changing my group

3. You will be assigned a label from 1 to 7. How often does your label change?

    - It is fixed and does not change
    - The computer changes it in each round
    - I can change it in each round

4. In Stage 2 (Actions) you only participate with your active connections. When is a connection active?

    - When I propose a connection to another player regardless of he/she proposing a connection to me
    - When another player proposes a connection to me regardless of me proposing a connection to him/her
    - When I propose a connection to a player who also proposes a connection to me

5. In Figure 2 above, how many active connections does player 7 have?

    - 5
    - 4
    - 3

[*page break*]

Points in Part 2

Stage 3: Results

In stage 3 for each round, you will see how many points you get in that round. The points are calculated by the gains of coordinating with your active connections minus the costs of your proposed connections. A coordination is an instance where you and another player with whom you have an active connection both chose the same action: mustard or black. The computer will do the calculations for you in each round, but below you can see how the points are calculated in 3 simple steps:

Step 1: Gains from coordination

Gains in each round depend on (i) your group, (ii) the action you choose, and (iii) the number of your active connections choosing the same (including yourself). You do not earn anything with a player if you do not have an active connection together and choose the same action.

If your group is ● "Cats" and you:

- choose mustard, you get **6 points** for each active connection also choosing mustard (including yourself) and **0 points** for any active connection choosing black.
- choose black, you get **4 points** for each active connection also choosing black (including yourself) and **0 points** for any active connection choosing mustard.

If your group is ▲ "Dogs" and you:

- choose black, you get **6 points** for each active connection also choosing black (including yourself) and **0 points** for any active connection choosing mustard.
- choose mustard, you get **4 points** for each active connection also choosing mustard (including yourself) and **0 points** for any active connection choosing black.

Step 2: Costs from proposals

Your costs in each round depend **only on the number of connections you propose**. You pay a cost of **2 points** for each proposal you make, independently of whether the other player also proposes a connection to you or not.

Step 3: Total points

Your points are calculated as the **gains from coordinating with your active connections** minus the **cost of each connection you propose**. The examples illustrate this:

*Example 1: Player 6 in group ● **"Cats"** chose mustard.*

1. Earnings: Player 6 formed **two** active connections (with players 2 and 3) and have **two** coordinations, (with player 2 and with himself). The two coordinations are multiplied by **6** (because his group is ● "Cats" and he chose mustard).

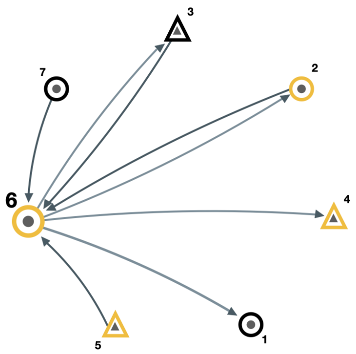


Fig. Example 1

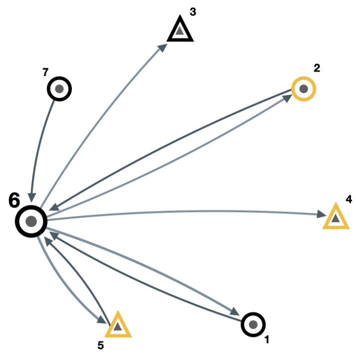


Fig. Example 2

Player 6 did not coordinate with player 3, therefore he does not earn points with that active connection. Player 6 coordinated with players 4 and 5, but did not form an active connection with either of them, so he cannot earn points with them. Player 6's earnings are: **12 points.**

2. Costs: Player 6 proposed **four** connections (to players 1, 2, 3 and 4), the cost of each proposal is 2 points, so he pays **8 points**.

3. Total Points: Player 6 gets 12 in gains minus 8 in costs: 12 - 8 = **4 points** in this round.

*Example 2: Player 6 in group ● **"Cats"** chose black.*

1. Earnings: Player 6 formed **three** active connections (with players 1, 2 and 5) and have **two** coordinations, (with player 1 and with himself). The two coordinations are multiplied by **4** (because his group is ● "Cats" and he chose black).

Player 6 did not coordinate with players 2 and 5, therefore he does not earn points with those active connection. Player 6 coordinated with players 3 and 7, but did not form an active connection with either of them, so he cannot earn points with them. Player 6's earnings are: **8 points.**

2. Costs: Player 6 proposed **five** connections (to players 1, 2, 3, 4 and 5), the cost of each proposal is 2 points, so he pays **10 points**.

3. Total Points: Player 6 gets 8 in gains minus 10 in costs: 8 - 10 = **-2 points** in this round.

[*page break*]

Comprehension Questions

Before you continue, please answer the following comprehension questions:

1. Imagine your group is ● "Cats", you chose action mustard and you have one active connection with a player who chose action mustard. What are the total points you get *from this connection* (not including what you get from coordinating with yourself)?
    - I gain 6 and pay the cost of 2 = 4 points in total
    - I gain 4 and pay the cost of 2 = 2 points in total
    - I gain 0 and pay the cost of 2 = -2 points in total
2. Imagine your group is ● "Cats", you chose action mustard and you have one active connection with a player who chose action black. What are the total points you get *from this connection* (not including what you get from coordinating with yourself)?
    - I gain 6 and pay the cost of 2 = 4 points in total
    - I gain 4 and pay the cost of 2 = 2 points in total
    - I gain 0 and pay the cost of 2 = -2 points in total

[*page break*]

Summary

The summary of the instructions (below) will be available in each round. To display it just click on the button **Show Summary** at the bottom of the screen.

Instructions Part 2

Groups, Appearances & Labels

In Part 2 you will interact with 6 other players. At the beginning of the experiment each of you was assigned to one of two groups: ● "Cats" or ▲ "Dogs" (internal symbol), and was assigned an appearance: ○ or △ (external symbol) and a numeric label (between 1 and 7). Your group and appearance are fixed for all 10 rounds in Part 2 while your numeric label and your position on the screen will change in each round. **There are 3 stages in each round.**

Stage 1: Connections

In stage 1, you can propose connections to the other players. A connection with another player is **active** if both you and the other player propose a connection to each other. Each connection you propose will cost you 2 points, independently of whether the other player also proposes a connection to you or not.

Stage 2: Actions

In stage 2, you will choose an action: mustard or black. You will earn points depending on your group, the action you choose, and the number of your active connections choosing **the same action** as you:

If your group is ● "Cats" and you:

- choose mustard, you get **6 points** for each active connection also choosing mustard (including yourself) and **0 points** for any active connection choosing black.
- choose black, you get **4 points** for each active connection also choosing black (including yourself) and **0 points** for any active connection choosing mustard.

If your group is ▲ "Dogs" and you:

- choose black, you get **6 points** for each active connection also choosing black (including yourself) and **0 points** for any active connection choosing mustard.
- choose mustard, you get **4 points** for each active connection also choosing mustard (including yourself) and **0 points** for any active connection choosing black.

Stage 3: Results

In stage 3, you will see how many points you get in that round. The points are calculated by the gains of coordinating with your active connections minus the costs of your proposed connections. The border color of each player's external symbol (◯ or △) shows the action that player chose.

Earnings in Part 2

You will participate for 10 rounds in Part 2. At the end of the experiment the computer **will randomly select one of the 10 rounds**. The number of points you got in the selected round will be added to the total number of points used to determine your payment.

[*page break*]

Instructions Part 3

Welcome to Part 3

You have concluded Part 2 and will now begin Part 3:

First, you will read the instructions for **Part 3**.

Then, you will participate in Part 3 for **10** rounds.

You will be paid for **one of the 10 rounds** in Part 3. As in Part 2, at the end of the study, the computer will randomly choose which round will be used for payment.

[*page break*]

Decisions in Part 3

In Part 3 you will participate for 10 rounds with the same 6 other players from Part 2. While, each round in Part 2 had 3 stages, each round in Part 3 will have 4 stages: **Group choice** (Stage 1), **Connections** (Stage 2), **Actions** (Stage 3) and **Results** (Stage 4).

The Connections, Actions and Results stages are the same as in Part 2. The **Group Choice** stage is **new** and is explained below.

Stage 1: Group Choice

At the beginning of each round, **before Stage 1**, all players will be assigned to the same group they belonged to in Part 2. If you were ● "Cats" in Part 2, you will begin each round in Part 3 as a ●. If you were ▲ "Dogs" in Part 2, you will begin each round in Part 3 as a ▲.

Then, in the **Group Choice** stage, each player will decide if he wants to stay in his group or if he wants to change it.

BASELINE: **The group symbol will indicate the group each player has chosen. A player's appearance will not change even if he changes his group**.

IGNORANCE: **The group symbol will not be visible regardless of whether a player changes or stays in his group. A player's appearance will not change even if he changes his group**.

SANCTIONS: **The group symbol will indicate the group each player has chosen. A player's appearance will not change even if he changes his group. Staying in the same group is free, but changing group has a fixed cost of 6 points, plus 2 points for each player in your group that decides not to change**

REVEAL: **The group symbol will indicate the group each player has chosen. But it will only be visible if the player has chosen to reveal it (see Stage 2). A player's appearance will not change even if he changes his group**.

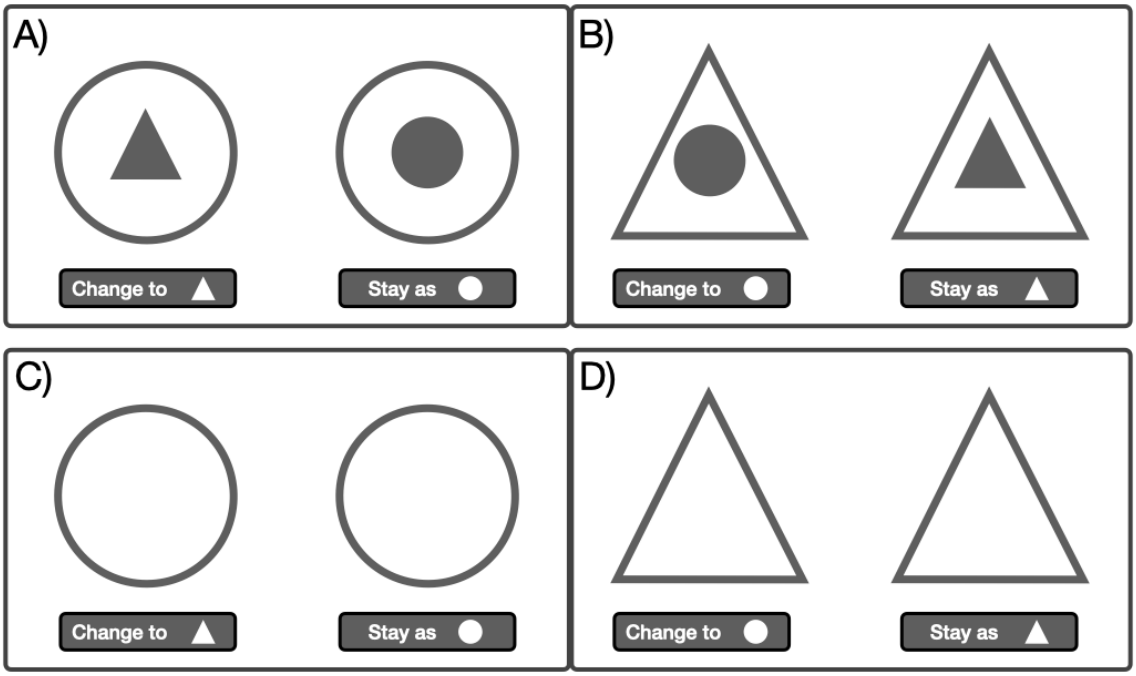
The example below shows the choices a player can make depending on his group at the beginning of each round. The image above each button indicates how other players will see him in the remaining three stages of that round (Connections, Actions and Results), depending on whether he changes or stays in his group.

Points in Part 3:

If a player in group ● "Cats" stays in the same group, he will continue earning points as a ● "Cats". That is, 6 points for each coordination on mustard and 4 points for each coordination on black.

However, if a player in group ● "Cats" changes to group ▲ "Dogs", he will earn points as a ▲ "Dogs". That is, 6 points for each coordination on black and 4 points for each coordination on mustard.

Similarly, if a player in group ▲ "Dogs" changes to group ● "Cats", he will earn points as a ● "Cats". That is, 6 points for each coordination on mustard and 4 points for each coordination on black.



Group choice for a player belonging to group ● "Cats" (Panel A) or to group ▲ "Dogs" (Panel B). And for players in treatment ignorance or reveal (Panel C and D).

But if he chooses to stay in group ▲ "Cats", he will continue earning points as a ▲ "Cats". That is, 6 points for each coordination on black and 4 points for each coordination on mustard.

Stage 2: Group revelation (*only in treatment* REVEAL)

After the **Group Choice** stage, each player will decide whether he wants to reveal the group he has chosen, so others see it or not. Not revealing your group is free, while revealing your group has a cost.

The cost has two elements. First, a fixed part of 6 points for revealing, independently of whether you chose to change or stay in your group. In addition, if you chose to change your group, there is a variable part that increases by 2 points for each person in your group that chooses not to change.

Stage 2: Connections

This Stage is the same as in Part 2. Each player can propose to others and pays a cost of 2 points for each connection proposed.

Stage 3: Actions

This Stage is the same as in Part 2. Each player chooses one of two actions: mustard or black and earns points by the number of his/her active connections choosing the same action.

Stage 4: Results

This Stage is the same as in Part 2. Each player is informed of the number of total points he gets in the round. Points are calculated by the gains from coordination (which depend on the group chosen in the current round, the action chosen, and the number of coordinations with active connections) minus the costs of proposing connections to others.

Comprehension Questions

Before you continue, please answer the following comprehension questions:

1. You have been assigned to a group: ● or ▲ (inner symbol). How often does your group change?
   - It is fixed and does not change
   - The computer changes it in each round
   - I can change it in each round
2. You have been given an appearance: ○ or △ (external symbol). How often does your appearance change?
   - It is fixed and does not change
   - The computer changes it in each round
   - I can change it in each round by changing my group
3. You will be assigned a label from 1 to 7. How often does your label change?
   - It is fixed and does not change
   - The computer changes it in each round

- I can change it in each round

4. Imagine you began in group ● "Cats", and you changed to group ▲ "Dogs". Then you chose action mustard and you have one active connection with a player who chose action mustard. What are the total points you get *from this connection* (not including what you get from coordinating with yourself)?
   - I gain 6 and pay the cost of 2 = 4 points in total
   - I gain 4 and pay the cost of 2 = 2 points in total
   - I gain 0 and pay the cost of 2 = -2 points in total

5. Imagine you began in group ● "Cats", and you stayed in your group. Then you chose action mustard and you have one active connection with a player who chose action mustard. What are the total points you get *from this connection* (not including what you get from coordinating with yourself)?
   - I gain 6 and pay the cost of 2 = 4 points in total
   - I gain 4 and pay the cost of 2 = 2 points in total
   - I gain 0 and pay the cost of 2 = -2 points in total

6. After the Group choice (Stage 1) what will other players see about you in the remaining three stages?
   - They can see the group I choose and my new appearance
   - They can see the group I choose and my appearance from Part 2
   - They cannot see the group I choose only my appearance from Part 2

[*page break*]

Summary

The summary of the instructions (below) will be available in each round. To display it just click on the button **Show Summary** at the bottom of the screen.

Instructions Part 3

In Part 3 you will interact with the same 6 players from Part 2. There are 4 stages in each round.

Stage 1: Group Choice

At the beginning of each round, **before Stage 1**, you and all other players will be assigned to the same group they belonged to in Part 2. In Stage 1 you will choose a group. You can stay in the same group you have been assigned to or you can change groups. **The group symbol will indicate the group each player has chosen. A player's appearance will not change when he changes his group.**

Stage 2: Connections

In this stage, you can propose connections to the other players at a cost of 2 points.

Stage 3: Actions

In this stage, you choose an action: mustard or black. You will earn points depending on the group you choose, the action you choose, and the number of your active connections choosing **the same action** as you:

If your choosen group is ● "Cats" and you:

- choose mustard, you get **6 points** for each active connection also choosing mustard (including yourself) and **0 points** for any active connection choosing black.
- choose black, you get **4 points** for each active connection also choosing black (including yourself) and **0 points** for any active connection choosing mustard.

If your chosen group is ▲ "Dogs" and you:

- choose black, you get **6 points** for each active connection also choosing black (including yourself) and **0 points** for any active connection choosing mustard.
- choose mustard, you get **4 points** for each active connection also choosing mustard (including yourself) and **0 points** for any active connection choosing black.

Stage 4: Results

In this stage, you will see how many points you get in that round. Points are calculated by the gains of coordinating with your active connections minus the costs of the connections you have proposed.

Earnings in Part 3

You will participate for 10 rounds in Part 3. At the end of the experiment the computer **will randomly select one of the 10 rounds**. The number of points you got in the selected round will be added to the total number of points used to determine your payment.

[*page break*]

Instructions Part 4

Allocation

This is the last part of the experiment and it is played for a single round. In this round, you will be endowed with 10 points and you have to decide how you want to distribute the points between two players who have been randomly matched with you: Players **A** and **B**.

Player **A** is one of the players assigned to group ● "Cats" in Parts 1 and 2. Player **B** is one of the players assigned to group ▲ "Dogs" in Parts 1 and 2.

The computer will randomly choose the allocation decision of one of the 7 players who have participated together in the experiment, and will implement the payment of that decision. For instance, if your decision is chosen, players **A** and **B** matched with you will receive the points you have allocated to them. None of the other 5 players (including yourself) will receive any points in this part. Similarly, if another player's decision is chosen and you are one of the players **A** or **B** matched with him, you will receive the points allocated to you.

Decision: You have 10 points to allocate between players **A** and **B**. Press the button with the allocation you prefer.

- 0 to A and 10 to B
- 1 to A and 9 to B

- 2 to A and 8 to B
- 3 to A and 7 to B
- 4 to A and 6 to B
- 5 to A and 5 to B
- 6 to A and 4 to B
- 7 to A and 3 to B
- 8 to A and 2 to B
- 9 to A and 1 to B
- 10 to A and 0 to B

[*page break*]

Your participation in this experiment ends here.

In the next screen you will be informed of your earnings. After that, please remain seated until you are called to receive your payment.

To conclude, please press the button indicating your gender:

- Female
- Male

[*End of the instructions.*]

B Regression tables

The data in the coordination game consists of the decisions made by 336 individuals who interact over two blocks of 10 periods in groups of 7 players in four treatments. The tables below report random effects GLS regressions with standard errors clustered on groups for (i) the setting with fixed identities, (ii) the setting with flexible identities, and (iii) differences between the two settings.

Table B-1 reports regressions testing differences in the level of assimilation of the minority and the majority across treatments. In all regressions the independent variables are dummies for each treatment, for which the BASELINE treatment is the omitted category. Moreover, I use group random effects in all regressions. The first dependent variable is the level of assimilation, representing the case in which a player changes her group identity, within the minority in column **I** and within the majority in column **II**. The subsequent dependent variable is the level of action alignment, representing the case in which a player chooses the action that aligns to her chosen identity, for the minority in column **III** and the majority in column **IV**. Finally, the last dependent variable is the level of efficiency, expressed as the fraction of payoffs normalized by the prediction for the *all-conform outcome*, for the minority in column **IV** and for the majority in column **V**.

Table B-1 The effect of barriers on assimilation and alignment

GLS regressions with standard errors clustered on groups (in parenthesis). The dependent variable is the level of assimilation of the minority in column I and of the majority in column II, and the level of alignment of the minority in column III and of the majority in column IV. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels.

	Assimilation		**Alignment**	
	Minor	*Major*	*Minor*	*Major*
	I	**II**	**III**	**IV**
IGNORANCE	-0.300^{***}	-0.002	-0.017	-0.004
	(0.112)	(0.006)	(0.012)	(0.007)
SANCTIONS	-0.358^{***}	-0.002	-0.064	0.000
	(0.091)	(0.006)	(0.047)	(0.004)
REVEAL	-0.250^{***}	0.010	-0.031	-0.006
	(0.090)	(0.009)	(0.024)	(0.007)
Constant	0.953^{***}	0.008^{*}	0.975^{***}	0.996^{***}
	(0.013)	(0.005)	(0.007)	(0.003)
χ^2	29.39^{***}	2.16	4.54	1.22
# Obs.	1440	1920	1440	1920
# Groups	48	48	48	48

Table B-2 reports regressions testing differences in collective outcomes (group-level variables) across treatments. In all regressions the independent variables are dummies for each treatment, for which the BASELINE treatment is the omitted category. Moreover, I use group random effects in all regressions. The first dependent variable in column **I** is the level of *fractured outcome*. This represents the case in which not all three minority players in a group assimilate. The subsequent dependent variable is the level of inter-group connectivity, measured by the fraction of links formed across groups. The next dependent variables is the level of inter-group connectivity conditional on having assimilated, for the *fractured outcome* in column **III**, and the *all-assimilate outcome* in column **IV**.

Table B-2 The effect of barriers on collective outcomes

GLS regressions with standard errors clustered on groups (in parenthesis). The dependent variable is the level of the *fractured outcome* in column I. The subsequent variable is the level of inter-group connectivity in column II. Then, inter-group connectivity of assimilators, in the *fractured outcome* in column III or *all-assimilate outcome* in column IV. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels.

	Fractured Outcomes *Minor*	**Inter-group contact** *Minor-Major*	**Connectivity (Fractured)** *Minor*	**Connectivity (All assim.)** *Minor*
	I	**II**	**III**	**IV**
IGNORANCE	1.333^{**}	-0.353^{***}	-0.156	-0.406^{***}
	(0.063)	(0.113)	(0.120)	(0.099)
SANCTIONS	0.358^{***}	-0.260^{**}	0.046	-0.062
	(0.104)	(0.105)	(0.089)	(0.106)
REVEAL	0.283^{***}	-0.292^{***}	-0.178^{*}	-0.085
	(0.109)	(0.093)	(0.097)	(0.109)
Constant	0.125^{***}	0.772^{***}	0.822^{**}	0.661^{***}
	(0.027)	(0.052)	(0.056)	(0.082)
χ^2	19.87^{***}	17.29^{***}	6.40^{*}	26.05^{***}
# Obs.	480	480	789	256
# Groups	48	48	34	99

Table B-3 reports regressions testing differences in inter-group proposals by players from the minority who assimilate. In all regressions the independent variable is a dummy for the *fractured outcome*, for which the outcome where all players assimilate is the omitted category. Moreover, I use group random effects in all regressions. The table reports regressions for the BASELINE in column **I**, IGNORANCE in column **II**, SANCTIONS column **III**, and REVEAL in column **IV**.

Table B-4 reports regressions testing differences in inter-group connections of players from the minority who assimilate. In all regressions the independent variable is a dummy for the *fractured outcome*, for which the outcome where all players assimilate is the omitted category. Moreover, I use group random effects in all regressions. The table reports regressions for the BASELINE in column **I**, IGNORANCE in column **II**, SANCTIONS column **III**, and REVEAL in column **IV**.

Table B-5 reports regressions testing differences in the randomization of groups by treatment, on the effect that fixed group identity (i.e., no assimilation) have on choices and

Table B-3 The effect of a fractured minority on inter-group proposals

GLS regressions with standard errors clustered on groups (in parenthesis). The dependent variable is the level of inter-group proposals from the minority to the majority, by treatment. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels.

	BASELINE *Minor-Major* I	IGNORANCE *Minor-Major* II	SANCTIONS *Minor-Major* III	REVEAL *Minor-Major* IV
Fractured minority	−0.041	0.018	−0.048	−0.022
	(0.056)	(0.067)	(0.029)	(0.050)
Constant	0.936***	0.813***	0.914***	0.849***
	(0.025)	(0.067)	(0.037)	(0.039)
χ^2	0.54	0.07	2.68	0.20
# Obs.	343	235	214	253
# Groups	12	12	11	12

Table B-4 The effect of a fractured minority on inter-group connections

GLS regressions with standard errors clustered on groups (in parenthesis). The dependent variable is the level of inter-group connections between the minority and the majority, by treatment. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels.

	BASELINE *Minor-Major* I	IGNORANCE *Minor-Major* II	SANCTIONS *Minor-Major* III	REVEAL *Minor-Major* IV
Fractured minority	−0.242***	−0.327***	−0.282***	−0.119**
	(0.069)	(0.126)	(0.059)	(0.059)
Constant	0.823***	0.663***	0.870***	0.648***
	(0.058)	(0.110)	(0.070)	(0.070)
χ^2	12.36***	6.70***	22.51***	4.10**
# Obs.	343	235	214	253
# Groups	12	12	11	12

payoffs. In all regressions the independent variables are dummies for each treatment, for which the BASELINE treatment is the omitted category. Moreover, I use group random effects in all regressions. The first dependent variable in column **I** is the level of inter-group connectivity, measured by the fraction of links formed across groups. The subsequent dependent variable is the level of alignment. For the fixed-identities game, this represents the case in which a player chooses the action that aligns with her assigned group identity, for the minority in column **II** and the majority in column **III**. Finally, the last dependent variable is the level of efficiency, expressed as the fraction of payoffs normalized by the prediction for the *all-conform outcome*, for the minority in column **IV** and for the majority in column **V**.

Table B-6 reports regressions testing differences in efficiency between the setting with flexible identities and the setting with fixed identities for the *minority*, to evaluate the effect of assimilation on payoffs. In all regressions the independent variable is a dummy

Table B-5 The effect of fixed identities on choices and payoffs

GLS regressions with standard errors clustered on groups (in parenthesis). The first dependent variable is the level of inter-group connectivity in column I. The second dependent variable is the level of alignment, for the minority in column II, and for the majority in column III. Finally, the level of efficiency for the minority in column IV and for the majority in column V. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels.

	Inter-group connectivity	**Alignment**		**Efficiency**	
	Min-Maj	*Minor*	*Major*	*Minor*	*Major*
	I	**II**	**III**	**IV**	**V**
IGNORANCE	0.004	0.022	−0.010	−0.006	−0.035
	(0.018)	(0.611)	(0.019)	(0.099)	(0.074)
SANCTIONS	0.012	−0.006	0.002	−0.002	0.063
	(0.033)	(0.087)	(0.007)	(0.103)	(0.039)
REVEAL	0.034	−0.009	−0.027	0.047	−0.068
	(0.033)	(0.089)	(0.021)	(0.100)	(0.055)
Constant	0.033***	0.894***	0.988***	0.752***	0.869***
	(0.011)	(0.055)	(0.006)	(0.078)	(0.030)
χ^2	1.11	0.46	2.31	0.46	7.54
# Obs.	480	1440	1920	1440	1920
# Groups	36	36	36	36	36

for the case with flexible identities, so that having fixed identities is the omitted category. I use group random effects in all regressions. The level of efficiency, i.e., payoffs as a fraction of the predicted gains in the *all-conform outcome*, is the dependent variable for treatment BASELINE in column **I**, for treatment IGNORANCE in column **II**, for SANCTIONS in column **III** and for REVEAL in column **IV**.

Table B-6 Effect of flexible identities on efficiency for the minority

GLS regressions with standard errors clustered on groups (in parenthesis). The dependent variable is the level of efficiency of the minority in BASELINE in column I, in IGNORANCE in column II, in SANCTIONS in column III, and in REVEAL in column IV. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels.

	BASELINE	IGNORANCE	SANCTIONS	REVEAL
	Minor	*Minor*	*Minor*	*Minor*
	I	**II**	**III**	**IV**
Flexible identities	0.928***	0.402***	0.162*	0.258**
	(0.064)	(0.109)	(0.090)	(0.130)
Constant	0.752***	0.746***	0.750***	0.799***
	(0.081)	(0.064)	(0.069)	(0.065)
χ^2	207.20***	13.61***	3.22*	3.96**
# Obs.	720	720	720	720
# Groups	12	12	12	12

Table B-7 reports regressions testing differences in efficiency between the setting with flexible identities and the setting with fixed identities for the *majority*, to evaluate the effect of assimilation on payoffs. In all regressions the independent variable is a dummy for the case with flexible identities, so that having fixed identities is the omitted category. I use group random effects in all regressions. The level of efficiency, i.e., payoffs as a fraction of the predicted gains in the *all-conform outcome*, is the dependent variable for treatment BASELINE in column **I**, for treatment IGNORANCE in column **II**, for SANCTIONS in column **III** and for REVEAL in column **IV**.

Table B-7 Effect of flexible identities on efficiency for the majority

GLS regressions with standard errors clustered on groups (in parenthesis). The dependent variable is the level of efficiency of the majority in BASELINE in column I, in IGNORANCE in column II, in SANCTIONS in column III, and in REVEAL in column IV. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels.

	BASELINE	IGNORANCE	SANCTIONS	REVEAL
	Major	*Major*	*Major*	*Major*
	I	**II**	**III**	**IV**
Flexible identities	0.554***	0.344***	0.345***	0.383***
	(0.051)	(0.072)	(0.066)	(0.079)
Constant	0.869***	0.835***	0.932***	0.801***
	(0.031)	(0.070)	(0.026)	(0.048)
χ^2	117.55***	22.58***	27.47***	23.24***
# Obs.	960	960	960	960
# Groups	12	12	12	12

Table B-8 reports regressions testing differences in efficiency for minority assimilators (and non-assimilators) by comparing the case where the minority acts collectively (all choose one same identity) or is fractured. In all regressions the dependent variable the level of normalized payoffs (efficiency) attained. The independent variables include a dummy on whether the minority player assimilated (conformists are the ommited category), also a dummy on whether the group was fractured (the omitted category is the case of collective unity where all choose the same identity), and treatment dummies where the reference category is the BASELINE treatment.

At the end of the experiment each participant makes an allocation choice once, by dividing 10-points between an in-group and an out-group receiver. Therefore, there are 108 observations for the minority and 144 for the majority, across treatments. I run OLS regressions with robust standard errors to test differences in the allocation bias towards the in-group, across treatments. Table B-9 reports these regressions for the minority in column **I** and the majority in column **II**. In all regressions the independent variables are dummies for each treatment, for which the BASELINE treatment is the omitted category.[25]

Table B-10 reports regressions testing differences in choices in the treatment REVEAL, contingent on whether a player revealed her chosen identity for the *minority*. In all re-

[25] Note there are 7 observations missing for the minority (101 remaining) and 11 for the majority (133 remaining). These were dropped because the allocation was biased towards the out-group, which is most likely due to a failure in understanding the instructions.

Table B-8 Effect of fractured minority on efficiency

GLS regressions with standard errors clustered on groups (in parenthesis). The dependent variable is the level of efficiency of the minority. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels.

	I	**II**	**III**
Assimilation	0.693^{***}	0.611^{***}	0.631^{***}
	(0.100)	(0.147)	(0.109)
Fractured minority	-0.499^{***}		-0.523^{***}
	(0.058)		(0.063)
Assimilation X Fractured	-0.486^{***}		-0.430^{***}
	(0.127)		(0.135)
IGNORANCE		-0.348^{***}	-0.247^{*}
		(0.148)	(0.132)
SANCTIONS		-0.549^{***}	-0.275^{**}
		(0.144)	(0.123)
Constant	0.959^{***}	$1.0.97^{***}$	1.176^{***}
	(0.061)	(0.156)	(0.127)
χ^2	239.08^{***}	315.82^{***}	27.47^{***}
# Obs.	1080	1080	1080
# Groups	36	36	36

gressions the independent variable is a dummy for the case where a player revealed her chosen identity, so that keeping the group identity concealed is the omitted category. I use group random effects in all regressions. The dependent variable is the level of assimilation in column **I**, the level of inter-group connectivity in column **II**, the level of efficiency in column **III** and the frequency of the *fractured outcome* in column **IV**.

Table B-11 reports regressions testing what drives participants in the minority to reveal their chosen identity in treatment REVEAL. In both regressions, the dependent variable is the revelation choice for the minority, without controls in column **I** and controlling for the number of assimilators in the in-group in column **II**. I use group random effects in all regressions.

Table B-9 The effect of assimilation on in-group allocation bias

OLS regressions with robust standard errors (in parenthesis). The dependent variable is the magnitude of the bias towards the in-group for the minority in column I and for the majority in column II. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels.

	In-group bias	
	Minor	*Major*
	I	**II**
IGNORANCE	0.148	0.032
	(0.090)	(0.091)
SANCTIONS	0.262^{**}	-0.055
	(0.103)	(0.087)
REVEAL	0.105	-0.099
	(0.097)	(0.084)
Constant	0.222^{***}	0.364^{***}
	(0.062)	(0.063)
Adjusted R^2	0.051	0.016
# Obs.	134	180

Table B-10 Effect of revelation of identities on choices and outcomes for the minority

GLS regressions with standard errors clustered on groups (in parenthesis). The dependent variable is the level of assimilation in column I, the level of inter-group connectivity in column II, the level of efficiency in column III and the frequency of the *fractured outcome* in column IV. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels.

	Assimilation	**Inter-group connectivity**	**Efficiency**	**Fractured minority**
	Minor	*Minor*	*Minor-Major*	*Minor*
	I	**II**	**III**	**IV**
Revealed identities	0.164^{***}	0.232^{***}	-0.250^{*}	0.223^{***}
	(0.055)	(0.078)	(0.149)	(0.039)
Constant	0.666^{***}	0.428^{***}	1.113^{***}	0.359^{***}
	(0.1000)	(0.084)	(0.131)	(0.108)
χ^2	8.84^{***}	8.80^{***}	2.82^{*}	33.05^{***}
# Obs.	360	360	360	360
# Groups	12	12	12	12

Table B-11 The effect of endogenous barriers on identity revelation

GLS regressions with standard errors clustered on groups (in parenthesis). In all regressions the dependent variable is the level of revelation for the minority, without controls in column I and controlling for the number of assimilators in the in-group in column II. ***, ** and * indicate statistical significance at the 0.01, 0.05 and 0.10 levels.

	Identity revelation	
	Minor	*Minor*
	I	**II**
Identity change	0.227^{***}	0.457^{***}
	(0.066)	(0.101)
1 adopter in subgroup		0.043
		(0.055)
2 adopters in subgroup		−0.013
		(0.037)
3 adopters in subgroup		-0.396^{***}
		(0.097)
Constant	0.063	0.085^{**}
	(0.049)	(0.038)
χ^2	11.86^{***}	57.81^{***}
# Obs.	360	360
# Groups	12	12

C Additional analysis

C.1 Description of the sample

Table C-1 reports the count and percentage of females subjects and males subjects in the experiment. Totals are reported separately by subgroup: majority or minority. I run non-parametric Wilcoxon-Mann-Whitney tests to evaluate the underlying distributions of females between treatments as well as between subgroups across treatments. The pairwise comparison shows no significant difference in the distribution of females between any pair of treatments, except for the BASELINE with respect to IGNORANCE ($p = 0.013$) and with respect to REVEAL ($p = 0.089$). By taking a deeper look into the subgroups, majority and minority, I find that the effect is driven by lower levels of females in the majority subgroup in the BASELINE, while there are no differences in the distribution of females in the minority subgroup across treatments.

Table C-1 Descriptive statistics of the sample

Frequency of female and male subjects by treatment. Counts are reported separately by subgroup: majority or minority. Percentages in parentheses.

	Treatments								Totals
	BASELINE		IGNORANCE		SANCTIONS		REVEAL		
	Major	*Minor*	*Major*	*Minor*	*Major*	Minor	*Major*	*Minor*	
Female	21	19	34	22	30	19	31	20	196
	(0.44)	(0.53)	(0.71)	(0.61)	(0.63)	(0.53)	(0.65)	(0.53)	(0.58)
Male	27	17	14	14	18	17	17	16	140
	(0.56)	(0.47)	(0.29)	(0.39)	(0.37)	(0.43)	(0.35)	(0.47)	(0.42)

Table C-2 reports the average payoffs per period in the game with fixed group identities, the game with flexible group identities, and the other-other allocation choice. Means are reported separately by gender (female and male) and by subgroup (majority and minority). As reported in the main text, there are no significant differences in earnings in the game with fixed group identities across treatments. On the contrary, in the game with flexible group identities, the minority gets significantly more in the BASELINE than in both IGNORANCE and SANCTIONS and there are no differences between the last two.

Table C-3 reports summary statistics of the main variables in the experiment for all four treatments (BASELINE, IGNORANCE, SANCTIONS, and REVEAL), separate for the majority and the minority. In complement to Table 2, the table reports *assimilation*, *within-group* proposals and links, *inter-group* proposals and links, *alignment*, *efficiency*, and *in-group bias*.

Table C-2 Payoffs and allocations by gender

Average payoffs per period in PART 2, average payoffs per period in PART 3, and average allocation to in-group in PART 4. Standard deviations in parenthesis.

	Treatments							
	BASELINE		IGNORANCE		SANCTIONS		REVEAL	
	Major	*Minor*	*Major*	*Minor*	*Major*	Minor	*Major*	*Minor*
Fixed identities								
Female	15.21	10.97	15.02	10.05	16.72	12.11	14.31	11.54
	(2.32)	(3.52)	(4.42)	(3.79)	(1.97)	(1.74)	(3.46)	(3.28)
Male	15.98	10.01	15.02	11.04	16.85	8.68	14.61	10.73
	(2.08)	(4.51)	(4.33)	(2.73)	(1.70)	(3.96)	(3.18)	(3.92)
Flexible identities								
Female	24.29	22.43	20.22	14.14	22.30	14.10	20.98	15.45
	(5.10)	(5.17)	(5.78)	(6.12)	(5.02)	(6.10)	(4.74)	(6.34)
Male	26.65	24.71	23.60	19.10	24.13	11.25	21.90	13.98
	(4.36)	(3.91)	(5.13)	(6.33)	(4.00)	(5.55)	(5.28)	(5.66)
Allocation								
Female	6.19	5.89	6.20	6.63	6.83	6.63	6.35	5.90
	(1.86)	(1.63)	(2.60)	(2.44)	(2.16)	(2.41)	(1.83)	(2.12)
Male	6.85	6.35	7.07	6.50	5.61	6.70	6.11	6.93
	(2.39)	(2.08)	(2.64)	(1.87)	(1.91)	(3.40)	(2.02)	(2.35)

Table C-3 Summary statistics in the social coordination game

Average fractions (percentages) for each of the main variables. Standard deviations in parenthesis. Observations are disaggregated by majority and minority; pooled for the game with *fixed* group identities and separated by treatment for the game with *flexible* group identities.

	Fixed identities		Flexible identities							
			BASELINE		IGNORANCE		SANCTIONS		REVEAL	
	Major	*Minor*	*Major*	*Minor*	*Major*	Minor	*Major*	*Minor*	*Major*	*Minor*
Assimilation	n/a	n/a	0.01	0.95	0.01	0.65	0.01	0.59	0.02	0.70
			(0.09)	(0.21)	(0.08)	(0.48)	(0.08)	(0.49)	(0.13)	(0.46)
WITHIN-GROUP										
proposals	0.94	0.92	0.95	0.91	0.96	0.88	0.98	0.78	0.96	0.81
	(0.17)	(0.24)	(0.16)	(0.23)	(0.15)	(0.29)	(0.09)	(0.32)	(0.15)	(0.32)
links	0.88	0.85	0.91	0.85	0.93	0.80	0.97	0.68	0.92	0.71
	(0.23)	(0.29)	(0.20)	(0.30)	(0.20)	(0.34)	(0.11)	(0.38)	(0.19)	(0.37)
INTER-GROUP										
proposals	0.15	0.16	0.82	0.91	0.50	0.59	0.57	0.63	0.55	0.66
	(0.29)	(0.29)	(0.20)	(0.23)	(0.06)	(0.16)	(0.33)	(0.57)	(0.41)	(0.43)
links	0.04		0.77		0.42		0.51		0.48	
	(0.14)		(0.33)		(0.44)		(0.43)		(0.40)	
Alignment	0.98	0.90	0.99	0.98	0.99	0.96	0.99	0.91	0.99	0.94
	(0.15)	(0.31)	(0.06)	(0.16)	(0.09)	(0.20)	(0.06)	(0.28)	(0.10)	(0.23)
Efficiency	0.86	0.76	1.42	1.68	1.18	1.15	1.28	0.91	1.18	1.06
	(0.26)	(0.35)	(0.33)	(0.57)	(0.39)	(0.71)	(0.35)	(0.67)	(0.37)	(0.66)
In-group bias	n/a	n/a	0.36	0.22	0.39	0.37	0.31	0.48	0.26	0.32
			(0.42)	(0.37)	(0.43)	(0.38)	(0.40)	(0.45)	(0.38)	(0.43)

C.2 Identity choices by treatment

In this section, I further explore the choices of assimilation of the minority across treatments. The first consideration is to see in which period minority players assimilation for the first time. Across the four treatments the median period in which assimilation starts is period 1. Figure 10A illustrates the fraction of participants that assimilate in each of the 10 periods in the game with flexible identities, by treatment. The bars on period 0 represent the fraction of participants that never assimilate, which are positive for all but BASELINE. By period 3, all participants in BASELINE had already assimilated at least once, while the latest period in which someone assimilated for the first time is 5 in REVEAL, 8 in IGNORANCE, and 10 in SANCTIONS.

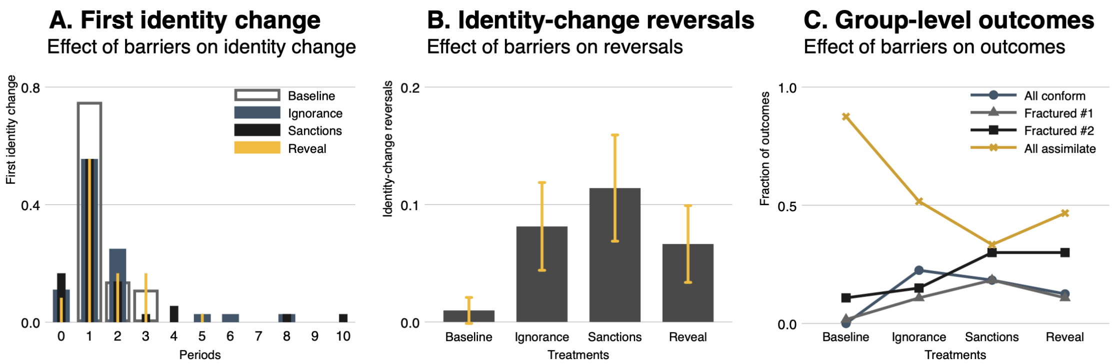


Figure 10 Identity choices

Panel A illustrates the fraction of minority members that assimilate for the first time per period. Panel B shows the fraction of subjects who reverse their identity-change choice by adopting back their original group identity in a subsequent period. In both panels choices are disaggregated by treatment. Panel C reports the frequency of collective outcomes by treatment.

The second consideration with respect to the assimilation choice is whether minority players reverse it. That is, after having assimilated, I look at the fraction of assimilators who choose back the minority's original group identity. This is illustrated in Figure 10B. I run non-parametric Wilcoxon-Mann-Whitney tests to evaluate the underlying distributions of choice reversals between treatments. The pairwise comparison shows that there is only a statistically significant difference between the BASELINE, where only 1% of the assimilators reverse, and the other three treatments; IGNORANCE where 8% reverse ($p < 0.001$), SANCTIONS where 11% ($p < 0.001$) and REVEAL where 7% reverse ($p < 0.001$).

Finally, I show in more detail the frequency of outcomes as a consequence of the assimilation choices of the minority (see Figure 10C). The *all-assimilate outcome* is most frequently observed in the BASELINE and least in SANCTIONS. On the other extreme, the *all-conform outcome* does not occur in the BASELINE, is most frequent in IGNORANCE and least in REVEAL. Finally, the *fractured outcome* is most frequent in SANCTIONS when two in the minority assimilate (which can lead to economic mobility), and also when only one minority player assimilates (which leaves all three players worse-off than in the *all-conform outcome*).

Thus, the analysis on identity choices suggest that the reason why the average assimilation level in SANCTIONS is lower than the BASELINE, and the frequency of the *fractured outcome* is highest in SANCTIONS than in any other treatment, is possibly due to: (i) minority players waiting to assimilate in SANCTIONS even to the last period, and (ii) minority players being somewhat more likely to reverse the assimilation choice in SANCTIONS than in the other treatments.

C.3 Network choices by treatment

After the assimilation choices have been made, players make network choices. Specifically, each participant makes linking proposals to both in-group and out-group members. This is illustrated in Figure 11, which depicts connectivity towards the in-group (Panel A) and towards the out-group (Panel B), for the minority (filled bars) and the majority (empty bars). The top of each range bar reports the proportion of proposals made, the bottom reports the proportion of proposals received (i.e., the links formed), and the height of each bar is the proportion of unreciprocated proposals (i.e., failed links).

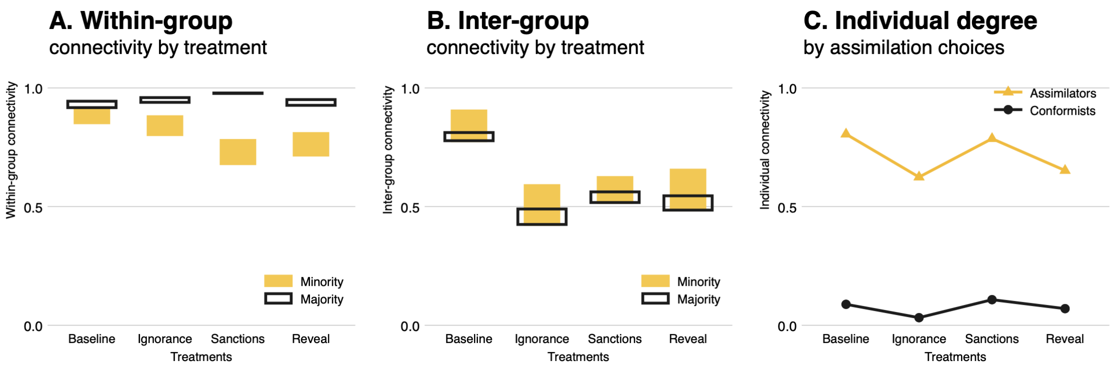


Figure 11 Network connectivity: Proposals and Links

Panel A illustrates the average level of connectivity within groups (minority with minority and majority with majority). Panel B illustrates the average level of connectivity across groups (minority with majority). Panel C illustrates the average level of connectivity of the minority, for players that assimilate and those that persist on conforming to their group identity. In all panels choices are disaggregated by treatment.

Finally, Figure 11C reports the average level of connectivity of minority players who assimilate versus minority players who conform, for each treatment. Notably, connectivity is mainly driven by assimilation choices, even when the choice is not visible, as in IGNORANCE and REVEAL.

D Theory

In this section, I model group identity change to evaluate collective outcomes through a coordination game where two social groups interact: a majority and a minority. Minority and majority members are guided by different group identities and, as such, have preferences to behave in opposing ways. Benefits from aligning to a group identity come from the number of others a player coordinates with. As such, the size of the group serves to model economic differences and inequality. That is, the identity of the majority group is more efficient because more people would want to conform to it (for related models see Hernandez et al. 2013; Ellwardt et al. 2016; Goyal et al. 2021). Minority players can, however, abandon their group's identity and assimilate that of the majority. Assimilation has the potential to foster network connections across groups, resulting in economic mobility for the minority. A particular interest is to evaluate if in equilibrium the minority adopts one same group identity collectively or if it fractures into assimilators and conformists.

D.1 The model

Suppose the population N is fixed, of size n, and composed of a majority group and a minority group. I denote by subscript M the majority group N_{M}, of size n_{M} and by subscript m the minority group N_{m}, of size n_{m} (with $n_{\mathsf{M}} > n_{\mathsf{m}}$ and $n_{\mathsf{M}} + n_{\mathsf{m}} = n$). The two groups can be *ex-ante* differentiated by appearance and identity markers. External markers are exogenous and fixed, such as skin color, hair, size, etc., which are exogenous to the individual (see e.g., Efferson et al. 2008). On the other hand, identities are expressed as normative prescriptions, e.g. preferences, views and values of the group, which are initially exogenous but can be affected by individual decisions. That is, before the game begins, a player i is exogenously assigned an identity θ_i = M if she belongs to the majority and θ_i = m if she belongs to the minority.

The game has three stages. In Stage 1, every minority player $i \in N_{\mathsf{m}}$ chooses which group identity to adhere to $\bar{\theta}_i \in \{\mathsf{M},\mathsf{m}\}$. This results in two types of minority players: a player i who changes her group identity by assimilating, $\bar{\theta}_i = \mathsf{M} \neq \theta_i$, and a player i who conforms by maintaining the status quo, $\bar{\theta}_i = \theta_i = \mathsf{m}$. For simplicity, majority players are assumed not change their group identity, so that $\bar{\theta}_i = \theta_i$ for all $i \in N_{\mathsf{M}}$. Thus, I define $\bar{N}_{\mathsf{M}} = \{i \in N : \bar{\theta}_i = \mathsf{M}\}$ of size $\bar{n}_{\mathsf{M}}$ and $\bar{N}_{\mathsf{m}} = \{i \in N : \bar{\theta}_i = \mathsf{m}\}$ of size $\bar{n}_{\mathsf{m}}$ as the sets of players choosing identity M and m, respectively. The resulting vector of identities $\bar{\theta} = \{\bar{\theta}_1, \ldots, \bar{\theta}_n\}$ is common knowledge.

In Stage 2, every player $i \in N$ chooses a set of link proposals g_i with others, g_i =

$(g_{i1}, \ldots, g_{ii-1}, g_{ii+1}, \ldots, g_{in})$, where $g_{ij} \in \{0,1\}$ for any $j \in N \backslash \{i\}$. Let $G_i = \{0,1\}^{n-1}$ define i's set of link proposals. The cost of proposing a link is $c > 0$, which is independent of whether a link was reciprocated or not. The induced network $g = (g_1, \ldots, g_n)$ is a directed graph. The closure of g is an undirected network denoted by $\bar{g}$, where $\bar{g}_{ij} = g_{ij}g_{ji}$ for every $i, j \in N$. The set of i's neighbors in the network $\bar{g}$ is $N_i(\bar{g}) = \{j \in N : \bar{g}_{ij} = 1\}$ with cardinality $|N_i(\bar{g})| = n_i$. Define the set of all undirected networks $\bar{g}$ as $\bar{G}$.

In Stage 3, every player i decides between two possible actions $a_i = \{\mathsf{M}, \mathsf{m}\}$, the same for all of her neighbors. Denote the vector of chosen actions as a. Building on Akerlof and Kranton (2000) a group identity specifies prescriptions and preferences on how to act. Accordingly, I model normative prescriptions through differences in the gains from choosing alike with others. A player choosing the majority's identity, i.e. either a majority player or a minority player who assimilates, prefers action M to action m, while a minority player who conforms prefers action m to M. Specifically, the return for i from coordinating on her prescribed action with a neighbor j is v. On the other hand, coordinating on the non-prescribed action returns $v - r$, where r represents a loss in utility i experiences for acting in misalignment with her adopted group identity. There are no returns from a connection if neighbors do not coordinate their actions. Given an outcome profile $(\bar{\theta}, g, a)$, the utility of player i is defined as:

$$u_i(\bar{\theta}_i, \bar{g}, a) = \lambda(1 + \sum_{j \in N_i(\bar{g})} \mathbb{1}_{\{a_j = a_i\}}) - cn_i \tag{A-1}$$

where $\mathbb{1}_{\{a_j = a_i\}}$ is the indicator function that yields 1 if player j chooses the same action as player i, $\lambda = v$ if the action i chooses is aligned with her group identity, $a_i = \bar{\theta}_i$, and $\lambda = v - r$ if i acts against her group identity. The number of proposals made is n_i, which is multiplied by the cost of making a link proposal c, where $0 < c < v - r$. [26]

This is a multi-stage game with simultaneous decisions in each stage. A player's strategy is a contingent plan, specifying a choice at each stage conditional on each possible previous decision profile. In the following section, I characterize equilibrium outcomes.

D.2 Equilibrium analysis

The aim of the equilibrium characterization is to identify if outcomes where the minority is fractured into assimilators and conformists are stable, in comparison to outcomes where the minority group collectively chooses one same group identity. For this, I look at

[26] Observe that if $v - r < c$, then no player will benefit from choosing an action in misalignment with their group's prescription. Moreover, if $v < c$, then no player benefits from proposing any link.

two settings with different assumptions of what it implies to assimilate. In the first, it is assumed that adherence to a group identity implies alignment on the action prescribed by such an identity (Schelling 1960; Bernheim 1994; Akerlof 1997). This means that when a player chooses the majority identity M in Stage 1, she will follow through by playing action M in Stage 3. The same holds for a player conforming to the minority identity m. In the alternative setting, I evaluate the assumption that alignment is not enforced and group identities are, thus, non-enforceable (Bicchieri 2006; Schram and Charness 2015; Fehr and Schurtenberger 2018). In a nutshell, I find that outcomes predicted under enforceable identities remain stable when alignment is not imposed. I report the results for the enforceable case first and then comment on the case of non-enforceability of group prescriptions. I discuss differences between the two settings in Section D.4.

The assumption that the prescriptions of behavior from adopting a group identity are binding transforms the 3-stage game into a 2-stage game. That is, players choose a group identity and action in Stage 1 and then form links in Stage 2. I now solve the two-stage game with identity (and action) choice and link formation. For this, I adapt the pairwise stability notion from Jackson and Wolinsky (1996) to my setting.[27] In the spirit of their definition, I say that a network is stable if no individual can profitably deviate either unilaterally or with one other individual.

Definition 1 *An identity-network pair* $(\bar{\theta}, \bar{g})$ *is pairwise stable if:*

- *for every* $\bar{g}_{ij} = 1, u_i(\bar{\theta}, \bar{g}) \geq u_i(\bar{\theta}, \bar{g} - \bar{g}_{ij})$ *and* $u_j(\bar{\theta}, \bar{g}) \geq u_j(\bar{\theta}, \bar{g} - \bar{g}_{ij})$.
- *for every* $\bar{g}_{ij} = 0, u_i(\bar{\theta}, \bar{g}) \geq u_i(\bar{\theta}, \bar{g} + \bar{g}_{ij})$ *or* $u_j(\bar{\theta}, \bar{g}) \geq u_j(\bar{\theta}, \bar{g} + \bar{g}_{ij})$

In this definition, part (1) says that no player can delete an existing link and profit, while part (2) says that no pair of players can form an additional link and increase their payoffs. This leads to the following proposition:

Proposition 1 *The pair* $(\bar{\theta}^*, \bar{g}^*)$ *is pairwise stable if one of the following obtains:*

(i) All minority players assimilate, so that $\bar{\theta}_i^* = \mathsf{M}$ *for all* $i \in N$ *and* $\bar{g}^*$ *is a complete network.*

(ii) Some or no minority players assimilate, so that the network $\bar{g}^*$ *is divided into two*

[27] I use pairwise stability as a solution concept instead of only Nash equilibrium because of the two-sided linking in the model, for which Nash equilibrium is too permissive and, thus, uninformative. For example, it is stable under Nash equilibrium for two players choosing the same group identity/action to not be connected with each other, despite both having incentives to connect.

complete components, $Q_{\mathsf{M}} = \bar{N}_{\mathsf{M}}$ and $Q_{\mathsf{m}} = \bar{N}_{\mathsf{m}}$, where every player in component Q_{M} follows identity M*, while every player in Q_{m} is a conformist following identity* m*.*

Proof. Player i's payoff from linking to a player j is $v - c$ if both choose the same group identity, and consequently the same action, or $-c$ if i chooses different from j. So, if $\bar{\theta}_i = \bar{\theta}_j$, both players have incentives to connect, and given deviations are bilateral they will. Otherwise, if their chosen identities are not the same, neither of them has incentives to connect. □

The intuition from Proposition 1 is that in a pairwise-stable outcome, players in the majority and the minority are willing to establish connections across groups when the latter assimilate, for it implies they will also choose the majority's action. This points to three types of equilibrium outcomes. Proposition 1(i) describes the *all-assimilate outcome*, where all minority players assimilate. Therefore, any pair of unconnected players i and j are better off forming a link and increasing their utility by v, given that both follow the same normative prescription. If all players assimilate the level of inter-group connectivity is maximal, as the complete network is the pairwise-stable outcome.

Proposition 1(ii) describes two outcomes, one when some minority players adopt the mainstream identity and the other when no one adopts it. The first is a *fractured outcome*, which implies that the minority is socially divided into a fraction that abandons the group's prescription to become mainstream and the rest that persist on it. When the minority is divided, adopters integrate with the majority players in Component Q_{M}, while non-adopters remain isolated from the majority in the network component Q_{m}. The second case is the *all-conform outcome*, where all players maintain their original group identity in unity. As such, there is no integration between groups because the minority collectively chooses the same identity. All minority players are in Q_{m} and there are only majority players in Q_{M}.

Now I turn to efficiency and social welfare, which I define as the sum of payoffs of all players. An outcome is said to be socially efficient if it maximizes aggregate welfare. I show that the *all-assimilate outcome* maximizes social welfare.

Proposition 2 *A complete network where all minority players change their group identity for the mainstream prescription is socially efficient.*

Proof. Let y and z be the number of players choosing action m in n_{M} and n_{m}, respectively. The sum of individual payoffs is

$$W(y, z) = (n - y - z)(v(n_{\mathsf{M}} - y) + (v - r)(n_{\mathsf{m}} - z)) + (y + z)((v - r)y + vz) \quad \text{(A-2)}$$

For fixed z, social welfare is decreasing in y if $y < y^*$ and increasing in y for $y > y^*$, where

$$y^* = \frac{(v-r)(n_{\mathsf{m}} - 2z) + v(n_{\mathsf{M}} - 2z) + v(n)}{2(2v-r)}. \tag{A-3}$$

Similarly, for any y, social welfare is decreasing in z if $z < z^*$, and increasing in z for $z > z^*$, where

$$z^* = \frac{v(n_{\mathsf{M}} - 2y) + (v-r)(n_{\mathsf{m}} - 2y) + (v-r)n}{2(2v-r)} \tag{A-4}$$

Since $0 \le y \le n_{\mathsf{M}}$ and $0 \le z \le n_{\mathsf{m}}$, it follows that $W(y,z)$ is maximized for some $y \in \{0, n_{\mathsf{M}}\}$ and some $z \in \{0, n_{\mathsf{m}}\}$. Note that $W(0, n_{\mathsf{m}}) = v(n_{\mathsf{M}}^2 + n_{\mathsf{m}}^2)$, and $W(n_{\mathsf{M}}, 0) = \beta(n_{\mathsf{M}}^2 + n_{\mathsf{m}}^2)$, which directly implies that $W(0, n_{\mathsf{m}}) > W(n_{\mathsf{M}}, 0)$ (because $r > 0$). Furthermore, since $W(0,0) = n(vn_{\mathsf{M}} + (v-r)n_{\mathsf{m}})$, then $W(0,0) > W(0, n_{\mathsf{m}})$ if and only if

$$\frac{n_{\mathsf{M}}}{n_{\mathsf{m}}} > \frac{r}{2v-r} \tag{A-5}$$

This inequality holds whenever $n_{\mathsf{M}} > n_{\mathsf{m}}$.

Similarly, since $W(n_{\mathsf{M}}, n_{\mathsf{m}}) = n((v-r)n_{\mathsf{M}} + vn_{\mathsf{m}})$, then $W(n_{\mathsf{M}}, n_{\mathsf{m}}) > W(0, n_{\mathsf{m}})$ if and only if

$$\frac{n_{\mathsf{m}}}{n_{\mathsf{M}}} > \frac{r}{2v-r} \tag{A-6}$$

This inequality holds whenever $n_{\mathsf{m}} > n_{\mathsf{M}}$. Furthermore, note that equations A-3 and A-4 hold for $n_{\mathsf{M}} = n_{\mathsf{m}}$ as long as $v > r$. To summarize, there is always either $W(0,0) > W(0, n_{\mathsf{m}})$ or $W(n_{\mathsf{M}}, n_{\mathsf{m}}) > W(0, n_{\mathsf{m}})$ as long as $n_{\mathsf{M}} \neq n_{\mathsf{m}}$ or $v > r$.

Now, consider the case where $y = n_{\mathsf{M}}$ and $z = n_{\mathsf{m}}$: this implies that $y + z = n$. Since $r > 0$, it can be shown that $W(0,0) > W(n_{\mathsf{M}}, n_{\mathsf{m}})$ so long as $n_{\mathsf{M}} > n_{\mathsf{m}}$. Moreover, $W(0,0) < W(n_{\mathsf{M}}, n_{\mathsf{m}})$ holds as long as $n_{\mathsf{M}} < n_{\mathsf{m}}$. Finally, $W(0,0) = W(n_{\mathsf{M}}, n_{\mathsf{m}})$ if $n_{\mathsf{M}} = n_{\mathsf{m}}$.

This shows that in a setting with fixed identities social welfare is maximized when the network is complete and all players choose action: M.

Consider now the setting with flexible identities. There are two different symmetric scenarios: (i) a case in which all minority players are conformists and maintain their initial group identity, $\bar{\theta}_i = \theta_i$, and (ii) a case where all players choose the same identity, because everyone in the minority assimilates.

The results so far have shown the efficient outcome in case (i). Compare this to the setting in which all players in the minority assimilate. Thus, $\bar{\theta}_i = \bar{\theta}_j = \mathsf{M}, \forall i, j \in N$. Thus, if the cost of assimilation is $\delta = 0$ then: $W(0,0) = n(vn_{\mathsf{M}} + (v-r)n_{\mathsf{m}}) < W(n_{\theta_i=\theta_j=\mathsf{M}}) = vn^2$ which is always true given that $r > 0$. This shows that if identities are flexible, it is efficient for all players to choose the group identity of the majority, as well as consistently choose the action prescribed to that identity. This outcome is better than a combination of identities, and thus it is more efficient than the case where only a subset of the minority assimilates while the rest conform.

Finally, consider the case in which the assimilation cost is positive: $W(0,0) = n(vn_{\mathsf{M}} + (v-r)n_{\mathsf{m}}) < W(n_{\theta_i=\theta_j=\mathsf{M}}) = vn^2 - \delta n_{\mathsf{M}}$ which is true as long as:

$$\delta < r \cdot n \tag{A-7}$$

This simply states that as long as the cost of assimilation is not too high, the efficient outcome is that in which all players choose the same group identity and the action that is prescribed with the chosen group identity, the *all-assimilate outcome*. If assimilation is costly the chosen identity/action is that of the majority, given that fewer players pay the cost. This completes the proof. □

Intuitively, this result states that a fraction of minority players conforming is never socially desirable. This is the case, because conformists divide and form a separate community from those who assimilate. As such, conformists are connected with fewer others and earn less than those who assimilate. However if everyone in the minority assimilates, all players in both the majority and the minority are better off. Figure 12 illustrates the utility of a majority player (solid line) and of a minority assimilator (long-dash line) as a function of the number of assimilators. Notably, without any identity change, minority players are worse-off than majority players. Such inequality is due to differences in group size (see the difference between lines at the intercept, when there are zero assimilators). However, once a player in the minority assimilates, her utility in the pairwise-stable outcome is the same as that of a majority player. Naturally, although not depicted in Figure 12, the utility of conformists decreases linearly on the number of assimilators, given their network shrinks. Therefore, in a *fractured outcome* inequality between assimilators and majority players disappears, while inequality between assimilators and conformists within the minority group emerges. However, when all minority players assimilate, every player $i \in N$ gets $u_i = vn - c(n-1)$, which maximizes efficiency and eliminates inequality within and between groups.

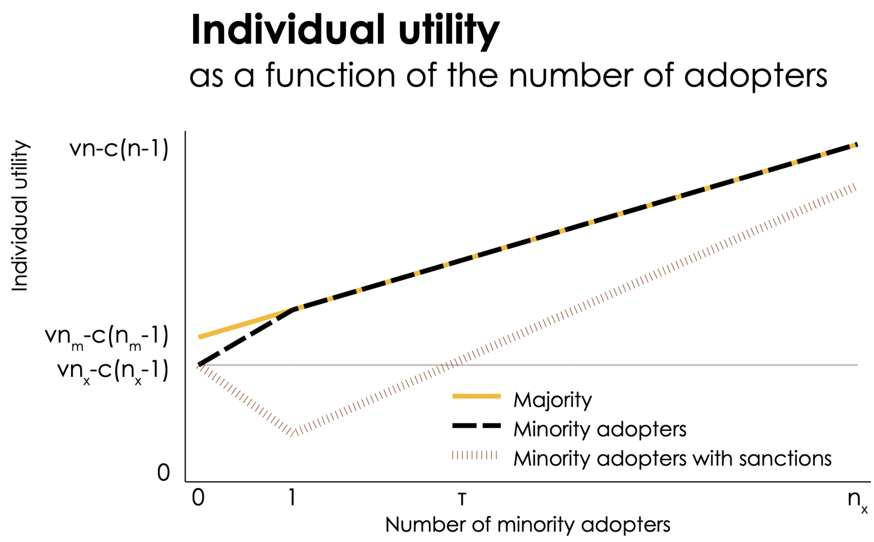


Figure 12 Individual utility with respect to the number of minority adopters.
The graph displays the utility of individual players in pairwise-stable outcomes (vertical axis) as a function of the number of assimilators (horizontal axis). The solid black line displays the utility for a majority player. The long-dashed line displays the utility for a minority player who assimilates. The dotted line displays the utility for a minority assimilator who is exposed to social sanctions, i.e., the cost of assimilation increases with the number of conformists. The horizontal gray line references the utility of a minority player in the *all-conform outcome*. Values below this line display cases where assimilation is detrimental and values above display cases where assimilation leads to economic mobility for the minority player.

D.3 Costs and social sanctions

In multiple settings identity change is not free. There are material and social costs for minority individuals who choose to assimilate. Let $\delta > 0$ be a positive cost that any minority player pays for assimilating.[28] The effect of the identity-change cost is presented in the following proposition:

Proposition 3 *If the assimilation cost δ is:*

(i) $\delta > vn - c(n-1)$, a minority player can never assimilate and be in a pairwise-stable outcome.

(ii) $0 < \delta < v(n_{\mathsf{M}} + 1) - cn_{\mathsf{M}}$, a minority player can assimilate and be in a pairwise-stable outcome, even when no one else in the minority assimilates.

(iii) $v(n_{\mathsf{M}} + 1) - cn_{\mathsf{M}} < \delta < (v - c)(n_{\mathsf{M}} + n_{\mathsf{m}} - \tau) + v$, a minority player can assimilate and be in a pairwise-stable outcome, as long as τ others in the minority assimilate as well.

[28] This transforms the payoff function in Equation A-1 into $u_i(\bar{\theta}_i, \bar{g}, a) = \lambda(1 + \sum_{j \in N_i(\bar{g})} \mathbb{1}_{\{a_j = a_i\}}) - (cn_i - \delta)$

Proof. Proposition 3 looks at the role of costs to assimilation. If a minority player abandons her group identity to assimilate that of the majority her utility is:

$$u_i = v(n_{\mathsf{M}} + \tau + 1) - c(n_{\mathsf{M}} + \tau) - \delta$$

where τ is the number of minority players who also assimilate, while $n_{\mathsf{m}} - \tau - 1$ is the number of minority players conforming. I look at three cases. First, the *all-assimilate outcome*, where all minority players assimilate. This leads to $\tau = n_{\mathsf{m}} - 1$ and $n_{\mathsf{M}} + \tau + 1 = n$. In this case, it is a best response for i to assimilate, if:

$$\delta < vn - c(n - 1) \tag{A-8}$$

The second case is a *fractured outcome* with full conformism in the in-group, so that only player i assimilates, while everyone else in the minority conforms to their group identity. This leads to $\tau = 0$. Thus, i only links to the set of majority players n_{M}. In this case, it is a best response for i to assimilate, if:

$$\delta < v(n_{\mathsf{M}} + 1) - c(n_{\mathsf{M}}) \tag{A-9}$$

The third case is a *fractured outcome* with some conformism, where a subset of minority players assimilates. This leads to $0 < \tau < n_{\mathsf{m}}$. Naturally, if δ is as in Equation D.3, player i assimilate independently of the size of τ. However, if δ is higher, then it is a best response for i to assimilate, if:

$$v(n_{\mathsf{M}} + 1) - c(n_{\mathsf{M}}) < \delta < (v - c)(n_{\mathsf{M}} - \tau) + v \tag{A-10}$$

This completes the proof. □

Intuitively, Proposition 3(i) states that if the cost of assimilation is higher than the payoff a minority player earns in the *all-assimilate outcome*, her utility from adopting the majority's identity would be always negative. As such, only the *all-conform outcome* is pairwise stable. Proposition 3(ii) gives a lower bound on the assimilation cost, such that any fraction of minority players assimilating can be sustained in a pairwise-stable outcome. For this, the cost must be lower than the payoffs that a minority player gets in the *fractured* *outcome* where she is the only assimilator. Finally, 3(iii) states that if the assimilation cost is within the two bounds mentioned above, a minority player only assimilates when there are τ other minority players assimilating, as well. Thus, the higher the value of δ, the larger the fraction of the minority that must assimilate to sustain a *fractured outcome*.

A particular case of Proposition 3(iii) takes place when the cost of assimilation is not fixed but rather decreases with the number of assimilators in the minority group. This type of social cost has been evidenced in settings where conformists exert peer pressure on those in-group members who assimilate (see e.g., Austen-Smith and Fryer 2005; Battu and Zenou 2010; Fryer and Torelli 2010). The aim of such social sanctions on assimilator is to police the group's boundaries and prevent peers from leaving. Consequently, instead of only having a fixed element, assimilation costs may increase as a function of the number of minority players conforming. For example, the cost of assimilation can be expressed by $\delta+\omega(n_{\mathsf{m}}-\tau)$, where ω is the rate at which the cost increases on the number of conformists, where $n_{\mathsf{m}}-\tau$ is the number of players in the minority minus those who assimilate.[29] This leads to the following corollary to Proposition 3:

Corollary 1 *In a setting of social sanctions, where the assimilation cost $\delta + \omega(n_{\mathsf{m}} - \tau)$ decreases with the number of assimilators, a minority player assimilates if $\tau \geq \frac{\delta+\omega n-v}{(v+\omega-c)} - n_{\mathsf{M}}$ of her peers assimilates as well.*

Proof. If the cost to assimilation is $\delta + \omega(n_{\mathsf{m}} - \tau)$, the utility of player i for changing her group identity is:

$$u_i = v(n_{\mathsf{M}} + \tau + 1) - c(n_{\mathsf{M}} + \tau) - (\delta + \omega(n_{\mathsf{m}} - \tau))$$

where $n_{\mathsf{m}} = n - n_{\mathsf{M}}$. It is a best response for i to assimilate, if:

$$\tau > \frac{\delta + \omega n - v}{v + \omega - c} - n_{\mathsf{M}} \tag{A-11}$$

so that the threshold necessary for i to assimilate is smaller the larger the size of the majority (n_{M}), also smaller the larger the gains from choosing the action in alignment with the chosen identity (v), and larger the size of the fixed cost δ and of the rate at which the variable cost increases (ω). □

[29] Naturally, there are many more functional forms one could use to express assimilation costs under social sanctions. However, the linear function is useful because of its simplicity and because the main goal of the example is to show how the cost can vary with the fraction of assimilators in the group.

The intuition from Corollary 1 is that if assimilation costs encompass social sanctions, utility for assimilators does not increase monotonically, as in the case without cost. Instead, it is a convex function in which a fractured minority leads to lower utility than no assimilation at all, when the number of assimilators is below the thresholds τ, and it leads to higher utility when it is above it. An example of this is illustrated by the dotted line in Figure 12.[30]

D.4 Discussion of the assumptions

I conclude the theory section with a discussion of the main assumptions underlying the analysis. I do so by addressing variations to the characterization presented above, which highlights the role played by some of its key features.

A1. *Assimilation choice for the minority alone*

In the model, I restrict assimilation choice to the minority. Assuming that majority members do not change their group identity is in line with other models in the literature, see for example Bisin et al. (2011). However, to further explore the implications of this assumption, I characterize pairwise-stable outcomes for the case in which majority players can also assimilate. I show that relaxing this assumption does not exclude any of the pairwise-stable outcomes characterized in Proposition 1. It does, however, allow for other outcomes to be pairwise stable. Among these is, for example, the *fractured outcome* where the entire majority gives up their identity and assimilates that of the minority, while the entire minority assimilates the identity of the majority. Although possible, outcomes like these are unlikely and not of central interest to my study on minority assimilation and economic mobility. Nonetheless, in the experiment I test the robustness of this assumption by allowing majority players to choose whether to keep or change their group identity. Previewing the findings, even when the assumption is relaxed and all players can change identities, majority players persist on keeping their group identity, irrespective of what minority players do (see Section **??**).

Proof. If the majority players can also change group identities, the intuition of the pairwise stable equilibria from Proposition 1 still holds, although other outcomes arise. This is described in the following corollary:

Corollary 2 *The pair $(\bar{\theta}^*, \bar{g}^*)$ is pairwise stable if one of the following obtains:*

[30] For this illustration I use the following parameters: $\omega = c$, $\delta = v$, and $N_M = N_m + 1$. Parameters are discussed in more detail in Section 3.

(i) All players adopt the same group identity, so that $\bar{\theta}_i^* = \{\mathsf{M}, \mathsf{m}\}$ for all $i \in N$ and $\bar{g}^*$ is a complete network.

(ii) The network $\bar{g}^*$ is divided into two complete components, $Q_{\mathsf{M}} = \bar{N}_{\mathsf{M}}$ and $Q_{\mathsf{m}} = \bar{N}_{\mathsf{m}}$, where every player in component Q_{M} adopts group identity M*, while every player in Q_{m} adopts group identity* m.

The proof for Corollary 2 is the same as that for 1. Importantly, Corollary 2 implies that the *all-assimilate outcome* can take two forms. Either all players adopt the identity of the majority, so that the minority assimilates but the majority conforms, or all players adopt the identity of the minority, so that the majority assimilates while the minority conforms. The second item, states that the *all-conform outcome* where no one assimilates remains a pairwise stable. It also states that a *fractured outcome* can take multiple forms. Not only those in which a fraction of the minority assimilates while no one in the majority does, but also some in which a fraction in each subgroup changes identities, and also one where all players change identities. □

A2. *Enforceable identities*

A key assumption in the equilibrium characterization above is that the group identity a player chooses in Stage 1 works as an enforceable commitment on the action she plays in Stage 3. Relaxing this assumption, by making the normative prescriptions associated with a group identity non-enforceable, turns the setting back into a 3-stage sequential game. Players are still potentially better-off aligning their action with the identity they choose, so that each gets v instead of $v - r$ for every connection in a pairwise-stable outcome. That is, all outcomes characterized with enforceable group identities remain pairwise stable if the prescription of a group identity is non-enforceable. However, with non-enforceability the action is chosen after the linking stage. As such, there are pairwise-stable outcomes in which a player best responds by aligning to her reference network rather than by aligning to her group identity. This is the case of a player who is mostly connected to neighbors choosing her disliked action, where she is better off acting against her group identity and benefiting from the large share of neighbors she can coordinate with. As such, in addition to the pairwise-stable outcomes described in Proposition 1, there are also outcomes where players choose an action in misalignment with their chosen group identity. In all cases, the same type of properties for the equilibrium outcomes hold: players are connected to all others choosing the same action, and to no one else.

I test the robustness of the enforceability assumption in the experiment, by allowing all players to choose actions independently of their chosen group identity. Previewing the findings, behavior in the experiment shows strong correlations between the *chosen* group identity in Stage 1 and the action chosen in Stage 3, for both majority and minority players.

Proof. If group identities are not enforceable, the game has three stages. I solve backwards starting with the choice of actions in a given network where players have also decided on their group identity. Let τ refer to the number of neighbors choosing the action that aligns to i's group identity. That is, if i's group identity is $\overline{\theta}_i = \mathsf{M}$ then the aligning action is $a_i = \mathsf{M}$, and the opposite when the chosen identity is m. This means that player i's payoff from choosing in alignment with her chosen identity is $v(\tau + 1)$ and from choosing in misalignment is $(v - r)(n - \tau + 1)$. So i is strictly better off choosing in alignment with her chosen group identity if and only if

$$v(\tau + 1) > (v - r)(n - \tau + 1) \tag{A-12}$$

This inequality can be rewritten as

$$\tau > \frac{v - r}{2v - r}n - \frac{r}{2v - r} \tag{A-13}$$

so that a player is better off selecting in alignment with her chosen group identity if and only if the proportion of her neighbors selecting the same action is sufficiently large.

From this, it follows that the pairwise stable networks are those in which the action profile is an equilibrium, and where players are linking as in Proposition 1. That is, no player connects to someone choosing a different action. And, only networks where all players choosing the same action are connected are pairwise stable.

Finally, given identity choice is non-enforceable and there are no direct benefits from choosing one group identity or the other in Stage 1 (naturally there are at the end of the game), the set of equilbria is such that: (i) there is a complete network where all players choose the same action, similarly to the *all-assimilate outcome*, but where players may or may not choose in alignment with their group identity, (ii) there is a segregated network where all players maintain their group identity and choose in alignment, as in the *all-conform outcome*, and (iii) there are segregated networks where a subset of players choose action M and the other subset chooses action m. In both cases, players may also be choosing the corresponding group identity or not. Naturally, the case where the action aligns to the social identity Pareto dominates the same action profile and network but where the vector of group identities is such that a player is not matching her group identity to the action in stage 3, given that she is earning v in the former and $v - r$ in the latter.

For more details on the equilibria for the 3-stage game see Goyal et al. (2021). $\square$

A3. *Visible group identities*

In the model there is complete information about the vector of chosen group identities. There are however, many settings of incomplete information where minority players who assimilate and minority players who conform are, a priori, indistinguishable (see e.g., Efferson et al. 2008). Such cases emerge when identities are not observable directly. I do not provide a formal characterization of this setting, but instead discuss some of its implications: If identity choice is unobservable, there is strategic uncertainty for everyone in the game. Even for those in the minority respect to others in the minority. Because of this, even if beliefs about the number of assimilators are accurate and the same for all players, there are high risks when only a fraction of the minority assimilates. For example, in a *fractured outcome*, if a single minority player assimilates, she is unable to show she is the assimilator. This can lead to losses for the assimilator if she proposes links to players in the majority. Also, for those in the majority if they mistakenly propose links to a minority player that did not assimilate, given they know for sure someone assimilated but cannot tell who among the minority players is the assimilator. It can also lead to losses for conformists in the minority because they may propose to the assimilator, instead of a conformists, for the same reasons mentioned above.

Naturally, with incomplete information about group-identity choices also includes the case in which players propose links correctly. As such, the outcomes described in Proposition 1 are also pairwise stable when group identities are not observable. Note also, that if beliefs are correct and symmetric (everyone has information about the number of assimilators but not their identities), the case where all conform or the case where everyone assimilates are equivalent to those of complete information, and so the risk is minimized compared to the *fractured outcome*.