Edutainment, social network, and ethnic discrimination in rural Bangladesh*

Pre-Analysis Plan

Abu Siddique† Michael Vlassopoulos‡ Yves Zenou§ Asad Islam¶

January 5, 2023

Abstract

In this document, we provide the research design and analysis plan for evaluating the impact of information dissemination through a documentary film that educates the ethnically dominant Bengalis about the ethnic minority Santals in Bangladeshi polyethnic villages. We want to evaluate whether exposure to information affects behavior and opinions of the ethnic majority towards minorities. More importantly, we want to test whether information diffusion is stronger when it is targeted to network-central ethnic majorities. We will measure outcomes using survey questions, administrative data, lab-in-the-field experiments, and a field experiment.

JEL: C93, D9, I31, J15, O12

Keywords: Ethnic discrimination, stereotypes, social network, edutainment, emotions, randomized experiment, Bangladesh.

*Ethical clearance for this project was received from the Monash University Human Research Ethics Committee (ref no. 35152). This project is funded by the Diligentia Foundation. We thank Kazi Robiul Alam and Golam Faruk Sarker from the University of Rajshahi, and Hopna Kisku from *Ashrai* NGO for their comments and guidance on the documentary film. We thank the production company *Chitrakkhi* for creating the documentary film, and the NGOs *SARCH* and *Ashrai* for their support in the fieldwork. We also thank Eliana La Ferrara, Cynthia Kinan, and Chris Roth for their comments on the experimental design.

†Department of Political Economy, King's College London, UK. Email: abu.b.siddique@kcl.ac.uk

‡Economics Department, Social Sciences, University of Southampton, UK, and IZA. Email: m.vlassopoulos@soton.ac.uk

§Department of Economics, Monash University, Australia, CEPR, and IZA. Email: yves.zenou@monash.edu

¶Department of Economics, Monash University, Australia, and CDES. Email: asadul.islam@monash.edu

1 Motivation

The objective of this project is to evaluate an innovative intervention aimed at updating the ethnic majority people's existing information regarding ethnic minority people, with a view to curb prejudice and foster interethnic harmony. More specifically, we propose to carry out a randomized experiment in rural Bangladesh to evaluate the effectiveness of local screenings of a documentary film that educates the ethnic majority Bengali about the lives and livelihoods of Santals (one of the largest ethnic minorities in Bangladesh) in reducing ethnic discrimination. We also leverage social relationships of the ethnic majority in villages and target information to network-central Bengalis to test if it reduces discrimination towards minorities further. Being exposed to entertainment media (edutainment) promises to change norms and behavior through the channels of providing both new information as well as making existing information attractive, as posited by social learning theory (Bandura, 1977; La Ferrara, 2016; Grady et al., 2021). However, information dissemination through central agents in a network can be more effective in spreading information than broadcasting information widely, such as via mass media (Banerjee et al., 2018). Therefore, in this study, we aim to provide the first empirical test of both theories in the context of ethnic discrimination.

2 Background

Bangladesh is a suitable place to study interethnic discrimination because it is home to about 45 different indigenous ethnic minority groups. These groups are culturally, racially, ethnically, and linguistically distinct from the majority Bengali population—the major ethnolinguistic group in Bangladesh. Conflicts and violence between the Bengali and these tribal ethnic minorities (known as Adivasis) are commonplace. There is a long history of Adivasis being stigmatized, marginalized, and discriminated against by the ethnic majority in many domains including access to basic social services such as health, food, and nutrition, education, employment, justice and politics (Roy, 2012).

The Northwestern region (Rajshahi and Naogaon districts) where this project will take place, is home to the second-largest ethnic minority community, the Santal. Like other ethnic minorities in Bangladesh, Santals also face poor economic conditions, lack of educational opportunities, poor health, and have experienced a loss of their agricultural lands to land grabbers. Santals are predominately landless farmers and have been found to be discriminated in agricultural markets (Siddique et al., 2023), and to have an aversion for interethnic competition (Siddique & Vlassopoulos, 2020). They live mostly

in remote multiethnic villages, but are segregated and rarely socialize with non-coethnic people.

3 The intervention

3.1 The documentary film and research design

In collaboration with a Bangladesh-based film production team, *Chitrakkhi*, which is operated by Bangladeshi film students, we produced a documentary film on the lives and livelihoods of Santals of Northwestern Bangladesh. This film has three main layers of storytelling: first, it introduces the Santali culture to the viewers with various details on the culture, rituals, cuisines, etc., that are often unknown to the ethnic majority (most of this filming took place during the *Baha* festival); second, the focus then shifts to the struggles and misfortune of Santals, such as housing, access to clean water, low income, professional barriers, educational challenges of young Santals, social barriers, etc.; and finally, viewers learn about the aspirations of Santals, their potential to be successful, and how various Santals are overcoming barriers and becoming successful in different fields, such as in education and profession. All of these stories are shown through a Santali perspective and all stories are told by the local Santals (i.e., non-actors). Importantly, the documentary film tries to capture the social problems and contexts naturally—without using any scripts—such that the film is an ethnographic documentation of the lives of Santals in Bangladesh. To retain this natural aspect, the film was made under the close supervision of a visual anthropologist, Kazi Robiul Alam, and economic anthropologist, Golam Faruk Sarker, both associate professors of anthropology at the University of Rajshahi, Bangladesh. We also sought comments and suggestions on the film from Hopna Kisku, the deputy director of *Ashrai* (an NGO that works on the well-being of ethnic minorities), who is also a Santal. The runtime of this film is about 45 minutes. See it here (please do not share without permission: https://www.youtube.com/watch?v=hWizDrLXLoc.

Using a cluster randomized control trial in the Rajshahi and Naogaon districts of Bangladesh, we will evaluate the effectiveness of this documentary film. We chose these districts because it is home to Santals, the second largest ethnic minority group in Bangladesh. Moreover, discrimination, conflicts, and violence between the ethnic majority (Bengali) and Santal are commonplace in this region. To carry out the fieldwork, we collaborated with two local NGOs, *Ashrai* and *SARCH*. From *Ashrai*, we obtained a list of about 150 multiethnic villages where the fraction of minorities is between 30-70 percent. In other words, if a multiethnic village has at least one-third Bengali or Santal

people, then they are considered for this RCT. We randomly selected 120 villages from this list, with 40 villages in each of the three treatment arms. The treatment arms are as follows:

- Treatment *Random* (40 villages, $N = 560$ Bengalis): 14 randomly selected households per village will watch the documentary film on Santals.
- Treatment *Central* (40 villages, $N = 560$ Bengalis): 7 most network-central households and 7 randomly selected households per village will watch the documentary film on Santals.
- Control (40 villages, $N = 560$ Bengalis): 14 randomly selected households per village will watch a documentary film on floriculture or flower farming in Bangladesh (same length as the documentary film on Santals).

We closely follow Banerjee et al. (2019) to find network-central residents in our Treatment *Central* villages. We rapidly surveyed 20 random people per village and asked them to name the 15 most suitable people (or households) in their villages who would be a good diffuser of information. Enumerators went to a central place per village, such as the village market, and randomly surveyed people to get this list, where each survey took about 5 minutes to complete. Specifically, we ask:

> *"If we want to spread information to everyone in the village about village events, immunization programs, a new loan project, or a fair that we would like to organize in your village, to whom should we speak? Please nominate 15 such people/households from your village."*

Based on this list, 7 most consistent and highly ranked names are selected for the *Central* treatment.

The screening will take place at each participants house, individually. From baseline, we will have mobile phone numbers of all participants. Enumerators will call and fix a day and time with the participants prior to each screening visit. The documentary film will be screened via 10 inch tablet PCs connected to USB speakers. This way participants and other household members will be able to watch the documentary film together. There will be a four-question quiz based on the film at the end of the screening. Answering each question correctly will allow participants to win 10 Taka (maximum 40 Taka). We will organize this incentivized quiz so that participants pay attention to the video content. Viewers will take a break at 25 minutes of the video.

Following the quiz, the enumerator will ask the following questions to the participants:

- Name five new things you learned from the video today.[1]

[1] Control group participants will be asked the same question.

"Now imagine a Santal in your village is in similar conditions as the Santals in the film that you just watched..."[2]

- What would you do in that situation?
- Why?
- What would you advise your neighbors to do in that situation?

We will ask the first post-screening question to understand to what extent the information through the documentary film is new to participants. The last three post-screening questions will allow participants to deeply think about this social problem and what they could do to mitigate it. All questions are open-ended, primarily to reduce response biases. The treatment delivery ends after the post-screening Q&A.

3.2 Data collection

We will collect data at baseline and endline. Baseline data collection will only involve surveys, while endline data collection will involve surveys, lab-in-the-field experiments, and a field experiment. Our surveys will be carried out in 120 multi-ethnic villages in Bangladesh (40 villages per treatment arm), and will be conducted on the following individuals:

- 1,680 main participants that are Bengali and have been treated with documentary films (560 Bengali per arm or 14 Bengali per village).
- 840 untreated Bengalis across the three treatment arms (280 Bengali per arm or 7 Bengali per village).
- 840 Santals across the three arms (280 Santals per arm or 7 Santals per village).

Therefore, we will survey 28 participants (one per household) per village, and 3,360 participants in total.

3.2.1 Baseline characteristics

At baseline, we will measure a range of individual and household characteristics. These questions include but not limited to: age, gender, education, income, household size, occupation, if they migrated from other villages recently, savings, informal borrowings, household assets, etc. All 3,360 households will be asked these questions at baseline.

[2] Control group participants were not asked this question because they watched a video on floriculture.

3.2.2 Survey outcomes

Index outcomes will be generated in the following way: aggregate all responses to create a score, subtract the average score of the control group, and then divide this difference by the standard deviation of score in the control group. Therefore, the control group will have a mean of zero and standard deviation of one.

Outcomes of Bengali

Interaction index. Based on the following 4 questions that capture how often Bengalis interact with their Santal neighbors (answered as *yes=1* or *no=0*, thus aggregated score will be between 0 and 4):

1. Do you offer them food when they visit you?
2. Do you offer them chair/seat when they visit you?
3. Do you invite them during festivals?
4. Do they invite you during festivals?

Number of ethnic minority close friends. Participants will be asked "Name your ten closest friends (full name)". Based on surnames, we will count the number of ethnic minority close friends they have. In addition to count, we will also code it as a proportion (number of minority close friend divided by 10).

Water bill charge. Participants will be asked "We know you need to pay water bills. How much do Santals pay per fetch when they come to fetch water at your house?". We will directly use this monetary amount as the outcome. We will also code it as a dummy that equals to 1 if they charge any positive amount and 0 if not charged.

Number of visits to Santal neighbors. Participants will be asked "How many times do you visit your Santal neighbors in a month?". We will directly use this frequency as the outcome. We will also code it as a dummy that equals to 1 if they the frequency is at least once and 0 otherwise.

Number of Santal visitors. Participants will be asked "How often do your Santal neighbors visit you in a month?". We will directly use this frequency as the outcome. We will also code it as a dummy that equals to 1 if they the frequency is at least once and 0 otherwise.

Discriminatory opinions index. Based on the following 7 questions that capture Bengalis' opinions about Santals (each answered on a 0-10 scale, where 10 means "completely agree", thus aggregated score will be between 0 and 70):

1. All Santals that I know are honest people *(reverse scoring required)*
2. I always eat food and drinks offered by Santals *(reverse scoring required)*
3. Schools should be separate for Bengali and Santal children

4. I think there should be more Santal teachers in my child's school *(reverse scoring required)*
5. I do not enjoy working/doing business with Santals
6. Some of my children's best friends are Santals *(reverse scoring required)*
7. One can easily trust a Santal person *(reverse scoring required)*

Mental health index. We measure depressive and general anxiety symptoms using the PHQ-4 questionnaire (Kroenke et al., 2009). We ask the following questions: "Over the last two weeks, how often have you been bothered by the following problems? *Not at all=0 / several days=1 / more than half the days=2 / nearly everyday=3*":

1. Feeling nervous, anxious, or on edge.[a]
2. Not being able to stop or control worrying.[a]
3. Feeling down, depressed, or hopeless.[d]
4. Little interest or pleasure in doing things.[d]

Questions with [a] measures general anxiety symptoms, and that with [d] measures depressive symptoms. We will create indices using the standardizing procedure explained before. We will also create two outcomes using the official cut-off: Total score ≥ 3 for first 2 questions suggests anxiety; thus, the dummy outcome variable would equal to 1 if the total score ≥ 3 and 0 otherwise. Similarly, the total score ≥ 3 for last 2 questions suggests depression; thus, the dummy outcome variable would equal to 1 if the total score ≥ 3 and 0 otherwise.

Subjective well-being index. We measure this outcome using the following 4 questions: "On a scale from 0 to 10, where 0 means *not at all* and 10 means *extremely or all the time*":

1. Overall, how satisfied are you with your life nowadays?
2. Overall, to what extent do you feel the things you do in your life are worthwhile?
3. Overall, how happy do you feel nowadays?
4. Overall, how anxious do you feel nowadays?

Economic preferences. Following Falk et al. (2018), we will measure altruism, trust, and solidarity using the following questions:

1. *Altruism:* Imagine the following situation—Today you unexpectedly received 5,000 Taka. How much of this amount would you donate to a Santal person in your village? We will directly use this monetary amount as the outcome. We will also code it as 1 if the donation is $\geq 2,500$ and 0 if less.
2. *Solidarity:* Imagine your neighbor, Hopna Kisku, lost his house during a heavy storm. Also, imagine that today you unexpectedly received 5,000 Taka. How much of this amount would you donate to Hopna Kisku and his family? We will

directly use this monetary amount as the outcome. We will also code it as 1 if the donation is $\geq 2,500$ and 0 if less.

3. *Trust:* Please indicate your answer on a scale from 0 to 10. A 0 means *does not describe me at all*, and a 10 means *describes me perfectly*. "I assume that Santals have only the best intentions." We will directly use this score as the outcome. We will also code it as 1 if the trust score is above the median value and 0 otherwise.

Survey measures of economic preferences will only be measured at baseline. At the endline, altruism and solidarity will be measured using incentivized games (see section 3.2.3), but trust will be measured using the same survey question as baseline.

Spillover effects on Bengali neighbors

Interaction index. Based on the following 4 questions that capture how often Bengalis interact with their Santal neighbors (answered as *yes=1* or *no=0*, thus aggregated score will be between 0 and 4):

1. Do you offer them food when they visit you?
2. Do you offer them chair/seat when they visit you?
3. Do you invite them during festivals?
4. Do they invite you during festivals?

Number of ethnic minority close friends. Participants will be asked "Name your ten closest friends (full name)". Based on surnames, we will count the number of ethnic minority close friends they have. In addition to count, we will also code it as a proportion (number of minority close friend divided by 10).

Water bill charge. Participants will be asked "We know you need to pay water bills. How much do Santals pay per fetch when they come to fetch water at your house?". We will directly use this monetary amount as the outcome. We will also code it as a dummy that equals to 1 if they charge any positive amount and 0 if not charged.

Number of visits to Santal neighbors. Participants will be asked "How many times do you visit your Santal neighbors in a month?". We will directly use this frequency as the outcome. We will also code it as a dummy that equals to 1 if they the frequency is at least once and 0 otherwise.

Number of Santal visitors. Participants will be asked "How often do your Santal neighbors visit you in a month?". We will directly use this frequency as the outcome. We will also code it as a dummy that equals to 1 if they the frequency is at least once and 0 otherwise.

Discriminatory opinions index. Based on the following 7 questions that capture Bengalis' opinions about Santals (each answered on a 0-10 scale, where 10 means

"completely agree", thus aggregated score will be between 0 and 70):

1. All Santals that I know are honest people *(reverse scoring required)*
2. I always eat food and drinks offered by Santals *(reverse scoring required)*
3. Schools should be separate for Bengali and Santal children
4. I think there should be more Santal teachers in my child's school *(reverse scoring required)*
5. I do not enjoy working/doing business with Santals
6. Some of my children's best friends are Santals *(reverse scoring required)*
7. One can easily trust a Santal person *(reverse scoring required)*

Economic preferences. Following Falk et al. (2018), we will measure altruism, trust, and solidarity using the following questions:

1. *Altruism:* Imagine the following situation—Today you unexpectedly received 5,000 Taka. How much of this amount would you donate to a Santal person in your village? We will directly use this monetary amount as the outcome. We will also code it as 1 if the donation is $\geq 2,500$ and 0 if less.
2. *Solidarity:* Imagine your neighbor, Hopna Kisku, lost his house during a heavy storm. Also, imagine that today you unexpectedly received 5,000 Taka. How much of this amount would you donate to Hopna Kisku and his family? We will directly use this monetary amount as the outcome. We will also code it as 1 if the donation is $\geq 2,500$ and 0 if less.
3. *Trust:* Please indicate your answer on a scale from 0 to 10. A 0 means *does not describe me at all*, and a 10 means *describes me perfectly*. "I assume that Santals have only the best intentions." We will directly use this score as the outcome. We will also code it as 1 if the trust score is above the median value and 0 otherwise.

Survey measures of economic preferences will only be measured at baseline. At the endline, altruism and solidarity will be measured using incentivized games (see section 3.2.3), but trust will be measured using the same survey question as baseline.

Outcomes of Santals

Interaction index. Based on the following 4 questions that capture how often Bengalis interact with their Santal neighbors (answered as *yes=1* or *no=0*, thus aggregated score will be between 0 and 4):

1. Do your Bengali neighbors offer you food when you visit them?
2. Do your Bengali neighbors offer you chair/seat when you visit them?
3. Do your Bengali neighbors invite you during festivals?
4. Do you invite them during festivals?

Number of ethnic majority close friends. Participants will be asked "Name your ten closest friends (full name)". Based on surnames, we will count the number of ethnic majority close friends they have. In addition to count, we will also code it as a proportion (number of majority close friend divided by 10).

Water bill charge. Participants will be asked "We know you need to fetch water from your Bengali neighbors. How much do you pay per fetch?". We will directly use this monetary amount as the outcome. We will also code it as a dummy that equals to 1 if they charge any positive amount and 0 if not charged.

Number of Bengali visitors. Participants will be asked "How often do your Bengali neighbors visit you in a month?". We will directly use this frequency as the outcome. We will also code it as a dummy that equals to 1 if they the frequency is at least once and 0 otherwise.

Mental health index. We measure depressive and general anxiety symptoms using the PHQ-4 questionnaire (Kroenke et al., 2009). We ask the following questions: "Over the last two weeks, how often have you been bothered by the following problems? *Not at all=0 / several days=1 / more than half the days=2 / nearly everyday=3*":

1. Feeling nervous, anxious, or on edge.[a]
2. Not being able to stop or control worrying.[a]
3. Feeling down, depressed, or hopeless.[d]
4. Little interest or pleasure in doing things.[d]

Questions with [a] measures general anxiety symptoms, and that with [d] measures depressive symptoms. We will create indices using the standardizing procedure explained before. We will also create two outcomes using the official cut-off: Total score ≥ 3 for first 2 questions suggests anxiety; thus, the dummy outcome variable would equal to 1 if the total score ≥ 3 and 0 otherwise. Similarly, the total score ≥ 3 for last 2 questions suggests depression; thus, the dummy outcome variable would equal to 1 if the total score ≥ 3 and 0 otherwise.

Subjective well-being index. We measure this outcome using the following 4 questions: "On a scale from 0 to 10, where 0 means *not at all* and 10 means *extremely or all the time*":

1. Overall, how satisfied are you with your life nowadays?
2. Overall, to what extent do you feel the things you do in your life are worthwhile?
3. Overall, how happy do you feel nowadays?
4. Overall, how anxious do you feel nowadays?

Economic preferences. Following Falk et al. (2018), we will measure altruism and trust using the following questions:

1. *Altruism:* Imagine the following situation—Today you unexpectedly received 5,000 Taka. How much of this amount would you donate to a Bengali person in your village? We will directly use this monetary amount as the outcome. We will also code it as 1 if the donation is $\geq 2,500$ and 0 if less.
2. *Solidarity:* Imagine your neighbor, Iqbal Rahman, lost his house during a heavy storm. Also, imagine that today you unexpectedly received 5,000 Taka. How much of this amount would you donate to Iqbal Rahman and his family? We will directly use this monetary amount as the outcome. We will also code it as 1 if the donation is $\geq 2,500$ and 0 if less.
3. *Trust:* Please indicate your answer on a scale from 0 to 10. A 0 means *does not describe me at all*, and a 10 means *describes me perfectly*. "I assume that Bengalis have only the best intentions." We will directly use this score as the outcome. We will also code it as 1 if the trust score is above the median value and 0 otherwise.

Survey measures of economic preferences will only be measured at baseline. At the endline, altruism and solidarity will be measured using incentivized games (see section 3.2.3), but trust will be measured using the same survey question as baseline.

Food security index. We use the USDA Household Food Security Survey Module to measure food security in the household, each question answered on a 3-point scale: *often true=2, sometimes true=1, never true=0*; thus, the score is between 0 and 6. We ask the following questions "In the last 6 months, can you tell me if these statements were true for you?:

1. We worried whether our food would run out before we got money to buy more.
2. The food that we bought just didn't last, and we didn't have money to get more.
3. We couldn't afford to eat balanced meals.

New employment. Did you start any new job recently that lets you earn more than before? Coded as *yes=1* and *no=0*.

Monthly income. Last months' household income. We will directly use this monetary amount and $ln(income)$ as the outcome.

3.2.3 Lab-in-the-field experiments and outcomes of Bengali only

Dictator game. We will use a standard dictator game to measure altruism towards Santals. In this game, each Bengali participant (Player A) will be matched with a Santal recipient (Player B). Player A will receive an endowment of 100 Taka (=$1), which s/he will be asked to divide between her/himself and Player B. The share given to Player B would measure altruism. Player B will be randomly selected from the pool of

Santals mentioned under section 3.2. This game will be played at Player A's home and all players will remain anonymous. After the game ends, shared money will be handed over to Player B at their homes by the enumerators.

We will code this outcome in three ways: (i) the portion of endowment given to the Santal player, i.e., allocation to Player B divided by 100; (ii) a dummy that equals to 1 if Player B receives any positive amount and 0 otherwise.

Solidarity game. This is a simplified two-person solidarity game between a Bengali/Santal (Player A) and a Santal/Bengali participant (Player B) (Selten & Ockenfels, 1998). So, all pairs will be non-coethnics and have one Bengali and one Santal. Decisions from this game measure how much solidarity Bengalis might show towards Santals in times of crises. This game will be played at Player A's home (while player B remains at her/his home) and both players will remain anonymous. Each player will be given an endowment of 100 Taka prior to the game. The enumerator (our experimenter) will do a coin toss in front of Player A, where 'heads' will destroy the entire endowment of Player B and 'tails' will keep Player B's endowment unharmed. Player A will be asked prior to the coin toss that, conditional on this shock, how much the participant is willing to give to Player B. Then Player A will put this money in an envelope and hand it to the RA, and await the coin toss. If 'heads' then the enumerator keeps the envelope and hands it over to Player B at their homes at a later time. But if 'tails' then Player A gets back the envelope. Therefore, the amount agreed to give to Santals (Bengalis) will measure Benaglis' (Santals') solidarity towards Santals (Bengalis) who lost her/his wealth due to a risk shock. Before the coin toss, Player A is asked "If I play this game with Player B, how much do you think they will be willing to give to you?".

We will code Player A's transfers in two ways: (i) the portion of endowment given to Player B, i.e., amount given to Player B divided by 100; (ii) a dummy that equals to 1 if Player A gives any positive amount to Players B to show solidarity and 0 if no money is offered. We will code Player A's expectation about Player B's transfer in two ways: (i) the portion of endowment that Player B is expected to transfer to Player A; (ii) a dummy that equals to 1 if Player B is expected to give any positive amount to Players A and 0 otherwise.

To eliminate wealth effect, we will randomly pick one game for the payment. We will also randomize the order of games. At endline, the lab-in-the-field experiment will be incorporated with the survey, where survey measured outcomes will be asked in one block and the incentivized experiments in another block. We will randomize the two blocks.

3.2.4 The field experiment and outcomes

Figure 1: Paper bags (*thongga*) making

Most of the participants in this experiment will be agricultural farmers. During the lean season (the time between planting and harvesting crops), there is little agricultural work and most agricultural farmers work as casual laborers until the harvesting season. We will carry out our field experiment during a lean season (February-March)—a period when farmers are available for casual labor and households are financially strained; thus, casual work opportunities and earnings are very attractive to these people.

We have designed a field experiment that allows a Bengali and a Santal (a non-coethnic pair) to work together in a small-scale manufacturing job. The job, which is a two-person task, will allow us to measure productivity and earnings of both participants. The job will involve making paper packets (locally known as *thongga*) for a local wholesaler that sells these packets in bulk to grocery shops and vendors. *Thongga*s are widely used all over the country and many low-income people make *thongga* for living. See Figure 1, where Bangladeshi women are making *thongga* from old newspapers.

The job will be a 2.5-hour casual work opportunity for the participants (2 hour for actual task and 30 minutes for instructions and breaks). Making *thongga* requires at

least two people, each responsible for one of the two main steps: the first step, performed by the first person (or *Preparer*), is cutting old newspapers and folding them into the required shape and size (thus, making sure the height and width are same for all packets); and the second step, performed by the second person (or *Finisher*), is checking the shape and size and then carefully gluing the side and the base of each packet, and letting it dry. If one person tries to do both steps then chances of sticking glue to and wasting newspapers are high.

At the beginning of the task, the wholesaler will train the participants to demonstrate the way paper packets should be made and his requirements on sizes and shapes. We will pay a flat rate of 50 Taka per worker for participating and completing the training. In this job, the productivity is the number of *thongga* produced during the work day that meets the standards of the wholesaler. Thus, faulty, odd-shaped, or torn/sticky packets will not be considered for payment. We will pay a piece-rate to both workers (2 Taka per packet, equally shared between two), so that participants' productivity directly affects their earnings. At the end, the total earnings will be equally divided between the two.

The task itself will be 2-hour long. During the first hour, one person will be randomly chosen to be *Preparer* and another to be *Finisher*. Then a 15 minutes break. After the break, during the second hour, they will switch roles (*Preparer* will be *Finisher* and *Finisher* will be *Preparer*). RAs will measure *Preparers* and *Finishers* productivity twice, during the break and at the end. This will allow us to measure productivity of both workers in both roles.

We will randomly invite Bengali—main and spillover participants—and Santal respondents that took part in our surveys for the field experiment. From each treatment arm, we will randomly invite 120 treated Bengali and 120 Santal participants; thus, each arm will have 120 non-coethnic pair of workers, and the total number of participants will be 720 casual workers (or 360 pairs).

In this experiment, we are mostly interested in the productivity of the Bengali participants. If Bengalis care about the well-being of Santals then they would show higher productivity in both roles. However, if a Bengali wish to 'harm' the Santal then s/he would be less productive, affecting the earnings of both Santal and her/himself. For example, as a *Preparer*, the Bengali will prepare less papers for the Santal *Finisher*. Similarly, as a *Finisher*, the Bengali will glue less paper bags, bringing down the pair's total productivity. Given the experiment will be conducted during the lean season, low productivity (and hence sacrificing earning) would be a strong signal for discrimination. Though Santals were not treated by our intervention, we will also assess Santal

participants' productivity in the same way.

We will construct the outcomes of Bengalis/Santals in four ways: (i) number of papers prepared in the first step (taking the log of this count variable); (ii) proportion of faulty/rejected papers from the first step; (iii) number of papers glued and dried in the second step (taking the log of this count variable); (iv) proportion of faulty packets from the second step. We will also use the total earning from piece-rate (and its log) as the outcome. We will use equation 1 to estimate these impacts.

3.3 Administrative data

We will use village-level inter-ethnic dispute data based on police complaints (from *Tanore* and *Godagari* police stations that cover our 120 villages) and complaints made to the local village councilors (from village councilors' offices, where each village has one elected councilor). We will use both village-level count variables as outcomes in equation 1 (where vector X will include village-level averages of characteristics and SE will not be clustered). Both data on complaints will also allow us to disaggregate complaints made by Bengalis from complaints made by Santals. We will use these count variables as outcomes.

Finally, instead of count variables, we will generate indicators for each (=1 for at least one complaint and =0 if no complaint made) and estimate treatment effects on these indicators.

3.3.1 Potential empathy channels using baseline data

We hypothesize that the documentary film will induce various emotions among the ethnic majority viewers, which as a result would curb their discriminatory opinions and preferences towards ethnic minority Santals. We will also explore potential informational channels, later discussed in section 3.3.3. Note that, by design, the role model (preferences) channel has been muted as the documentary film does not highlight or mention any ethnic majority role models.

To measure individual empathy, we use the Interpersonal Reactivity Index by Davis (1983). We measure three dimensions of empathy to understand which channels can get activated from the documentary film: perspective-taking or PT (the tendency to spontaneously adopt the psychological view of others in everyday life), empathic concern or EC (the tendency to experience feelings of sympathy or compassion for unfortunate others), and personal distress or PD (the tendency to experience distress or discomfort in response to extreme distress in others). We ask the following:

"The following statements inquire about your thoughts and feelings in a variety of situations. For each item, indicate how well it describes you by choosing between 0 and 4, where 0 means *does not describe me at all* and 4 means *describes very well*. Thus, the higher the number, the more it describes your thoughts and feelings. Answer as honestly as you can.":

1. I often have tender, concerned feelings for people less fortunate than me. (EC)
2. I sometimes find it difficult to see things from the "other guy's" point of view. (PT) (-)
3. Sometimes I don't feel very sorry for other people when they are having problems. (EC) (-)
4. In emergency situations, I feel apprehensive and ill-at-ease. (PD)
5. I try to look at everybody's side of a disagreement before I make a decision. (PT)
6. When I see someone being taken advantage of, I feel kind of protective towards them. (EC)
7. I sometimes feel helpless when I am in the middle of a very emotional situation. (PD)
8. I sometimes try to understand my friends better by imagining how things look from their perspective. (PT)
9. When I see someone get hurt, I tend to remain calm. (PD) (-)
10. Other people's misfortunes do not usually disturb me a great deal. (EC) (-)
11. If I'm sure I'm right about something, I don't waste much time listening to other people's arguments. (PT) (-)
12. Being in a tense emotional situation scares me. (PD)
13. When I see someone being treated unfairly, I sometimes don't feel very much pity for them. (EC) (-)
14. I am usually pretty effective in dealing with emergencies. (PD) (-)
15. I am often quite touched by things that I see happen. (EC)
16. I believe that there are two sides to every question and try to look at them both. (PT)
17. I would describe myself as a pretty soft-hearted person. (EC)
18. I tend to lose control during emergencies. (PD)
19. When I'm upset at someone, I usually try to "put myself in his shoes" for a while. (PT)
20. When I see someone who badly needs help in an emergency, I go to pieces. (PD)
21. Before criticizing somebody, I try to imagine how I would feel if I were in their place. (PT)

Statements with (-) requires reverse scoring. There are 7 questions on each of the three dimensions, so the score for each would be between 0 and 28. Thus, higher aggregate score corresponds to someone being more empathetic. Using the continuous score of each dimension measured at baseline, we would carry out a heterogeneity analysis using equation 5 to check which empathetic types benefited the most. We will also code it as dummy: =1 if score is higher than the median value and 0 otherwise. We will also aggregate all dimensions into a single empathy index (and a dummy =1 if score if higher than median value and 0 otherwise) and carry out a heterogeneity analysis using equation 5.

3.3.2 Potential emotional channels using immediate endline data

To measure emotions at play right after watching the documentary, we will ask: "Think carefully about the film you just watched. Now answer the following questions as accurately as possible: On a scale of 0-10, where 0 means *not at all* and 10 means *very...*"

1. How happy do you feel?
2. How sad do you feel?
3. How much do you want to help others?
4. How angry do you feel?
5. How disgusted do you feel?

See section 3.2.2 on how we will construct each index. We will use equation 2 to estimate the impact of the intervention on this intermediate outcome.

We will also measure participants emotions more objectively. During the screening, we will also take photos of each participant's face (with consent, but candid photos) at a random time during the screening. Using the Emotimeter software, that uses machine learning to detect emotions in portrait photos, we will analyse anger, sadness, disgust, happiness, and other emotions of each participant. This software gives a score between 0% and 100% to each emotion category, where a higher % corresponds to higher triggering of that particular emotion.

We will directly use this percentage as the outcome variable. As we expect the documentary to trigger sadness among participants, we will also construct a dummy outcome that equals to 1 if sadness is the most triggered emotion out of the four and 0 otherwise. We will use equation 2 to estimate the impact of the intervention on these intermediate outcomes.

3.3.3 Statistical versus taste-based (information channel)

At baseline, we collect responses of Bengalis to the following statements (each answered on a 0-10 scale, where 10 means "completely agree"):

1. Santals are often unclean/unhygienic
2. Santals would make very good doctors *(reverse scoring required)*
3. I have not met or known any Santals who have established themselves or made a mark
4. Santals do not make very good school teachers.
5. Santals do not continue beyond schools
6. Santals should continue working in the agricultural sector

Using these questions, we will create a "Stereotype Index" by first aggregating all 6 responses and divide it by 60, so that the index is normalized between 0 and 1. Here a higher score will correspond to someone having a stronger negative stereotypes about Santals. We will examine heterogeneity in treatment effects by this index (using equation 5). If we find treatment effects to vary by this index (i.e., someone with strong negative stereotypes at baseline discriminated least at endline), then that would be evidence of discrimination being statistical. This is because, the documentary highlights some success stories of Santals, which should be correcting various negative stereotypes associated with their education and profession. As a robustness check, we will also construct a dummy variable using this index, where it will equal to 1 if the index value is greater than the median index value, and 0 otherwise.

We will also re-do the above heterogeneity analysis using the intercultural competence index developed by Fantini (2000), which we measure at baseline. The simplified version of this index (4-items), used in Siddique & Vlassopoulos (2020), captures a Bengali's awareness of Santals and their culture, and can also be a proxy for how often Bengalis interact with Santals. The four questions are as follows:

1. What is the language spoken by Santals?
2. Do you speak or understand that language?
3. What is their major religion?
4. Name a major Santal festival?

We will produce an index from 0 to 1 for each Bengali participant, by assigning 0.25 to each correct answer, such that the index is normalized between 0 and 1. If we find that participants that score low in this index show most improvements after the intervention, then that would suggest discrimination was statistical.

We will also measure intercultural competence at endline. Thus, using equation 2, we check if the intervention improved the intercultural competence level of Bengalis. If

it significantly improves then the intervention must have provided new information to Bengalis, allowing them to update their beliefs about Santals (suggesting discrimination was statistical).

3.4 Our hypotheses

Our intervention is designed to test the following hypotheses:

1. Treatments *Random* and *Central* will reduce discrimination towards the minority (measured using survey questions, lab-in-the-field experiments, a field experiment, and administrative data) relative to the control group. We will test this using equations 1 and 2.
2. Treatment *Central* will be more effective than Treatment *Random*. We will test this using equations 1 and 2, and then run a Wald-test to compare estimated coefficients on *Random* and *Central* indicators.
3. Treatment *Central* will generate more positive spillovers on neighboring (untreated) Bengalis relative to Treatment *Random*. In other words, untreated Bengalis in Treatment *Central* villages will also show less discriminatory attitudes towards Santals relative to untreated Bengalis in Treatment *Random*. We will test this using equations 1.
4. Reduced discrimination among Bengalis will positively impact the economic, psychological, and subjective well-being of their neighboring Santals. We will test this using equations 1.

4 Empirical analyses

4.1 Treatment effects on discrimination

We will begin by estimating the intent-to-treat (ITT) effects with the following regression specification:

$$Y_{1ijc} = \alpha + \beta_1 Random_{ijc} + \beta_2 Central_{ijc} + \zeta Y_{0ijc} + \Gamma' X_{ijc} + \nu_c + \epsilon_{ijc} \tag{1}$$

where Y_{1ijc} is the outcome of individual i living in village j, in union council c, measured at the endline. $Random_{ijc}$ and $Central_{ijc}$ are indicators for Treatments *Random* (where random Bengalis are targeted) and *Central* (where influencial Bengalis are targeted), respectively. Y_{0ijc} is the baseline analogue of the outcome. X is a vector of baseline controls that will be selected using the post-double-selection LASSO procedure (Belloni et al., 2014). ν is union council fixed effects. We would cluster standard errors at the

village level (our unit of randomization).

As research assistants will visit ethnic majority households in person to screen the documentary film, we expect participation to be high. Moreover, we will call households (since we will collect their mobile numbers at baseline) to fix a day and time for the screening visit, which should reduce non-participation due to unavailability and increase take-up. In case of low compliance/participation to the screening, we will estimate the local average treatment effect (LATE) by using the assignment to treatments as instruments, such that $Random_{ijc}$ and $Central_{ijc}$ only affects Y_{1ijc} through the endogenous participation dummies $D1_{ijc}$ and $D2_{ijc}$ (=1 if participated in the screening). Therefore, the second stage in two stage least squares estimation would be:

$$Y_{1ijc} = \delta + \pi_1 D^R_{ijc} + \pi_2 D^C_{ijc} + \lambda Y_{0ijc} + \Theta' X_{ijc} + \nu_c + \sigma_{ijc} \tag{2}$$

where $Random_{ijc}$ and $Central_{ijc}$ are excluded instruments for indicators D^R_{ijc} and D^C_{ijc} of compliance. Note that LATE equals treatment-on-treated (TOT) effects because there are no 'always-takers' of ethnic minority documentary film in the control arm. We will use equation 2 to estimate treatment effects on the ethnic majority only, as they are the one that watch the documentary film.

We will estimate spillover effects on the neighboring (untreated) Bengalis and Santals using equation 1, where Y_{ij} will be the outcomes of untreated Bengalis/Santals within treatment arms.

4.2 Correlation between incentivized and survey measures of preference

We will show the correlation between altruism and solidarity measured via an incentivized game and the survey by running the following equation:

$$Incentivized_{ijc} = \alpha + \beta Survey_{ijc} + \Gamma' X_{ijc} + \nu_c + \epsilon_{ijc} \tag{3}$$

where $Incentivized$ is the proportion of endowment transferred in the incentivized game, and $Survey$ is the survey analogue of altruism/solidarity. In addition, we would also report the pairwise correlations between the two.

4.3 If spillovers in *Central* treatment, why?

Because we have both network central (7 per village) and randomly selected participants (7 per village) in the *Central* arm, we collect data on how connected each

Bengali spillover respondent (7 per village) is to each network central and randomly selected participant in this treatment arm. Specifically, we ask: "How often did you visit .....'s house or s/he visited your house last month?", where ..... is the name of network central/randomly selected participant. Then we ask: "How long does it take you (in minutes) to walk to .....'s house?".[3] We ask these questions 14 times to each Bengali spillover respondent, one for each participant in the *Central* treatment arm. Using these responses, we will create four averages for each Bengali spillover respondents: average visits to network-central, average visits to random, average distance to network-central, and average distance to random. We will create a bar chart to show the raw difference in visits and distances.

4.4 Inference and multiple hypotheses testing

Since we will test many hypotheses, we will correct p-values using two methods: (i) List-Shaikh-Vayalinkal (LSV) and (ii) Westfall-Young (WY) adjustments (List et al., 2021; Westfall & Young, 1993). We consider all survey outcomes as one family of hypotheses and all objective outcomes (lab-in-the-field, field experiments, administrative data) as another family of hypotheses. For both (i)-(ii), we will use 1,000 bootstrap resampling. We will also compute p-values by using randomized-based inference (RI) with randomization permuted at the village level (Young, 2019, with 1,000 replications). In the main tables, we will report both p-values. However, if RI p-values and conventional p-values appear nearly identical, then we will only report the FWER p-values in the main tables.

4.5 Experimenter demand effects

In our intervention, participants in all three treatment arms receive a treatment: the main two arms receive information through the documentary we produced and the control arm receives a placebo documentary on flowers. Thus, experimenter demand effects in survey responses should be constant across all arms. However, the closeness between the Santal film contents and survey questions might still induce some biases among respondents, especially among those that have a higher tendency to give socially desirable responses to survey questions.

To address social desirability bias (SDB) concerns pertaining to self-reported outcomes, we will closely follow Dhar et al. (2022). Using this 13-point SDB scale (each item is answered as agree/disagree, where socially desirable answers are coded as 1 and 0

[3] We wanted to record GPS coordinates of each respondent during the survey but we could not do it because of poor 3G internet coverage in these villages.

otherwise), we will aggregate all 13 responses such that the value range is between 0 and 13, where 13 means the respondent is most likely to give socially desirable responses. We will also create a dummy for high SDB, which will equal to 1 if the SDB score is above the median value and 0 otherwise. We will carry out a heterogeneity analysis by this SDB scale/dummy on survey outcomes using equation 5. If the coefficients on the interaction term become statistically significant then that would be evidence for social desirability bias influencing results. However, if $\hat{\beta}_1$ and $\hat{\beta}_2$ from equation 5 remains statistically significant regardless of the significance of interactions, then that would imply SDB are not driving the entire effects found. Results using both variables will be reported in the appendix. We will ask the following questions: "Please answer as accurately as possible if the following characteristics describe you or not. Please answer as either *agree* or *disagree*":

1. It is sometimes hard for me to go on with my work if I am not encouraged (Disagree)
2. I sometimes feel resentful when I don't get my way (Disagree)
3. On a few occasions, I have given up doing something because I thought too little of my ability (Disagree)
4. There have been times when I felt like rebelling against people in authority even though I knew they were right (Disagree)
5. No matter who I'm talking to, I'm always a good listener (Agree)
6. There have been occasions when I took advantage of someone (Disagree)
7. I'm always willing to admit it when I make a mistake (Agree)
8. I sometimes try to get even rather than forgive and forget (Disagree)
9. I am always courteous, even to people who are disagreeable (Agree)
10. I have never been irked when people expressed ideas very different from my own (Agree)
11. There have times when I was quite jealous of the good fortune of others (Disagree)
12. I am sometimes irritated by people who ask favors of me (Disagree)
13. I have deliberately said something that hurt someone's feelings (Disagree)

The socially desirable responses are given in parentheses. Therefore, for each question, a respondent gets a point if their response matches with the socially desirable response and 0 otherwise.

To address experimenter demand effect concerns related to subjective outcomes, we will closely follow Chopra et al. (2022) and ask participants at the end of endline survey the following open-ended questions: "If you had to guess, what would you say was the purpose of this study? You will only get one guess and if you guess it correctly, you will get 10 Taka." Enumerators will be given eight common options to tick from:

(1) how often Bengalis and Santals visit or talk to each other, (2) understanding Santals' (Bengalis') lives in the villages, (3) whether documentary film improves Bengalis' attitudes toward Santals, (4) how to donate more money to Santals (Bengalis), (5) how poor people are in these villages, (6) understanding how much financial help Santals (Bengalis) need, (7) how good are Santals (Bengalis) as neighbors, (8) other reasons. Answering option (3) would suggest participants could correctly guess the purpose of this study.

Using this count variable (values 1-8), we will create a distribution plot to show which answers were more common. We will create a dummy variable where it will be =1 if answered (3) and =0 if answered otherwise. We will recreate the table with subjective outcomes but will report the treatment effects separately on those that are in the '=1' category and '=0' category. If treatment effects persist even among the '=0' category, then that would suggest experimenter demand effects were not driving those results.

4.6 Attrition at endline

We will be addressing attrition bias concerns only if the attrition rate is found to be differential at the endline (e.g., significantly more attrition in the control arm relative to treatment arms). In the case of differential attrition, we will use Inverse Probability Weighting and Lee (2009) bounds to address attrition bias concerns. We will check if attrition is differential at the endline by running the following OLS regression:

$$Attrited_{ijc} = \alpha + \beta_1 Random_{ijc} + \beta_2 Central_{ijc} + \nu_c + \epsilon_{ijc} \tag{4}$$

where *Attrited* equals to 1 if the person is missing at the endline and 0 if present. Rejecting $\beta_k = 0$ would imply differential attrition.

5 Heterogeneity analysis

We will examine heterogeneity in treatment effects using both machine learning (Chernozhukov et al., 2020, such as Random Forest) and the more traditional method with interactions. As Chernozhukov et al. (2020) takes all baseline covariates into account, we will only report heterogeneity results by the following baseline characteristics:

- Monthly household income
- Education of participants
- Age of participant
- Gender of participant

- Baseline discriminatory opinions index

We will only report the classification analysis (CLAN) results in the main paper, which compares the average characteristics of those most affected by the intervention (top 25%) versus those least affected (bottom 25%).

To examine heterogeneity using interactions, we will estimate the following equation:

$$\begin{aligned} Y_{1ijc} = \alpha + \beta_1 Random_{ijc} + \beta_2 Central_{ijc} + \beta_3 H_{ijc} + \beta_4 (Random \times H)_{ijc} \\ + \beta_5 (Central \times H)_{ijc} + \zeta Y_{0ijc} + \Gamma' X_{ijc} + \nu_c + \epsilon_{ijc} \end{aligned} \tag{5}$$

where H_{ijc} is alternatively (i) monthly household income, (ii) education, (iii) age, (iv) gender, (v) discriminatory opinions index, all measured at baseline. Statistically significant $\hat{\beta}_4$ and/or $\hat{\beta}_5$ would mean treatment effects differ by some baseline characteristics.

References

Bandura, A. (1977). *Social learning theory*, volume 1. Englewood cliffs Prentice Hall.

Banerjee, A., Breza, E., Chandrasekhar, A. G., & Golub, B. (2018). When less is more: Experimental evidence on information delivery during india's demonetization. *NBER working paper no. 24679*.

Banerjee, A., Chandrasekhar, A. G., Duflo, E., & Jackson, M. O. (2019). Using gossips to spread information: Theory and evidence from two randomized controlled trials. *Review of Economic Studies*, *86*(6), 2453–2490.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies*, *81*(2), 608–650.

Chernozhukov, V., Demirer, M., Duflo, E., & Fernandez-Val, I. (2020). Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in india. *NBER working paper no. 24678*.

Chopra, F., Haaland, I., & Roth, C. (2022). The demand for news: Accuracy concerns versus belief confirmation motives. *CEPR Discussion Paper No. DP17169*.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*(1), 113.

Dhar, D., Jain, T., & Jayachandran, S. (2022). Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in india. *American Economic Review*, *112*(3), 899–927.

Falk, A., Becker, A., Dohmen, T., Enke, B., Huffman, D., & Sunde, U. (2018). Global evidence on economic preferences. *Quarterly Journal of Economics*, *133*(4), 1645–1692.

Fantini, A. E. (2000). A central concern: Developing intercultural competence. *SIT Occasional Paper Series*, *1*, 25–42.

Grady, C., Iannantuoni, A., & Winters, M. S. (2021). Influencing the means but not the ends: The role of entertainment-education interventions in development. *World Development*, *138*, 105200.

Kroenke, K., Spitzer, R. L., Williams, J. B., & Löwe, B. (2009). An ultra-brief screening scale for anxiety and depression: the phq–4. *Psychosomatics*, *50*(6), 613–621.

La Ferrara, E. (2016). Mass media and social change: Can we use television to fight poverty? *Journal of the European Economic Association*, *14*(4), 791–827.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, *76*(3), 1071–1102.

List, J. A., Shaikh, A. M., & Vayalinkal, A. (2021). Multiple testing with covariate adjustment in experimental economics. *Working paper*.

Roy, R. D. (2012). Country technical notes on indigenous peoples' issues: People's republic of bangladesh. *IFAD Country Report*.

Selten, R. & Ockenfels, A. (1998). An experimental solidarity game. *Journal of Economic Behavior & Organization*, *34*(4), 517–539.

Siddique, A. & Vlassopoulos, M. (2020). Competitive preferences and ethnicity: Experimental evidence from bangladesh. *The Economic Journal*, *130*(627), 793–821.

Siddique, A., Vlassopoulos, M., & Zenou, Y. (2023). Market competition and discrimination. *Forthcoming at European Economic Reveiew*.

Westfall, P. H. & Young, S. S. (1993). *Resampling-based multiple testing: Examples and methods for p-value adjustment*, volume 279. John Wiley & Sons.

Young, A. (2019). Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. *Quarterly Journal of Economics*, *134*(2), 557–598.