Identity Uncertainty*

Anujit Chakraborty† Arkadev Ghosh‡ Matt Lowe§ Gareth Nellis¶

Abstract

In contexts with persistent social inequality, members of stigmatized groups may choose or feel compelled to withhold identity markers in order to avoid discrimination. Whether concealing group identity improves or worsens social and economic relations remains unclear. We study this in the context of caste in north India, where we form pairs of high- and low-caste men for data entry tasks. We randomly assigned pairs to interact with knowledge of full names, including caste-distinctive surnames (*Revealed*), first names only (*First Name Only*), or to a condition with instructions to keep caste identity hidden (*Hidden*). As expected, the *First Name Only* and *Hidden* conditions have strong negative effects on whether a participant correctly guesses their partner's caste at endline, as well as confidence in their guess. Surprisingly, we find that these identity-concealment conditions reduce trust and weaken social relations, with mixed effects on productivity. Our findings are contrary to the conventional wisdom that group boundaries hinder social relations — we find that disclosure of group identity brings partners together. Additional analyses suggest that the key mechanism is uncertainty. Relations are strongest in same-caste pairs with identity revealed, intermediate in mixed-caste pairs with identity revealed, and worst in mixed-caste pairs with identity shrouded.

*This project was pre-registered in the AEA RCT Registry with ID AEARCTR-0015151. We thank Suryashis Ghosh, Ujjwal Kumar and Ved Sharma for excellent research assistance.

†University of California, Davis. Email: chakraborty@ucdavis.edu.

‡Duke University. Email: arkadev.ghosh@duke.edu.

§University of British Columbia. Email: matt.lowe@ubc.ca.

¶University of Pennsylvania. Email: gnellis@upenn.sas.edu.

1 Introduction

A sizable body of research highlights the importance of social-group membership for development outcomes. Identities such as race, religion, and ethnicity have been shown to underpin discrimination, levels of public goods provision, workplace productivity, and conflict (Bertrand and Mullainathan, 2004; Alesina and La Ferrara, 2005; Horowitz, 2000; Hjort, 2014). In much of this literature, identity is treated as known and fixed in social interactions: individuals' group memberships are fully observed and readily apparent to all parties.

Yet in many real-world scenarios, identity is only imperfectly legible. People must rely on imperfect cues—such as names, language, dress, or institutional environment—to infer the social group affiliation of those around them. As a result, *identity uncertainty* can arise because identity markers are absent, because institutions suppress them (Abdelgadir and Fouka, 2020; Goldin and Rouse, 2000), or because individuals choose not to disclose them (Dahis et al., 2019; Pachankis et al., 2018). This uncertainty is often two-sided: individuals may be unsure not only of a partner's identity, but also of whether their own identity has been inferred and how it will be received if revealed. Disclosure can invite reciprocity and smoother coordination, but it can also expose individuals to exclusion, hostility, or retaliation. This raises a distinct question: how does the legibility of identity itself shape trust, coordination, collaboration, and the decision to reveal?

We study these questions in the context of caste in India. Caste is one of the most enduring and pervasive systems of social stratification worldwide, informing occupation, marriage, social networks, trade, residence, and vote choice (Munshi, 2019; Oh, 2023; Cassan et al., 2021; Shukla, 2025; Banerjee et al., 2013; Lowe, 2021; Anderson, 2011; Asher et al., 2026; Chandra, 2007). At the same time, caste is not always fully transparent in everyday interaction. It can sometimes be inferred from surnames, speech, appearance, or demeanor, but these signals can be unevenly informative.[1] India therefore provides a valuable setting in which identity is both highly consequential and only partially legible, allowing us to study how variation in identity visibility affects interpersonal relations and economic collaboration.

To study the effects of identity legibility, we ran an experiment in Uttar Pradesh in which young men from different caste backgrounds were randomly paired to work together for a full day on paid data-entry and problem-solving tasks. Cross-caste pairs were assigned to one of three information environments. In the *Revealed* condition, enumerators publicly announced both partners' full names, including caste-distinctive surnames. In the *First Name Only* condition, participants were introduced only by first names, leaving caste potentially inferable over the time

[1] Qualitative evidence shows that many individuals, especially those from marginalized ("lower") caste groups, seek to obscure their caste identities in day-to-day settings (Valmiki, 2008; Satyanarayanan and Lee, 2023). Such efforts range from the use of socially neutral surnames such as Singh or Kaur in Sikh communities, to "Sanskritization," whereby members of disadvantaged groups adopt the ritual practices and customs of higher castes (Srinivas, 1966).

of interaction but uncertain. In the *Hidden* condition, participants were introduced by first names and instructed not to communicate information that could reveal caste, thereby further reducing the legibility of identity. These subtle differences in the informational environment allow us to compare interactions under full identity revelation, endogenously partial inferability, and complete concealment, without making caste itself the explicit focus of the study.

In particular, the *First name only* condition allows us to study identity-hiding behaviors when participants are free to optimize, choosing whether or not to reveal their caste, and whether or not to seek caste information of their partner. In contrast, the *Hidden* condition mimics a more paternalistic workplace policy that aims to prevent unraveling, by instructing all castes to keep caste identity private. Anti-caste political movements have pushed to eliminate surnames altogether, as in parts of Tamil Nadu, in order to weaken caste markers at their source. We implement our treatment conditions through subtle variations in wording, without participants' awareness of the study purpose. As a result, only 4% of participants guess that the study purpose involves caste and social relations, with this number similar across treatment groups.

The design allows us to observe and measure both the legibility of caste identity and its downstream consequences. Pairs worked together for roughly four hours across five collaborative tasks, with detailed measures of productivity, role allocation, and accuracy recorded throughout. At the end of the session, each participant makes guesses about the attributes of his partner, including his caste, allowing us to characterize first stage effects on de facto concealment. Participants also complete a range of incentivized and self-reported measures of social relations, trust, and well-being at work.

To guide our empirical analysis, we develop a simple model of strategic identity disclosure for the *First name only* condition. Individuals from two caste groups interact in pairs and decide whether to reveal or conceal their caste identity. We make the following assumptions: one-sided revelation results in an awkward interaction. Mutual revelation can improve interaction quality by providing a shared social context that facilitates communication and coordination. However, for members of disadvantaged castes revelation may also expose them to discrimination, generating heterogeneous payoffs from disclosure. As a result, lower-caste individuals differ in their willingness to reveal their identity.

The model admits two equilibria: one in which both parties conceal their identity, because they are uncertain if the other side will reveal, and they want to avoid the awkwardness of one-sided revelation. And another in which high-caste individuals reveal while lower-caste individuals reveal only when the expected interaction payoff is sufficiently high. Our framework generates three key predictions. First, policies that suppress identity information can lead to inefficient concealment even when mutual revelation would benefit both parties. Second, allowing voluntary

disclosure leads to selective revelation by individuals who expect positive interactions. Third, environments that generate uncertainty about identity may reduce interaction quality by making it harder for individuals to coordinate their behavior.

We report four sets of findings. First, we find meaningful first stage effects of our treatments on de facto concealment, with no evidence of unraveling. In the *Revealed* condition, 97% of participants correctly guess their partner's caste category (from five options: General, Scheduled Caste [SC], Scheduled Tribe [ST], Other Backward Caste [OBC], Muslim). The correct guess rate is 14 and 21 percentage points lower in the *First name only* and *Hidden* conditions, respectively. First stage effects are even larger on correctly guessing the partner's specific caste (20 and 34 pp), reductions in confidence in the caste category guess made (13 and 22 pp), and reductions in the perceived likelihood that the partner knows the respondent's caste (15 and 35 pp). For all of these first stage outcomes, we reject equality of *Revealed* with *First Name Only*, and of *First Name Only* with *Hidden*, at the 1% level. These results show that a person's last name is not fully substituted by other cues, like appearance and language, when inferring a partner's caste in our setting. They also show that our two identity-hiding treatments do more than conceal group identity: they increase uncertainty in perceptions about the partner's group identity, and also in meta-perceptions. Participants know their partner less, and they believe their partner to know them less too.

While we find de facto concealment of caste overall, the extent of concealment varies by caste background. Low-caste individuals guess their partners' caste with similar accuracy in the *First Name Only* and *Hidden* conditions. But high-caste individuals perform much worse in *Hidden*: their accuracy rate under *Hidden* is about 15 pp lower than in *First Name Only*, making for a 29-point gap ($p < 0.01$) between *Revealed* and *Hidden* when high castes guess low-caste partners. This suggests that low-caste identities are either easier to conceal, or are deliberately concealed more often (when individuals are primed to conceal), whereas high-caste identities are more readily inferred from cues such as speech, appearance, or demeanor, or are deliberately revealed by high-caste members more often. For the most part, second-order beliefs—beliefs about whether my partner can correctly identify my caste—move in parallel with the effects on first-order beliefs—actual knowledge of the partner's caste. The one exception is that while low-castes make similarly accurate guesses in *First Name Only* as in *Hidden*, their high-caste partners think that their partners' guesses will be considerably less accurate in *Hidden*. It follows that high-caste individuals over-report or overestimate their ability to hide their caste when instructed to do so. The greater sophistication of low-caste individuals here may be a result of hiding behaviors being more payoff-relevant, and thus strategically important, for marginalized groups.

Our second set of findings concern partner relations at work. We find that the non-revelation of last names reduces the quality of interpersonal interactions relative to when they are revealed. Subjects report feeling less close

to their partners (0.5σ, $p < 0.01$), being less able to be their true selves (0.35σ, $p < 0.01$), and are less willing to meet their partner in the future (0.44σ, $p < 0.01$). In measures of more direct frictions—such as partner rudeness, tension, and bossiness—the effects are smaller and are not statistically significant.

Our third set of findings show that the reduced quality of interactions during the work day carries through to downstream outcomes. Subjects send lower amounts to their partner in an incentivized trust game (0.29σ, $p < 0.01$) and expect lower reciprocity (0.23σ, $p < 0.05$), a belief that is borne out in the data relative to the identity-revelation condition (subjects in the *Hidden* condition return roughly 0.25σ less to their partners than under *Revealed*, $p = 0.01$). Consistent with this, when last names are hidden, subjects exhibit a higher reservation wage to work with their partner than with a stranger.

Across all results, we observe a consistent pattern: outcomes in the *First Name Only* condition lie between those in the *Hidden* and *Revealed* conditions. This is consistent with respondents having some scope to learn their partner's identity, which our first-stage results suggest they indeed ended up doing (compared to the *Hidden* condition).

With discrimination, one might expect that suppressing caste-identifying names would facilitate smoother inter-group interactions by reducing the salience of caste. Yet, we find the opposite. What drives these negative effects of concealment? We find evidence for two related channels. First, caste identity is so deeply ingrained in Indian society and culture that it anchors behavior in social interactions and provides shared familiarity and context for conversations. When caste identity is concealed, the interaction environment is stripped of this social context, and of the cues that individuals typically draw on to build ties. Second, and closely related, concealment generates *uncertainty* — not just about a partner's caste, but likely also about how to behave, what topics are safe to raise, and how to interpret the other person's signals. Together, these forces reduce the quality of social interactions leading to more negative effects on downstream outcomes.

Consistent with this, we find that in the *Hidden* condition, subjects not only bring up caste less in their conversations, but also discuss a range of other characteristics less frequently, including education, income, religion, age, and marital status. This is likely because many of these characteristics may indirectly reveal caste, or open up conversations that do so. As a result, successfully concealing caste may require subjects to avoid discussing broader demographic characteristics, which in turn has negative consequences for tie-building. Crucially, our evidence suggests that uncertainty is the primary driver, rather than the reduction in shared social context alone. In the *First Name Only* condition subjects discuss these characteristics at broadly similar rates to *Revealed*, yet uncertainty around their partner's identity remains, and social outcomes are depressed (of course less so than in *Hidden*). This suggests it is not the absence of conversation topics per se, but the uncertainty about who one's partner is, that drives the negative effects: once we control for subjects' confidence in their guess about their partner's caste, the negative treatment

effects on trust, identity fusion, and authenticity largely disappear, while confidence itself is a strong and significant predictor of better social outcomes across all measures.

Once again, we find that this negative effect is much less prominent in the *First Name Only* condition. In fact, subjects ask about their partner's caste significantly more in this condition than in *Hidden*, and even slightly more than in *Revealed* (where caste is, of course, already known). For other characteristics, *First Name Only* subjects either discuss them slightly less or at similar rates relative to *Revealed*. Taken together, this revealed preference suggests that learning identity has benefits in this context. Indeed, correlating whether subjects asked their partner about their identity — or were themselves asked — shows more positive effects on trust within the pair when caste was known.

To our knowledge, we are the first to document a cost of reduced identity legibility in the form of weaker social ties. These costs may seem difficult to reconcile with the prevalence of identity concealment in India, but they need not be. Concealment may reflect a rational tradeoff in which members of marginalized groups forego some of the gains from trust and social connection in order to insure themselves against severe discrimination. Consistent with this view, we find that the negative effects of *Hidden* are concentrated in the upper part of the outcome distribution, while the effects are muted and not statistically significant in the left tail. This pattern suggests that concealment is especially costly when interactions would otherwise have gone well, while offering some indication that it may be less costly, and perhaps even protective, in cases where interactions are more likely to go poorly. Although our design (in principle) is capable of detecting such protective benefits on the margins we study, we do not find direct evidence of positive effects in our setting. Still, the distributional pattern is consistent with the idea that concealment is most valuable in environments that are difficult to study experimentally for ethical reasons. Our results, however, identify a clear relational cost of reduced identity legibility in ordinary workplace interactions.

We find similar negative effects for the *First Name Only* condition at more positive values of the outcome distribution, though these effects are smaller than in the *Hidden* condition. This is intuitive: when subjects anticipate positive returns from identity revelation and have the option to reveal (unlike in *Hidden*), they choose to do so. Indeed, at very high outcome levels, the effects in *First Name Only* are very close to full revelation, suggesting that when outcomes are especially positive, subjects in this condition reveal at very high rates, whereas in the middle of the distribution uncertainty about the gains from revealing makes this harder.

Lastly, our productivity results are less clear-cut and somewhat noisy: the *First Name Only* condition performs marginally better than either *Revealed* or *Hidden*. One plausible explanation is that while *Hidden* undermines social ties, *Revealed* reinforces misperceptions about group productivity (Ghosh, 2025). In this context, the option value of choosing whether to reveal one's identity (based on initial interactions and match characteristics) may thus be

optimal.

Related literature A small literature in economics has begun to grapple with the fact that social identities are not innate traits, but rather social categories that are actively constructed and reconstructed by both the individual and their audience (Akerlof and Kranton, 2000; Rose, 2023). Recent work shows that conditional on skin tone, US census workers were more likely to categorize wealthier and literate individuals as White or Mulatto (Adukia et al., 2025); more-educated Roma people are more likely to "pass" for non-Roma in Romanian census data (Mitrut et al., 2025); and many Indians manipulated their caste identity to secure eligibility for land acquisition (Cassan, 2015). Each of these papers speak to the question of which types of people are able to hide their identity, passing for members of other groups. Taking this a step further, Dahis et al. (2026) explore the *consequences* of passing, showing that for Black men in the US, passing is associated with large income gains, though at the expense of family ties. We bring an experimental approach to this mostly historical and descriptive literature. We contribute by (i) estimating *causal* effects of identity revelation, (ii) measuring both first- and second-order beliefs to characterize the effectiveness of passing, and (iii) eliciting rich measures of social relations in mixed groups to explore the effects of identity revelation on group outcomes, rather than only individual outcomes. Furthermore, our experiment design links more directly to policy design, in that our treatments correspond to concrete workplace policies that could be adopted with relative ease.

Second, we contribute to work that uses experimental methods to measure discrimination (Bertrand and Duflo, 2017). Consider racial discrimination in the labor market. Audit studies recruit pairs of White and Black "auditors" that are matched on other personal characteristics, and have them apply for jobs (Fix and Struyk, 1993). Correspondence studies instead randomize whether a job application comes from a White-sounding or Black-sounding name (Bertrand and Mullainathan, 2004; Kline et al., 2022). Correspondence studies have the advantage that the randomization is cleaner—in audit studies, it is difficult to match White and Black auditors on all relevant characteristics. Audit studies have the advantage that we can observe effects beyond the initial employer's response, potentially tracing through effects all the way through to job acceptance. We offer a new experiment design that combines advantages of both, and that applies to social identities that are at least partly concealable: by randomizing the revelation of group identity in an unobtrusive way, we are able cleanly identify the effects of perceived group identity on in-person behaviors. And while findings from audit and correspondence studies have motivated blinded hiring processes (Goldin and Rouse, 2000; Behaghel et al., 2015; Boring et al., 2025), our work relates to the more overlooked question of whether some form of blinding should continue to apply after candidates have been hired.[2]

[2] A related literature considers how individuals respond to the prospect of expected discrimination during an application process (Kang et al., 2016; Charness et al., 2020; Lepage et al., 2022; Ruebeck, 2023; Avery et al., 2024; Del Carpio and Fujiwara, 2026; Angeli et al., 2026). For example, Angeli et al. (2026) show that favela residents in Brazil perform worse at interview when they expect the interviewer to

A parallel line of research emphasizes that identity is not only a basis for bias, but also a key input into social interaction and coordination. Economic models of identity argue that social categories shape norms, expectations, and preferences, affecting behavior even when material incentives are held fixed (Akerlof and Kranton, 2010). Social-psychological theories similarly stress that shared identity provides interpretive frameworks that allow individuals to predict one another's behavior and build trust (Tajfel et al., 1971; Pettigrew and Tropp, 2006). This work, and the associated work on intergroup contact, emphasizes interaction under common knowledge of group identity (Mousa, 2018; Lowe, 2021; Ghosh, 2025; Ghosh et al., 2025). But much less is known about how cooperation is affected when identity information is withheld altogether. This is despite the fact that core theories in psychology posit important (and opposing) roles for the salience of group identity in promoting improved intergroup relations (Brewer and Miller, 1984; Hewstone and Brown, 1986; Miller, 2002; Dovidio et al., 2003). We make progress here by randomizing whether partners' group identity is known versus unknown, rather than the more usual randomization of whether a partner belongs to the ingroup or the outgroup. In turn, we show that reducing group identity salience *negatively* affects social relations.

2 Experimental design and data

Setting The experiment was implemented in rural Uttar Pradesh, in the districts of Barabanki and Ayodhya. These districts are characterized by high levels of residential segregation by caste, limited formal-sector employment, and frequent reliance on casual labor markets and local patronage networks. In this setting, caste is a salient social category and is widely inferable from surnames, as well as other markers such as speech patterns, neighborhood of residence, and family occupation. At the same time, a substantial share of interactions relevant for economic life occur in settings where individuals have weak prior familiarity with one another, such as short-term work arrangements, hiring through informal networks, or transactions in nearby towns. This combination makes the setting well suited for studying the feasibility and consequences of identity uncertainty.

The study was conducted in temporary work sites located near the study villages. Sessions were run as structured workdays. Participants arrived individually, were seated apart, and were monitored to limit communication prior to pairing. Pairs then worked in private booths on a sequence of paid tasks designed to emulate routine clerical work under time pressure, with a clear within-pair status hierarchy induced by role assignment (tablet operator versus reader). The interaction was therefore prolonged, goal-oriented, and economically consequential, while remaining sufficiently standardized to allow a clean comparison across treatments. The design also limits spillovers across

know their favela status. In addition, work on stereotype threat finds that the revelation of group identity can harm individual performance. Most close to us, Hoff and Pandey (2006) find that the public revelation of caste led to a decline in task performance—solving mazes—by low-caste subjects, when grouped with high-caste subjects. This effect leads a caste performance gap to emerge.

pairs by restricting cross-booth communication and by instructing participants to leave immediately after endline. We allowed 15-minute breaks between tasks, giving participants time to talk informally, get to know each other, and potentially form social ties.

Recruitment and baseline We recruited subjects in Barabanki and Ayodhya through door-to-door campaigns in villages located near the study sites. Participation required full, informed consent. Eligibility was determined by three criteria. First, respondents had to be men between the ages of 18 and 40. Second, they were required to have a last name drawn from a pre-specified list of distinctive caste surnames, names that a prior scoping survey had shown to be almost exclusively associated with a single caste category (see Online Appendix Section D.1). Enumerators further confirmed that the respondent's reported caste category matched the caste typically linked to their surname. We primarily recruited participants from Scheduled Caste ("Low") or General ("High") caste backgrounds, though we included a small number of OBCs (Other Backwards Classes) for obfuscation purposes, which we explain below. Third, respondents completed short, timed exercises testing basic competencies in typing and document reading; only those who demonstrated at least a minimal level of proficiency were invited to continue.

Individuals that met the three criteria and completed a baseline survey were invited to participate in the study. The baseline survey gathered information on demographics, including caste category and caste, social and political attitudes, and psychological well-being. Eligible participants were informed of the study location and given a tentative participation date. They were also provided with a unique identification code, which they were told to bring to the study location. Enumerators followed up with reminder phone calls on the evening prior to their assigned session.

Flow of the experiment Each iteration of the experiment lasted one full day and ran as follows:

1. *Arrival.* Participants arrived at the study site and checked in with facilitators using their unique identification code. Between a minimum of four and a maximum of 32 individuals could participate on a given day. Upon arrival, participants were seated in widely-spaced chairs and instructed not to interact with one another, with enumerators monitoring compliance.

2. *Random assignment.* Participants were then randomly assigned to pairs, and pairs were randomly allocated to one of the three conditions, explained below.

3. *Professionalism instructions and survey.* Each participant was individually called to complete a short "professionalism" survey, which presented a set of workplace rules and checked comprehension of those rules.

4. *Tasks in pairs.* Participants were escorted to private booths, where they worked in pairs on a sequence of five collaborative tasks lasting approximately four hours.

5. *Endline survey.* After completing the tasks, participants individually responded to an endline survey, facilitated by an enumerator.

6. *Exit.* Immediately upon finishing the endline survey, participants were instructed to leave the study site. Enumerators ensured that participants did not interact across booths before departure.

7. *Incentive payments.* Participants were informed about their team earnings and senior bonuses, with base payments made on the same day. Any delayed bonus payments were disbursed within three days.

Randomization At the start of each daily session, participants provided enumerators with their unique identification codes. These codes were used to retrieve background information on the participant's caste from a pre-loaded database, minimizing the salience of caste during the session itself. A research assistant then entered the codes of attending participants into a randomization platform (an `R` Shiny app), which used the preloaded baseline data to conduct that day's randomization. The initial step of the randomization involved assigning all participants present at that day's session into one of three types of dyads, formed through random pairing:

- *Mixed High–Low pairs (primary sample, $N = 532$ pairs).* Most participants were allocated to high-caste-with-low-caste (General-SC) dyads, since our primary research question concerns how intergroup interactions are affected by identity revelation.

- *Homogeneous High–High or Low–Low pair (obfuscation sample, $N = 49$ pairs).* For each daily session, one caste group (either High or Low) was designated to form that day's homogeneous pair. On even-numbered days (e.g., July 2), two High-caste participants were randomly selected to constitute a homogeneous dyad, while on odd-numbered days (e.g., July 3), two Low-caste participants were selected instead.[3]

- *Mixed High–OBC or Low–OBC pair (obfuscation sample, $N = 18$ pairs).* If at least one OBC participant was present at a given session, exactly one was randomly selected and paired with a participant from the caste opposite to that day's designated homogeneous group. Thus, on High-designated (even-numbered) days, the OBC was paired with a Low-caste participant, and on Low-designated (odd-numbered) days with a High-caste participant. These dyads added variation in caste pairings but, like the homogeneous pairs, were included mainly to obscure the structure of the primary sample.

[3] This dyad type and the next were introduced partway through the experiment to enhance obfuscation.

The inclusion of homogeneous and OBC-containing dyads served an important masking function. If, following the experiment, participants discussed their assignments outside the experimental setting, the presence of these additional pair types helped ensure that the study design did not appear to be focused exclusively on cross–caste High-Low pairings. This reduced the likelihood that the intergroup relations angle of the experiment would be obvious to future participants.

After the randomization was implemented, any unmatched individuals were placed on a waitlist and asked to return on another day at their convenience. In practice, enumerators sought to invite participants in the appropriate proportions each day so as to minimize the need for waitlisting.[4]

Dyads served as the unit of randomization. All cross-caste pairs were independently assigned, with equal probability, to one of three conditions—*Revealed*, *First Name Only*, or *Hidden*. Homogeneous dyads were always assigned to the *Revealed* condition, as were dyads involving an OBC participant (when formed). We did this to aid our goal of obfuscation—the purpose of these dyads was for participants to learn that non-High-Low pairings were possible. To aid in this learning, we always revealed identity in these pairs.

Lastly, we randomly assigned dyads to booths to prevent any systematic association between treatment status and physical work location.

Treatment conditions The central intervention manipulated the legibility of caste identity through the revelation or concealment of participants' names, along with instructions not to reveal caste identity. Each of the conditions was deliberately "light touch." We randomly assigned High-Low dyads to one of three treatment conditions, with equal probability:

- *Revealed ($N = 176$ pairs).* Prior to partners being introduced, participants privately complete a professionalism survey, which lists some neutral workplace guidelines (e.g., not communicating across booths, limiting phone use, coordinating short breaks). Participants were required to affirm their understanding, and comprehension was checked through four follow-up questions. At the start of the task survey, with both partners present in their assigned booth, an enumerator read aloud each participant's full (i.e., first and last) name twice, purportedly to confirm that the correct individuals were in the correct booth. Full names—preloaded by the enumerator—were then also displayed on the tablet screen throughout the tasks and were written on slips of paper placed on the desk. Participants were asked to use these slips to record who served as the senior partner

[4]The randomization tool allowed a maximum of 32 participants to be entered in a single session. For the procedure to execute without error, a minimum number of participants was required, which depended on caste composition each day. In the absence of any OBC participants, at least four individuals had to be entered—two from the day's selected caste to form a homogeneous pair and at least one each from the remaining H and L groups to form a heterogeneous pair. When an OBC participant was included, the minimum rose to six participants, reflecting the additional requirement of one OBC individual and one opposite-caste partner, while still ensuring the formation of at least one heterogeneous H–L pair. Subjects that remained (randomly) unmatched under these constraints were waitlisted for the following day.

for each task. Combined with the fact that we selected participants whose last names were highly reliable caste markers, this collection of interventions meant that caste identity was revealed, and this was *common knowledge* in the *Revealed* condition.

- *First Name Only ($N = 179$ pairs).* Identical to *Revealed*, other than the fact that participants were introduced only by first names at the start of the task survey, and these first names appeared on screen and on the desk slips. Caste identity was neither explicitly revealed through names nor actively suppressed by the study protocol, leaving participants scope to infer or disclose caste through subsequent interaction.

- *Hidden ($N = 177$ pairs).* This condition was identical to *First Name Only*, except that the professionalism survey additionally instructed participants not to disclose their last names or any other information that could reveal their caste. Respect for one another's privacy was given as the rationale for this instruction. Participants answered one additional comprehension question concerning the added instruction. Thus, this condition attempted to remove the legibility of caste identity by discouraging the exchange of clear caste-revealing cues during the interaction.

The exact wording for the the various manipulations are provided in Table 1.

Table 1: Experimental manipulations, by treatment condition.

	Professionalization survey	Tasks survey		
		Enumerator introductory checks	Tablet display, appeared after every task	Record sheet, sitting on desk throughout
Revealed	We wish to create a positive and productive work environment. To contribute to this environment, you must follow these factory guidelines ... You can take breaks when you need to, but do not make them too long. Ideally, coordinate short breaks together with your partner to maintain a balanced workflow and avoid interruptions for your partner. Do not communicate with workers in the other booths at all over the course of the work day. Please limit the use of your mobile phone while you are working.	Before you begin the work tasks, I must confirm that the correct partners are in the booth. Is ***Pramod Gautam*** in this booth? To repeat, is ***Pramod Gautam*** in this booth? [CONFIRM YES] Thank you. Is ***Anup Pandey*** in this booth? To repeat, is ***Anup Pandey*** in this booth? [CONFIRM YES]	Who do you select to be the senior partner for this task (enters information on the tablet)? Remember, the person chosen as senior partner will be paid 50 INR extra for this task. PARTICIPANTS SELECT: ***Pramod Gautam*** OR ***Anup Pandey***	Number of times ***Pramod Gautam*** was senior partner [PARTICIPANTS WRITE NUMBER]; Number of times ***Anup Pandey*** was senior partner [ENTER NUMBER]
First Name Only	« Same as *Revealed* »	Before you begin the work tasks, I must confirm that the correct partners are in the booth. Is ***Pramod*** in this booth? To repeat, is ***Pramod*** in this booth? [CONFIRM YES] Thank you. Is ***Anup*** in this booth? To repeat, is ***Anup*** in this booth? [CONFIRM YES]	Who do you select to be the senior partner for this task (enters information on the tablet)? Remember, the person chosen as senior partner will be paid 50 INR extra for this task. PARTICIPANTS SELECT: ***Pramod*** OR ***Anup***	Number of times ***Pramod*** was senior partner [PARTICIPANTS WRITE NUMBER]; Number of times ***Anup*** was senior partner [ENTER NUMBER]
Hidden	We wish to create a positive and productive work environment. To contribute to this environment, you must follow these factory guidelines ... You can take breaks when you need to, but do not make them too long. Ideally, coordinate short breaks together with your partner to maintain a balanced workflow and avoid interruptions for your partner. Do not communicate with workers in the other booths at all over the course of the work day. ***Do not tell your last name to your work partner, or provide information that could reveal your caste. It is important to respect each other's privacy.*** Please limit the use of your mobile phone while you are working.	« Same as *First Name Only* »	« Same as *First Name Only* »	« Same as *First Name Only* »

Notes: This table presents excerpts from the professionalism and tasks surveys, in which the key experimental manipulations were delivered.

Data-entry tasks and piece-rate payments We instructed pairs to perform a sequence of five paid collaborative tasks. Each task lasted up to 45 minutes, with 15-minute breaks in between, and the entire exercise took roughly four hours to complete. The tasks were structured to reflect a clear status hierarchy: one partner, designated the *senior partner*, operated the computer to input information, while the other, designated the *junior partner*, read aloud data from paper forms contained in labeled envelopes. The tasks varied in content and difficulty but shared a common emphasis on coordination under time pressure. Pairs began with the two simplest tasks, to build familiarity with the technology and procedures. They carry out the two tasks in the following order:

1. *Ordered Data Entry.* Participants received a list of names and corresponding grades. On their screen, the names appeared in the same order as in the paper list. The task was to correctly type each grade next to

the matching name. Because the entries were aligned, the main challenge lay in accuracy and sustained concentration rather than searching. This task was easy.

2. *Unordered Data Entry.* Participants again received a list of names, this time with dates of birth. On the screen, the names were shuffled relative to the paper list. The task was to find each name in the paper list and type the correct year of birth next to it, always in a four-digit format (e.g., 1999). This required more searching and cross-checking, making the task moderately difficult.

Next, pairs carry out three more demanding tasks, in random order, allowing us to purge order effects in the analysis:

3. *Cross-Validation.* Here, the envelope contained two separate lists of names with grades. The screen displayed names from List 1 in the same order. Participants had to check whether the grade in List 1 matched the grade in List 2, and then enter "0" if the name appeared in both lists but the grades differed, "1" if the name and grade matched across both lists, and "2" if the name from List 1 was missing from List 2. This task demanded detailed comparison and careful record-keeping, and was rated hard.

4. *Receipts.* In this task, participants were given mock shop receipts showing purchased items and prices. On screen, they had to enter totals or check whether the items and prices matched across versions. The job simulated everyday record-keeping in workplace or market contexts, and required attention to numbers, calculation, and consistency. The difficulty level lay between medium and hard, depending on the volume of receipts.

5. *Ravens.* This task involved a Raven's Progressive Matrices–style cognitive test, unrelated to data entry. Participants were shown visual patterns with one piece missing and had to choose the correct option to complete the sequence. Unlike the other tasks, this measured abstract reasoning ability rather than clerical accuracy. It was typically regarded as hard because it tested problem-solving under time pressure.

We compensated participants based on the joint accuracy of their performance across tasks. Team earnings ranged from 150 to 450 INR per person, with higher accuracy and greater completion leading to higher pay. In addition, the partner designated as the *senior* for a given task received a personal bonus of 50 INR. This senior bonus was awarded at the end of each task, except in one task where the role was randomly assigned by lottery. In total, a participant who served as senior for all five tasks could earn up to 250 INR in bonuses, on top of the team-based payment. Team payments were disbursed immediately after the session, while senior bonuses were distributed within three days.

Context-sensitivity and external validity Two features of the design matter for external validity. First, the intervention alters the visibility of a salient identity marker—surnames—and the norms surrounding disclosure, without

changing incentives, payoffs, or the structure of the task itself. This maps onto many real-world settings in India where surnames are observed in hiring, schooling, or workplace interactions, but where there is also a tendency, especially in more informal contexts, to suppress caste-linked cues, for example by using first names only or adopting neutral surnames. In fact, some formal workplaces have begun to adopt first-name-only policies precisely to reduce the salience of identity and hierarchy. For example, there have been documented shifts in Indian corporate workplaces away from "Sir/Madam" honorifics toward first-name address, motivated by efforts to flatten hierarchy and normalize more neutral interaction.[5] In these contexts, discretion primarily concerns whether and when identity is revealed; once a surname is observed, the scope for concealment is significantly reduced. The *First Name Only* condition is especially relevant in this respect, as it captures environments in which identity is initially masked but may be revealed endogenously over the course of interaction.

Second, the design features repeated interaction with a single partner around a shared production task. This captures an important class of relationships—short-term workplace collaborations or team-based tasks—in which trust and relational capital matter. At the same time, the experiment is not designed to speak to one-shot selection margins, such as hiring callbacks, or to fleeting and anonymous interactions. As a result, the tradeoff between concealment and disclosure may differ in settings where the primary motivation for hiding identity is to avoid outright exclusion rather than to manage ongoing coordination and rapport.

Finally, we conducted the experiment in a relatively controlled environment, with supervision and explicit workplace rules. These features limit extreme conflict and likely compress the lower tail of outcomes. We therefore interpret the results as capturing how identity concealment shapes everyday relational dynamics, rather than as a test of its ability to insure against rare but severe discriminatory episodes.

Outcomes We pre-specified five families of primary outcomes (AEA #15151), covering beliefs about group identity, social relations, trust, well-being at work, and productivity and task allocation. The paragraphs that follow denote the exhaustive set of pre-specified primary outcomes. We also refer to some pre-specified secondary outcomes, explicitly flagging these outcomes as secondary. We denote minor deviations from our pre-registration in Appendix D.2.

Beliefs about group identity. At the end of the work session, we asked each worker to guess both the caste category (SC, ST, OBC, General, or Muslim) and the specific caste (e.g. Brahmin, Chamar) of their partner. We then classify their answer as correct or not by matching their guess with the baseline data for their partner. Participants reported their level of confidence in the caste category guess on a 0 to 100% scale, with the answer reflecting the participant's

[5] Harsh Mariwala on dismantling "Sir ji" culture at Marico; SiliconIndia on Indian IT firms moving away from "Sir/Madam" norms.

perceived probability that they guessed correctly, and participants were additionally asked whether they thought their partner would be able to correctly identify their own caste—a second-order belief. This allows us to characterize not only raw accuracy, but also workers' "meta-beliefs" about their ability to conceal or reveal caste identity. To guard against strategic behavior, we did not incentivize guesses.[6]

Social relations. We measured how participants valued future interaction with their assigned partner. First, in a series of incentivized willingness-to-pay (WTP) and willingness-to-accept (WTA) choices, participants decided whether they would forgo higher wages in order to work again with their current partner rather than a randomly selected co-worker of the same caste. Participants also reported whether they expected to meet up with their partner in the future outside the study context. Finally, we used an overlapping circles (inclusion-of-other-in-self) measure to assess subjective identity fusion with the partner on a scale of 0 to 4.

Trust. Pairs played a standard incentivized trust game (Berg et al., 1995). In this game, Player A (the sender) received Rs. 50 and decided how much to send to their work partner. Transfers were tripled, and Player B (the receiver) then chose how much to return. Participants stated their intended transfer amounts if chosen as Player A, their expectations about how much their partner would transfer if assigned as Player A, and their conditional return amounts for each possible transfer if chosen as Player B (using the strategy method). These data allow us to capture multiple dimensions of trust and trustworthiness: willingness to place resources at risk, expectations of reciprocity, and actual reciprocity when endowed with power.

Well-being at work. Finally, we measured workers' affective experience during the session. Respondents rated the extent to which they felt happy, self-conscious, confident, relaxed, and anxious, each on a scale from 0 = not at all to 3 = very. As a secondary outcome, we also asked workers to report to what degree they were able to show their "true self" to their partner on a scale from 1 = not at all to 10 = completely.

Productivity and task (mis)allocation. We recorded detailed measures of performance on the joint data-entry tasks: the number of fields entered (quantity), the fraction of fields correctly entered (quality), and the number of fields correctly entered (a joint measure of quantity and quality). Because pairs had to assign one worker to the "senior partner" tablet role and the other to read aloud from forms, we also track the identity of the partner who assumed the higher-status task on each task. We complement these behavioral measures with three secondary outcomes: self-reports on whether workers were satisfied with the allocation of tasks (three-point scale: happy with allocation / would prefer to do more tablet work / would prefer to do less tablet work), whether they believed their preferred

[6]For example, if we incentivized guesses, and new participants learned from previous participants that this was the case, the new participants might exert more effort to learn the caste of their partner in the *First Name Only* and *Hidden* conditions. This would attenuate our first stage effects.

allocation would have improved performance, and perceptions of own and partner's performance on a scale of 1 to 10.

3 Model: The *First name only* treatment as a Bayesian game

In this section, we develop a two-player Bayesian game that models the *First Name Only* treatment, where identity revelation is an endogenous choice. The model rationalizes our core empirical findings across treatments and generates additional predictions that we test in the data.

Each player $i \in \{1,2\}$ chooses an action $a_i \in \{R,H\}$, where R denotes revealing caste identity and H denotes hiding it. Action R captures both explicit and implicit ways of communicating caste identity during the interaction. The experimental treatments correspond to restrictions on this choice set: in the *Hidden* treatment players cannot choose R, while in the *Revealed* treatment they cannot choose H.

We model a player i's private type as

$$t_i = (c_i, y_i),$$

where $c_i \in \{G,L\}$ denotes caste identity (General or SC) and y_i is i's ex-ante expected payoff from a cross-caste interaction, given i's belief about facing caste-domination at work. We assume each caste occurs with probability $1/2$ and this is common knowledge.

Under the restriction $x > 0 \geq z$, the payoffs from the game are given by

$$u_i(a_i,a_j;t_i,t_j) = \begin{cases} 0 & \text{if } a_i = a_j = H \quad \text{(no revelation outcome is normalized to "0")}, \\ x & \text{if } a_i = a_j = R \text{ and } c_i = c_j \quad \text{(mutual revelation is beneficial for same-caste matches)}, \\ y_i & \text{if } a_i = a_j = R \text{ and } c_i \neq c_j \quad \text{(...but in cross-caste matches, it depends on your payoff-type)}, \\ z & \text{if } (a_i,a_j) \in \{(R,H),(H,R)\} \quad \text{(one-sided sharing is awkward, hence worse than "0")}. \end{cases}$$

We assume that one-sided identity sharing is worse than mutual hiding and this has a non-trivial testable prediction. Given participants can communicate before choosing whether to reveal or hide during the session, they should attempt to avoid reaching the one-sided sharing outcome – perhaps in favor of the pareto-dominating mutual hiding outcome – whenever possible. Thus, we should see few such outcomes.

We also assume that the payoff $z \leq 0$ applies to both to the partner who revealed and who did not reveal - this is also for simplicity. Assuming that the outcome is worse for the person one-sidedly sharing their identity - $z_s < z_{ns} \leq 0$

- would result is the same qualitative predictions as we currently do.

Assumptions about y_i General-caste players do not anticipate facing caste-domination, and hence $y_i = x$ deterministically. But, for lower-caste players, $y_i \sim F$ with support $[\underline{y}, x]$ where $\underline{y} < 0$. Lower-caste players with low y_i anticipate negative payoffs from mutual revelation in cross-caste interactions. This heterogeneity captures differences in anticipated interaction quality based on past experiences, social information, and contextual cues. Lower-caste individuals may draw on prior workplace interactions, accounts from peers or media, or signals received during the interaction itself when forming expectations about how a cross-caste interaction might unfold. As a result, some lower-caste players expect positive interactions following revelation, while others anticipate negative outcomes.

For everyone else, mutual revelation is beneficial because it provides a shared social context in which individuals can interact and communicate more easily. Knowledge of a partner's social identity helps anchor expectations about norms, language, and appropriate modes of interaction, and can facilitate coordination in joint tasks. When identity is known, individuals may find it easier to interpret behavior, establish conversational common ground, and anticipate how their partner will respond. In this sense, identity revelation can reduce social uncertainty and help interactions proceed more smoothly.

A pure strategy for a General-caste player is a choice $s_G \in \{R, H\}$, since General-caste players have no payoff-relevant heterogeneity. By contrast, a pure strategy for a lower-caste player is a measurable map

$$s_L : [\underline{y}, x] \to \{R, H\},$$

which assigns an action to each payoff-type y. Given a lower-caste strategy s_L, let

$$p_L \equiv \Pr(s_L(y) = R) = \int \mathbf{1}\{s_L(y) = R\}\, dF(y)$$

denote the proportion of lower-caste players who reveal. Thus p_L is the endogenous disclosure rate induced by the strategy s_L.

Two opposing considerations shape disclosure dynamics in this game, as formalized in Proposition 1. First, unilateral revelation is socially costly: revealing caste when the partner does not do so yields payoff $z < 0$. Thus individuals prefer not to reveal if they expect others to conceal. As a result, a no-revelation outcome – where both sides hide caste identity – is always sustainable as an equilibrium.

Second, mutual revelation can generate benefits by clarifying the social context of the interaction. For example, if General-caste players reveal their caste, lower-caste players may also benefit from revealing, depending on their

type y_i. Individuals with sufficiently low y_i anticipate unfavorable cross-caste interactions and prefer concealment, whereas those with higher y_i are willing to reveal.

These forces lead to two possible interaction regimes in pairs of worker groups. In one regime, no one reveals caste identity. In the other, General-caste players reveal, while lower-caste players reveal only if their type exceeds a threshold.

Proposition 1 (Equilibrium characterization)**.** *Assume $\underline{y} < -x+2z$. Then the game admits exactly two pure-strategy symmetric Bayesian Nash equilibria.*

(i) No revelation. $s_G = H, \quad s_L(y) = H \ \ \forall y.$

(ii) Partial lower-caste revelation. $s_G = R$ *and there exists a cutoff* y^* *such that*

$$s_L(y) = \begin{cases} H & y < y^* \\ R & y \geq y^* \end{cases}$$

where $y^* = -p_L(x-2z)$ *and* p_L *solves*

$$p_L = 1 - F\big(-p_L(x-2z)\big).$$

The condition $\underline{y} < -x+2z$ implies that some L types prefer concealment in the candidate full-revelation equilibrium, which is sufficient to rule that equilibrium out. All proofs are relegated to Section A.

Next, to connect the disclosure game to observed workplace outcomes, suppose realized interaction quality in a mixed-caste pair is

$$Q = \begin{cases} q(y_i) & \text{if } (R,R), \\ q_{HH} & \text{if } (H,H), \\ q_{RH} & \text{if } (R,H) \text{ or } (H,R), \end{cases}$$

where $q(\cdot)$ is continuously increasing in the lower-caste type $y_i \in [\underline{y}, x]$. Assume

$$q(x) > q_{HH} \geq q_{RH} > q(\underline{y}).$$

Thus mutual revelation improves interaction quality for sufficiently high-y lower-caste individuals, but reduces it for sufficiently low-y individuals relative to mutual concealment.

Our next result highlights a key difference between the *First Name Only* and *Hidden* treatments. Under *Hidden*, mutual concealment is imposed for all pairs. Under *First Name Only*, by contrast, lower-caste individuals reveal only

if their type exceeds a threshold, leading to positive selection into mutually revealing interactions among mixed-caste pairs. As a result, interactions with the greatest potential gains from revelation are the ones most likely to realize it.

To formalize this comparison, suppose mixed-caste pairs in the *First Name Only* treatment coordinate on the selective-revelation equilibrium with probability $s \in (0,1)$ and on the no-revelation equilibrium with probability $1-s$. In the selective-revelation equilibrium, General-caste players always reveal and lower-caste players reveal if and only if $y_i \geq y^*$, so the conditional lower-caste revelation rate is

$$p_L = 1 - F(y^*).$$

For a mixed-caste pair it follows that under *First Name Only*

$$P(R,R) = sp_L, \qquad P(H,H) = 1-s, \qquad P(R,H) + P(H,R) = s(1-p_L),$$

The comparison between *First Name Only* and *Revealed*, however, is ambiguous. Under *Revealed*, mutual revelation is imposed for all pairs. On the one hand, *Revealed* eliminates the inefficient mutual-hiding equilibrium, which benefits mixed-caste pairs with high y for whom mutual revelation improves interaction quality but who might otherwise have coordinated on mutual concealment. On the other hand, *Revealed* forces mutual disclosure even for lower-caste individuals with low y, for whom mutual revelation yields lower interaction quality than mutual concealment. Which force dominates is therefore an empirical question.[7]

Remark 1. *Suppose the First Name Only treatment selects the selective-revelation equilibrium with probability $s \in (0,1)$ and the no-revelation equilibrium with probability $1-s$. Then, for any $s \in (0,1)$, the expected interaction quality in mixed-caste pairs is weakly higher under First Name Only than under Hidden whenever*

$$p_L\big(\mathbb{E}[q(y_i) \mid y_i \geq y^*] - q_{HH}\big) \geq (1-p_L)(q_{HH} - q_{RH}).$$

The comparison between First Name Only and Revealed is ambiguous.

The condition in the remark states that *First Name Only* improves on *Hidden* when the gains from mutual revelation among positively selected high-y mixed-caste pairs outweigh any losses generated by asymmetric disclosure outcomes. For example, this condition holds if $q_{HH} \approx q_{RH}$. The comparison between *First Name Only* and *Revealed*, however, depends on the equilibrium-selection parameter s with which play selects the selective-revelation equilibrium under *First Name Only*. In particular, when s is low, the coordination failure under *First Name Only* is frequent

[7] This incomparability holds irrespective of any assumption about interaction quality in asymmetric revelation outcomes (R,H) or (H,R).

and *Revealed* tends to generate higher interaction quality; when s is high, the ability of *First Name Only* to avoid forcing revelation on low-y individuals may dominate.

4 Estimation

For our analysis, we estimate our pre-specified intent-to-treat specification:

$$Y_i = \alpha + \beta_1 \textit{First Name Only}_{p(i)} + \beta_2 \textit{Hidden}_{p(i)} + \gamma X_i + \delta Y_{i0} + \lambda_s + \varepsilon_i, \tag{1}$$

where Y_i denotes the outcome of interest for individual i. The indicators $\textit{First Name Only}_{p(i)}$ and $\textit{Hidden}_{p(i)}$ denote assignment of pair p to the *First Name Only* or *Hidden* conditions, with the *Revealed* condition serving as the omitted reference category. X_i is a vector of individual-level controls (age and fixed effects for the participant's level of education), Y_{i0} is the baseline measure of the outcome to improve precision (if available), and λ_s are randomization strata fixed effects (day-by-caste, where caste is High/General or Low/SC). We cluster standard errors at the pair level for outcomes measured at the individual level, and we estimate heteroskedastic-robust standard errors in specifications that include only one of the two caste categories.

Summary Statistics Table A1 presents summary statistics on the characteristics of our experimental sample. The average participant is 23 years old. Additionally, 73% of the sample having completed 12th grade. College completion is substantially lower at around 20%, partly because many individuals have not yet reached college completion age. The average monthly income is Rs. 10,700 (roughly 120 USD), indicating that the participants are relatively low-income. Finally, in terms of identity rankings, participants most commonly identify as "Indian" first, followed by personal relationships (e.g., being a father, son, or brother), and then religion, caste, and state (Uttar Pradesh) identity, all of which receive similar rankings.

What is perhaps more interesting are the differences between high- and low-caste individuals (see Table A2). As expected, high-caste participants are, on average, more educated, come from wealthier families, and report having fewer siblings (consistent with lower fertility in this group). When it comes to identity rankings, however, high castes place their Indian identity relatively higher than low castes do, and their caste identity relatively lower (lower scores denote higher ranks). At first glance, this might seem counterintuitive given the historical dominance and better socioeconomic outcomes of high castes. But recall that participants were asked to provide relative rankings of identities. One plausible explanation is that high castes may feel a stronger sense of pride in the nation, perceiving themselves as more central to the national identity and its historical narrative, whereas low castes may place greater

emphasis on caste identity as a counter-response to discrimination and marginalization.

Selection We approached a total of 1,677 potential subjects in Barabanki and Ayodhya through door-to-door campaigns to arrive at our final experimental sample of 1,198 men. We do not find large or systematic differences between those who eventually participated in the experiment and those who did not (Table A3). The experimental sample is slightly younger, and consequently somewhat less likely to have completed college, and ranks their religious identity marginally lower and personal identities (e.g., father, son, brother) marginally higher, while ranking their caste identity similarly. These differences likely reflect that participants were more often younger, college-going men who may have had greater scheduling flexibility to participate. This younger demographic may also be associated with placing somewhat less importance on religious identity relative to personal identities. Crucially, however, we do not observe selection on the importance of caste identity, which was our primary concern.

Balance tests Table A4 indicates that the randomization was successful. We estimate specification 1 with participant characteristics as the dependent variable and report p-values for the full set of pairwise comparisons between treatment groups. Across all comparisons, the estimated coefficients are small in magnitude and not statistically significant.

Attrition 72% of all individuals we approached participated in the study (1,198 men participated out of a total of 1,677 we approached). Since the endline survey was performed immediately after the end of the task completion, there is no attrition between the intervention and the endline.

5 Results

We begin by showing that in the *First name only* and *Hidden* conditions, participants were substantially less likely to correctly identify their partner's caste (compared to *Revealed*), meaning that we have a strong first stage with which to estimate the downstream effects of identity knowledge and uncertainty. We go on to examine how the absence of clear identity information impacted social relations and trust, as well as productivity and task allocations.

First stage: beliefs about group identity Our *Revealed* treatment condition works as intended: 97% of participants in that condition correctly guess their partner's caste category at the end of the workday, 88% correctly guess their partner's specific subcaste (*jati*), participants report that they are 89% confident in their caste category guess on average, 55% report being completely confident in their caste-category guess, and 97% predict that their partner will

guess their caste category accurately (omitted category mean for columns 1 to 5, Panel A, Table 2). These numbers confirm that caste is common knowledge in the *Revealed* condition.

The *First name only* and *Hidden* conditions sharply erode caste knowledge. Relative to the *Revealed* condition, correct caste category guesses fall by 14 and 22 percentage points (Table 2, Panel A, Column 1), correct *jati* guesses by 20 and 34 points (Column 2), confidence by 13 and 22 points (Column 3), and beliefs that one's partner can correctly infer one's caste category by 15 and 35 points (Column 5). All pairwise differences are statistically significant at the 1% level, implying that *First name only* weakens caste knowledge relative to *Revealed*, and *Hidden* reduces it still further. Effects are much larger for a stricter measure, namely, reporting 100% confidence in one's caste category guess, which declines by 28 and 41 percentage points, respectively (Column 4).

We note that even these estimates likely understate the effective first stage. By design, participants in *Revealed* learn their partner's caste immediately, generating four full hours of common knowledge. In the other conditions, by contrast, information likely arrives gradually over the course of the workday; thus, it is reasonable to suppose that although end-of-day guess rates are relatively high, participants spend much of the interaction under conditions of uncertainty, with only partial and evolving signals about their partner's identity.

Our first-stage effects remain strong when considering high and low castes separately (Table 2, Panels B and C). Still, we observe marked heterogeneity by caste status when comparing the *Hidden* condition to *First name only*. SC participants are able to guess their partner's caste category with similar accuracy in these two conditions (Panel B, column 1). By contrast, for General-caste participants, there is a large gap: their accuracy in *Hidden* is 15 percentage points lower than in *First name only* ($p < 0.01$), indicating a nearly 30 percentage point loss in accuracy relative to the *Revealed* condition. This pattern suggests that low-caste individuals are more effective at concealing their identity when prompted to do so: either because they are more practiced at concealing, or because high-caste individuals reveal their identity more readily (deliberately or inadvertantly) through other social cues.[8]

To understand which of these possibilities are more likely, one can look at the second-order beliefs. Both lower and higher castes report that their partners would be less likely to guess their caste category in the *Hidden* than in *First name only* treatment. Under the assumption that such perceptions are based on the social cues they revealed during the sessions, or equivalently, on the effort they exerted in hiding their caste in the *Hidden* condition, one can interpret the previous result as lower castes having a higher ability at hiding their caste identity when needed.

As intended, our treatments also change the strategies by which participants infer the caste of their partner. In *Revealed*, 74% report having used the partner's last name to guess their caste, versus only 17% in *First name only* and 8% in *Hidden* (Table 3, Panel A, Column 2). In *First name only*, the most common explanations for how the

[8]Note, we cannot distinguish between these mechanisms, but the result is consistent with expectations in a hierarchical society: the identity of the privileged group is more easily discernible than that of the less privileged group.

Table 2: First stage: Guessing partner's caste identity by treatment and caste group.

	Correctly guessed caste category (1)	Correctly guessed jati (2)	Confidence in caste category guess (3)	100% confidence in caste category guess (4)	Partner can guess my caste category (5)	Second-order belief accuracy (6)
A. Full sample						
First name only	-0.14***	-0.20***	-12.67***	-0.28***	-0.15***	-0.12***
	(0.03)	(0.04)	(1.90)	(0.05)	(0.03)	(0.03)
Hidden	-0.22***	-0.34***	-22.06***	-0.41***	-0.35***	-0.23***
	(0.03)	(0.03)	(1.86)	(0.05)	(0.04)	(0.03)
First name only vs. Hidden p-value	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Revealed mean	0.97	0.88	89.10	0.53	0.97	0.94
N	1,013	1,013	1,013	1,013	1,013	1,013
B. SC sample						
First name only	-0.14***	-0.20***	-12.27***	-0.24***	-0.17***	-0.12***
	(0.03)	(0.05)	(2.05)	(0.05)	(0.04)	(0.04)
Hidden	-0.16***	-0.28***	-22.31***	-0.37***	-0.34***	-0.24***
	(0.03)	(0.04)	(2.08)	(0.05)	(0.04)	(0.04)
First name only vs. Hidden p-value	0.66	0.13	<0.01	<0.01	<0.01	0.01
Revealed mean	0.98	0.90	89.06	0.51	0.96	0.93
N	506	506	506	506	506	506
C. General sample						
First name only	-0.13***	-0.20***	-13.31***	-0.31***	-0.15***	-0.12***
	(0.03)	(0.04)	(2.09)	(0.05)	(0.03)	(0.03)
Hidden	-0.28***	-0.40***	-22.00***	-0.46***	-0.36***	-0.23***
	(0.04)	(0.05)	(2.01)	(0.05)	(0.04)	(0.04)
First name only vs. Hidden p-value	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Revealed mean	0.95	0.86	89.13	0.55	0.98	0.96
N	507	507	507	507	507	507
D. Difference in differences						
First name only × General (vs. SC) p-value	0.81	0.92	0.52	0.03	0.51	0.90
Hidden × General (vs. SC) p-value	<0.01	0.03	0.81	<0.01	0.64	0.89

Notes: Each column reports estimates from an OLS regression of the indicated outcome on indicators for assignment to the *First name only* and *Hidden* treatments, with *Revealed* as the omitted category. Panel D reports difference-in-differences p-values from regressions that interact each treatment indicator with an indicator for General caste (SC is the omitted group); the reported p-values correspond to the interaction terms. Outcomes are defined as follows: (1) equals 1 if the participant correctly guessed their partner's caste category; (2) equals 1 if the participant correctly guessed their partner's jati; (3) is confidence in the caste-category guess on a 0–100% scale; (4) equals 1 if the participant reported they were 100% confident in their caste-category guess; (5) equals 1 if the participant believes their partner can correctly guess their caste category; (6) equals 1 if the participant correctly predicted whether their partner would guess their caste category correctly. Standard errors are reported in parentheses. In Panels A and D, standard errors are clustered at the pair level. In Panels B and C (the SC and General subsamples), standard errors are heteroskedasticity-robust. All specifications include randomization strata fixed effects and pre-specified controls for participants' education level and age. $*p < 0.1$; $**p < 0.05$; $***p < 0.01$.

participant guessed the caste are that the partner told them their caste directly (42%), or that they guessed based on the way the partner spoke (36%), or based on where the partner lives (35%). In *Hidden*, we see some evidence that participants followed our instructions—far fewer report that their partner told them their caste directly (15%). The most common answers are that they guessed based on the way the partner spoke (49%) or based on where the partner lives (32%). The frequency of participants organically revealing their caste in *First name only* might suggest that the *Hidden* treatment is costly in its paternalism: it prevents a substantial fraction of participants from playing their "best response," by revealing their caste. We explore this by estimating effects on downstream outcomes in the subsequent sub-sections.

Table 3: Reported cues used to infer partner's caste identity, by treatment and caste group.

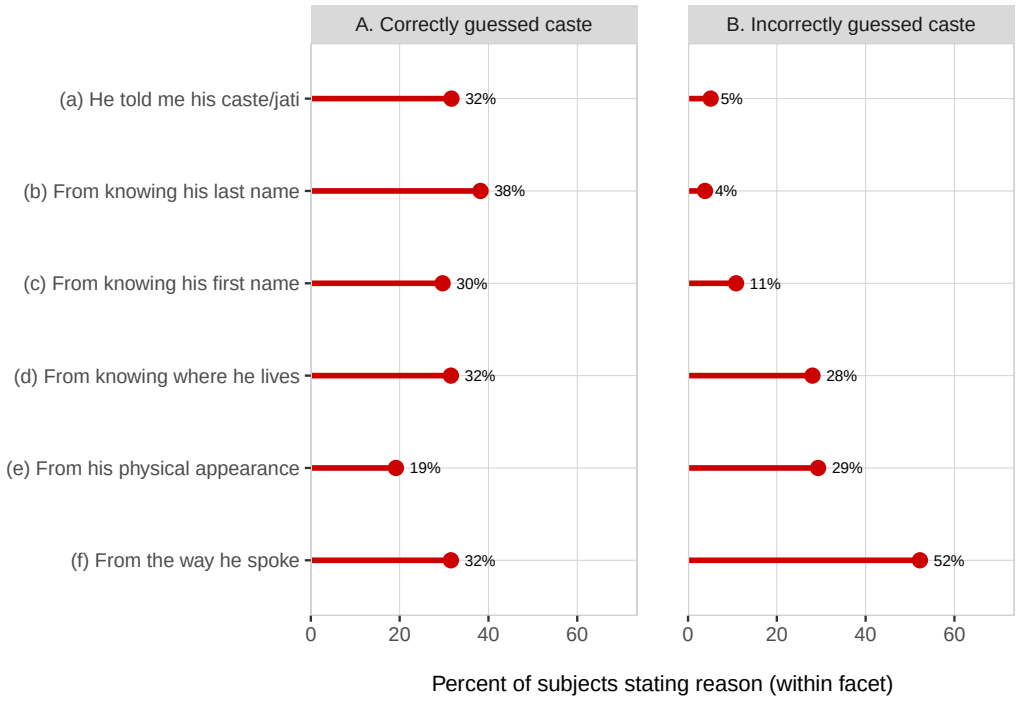
	He told me (1)	Used last name (2)	Used first name (3)	Used speaking (4)	Used appearance (5)	Used where lives (6)	Used other (7)
A. Full sample							
First name only	0.17***	-0.57***	-0.17***	0.18***	0.13***	0.09**	0.33***
	(0.05)	(0.04)	(0.05)	(0.04)	(0.03)	(0.05)	(0.04)
Hidden	-0.10**	-0.66***	-0.28***	0.29***	0.20***	0.07	0.48***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)
First name only vs. Hidden p-value	<0.01	<0.01	<0.01	0.01	0.08	0.60	<0.01
Revealed mean	0.25	0.74	0.42	0.19	0.10	0.26	0.45
N	1,013	1,013	1,013	1,013	1,013	1,013	1,013
B. SC sample							
First name only	0.14***	-0.58***	-0.15***	0.15***	0.14***	0.09*	0.32***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)
Hidden	-0.12***	-0.66***	-0.27***	0.29***	0.22***	0.04	0.46***
	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)
First name only vs. Hidden p-value	<0.01	0.04	<0.01	0.01	0.12	0.28	<0.01
Revealed mean	0.27	0.76	0.41	0.19	0.13	0.27	0.46
N	506	506	506	506	506	506	506
C. General sample							
First name only	0.20***	-0.56***	-0.18***	0.20***	0.13***	0.09*	0.35***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)
Hidden	-0.08*	-0.66***	-0.29***	0.29***	0.18***	0.11**	0.50***
	(0.05)	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.04)
First name only vs. Hidden p-value	<0.01	<0.01	0.02	0.09	0.25	0.82	<0.01
Revealed mean	0.23	0.73	0.43	0.20	0.06	0.25	0.45
N	507	507	507	507	507	507	507
D. Difference in differences							
First name only × General (vs. SC) p-value	0.14	0.56	0.37	0.36	0.77	0.93	0.62
Hidden × General (vs. SC) p-value	0.25	0.91	0.50	0.94	0.52	0.13	0.34

Notes: Each column reports estimates from an OLS regression of the indicated outcome on indicators for assignment to the *First name only* and *Hidden* treatments, with *Revealed* as the omitted category. Panel D reports difference-in-differences p-values from regressions that interact each treatment indicator with an indicator for General caste (SC is the omitted group); the reported p-values correspond to the interaction terms. The outcomes are binary indicators for whether respondents selected each stated reason in response to the question: "How were you able to guess your partner's caste category or jati? Even if you are not completely sure about your guess, please tell us what made you guess the way you did. Select all that apply": (1) being told directly ("he told me his caste/jati"); (2) from knowing his last name; (3) from knowing his first name; (4) from the way he spoke; (5) from his physical appearance; (6) from knowing where he lives; and (7) other reasons. Standard errors are reported in parentheses. In Panels A and D, standard errors are clustered at the pair level. In Panels B and C (the SC and General subsamples), standard errors are heteroskedasticity-robust. All specifications include randomization strata fixed effects and pre-specified controls for participants' education level and age. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Figure 1 breaks down the reported reasons for caste guesses, according to the ultimate accuracy of those guesses. Accurate caste guesses rely disproportionately on direct or high-quality signals, whereas incorrect guesses lean more heavily on noisier, indirect cues. Among correct guesses, participants frequently cite knowing their partner's last name (38%), and being told caste directly (32%). By contrast, incorrect guesses are much more likely to be based on speech, physical appearance, or residential location, all of which are more ambiguous signals of caste. This suggests

that when clear identifiers are unavailable, participants substitute toward weaker heuristics.

Figure 1: Reported reasons for caste guesses, by accuracy of guess.



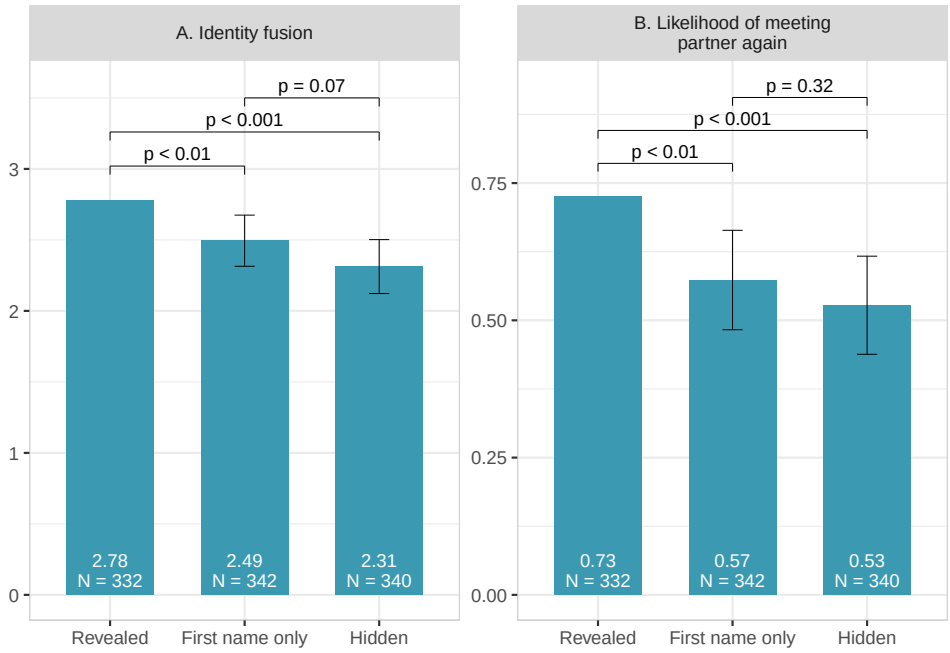
Notes: The unit of analysis is the participant. The outcomes are binary indicators for whether respondents selected each stated reason in response to the question: "How were you able to guess your partner's caste category or jati? Even if you are not completely sure about your guess, please tell us what made you guess the way you did. Select all that apply." Within each panel, the figure shows the percentage of participants citing each stated reason, conditional on whether the participant correctly or incorrectly guessed their partner's caste.

Social relations Our two identity-hiding conditions substantially worsen social relations between partner pairs (Figure 2). Compared to *Revealed*, participants in *First name only* are less likely to report feeling close with their partner, using the 0 to 4 scale of identity fusion (Panel A, $\hat{\beta}_1 = -0.29$ or 0.32σ, $p < 0.01$), and 15 percentage points less likely to report expecting to meet up with their partner again ($p < 0.01$), off a high base of 73% (Panel B).[9] These are large effects, especially given that (i) the treatment is subtle, merely the use of full names rather than first names, and (ii) participants can in principle replicate the *Revealed* condition, by simply revealing their full name to their partner. Social relations deteriorate still further in the *Hidden* condition, consistent with the stronger effects of *Hidden* on caste knowledge, although the magnitudes are smaller than the effects of *First name only* versus *Hidden*, with correspondingly larger p-values ($p = 0.07$ for identity fusion and $p = 0.32$ for expectation of future interaction).

How does differential awareness of partner identity affect the wage that participants are willing to accept to work with their partner again in future? To understand this, we offered subjects the choice between working another day

[9]This high base cannot be attributed to pre-existing relationships: only 3% of participants in the sample reported encountering or having any knowledge of their work partner in the past.

Figure 2: Estimated treatment effects on social relations.



Notes: The unit of analysis is the participant. The outcomes are: (A) Which picture best describes your relationship with your work partner today? (0 = self and partner circles fully apart to 4 = self and partner circles fully overlapping); (B) Do you think you will meet up with your partner again in future? (0 = No, 1 = Yes). The *Revealed* bar in each panel shows the mean of the outcome variable for *Revealed* participants. The *First name only* and *Hidden* bars add the corresponding treatment effects, estimated from regressions that control linearly for participants' age, and include randomization-strata and education-bin fixed effects. The 95% confidence intervals are based on standard errors clustered at the pair level. Brackets report two-sided *p*-values for pairwise differences across arms.

with their partner versus a stranger who had participated previously in the research, and varied the wage offered across these two choices. We started with a wage of Rs. 400 for both choices and then gradually increased the offered wage for the choice that they initially did *not* make, to ascertain their switching point. Table 4 presents Tobit results where the outcome variable is the relative willingness to pay for a partner versus a stranger, coded as the point of transition from partner to stranger (or stranger to partner). If a subject reported indifference at a given price, we take the midpoint between the price at which they chose an option and the price at which they reported indifference.

Table 4: Estimated treatment effects on the daily wage participants are willing to sacrifice to work again with partner rather than same-caste stranger.

	Full sample (1)	General only (2)	SC only (3)
First-name only	-4.07	-7.21	-0.56
	(21.59)	(25.16)	(22.76)
Hiding	-36.88*	-43.38*	-30.48
	(21.33)	(24.40)	(22.64)
First-name only vs. Hiding p-value	0.12	0.14	0.19
Revealed mean	173	171	175
N	1,012	507	505

Notes: The unit of analysis is the participant. The outcome is the wage (in Rs.) that a participant would be willing to accept to work with their partner again, rather than with a same-caste stranger. Tobit regressions control linearly for participants' age, and include randomization strata and education-bin fixed effects. Standard errors are clustered at the pair level. $*p < 0.1$; $**p < 0.05$; $***p < 0.01$

The results show that subjects in the *Hidden* condition are willing to sacrifice around Rs. 37 ($p = 0.084$) to work with a stranger rather than their partner. This is driven primarily by General caste subjects (Column 2, Rs. 43.38, $p = 0.076$) and less so by SCs (Rs. 30.48, $p = 0.179$). The effects for *First name only* are relatively small and not statistically significant, perhaps because subjects retain the possibility of learning about their partner's identity—much as they might with a stranger—making them more willing to continue working with their partner relative to the *Hidden* condition, where this opportunity is foreclosed entirely.

Overall, uncertainty over a partner's identity leads to worse social relations, with the effects being especially strong when the option to reveal is taken away (as in the *Hidden* condition relative to *First name only*).

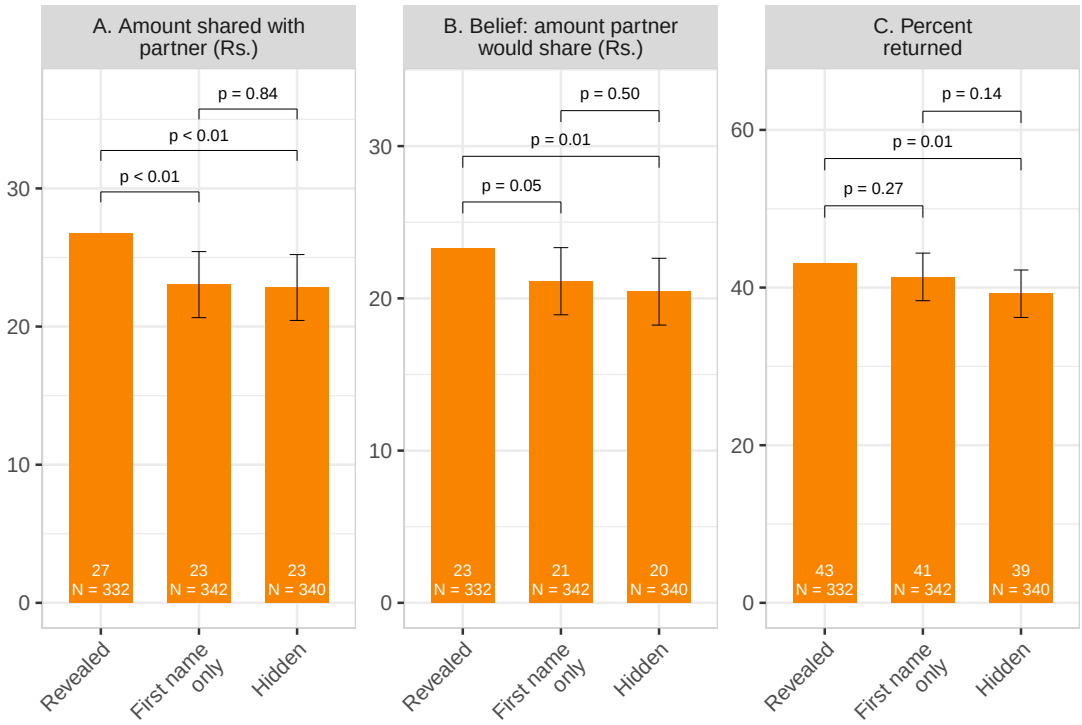
Trust Effects of our caste-hiding treatments on trust game behaviors are similar to those on identity fusion and anticipated future interaction. In Figure 3, Panel A, we see that participants in the *Revealed* condition share Rs. 27 (out of 50) of their original endowment on average.[10] Both the *First name only* and *Hidden* conditions reduce giving by approximately Rs. 4—about a 15% reduction relative to *Revealed*, with both differences highly statistically significant. Meanwhile, the *First name only* and *Hidden* arms are statistically indistinguishable from one another ($p = 0.84$). Failure to disclose full, caste-identifying information thus substantially reduces Sender transfer amounts.

The similarity of the effects produced by the *First name only* and *Hidden* manipulations is quite telling. One concern is that directing participants to actively conceal their identity creates an artificially cagey setting, and that this is what dampens trust. But the *First name only* treatment involves no such instruction: it withholds full caste-identifying information in a natural, unobtrusive way, yet still the reduction in trust is near-identical. This suggests

[10]There is no statistically significant difference in the amounts sent by SC- and General-caste participants.

that what propels changes in trust is not the framing of concealment, but the absence of clear caste signals.

Figure 3: Estimated effects on trust game outcomes.



Notes: The unit of analysis is the participant. The outcomes are: (A) the amount (Rs. 0–50) sent to the partner when the participant is Player A; (B) the participant's belief about the amount (Rs. 0–50) the partner would send if selected as Player A; and (C) the percentage of the tripled amount the participant (as Player B) would return to their partner, averaged across all possible transfer levels. The *Revealed* bar in each panel shows the mean of the outcome variable for *Revealed* participants. The *First name only* and *Hidden* bars add the corresponding treatment effects, estimated from regressions that control linearly for participants' age and include randomization-strata and education-bin fixed effects. The 95% confidence intervals are based on standard errors clustered at the pair level. Brackets report two-sided p-values for pairwise differences across arms.

Negative effects on Sender behavior could be driven by beliefs (expecting the partner to return less) or preferences (feeling less altruistic toward the partner). To unpack these channels, we turn to our measures of beliefs and Receiver behavior in the trust game. Participants in the *Revealed* condition expect their partner to share Rs. 23 of their endowment if selected as the Sender, compared to Rs. 21 and Rs. 20 in the *First name only* and *Hidden* conditions, respectively ($p = 0.05$ and $p = 0.01$ for the comparisons with *Revealed*). Once again, there is no significant difference between the *First name only* and *Hidden* treatment arms. These results suggest that the decline in transfers when identities are more opaque is at least partly due to diminished expectations of reciprocity. Nevertheless, we also see a role for preferences: Panel C of Figure 3 examines treatment effects on actual reciprocity, measured as the percentage of the tripled amount that participants would return, averaged across all possible transfer levels. Participants in *Revealed* return 43 percent on average, compared to 41 percent in *First name only* and 39 percent in *Hidden*. The reduction in returns is statistically significant for the *Revealed* versus *Hidden* comparison ($p = 0.01$)

but not for the other two possible comparisons. For the case of reciprocal behavior, therefore, only the explicitly instructed hiding condition produces a detectable decline.

Taken as a whole, we see that treatments that mask caste identity have measurable efficiency costs: reducing the total surplus generated in the trust game by around 15%. Lower transfers directly shrink the size of the tripled surplus available for division, and reciprocity also declines, with participants returning a smaller share of a smaller pie.

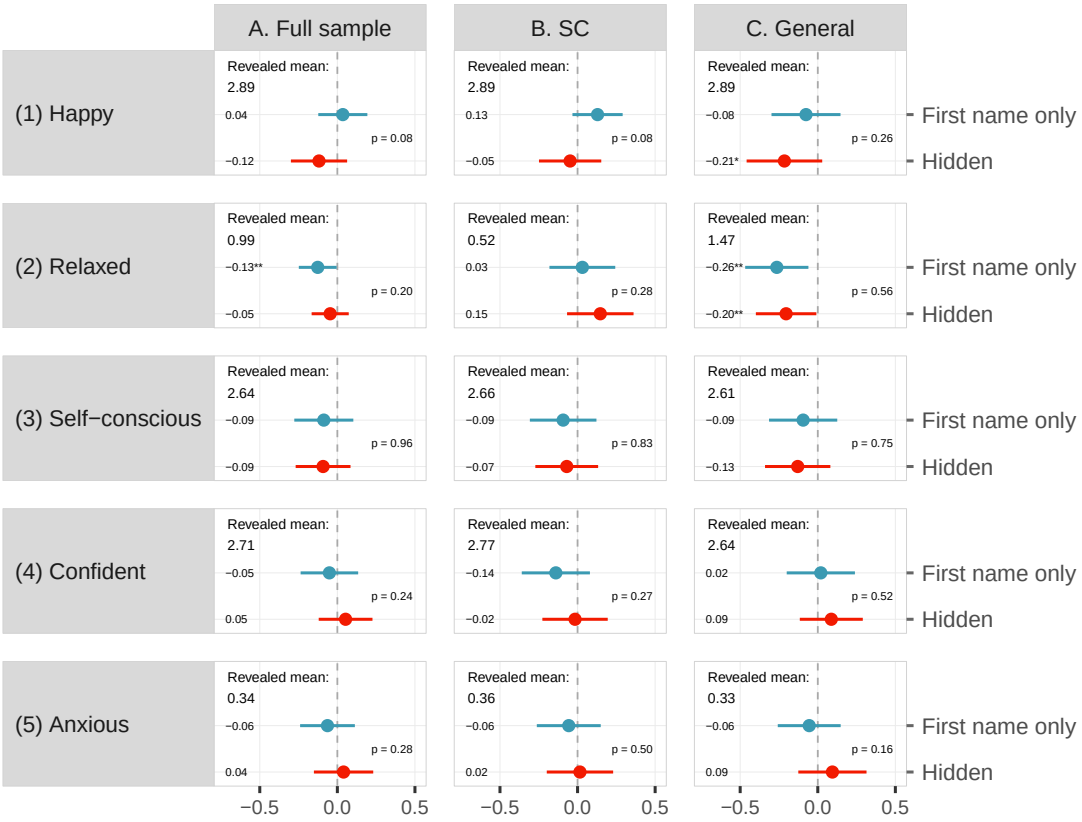
Well-being at work Figure 4 considers the effects of identity knowledge on well-being, as experienced during the four-hour workplace interaction. Across most outcomes and samples, treatment estimates on outcomes that gauge well-being are relatively small and imprecise, with few consistent differences between the *First name only* and *Hidden* conditions relative either to each other or to *Revealed*. Therefore, concealing caste identity does not substantially degrade (or improve) participants' subjective experience of the interaction along these affective dimensions.

One notable exception emerges for relaxation among General caste participants (Figure 4, Panel C, Row 2). In this group, both the *First name only* and *Hidden* conditions reduce feelings of being relaxed at work, with negative and statistically measurable effects compared to *Revealed*. This is striking. If anything, one might have expected that obscuring identity would dampen the psychological strain of cross-group interaction by reducing salience of social divisions. Instead, concealment may introduce a subtle form of discomfort, perhaps by generating ambiguity or preventing the development of interpersonal ease. Importantly, however, this effect appears specific to relaxation and does not extend to other well-being outcomes.

Productivity Figure 5 reports treatment effects on team productivity across the five paired tasks, and on an index measure summarizing those five outcomes. In contrast to the clear and consistently negative effects seen for social ties and trust, the productivity results are noisier and generally null. Neither the *First name only* nor *Hidden* condition produces a statistically significant change in overall productivity, and the difference between the two treatments is small statistically and in magnitude (Panel A). Disaggregating into task components (Panel B) reveals some limited variation. The *First name only* condition has a positive effect on productivity in the ordered data entry task, which was this simplest task; this is true relative to both *Revealed* and *Hidden* groups. For all other tasks, however, the estimated treatment effects are negligible.

This relative lack of effects on realized productivity is perhaps surprising in light of Figure 6, which demonstrates large negative impacts of the treatments on perceptions of productivity—according with the negative effects seen on social relations. Participants in both the *First name only* and *Hidden* conditions report substantially lower assessments of both their own and their partner's productivity, on the order of one-third to one-half of a standard

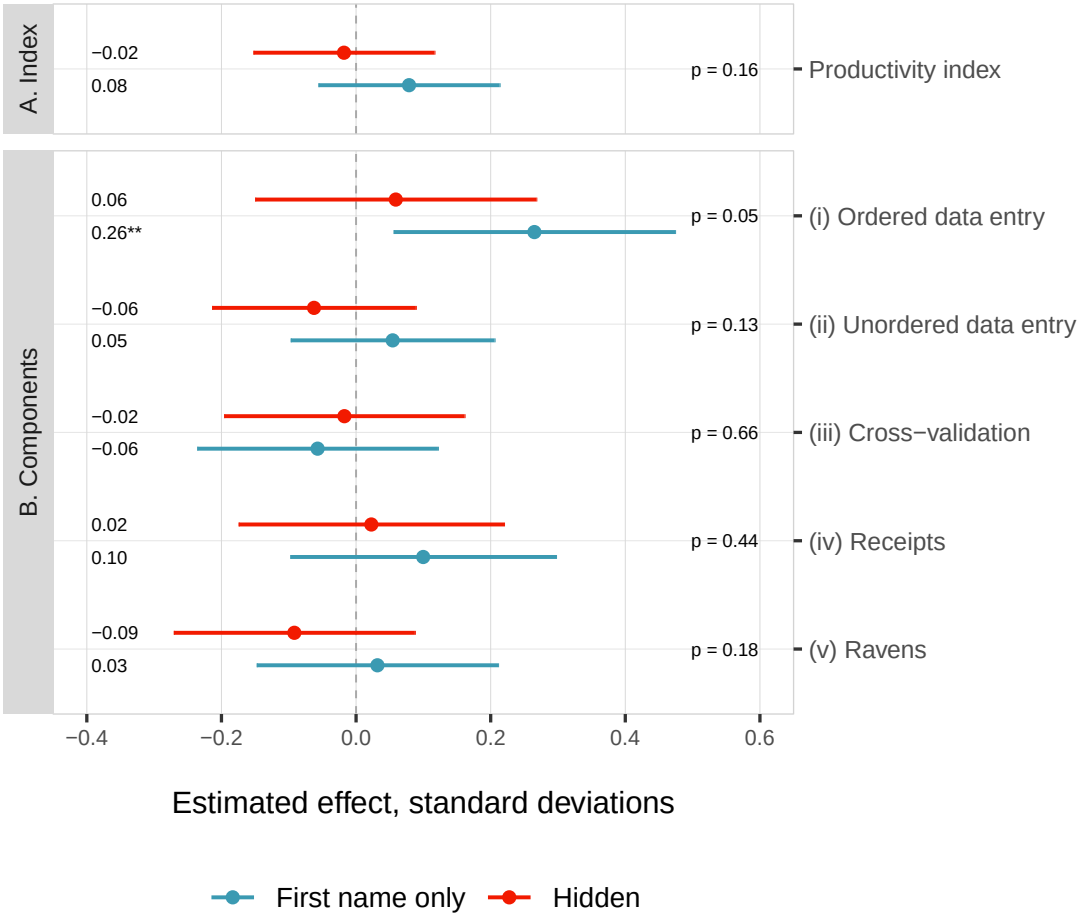
Figure 4: Estimated treatment effects on well-being.



Notes: Each facet panel reports estimates from an OLS regression of an outcome on indicators for assignment to the *First name only* and *Hidden* treatments, with *Revealed* as the omitted category. Outcomes are based on responses to the question, "At work today, to what extent did you feel...[word indicated on each row]" (0 = Not at all, 1 = A little, 2 = Some, 3 = Very). In each facet, "Revealed mean" shows the average score on the original 0–3 point scale for the *Revealed* group. Treatment effects are standardized. In Panels A, standard errors are clustered at the pair level; in Panels B and C (the SC and General subsamples), standard errors are heteroskedasticity-robust. All specifications include randomization strata fixed effects and pre-specified controls for participants' education level and age. $*p < 0.1$; $**p < 0.05$; $***p < 0.01$.

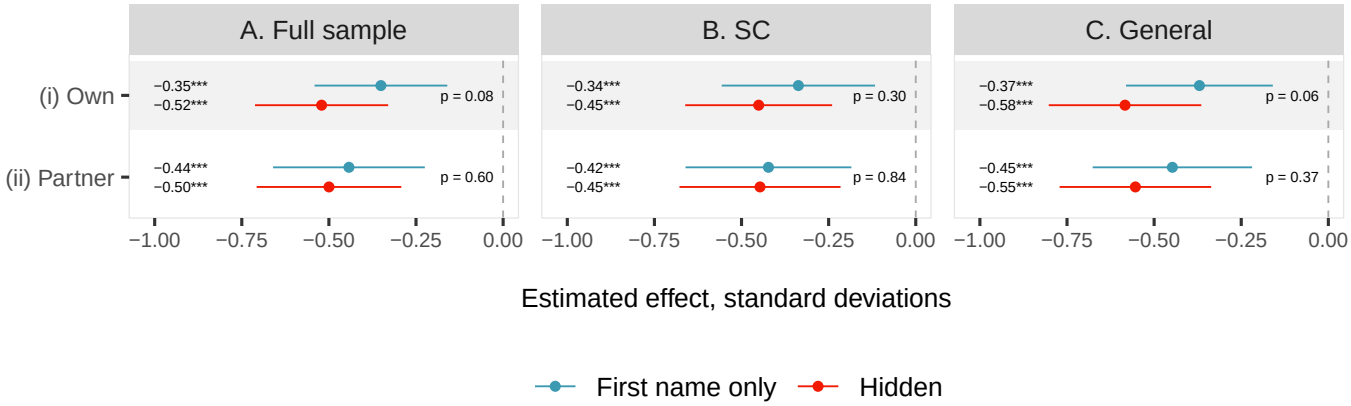
deviation, relative to the *Revealed* condition, with especially pronounced effects among General-caste participants.

Figure 5: Estimated treatment effects on team production.



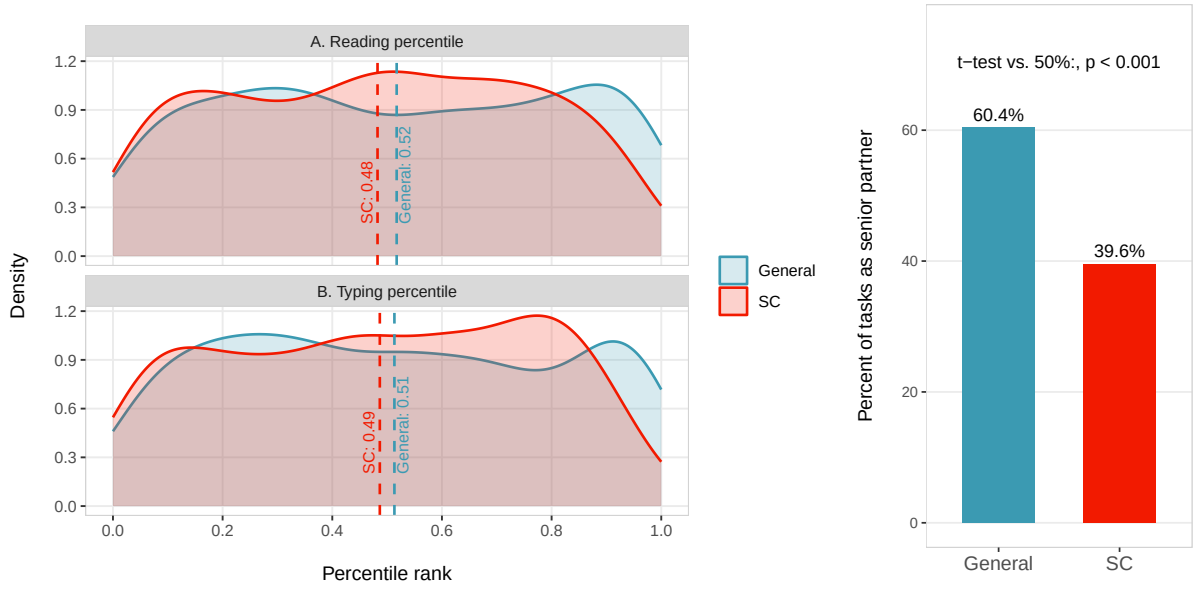
Notes: This figure plots treatment effects on team productivity. All outcomes are counts that measure the number of correctly inputted items for each task. Counts are standardized to z-scores using the mean and standard deviation of the variable in the *Revealed* condition. The index is the simple average of the five standardized component outcomes. All regressions include date and education fixed effects, the average age of the pair, as well as mean baseline levels of productivity for the pair in both reading and typing. Standard errors are clustered at the pair level. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Figure 6: Estimated treatment effects on perceptions of own and partner productivity.



Notes: The figure reports estimated treatment effects on perceived productivity, measured separately for respondents' own performance and that of their partner. Outcomes are standardized to z-scores using the mean and standard deviation of the variable in the *Revealed* condition. All regressions control linearly for participants' age and include randomization-strata and education-bin fixed effects. Standard errors are clustered at the pair level. Horizontal bars denote 95% confidence intervals. Reported p-values correspond to tests of equality between the *First name only* and *Hidden* coefficients. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Figure 7: Baseline productivity differences and senior partner task assignment, by caste.



(a) Baseline productivity distributions

(b) Senior partner role allocation

Notes: Panel (a) shows kernel density estimates of reading and typing percentile ranks for General and SC participants using data from the baseline productivity tasks. Percentiles are constructed within task, with higher values indicating greater productivity. Vertical dashed lines indicate group means. Panel (b) shows the percent of tasks in which individuals took the senior partner role, by caste group. The task for which the senior partner role was randomly assigned for each pair is excluded. The *p*-value is from a two-sided one-sample *t*-test of whether the mean share of senior assignments for General participants differs from 50 percent.

Next, we turn to baseline productivity differences and the allocation of the senior partner role across caste groups (Figure 7). Panel (a) plots the densities of within-task percentile ranks for the baseline reading and typing tasks—tasks that were designed to test the skills needed in the final workplace environment. The distributions for General and SC participants substantially overlap, with General participants having only modestly higher mean percentiles for both tasks (reading: 0.52 vs. 0.48; typing: 0.51 vs. 0.49). Panel (b), meanwhile, examines how the senior partner role is allocated in practice. General participants take the senior role in 60% of tasks, significantly above the 50 percent "fair split" benchmark, while SC participants do so in only 40% percent of tasks. This imbalance arises despite the modest group-wise differences in baseline productivity, suggesting that role assignment reflects factors beyond underlying task ability.

Finally, Table 5 examines whether treatment conditions affect which partner—General or SC—is assigned to the senior partner role. We find no evidence of treatment effects. The coefficients on both the *First name only* and *Hidden* conditions are small in magnitude and statistically indistinguishable from zero across two outcomes: (i) an indicator for whether the General-caste participant takes the senior role (column 1), and (ii) an indicator for

Table 5: Estimated treatment effects on senior partner role assignment.

	General senior (1)	Optimal typer senior (2)
First name only	0.01	0.02
	(0.03)	(0.04)
Hidden	-0.00	0.01
	(0.03)	(0.04)
First name only vs. Hidden p-value	0.83	0.81
Revealed mean	0.60	0.49
N	2,024	2,024

Notes: Each column reports estimates from a linear regression. The unit of analysis is the pair/task; tasks for which the senior role was randomly assigned are excluded. The outcome in column (1) is an indicator for whether the General-caste participant took the senior partner role. The outcome in column (2) is an indicator for whether the worker assigned to the senior (typing) role is the optimal typist. The optimal allocation assigns the senior role to the worker who maximizes total team productivity, defined as the assignment that yields the higher sum of within-sample percentile ranks in baseline reading and typing performance across the pair. All specifications include date fixed effects. Heteroskedasticity-robust standard errors are reported in parentheses. $p < 0.1$, $p < 0.05$, $p < 0.01$.

whether the typist (senior partner) role is assigned "optimally" based on baseline assessments (column 2). Common knowledge regarding identity, it seems, does not alter how roles are allocated within teams.

5.1 Heterogeneity by caste

Our first stage results show that lower caste individuals conceal their identity more successfully than higher caste individuals. This is consistent with concealment being more payoff-relevant for marginalized groups (as in our model), making it more strategically practiced by them. The model offers a complementary prediction: since General caste players always anticipate positive payoffs from mutual revelation, they stand to lose the most when it is prevented or uncertain. We therefore expect the costs of concealment to be higher for General caste individuals in downstream outcomes — even as both groups are negatively affected overall.

In Table A5 we present heterogeneous treatment effects on our core outcomes of identity fusion, willingness to meet partner again, and trust amount sent, separately for Scheduled Caste and General Caste participants. The results show a more negative reaction of concealment among higher-caste individuals than among Scheduled Castes. General caste subjects report lower identity fusion than SCs in both *First Name Only* (p = 0.089) and *Hidden* (p = 0.105), are less willing to meet their partner again (p = 0.04 in *First Name Only* and p = 0.61 in *Hidden*), and send less in the trust game (Rs. -4.63 versus Rs. -2.92 under *Hidden*, p = 0.18, and Rs. -4.02 versus Rs. -3.14 under *First Name Only*, p = 0.49). While we do not always have the statistical power to distinguish these effects, the pattern is clear — General caste individuals penalize identity concealment more than Scheduled Caste participants, consistent

with the model's prediction that those who always anticipate positive payoffs from mutual revelation stand to lose the most when it is prevented or uncertain.

6 Mechanisms

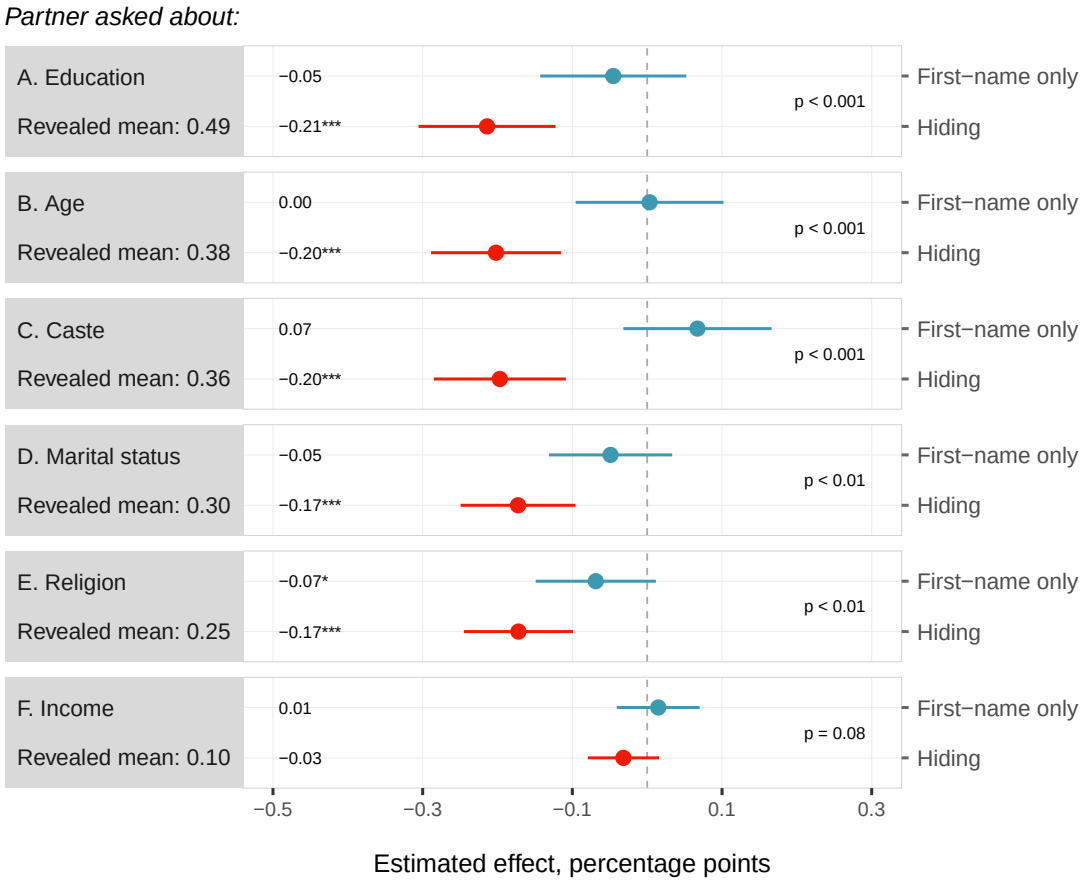
Our findings prompt two main questions. First, in Section 6.1, we ask why does identity-hiding harm social relations and trust? Our analyses suggest that reduced identity legibility harms social relations primarily by increasing uncertainty about who their partner is, and how the interaction should be navigated. We further rule out that these results are not driven by beliefs about team composition. Next, in Section 6.2, we answer why participants don't fully reveal caste in *First Name Only*, to replicate the more positive relations in *Revealed*. We show that concealment may reflect a rational tradeoff in which members of marginalized groups forego some of the gains from trust and social connection in order to insure themselves against severe discrimination.

6.1 Why does identity-hiding harm social relations and trust?

Identity uncertainty We hypothesize that caste identity is synonymous with a host of other observable characteristics in this context, so hiding caste in a conversation effectively removes many of the social cues and levers that people typically use to form bonds. In other words, once caste is suppressed, individuals may no longer be able to converse about topics correlated with caste identity, that they normally rely on to initiate, guide, and strengthen social connections. By contrast, revealing caste restores those signals, which can in some cases help people find common ground and form ties.

Another related possibility with similar implications is that when caste is hidden, people may be more worried about saying something inappropriate or misreading the other person, which introduces uncertainty and awkwardness into the interaction. This kind of elevated uncertainty can itself dampen the quality of the relationship, generating the pattern of negative effects we observe.

Figure 8: *Hiding* diminishes the quality of social interactions.



Notes: The unit of analysis is the participant. The outcomes measure whether the participant reports that their partner attempted to discover specific personal characteristics during the work session. In particular, participants were asked whether their partner tried to discover their age, education level, family income, caste (e.g., by asking their last name), marital status, or religion (0 = No, 1 = Yes). Each point reports the estimated treatment effect relative to the *Revealed* condition. The *Revealed mean* is the raw mean of the outcome for participants in the *Revealed* condition. The *First Name Only* and *Hiding* estimates are coefficients from regressions that control linearly for participants' age and include randomization-strata and education-bin fixed effects. Horizontal bars show 95% confidence intervals based on standard errors clustered at the pair level. Reported p-values test the null hypothesis that the *First Name Only* and *Hiding* treatment effects are equal. $*p < 0.1$; $**p < 0.05$; $***p < 0.01$.

To examine this, we test whether concealing caste by sanction also censors other identity cues that participants typically rely on to form social bonds in this setting. Figure 8 presents evidence consistent with this mechanism. In the *Hidden* treatment, partners not only ask fewer questions about caste itself, but also ask less about other identity markers and characteristics that could indirectly reveal caste yet are commonly used to build rapport—such as questions about education, marital status, or family relations, which are ubiquitous conversational anchors in the Indian context.

Furthermore, in the *First Name Only* condition, there is a clear tendency to ask about caste: both more than in the *Revealed* condition (which is expected given that caste is known almost immediately through names), but significantly more than in the *Hidden* condition. This pattern suggests that when individuals are given even a limited

opportunity to learn about identity, they choose to do so—prima facie evidence that these signals are valuable and helpful for navigating social interactions in this environment.

It is worth noting from Figure 8 that the frequency with which partners discuss personal characteristics is quite similar in the *First Name Only* and *Revealed* conditions. One possible interpretation is that *First Name Only* opens up more potential conversation points than *Hidden*, since subjects may be less inclined to actively avoid topics that could reveal caste identity. However, *First Name Only* still lacks the certainty that comes with explicit revelation of identity (as shown in Table 2). Consistent with this, Figure A1 shows that participants feel less like their true self and report their partner as less friendly in the *First Name Only* condition.[11] The fact that downstream outcomes remain worse in *First Name Only* than in *Revealed* therefore suggests that broader scope for conversation alone is not enough: certainty about identity appears to matter in its own right. In this sense, the negative effects under *Hidden* may reflect both constrained conversation and the absence of identity certainty, whereas the gap between *First Name Only* and *Revealed* is more likely to capture this certainty channel.

In order to test this further, in Table A7 we report treatment effects on four social interaction outcomes, before and after controlling for participants' caste guess confidence and their beliefs about their partner's caste. Panel A replicates the main experimental results: relative to the Revealed arm, both Hiding and First Name Only substantially and significantly reduce trust sent, identity fusion, willingness to meet again, and feelings of authenticity. Panel B then adds controls for caste guess confidence (rescaled 0–1). Two findings stand out. First, caste confidence is strongly associated with better outcomes across all four measures — a participant who moves from zero to full confidence in their caste guess sends roughly Rs. 17.5 ($p < 0.01$) more in the trust game, reports 1.4 more units of identity fusion ($p < 0.01$) , and scores 5.6 points ($p < 0.01$) higher on feeling like their true self.

Second, and strikingly, the treatment dummies largely disappear once confidence is controlled for: the *Hidden* and *First Name Only* coefficients shrink toward zero and lose significance for each of the outcomes, suggesting that the negative effect of these treatments on social relations operates primarily through the uncertainty they generate about a partner's caste identity. While this evidence is correlational — caste confidence is itself a product of the treatment — but it is consistent with uncertainty being the primary mechanism linking caste concealment to worse social interactions.

Uncertainty over partner identity or beliefs about team composition? Our *Hidden* and *First Name Only* treatments reduce certainty about a partner's caste identity: relative to the *Revealed* arm, participants in the *Hiding* and

[11] Figure A1 also shows that concealment does not increase overtly negative partner behavior. On outcomes such as partner rudeness, bossiness, and felt tension, point estimates are near zero or slightly positive relative to Revealed (but not statistically significant). In other words, while concealment weakens positive relational outcomes (closeness, willingness to re-engage, self-expression), it does not translate into higher reports of mistreatment in the booth. A useful interpretation is that interactions under concealment look more distant rather than more hostile: pairs remain professional, but are less likely to develop rapport.

First Name Only arms are substantially less confident in their caste guess and expect their partner to be less certain too. Second, and more subtly, they shift beliefs about whether one's partner shares one's own caste. In the Revealed arm, participants know with certainty that they are in a cross-caste pair — by design, all HL pairs are heterogeneous. In the Hiding and First Name Only arms, however, participants have no such knowledge, and may reasonably assign non-trivial probability to being paired with a same-caste partner. Could our results be driven by the latter? It seems unlikely because if anything one would expect the effect of caste homogeneity to be positive for our outcomes, and *Hidden* and *First Name Only* have negative effects. But we are able to test this using the homogeneous pairs we have in our sample. By comparing outcomes across teams where identity is revealed but the difference is in composition, we can test the direction of the composition effect. As expected, Table A6 shows that homogeneous teams have more trust and greater identity fusion, though the effect on willingness to meet again is not significant. This suggests that the effect of expecting — or having — a same-caste partner is if anything positive, and therefore cannot explain our core results.

Does asking partner's identity increase trust? If identity uncertainty leads to negative effects on social relationships, reduced uncertainty should positively affect outcomes. This is clearly the case when comparing the *Revealed* arm to our other conditions, but is it also true when revelation is endogenous or when identity is confirmed by partners asking each other?

Although subjects were randomly assigned to treatments with different instructions about revealing or concealing their identity, they were neither forced nor penalized for non-compliance. Even so, we do observe a strong first stage in terms of treatment compliance as shown in Table 2.

This naturally raises the question: among those who ended up learning their partner's caste—whether by asking their partner's last name or by asking about other attributes that implicitly revealed caste—did having this extra information actually improve their relationship with the partner? To examine this, we correlate (i) asking about the partner's caste (either explicitly through last name or indirectly through questions about neighborhood, occupation, schooling, etc.) and (ii) being asked about one's own caste, with the trust expressed toward the partner at endline.

Table 6: Asking about identity increases trust

	General Caste		Scheduled Caste	
	(1)	(2)	(3)	(4)
	Trust: Sent	Trust: Belief	Trust: Sent	Trust: Belief
Partner asked your caste?	1.004	-1.147	4.014**	-0.078
	(2.011)	(2.057)	(2.011)	(1.857)
Asked partner their caste?	4.057**	4.463**	1.253	2.174
	(2.052)	(2.105)	(1.925)	(1.779)
Mean Dep. Var.	25.03	22.5	24.3	21.5
Observations	564	564	563	563
R-squared	0.128	0.124	0.149	0.133

Notes: Robust standard errors reported in parenthesis. Columns (1)–(2) restrict the sample to General-caste participants and columns (3)–(4) to SC participants. $e_did_p_ask_caste = 1$ if the partner asked about caste; $e_did_ask_p_caste = 1$ if the respondent asked the partner's caste. All specifications include date fixed effects and pre-specified controls for baseline education and age. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 6 shows that this is indeed the case. Participants who report having asked their partner's caste seem to trust them more (Columns 1 and 3), and also believe that their partner would trust them more (Columns 2 and 4, but these effects are not statistically significant). These effects are stronger for General Castes than for Scheduled Castes: which suggests that that higher-caste participants may especially dislike the uncertainty of not knowing their partner's identity. The results also suggest that for lower castes even being asked about their caste is somewhat helpful (Column 1): they tend to send larger amounts to higher-caste partners once the partner learns their identity, perhaps because discovering that revealing did not backfire in this setting reassures them. We cannot, of course, distinguish learning through last names from learning through other attributes, which may also function as social cues or increase familiarity, but the overall pattern closely aligns with our causal findings.

6.2 Why don't participants fully reveal caste in *First Name Only*?

Why is identity concealment prevalent in many parts of India, if it results in negative outcomes in social interactions? Relatedly, why don't participants fully reveal their identity in the *First Name Only* condition, to gain benefits of disclosure?

One possibility is that concealing identity may be optimal in less intimate, low-stakes interactions than those we study—fleeting encounters with many individuals in rotating labor markets. In such settings, the potential downside of discrimination may loom larger than the benefits of disclosure, making concealment a more attractive strategy.

However, anecdotal evidence suggests that concealment is common even in more intimate settings (Satyanarayanan and Lee, 2023).

A second possibility, driven by our theory, is that the costs of concealment are heterogeneous across the outcome distribution. Under *First Name Only*, positive selection into mutual revelation means that pairs with the highest potential gains from disclosure are most likely to realize it. This suggests that the *First Name Only* condition should approach *Revealed* at the upper tail. Meanwhile, since low-y lower-caste individuals prefer concealment, *Hidden* should have the smallest costs and possibly benefits (relative to *Revealed*) precisely where interactions would otherwise have been poor.

We test these predictions using a quantile regression framework, which allows us to examine how treatment effects vary across the outcome distribution rather than focusing on average effects alone. If this channel were true, one would expect the negative effects under the *Hidden* condition to be most prominent when interactions are generally positive – where disclosure would have helped strengthen relationships. Whereas concealment may offer protection or insurance in cases where interactions are more hostile.

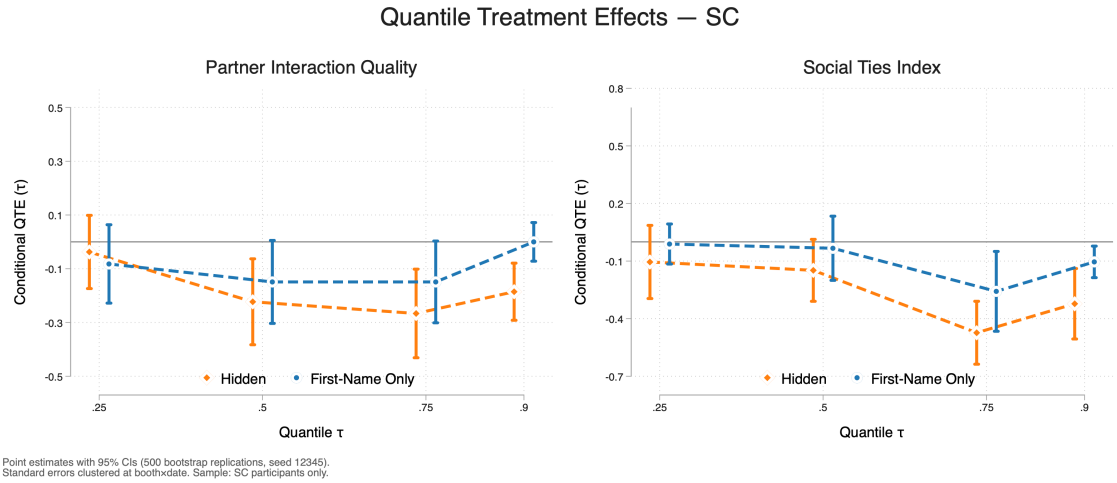
Figure 9 reports quantile treatment effects on an index of (a) interaction quality with the partner and (b) downstream outcomes—trust, willingness to meet, and reservation wage for working together. Both sets of results paint a consistent picture. In the lower quantiles of the outcome distributions, the *Hidden* (and *First Name Only*) condition has relatively little downside: interaction quality is not made more negative at these points, and the adverse effects on trust and cooperation are muted and not statistically significant. But concealment becomes costly when interactions are relatively more positive. The negative effects on social ties are stronger for General Castes, consistent with them being less comfortable with non-disclosure.

The results also display a clear non-monotonicity for the *First Name Only* condition relative to *Revealed*, reinforcing this interpretation. The *First Name Only* condition consistently produces smaller negative effects than *Hidden*, suggesting that simply having the option to reveal meaningfully reduces the costs of concealment—consistent with the correlations documented earlier. On the right tail of the outcome distribution, the effects of *First Name Only* actually become less negative (similar to the left-tail), indicating that when subjects expect particularly positive interactions, they perhaps tend to reveal more organically – to benefit from stronger social ties.[12] Taken together, these patterns align with the idea of a tension: concealing identity can help avoid negative outcomes, but it also limits access to social levers that build stronger ties and generate both pecuniary and non-pecuniary benefits. The option to reveal or conceal may be very important, and the optimal choice may thus depend on the underlying risk expectation or environment.

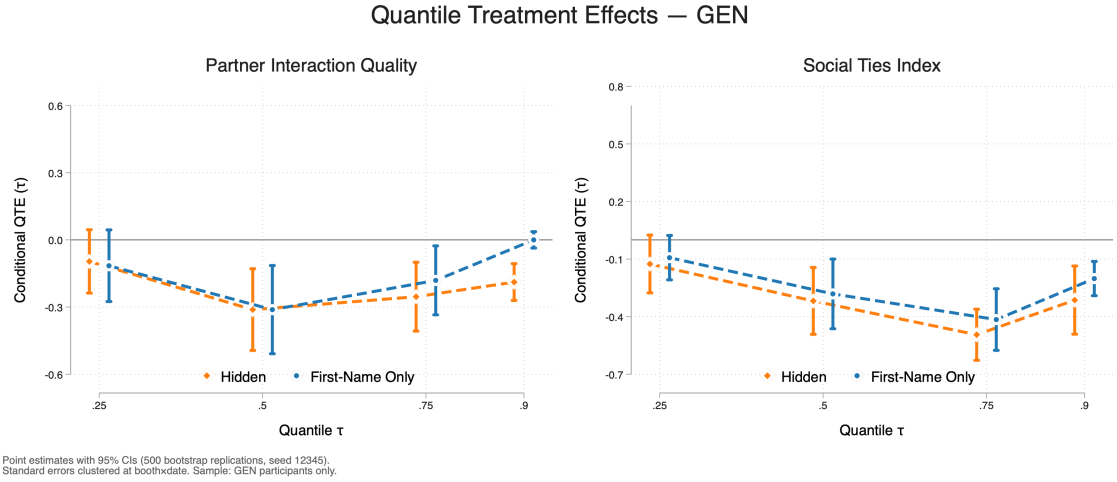
[12]This is consistent with partner-friendliness being positively associated with identity revelation.

Figure 9: Quantile treatment effects of identity-*Hidden* treatments on interaction quality and downstream outcomes

Panel A: Scheduled Caste (SC)



Panel B: General Caste (GEN)



Notes: Each panel reports conditional quantile treatment effects (QTEs) of the *Hidden* and *First-name only* treatments relative to the Revealed group at different quantiles τ of the outcome distribution. Vertical lines denote 95% confidence intervals. Panel A shows results for Scheduled Caste (SC) participants; Panel B for General caste participants. The left column reports QTEs for an index summarizing the quality of interaction with the partner, constructed from responses on perceived friendliness, rudeness, tension, comfort in being one's true self, and degree of identity fusion with the partner. The right column shows QTEs for an index capturing downstream trust and cooperative behavior (trust giving, trust returned, willingness to have a meal with partner, and reservation wage to work with partner in the future). Positive values indicate better outcomes. Point estimates with 95% confidence intervals based on 500 bootstrap replications, clustered at booth$\times$date.

Finally, the fact that the negative effects are concentrated in the upper quantiles in the *Hidden* and *First-Name Only* conditions, without offsetting gains in the lower tail, likely reflects the fact that our experiment captures ordinary day-to-day relational dynamics rather than the very $low - y$ interactions the model admits – where individuals anticipate severe negative payoffs from cross-caste revelation. In less controlled real-world settings, where revealing one's identity may sometimes provoke severe negative reactions, the protective value of concealing identity could be even larger. Our results therefore highlight the basic relational mechanism, while its real-world effects may be even stronger.

7 Conclusion

This paper documents an unexpected set of effects of identity concealment. In contrast to the conventional view that suppressing salient social identities may facilitate smoother interaction, we find that *Hidden* caste identity substantially weakens social ties in Uttar Pradesh, India. In our field experiment involving joint production in pairs, participants who conceal identity report lower closeness, reduced authenticity, and diminished willingness to interact again — and they also exhibit lower trust and cooperation in incentivized outcomes post work. These relational costs are large and robust, while effects on team productivity are mixed and do not point to systematic gains from concealment.

Our quantile treatment effects shed light on why identity *Hidden* may nonetheless be prevalent in practice. The negative effects of concealment are concentrated in the upper part of the outcome distribution—where interactions, trust, and cooperation would otherwise have been relatively positive. In contrast, we find little evidence of harm in the lower tail of the distribution. The institutional setting in our experiment naturally does not allow for extreme discriminatory behaviors by participants; but the patterns are consistent with a tradeoff in which concealment limits the upside of social interaction while offering insurance against relatively rate but damaging experience. Perhaps under the possibility of extreme discrimination—as it plausibly is in many settings — identity concealment could even yield benefits in the left tail that our experiment cannot capture.

This is likely to be a productive direction for future work on this topic. Experimental designs are perhaps not well suited to observing rare but extreme discriminatory events, making it difficult to directly test the insurance value of concealment. While such events may be more prevalent in observational settings (for example, within firms or labor markets), it is often difficult in those contexts to recover individuals' underlying motivations without simultaneously collecting rich survey data. We therefore see an important role for complementary descriptive and anthropological approaches—such as surveys and qualitative interviews that document when and why individuals choose to hide their identities, what risks they seek to avoid, and what costs they perceive. Such approaches may help clarify both

the motivations underlying concealment and the conditions under which it is welfare-improving.

References

Abdelgadir, Aala and Vasiliki Fouka, "Political secularism and Muslim integration in the West: Assessing the effects of the French headscarf ban," *American Political Science Review*, 2020, *114* (3), 707–723.

Adukia, Anjali, Richard Hornbeck, Daniel Keniston, and Benjamin Lualdi, "The social construction of race during reconstruction," Technical Report, National Bureau of Economic Research 2025.

Akerlof, George A and Rachel E Kranton, "Economics and identity," *The quarterly journal of economics*, 2000, *115* (3), 715–753.

_ and _, "Identity economics: How our identities shape our work, wages, and well-being," in "Identity Economics," Princeton University Press, 2010.

Alesina, Alberto and Eliana La Ferrara, "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, 2005, *43* (3), 762–800.

Anderson, Siwan, "Caste as an Impediment to Trade," *American Economic Journal: Applied Economics*, 2011, *3* (1), 239–263.

Angeli, Deivis, Ieda Matavelli, and Fernando Secco, "Expected discrimination and Job Search," Technical Report, Working Paper 2026.

Asher, Sam, Kritarth Jha, Paul Novosad, Anjali Adukia, and Brandon Tan, "Residential segregation and unequal access to local public services in India: Evidence from 1.5 m neighborhoods," Technical Report, National Bureau of Economic Research 2026.

Avery, Mallory, Andreas Leibbrandt, and Joseph Vecci, "Does artificial intelligence help or hurt gender diversity? Evidence from two field experiments on recruitment in tech," 2024.

Banerjee, Abhijit, Esther Duflo, Maitreesh Ghatak, and Jeanne Lafortune, "Marry for what? Caste and mate selection in modern India," *American Economic Journal: Microeconomics*, 2013, *5* (2), 33–72.

Behaghel, Luc, Bruno Crépon, and Thomas Le Barbanchon, "Unintended Effects of Anonymous Résumés," *American Economic Journal: Applied Economics*, July 2015, *7* (3), 1–27.

Berg, Joyce, John Dickhaut, and Kevin McCabe, "Trust, reciprocity, and social history," *Games and economic behavior*, 1995, *10* (1), 122–142.

Bertrand, M. and E. Duflo, "Field Experiments on Discrimination," *Handbook of Economic Field Experiments*, January 2017, *1*, 309–393.

Bertrand, Marianne and Sendhil Mullainathan, "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination," *American Economic Review*, September 2004, *94* (4), 991–1013.

Boring, Anne, Katherine Coffman, Dylan Glover, and Marıa José González-Fuentes, "Discrimination, Rejection, and Willingness to Apply: Effects of Blind Hiring Processes," Technical Report, Working Paper 2025.

Brewer, Marilynn B and N Miller, "Beyond the contact hypothesis: Theoretical perspectives on desegregation," *Groups in contact: The psychology of desegregation*, 1984, *281*.

Carpio, Lucia Del and Thomas Fujiwara, "Spillovers and Scalability in Job Ad Experiments: Evidence from Gender-Neutral Language," Technical Report, Working Paper 2026.

Cassan, Guilhem, "Identity-based policies and identity manipulation: Evidence from colonial Punjab," *American Economic Journal: Economic Policy*, 2015, *7* (4), 103–131.

_ , Daniel Keniston, and Tatjana Kleineberg, "A division of laborers: Identity and efficiency in India," Technical Report, National Bureau of Economic Research 2021.

Chandra, Kanchan, *Why ethnic parties succeed: Patronage and ethnic head counts in India*, Cambridge University Press, 2007.

Charness, Gary, Ramón Cobo-Reyes, Simone Meraglia, and Ángela Sánchez, "Anticipated discrimination, choices, and performance: Experimental evidence," *European Economic Review*, 2020, *127*, 103473.

Dahis, Ricardo, Emily Nix, and Nancy Qian, "Choosing racial identity in the United States, 1880-1940," Technical Report, National Bureau of Economic Research 2019.

_ , Ludovica Mosillo, Emily Nix, and Nancy Qian, "Black-White Income Inequality During Jim Crow: Evidence from "Passing" for White," Technical Report 2026.

Dovidio, John F, Samuel L Gaertner, and Kerry Kawakami, "Intergroup contact: The past, present, and the future," *Group processes & intergroup relations*, 2003, *6* (1), 5–21.

Fix, Michael and Raymond J. Struyk, *Clear and Convincing Evidence: Measurement of Discrimination in America*, Washington, DC: Urban Institute Press, 1993.

Ghosh, Arkadev, "Religious Divisions and Production Technology: Experimental Evidence from India," *Journal of Political Economy*, 2025.

_ , Prerna Kundu, Matt Lowe, and Gareth Nellis, "Creating cohesive communities: A youth camp experiment in india," *Review of Economic Studies*, 2025, p. rdaf026.

Goldin, Claudia and Cecilia Rouse, "Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians," *American Economic Review*, 2000, *90* (4), 715–741.

Hewstone, Miles and Rupert Brown, "Contact is Not Enough: An Intergroup Perspective on the 'Contact Hypothesis'," in Miles Hewstone and Rupert Brown, eds., *Contact and Conflict in Intergroup Encounters*, Oxford: Basil Blackwell, 1986, pp. 1–44.

Hjort, Jonas, "Ethnic Divisions and Production in Firms," *Quarterly Journal of Economics*, 2014, *129* (4), 1899–1946.

Hoff, Karla and Priyanka Pandey, "Discrimination, social identity, and durable inequalities," *American economic review*, 2006, *96* (2), 206–211.

Horowitz, Donald L, *Ethnic groups in conflict, updated edition with a new preface*, Univ of California Press, 2000.

Kang, Sonia K, Katherine A DeCelles, András Tilcsik, and Sora Jun, "Whitened résumés: Race and self-presentation in the labor market," *Administrative Science Quarterly*, 2016, *61* (3), 469–502.

Kline, Patrick, Evan K Rose, and Christopher R Walters, "Systemic discrimination among large US employers," *Quarterly Journal of Economics*, 2022, *137* (4), 1963–2036.

Lepage, Louis-Pierre, Xiaomeng Li, and Basit Zafar, "Anticipated gender discrimination and grade disclosure," Technical Report, National Bureau of Economic Research 2022.

Lowe, Matt, "Types of Contact: A Field Experiment on Collaborative and Adversarial Caste Integration," *American Economic Review*, 2021, *111* (6), 1807–1844.

Miller, Norman, "Personalization and the promise of contact theory," *Journal of Social Issues*, 2002, *58* (2), 387–410.

Mitrut, Andreea, Gabriel Kreindler, Margareta Matache, Andrei Munteanu, and Cristian Pop-Eleches, "Education and Selection into Ethnic Identification: Evidence from Roma People in Romania," Technical Report, National Bureau of Economic Research 2025.

Mousa, Salma, "Overcoming the Trust Deficit: Associational Life and Inter-Group Contact in Post-ISIS Iraq," *Mimeo*, 2018.

Munshi, Kaivan, "Caste and the Indian economy," *Journal of Economic Literature*, 2019, *57* (4), 781–834.

Oh, Suanna, "Does identity affect labor supply?," *American Economic Review*, 2023, *113* (8), 2055–2083.

Pachankis, John E, Mark L Hatzenbuehler, Katie Wang, Charles L Burton, Forrest W Crawford, Jo C Phelan, and Bruce G Link, "The burden of stigma on health and well-being: A taxonomy of concealment, course, disruptiveness, aesthetics, origin, and peril across 93 stigmas," *Personality and Social Psychology Bulletin*, 2018, *44* (4), 451–474.

Pettigrew, Thomas and Linda Tropp, "A Meta-Analytic Test of Intergroup Contact Theory," *Journal of Personality and Social Psychology*, 2006, *90* (5), 751.

Rose, Evan K, "A constructivist perspective on empirical discrimination research," *Journal of Economic Literature*, 2023, *61* (3), 906–923.

Ruebeck, Hannah, "Perceived discrimination at work," Technical Report, Working Paper 2023.

Satyanarayanan, Kusuma and Joel Lee, *Concealing caste: Narratives of passing and personhood in dalit literature*, Oxford University Press, 2023.

Shukla, Soumitra, "Making the Elite: Class Discrimination at Multinationals," *Working Paper*, 2025.

Srinivas, M. N., *Social Change in Modern India*, Berkeley and Los Angeles: University of California Press, 1966.

Tajfel, Henri, M. G. Billig, R. P. Bundy, and Claude Flament, "Social categorization and intergroup behaviour," *European Journal of Social Psychology*, April 1971, *1* (2), 149–178.

Valmiki, Omprakash, *Joothan: an Untouchable's life*, Columbia University Press, 2008.

A Proofs

Proof. (Proposition 1)

Since caste is hidden in the First Name Only treatment, players evaluate their actions using the prior distribution over the partner's caste and disclosure behavior.

Consider first a lower-caste player of type y_i. If he reveals, then:

- with probability $\frac{1}{2}p_L$, the partner is lower-caste and reveals, yielding payoff x;
- with probability $\frac{1}{2}p_G$, the partner is General-caste and reveals, yielding payoff y_i;
- otherwise, mutual revelation does not occur, and payoff is z.

Hence

$$U_L(R \mid y_i) = \frac{1}{2}p_L x + \frac{1}{2}p_G y_i + \left(1 - \frac{1}{2}p_G - \frac{1}{2}p_L\right) z.$$

If instead he hides, then he receives payoff z whenever the partner reveals and payoff 0 otherwise:

$$U_L(H \mid y_i) = \left(\frac{1}{2}p_G + \frac{1}{2}p_L\right) z.$$

Therefore a lower-caste player reveals if and only if

$$\frac{1}{2}p_L x + \frac{1}{2}p_G y_i + \left(1 - \frac{1}{2}p_G - \frac{1}{2}p_L\right) z \geq \left(\frac{1}{2}p_G + \frac{1}{2}p_L\right) z.$$

Rearranging,

$$\frac{1}{2}p_G y_i + \frac{1}{2}p_L x + z(1 - p_G - p_L) \geq 0.$$

If $p_G > 0$, this is equivalent to

$$y_i \geq y^*(p_G, p_L) \equiv -\frac{p_L}{p_G}x - \frac{2z(1 - p_G - p_L)}{p_G}.$$

The expected payoff comparison implies that, the gain from revealing is strictly increasing in y whenever $p_G > 0$. Therefore the lower-caste best response is monotone: there exists a threshold $y^* \in [\underline{y}, x]$ such that

$$s_L(y) = \begin{cases} H & \text{if } y < y^*, \\ R & \text{if } y \geq y^*. \end{cases}$$

First consider $s_G = H$. Thus $p_G = 0$, and the payoff difference for a lower-caste player is

$$U_L(R|y) - U_L(H|y) = \frac{1}{2}p_L x + (1 - p_L)z.$$

This expression does not depend on y, so all lower-caste players must take the same action. If all revealed ($p_L = 1$), then General-caste players would strictly prefer revealing since

$$U_G(R) - U_G(H) = \frac{1}{2}x > 0,$$

contradicting $p_G = 0$. Therefore the only equilibrium with $p_G = 0$ is

$$s_G = H, \qquad s_L(y) = H.$$

Next consider $s_G = R$, so $p_G = 1$. From the payoff comparison derived earlier,

$$U_G(R) - U_G(H) = \frac{x + p_L(x-2z)}{2} > 0$$

for all $p_L \in [0,1]$, so revealing is always optimal for General-caste players.

For a lower-caste player of type y,

$$U_L(R|y) - U_L(H|y) = \frac{1}{2}\big(y + p_L(x-2z)\big),$$

which is strictly increasing in y. Hence lower-caste best responses take cutoff form with threshold

$$y^* = -p_L(x-2z).$$

Consistency requires

$$p_L = 1 - F(y^*) = 1 - F\big(-p_L(x-2z)\big).$$

Define

$$\phi(p) = p - \big(1 - F(-p(x-2z))\big).$$

Since F is continuous, ϕ is continuous on $[0,1]$. Moreover

$$\phi(0) = F(0) - 1 \le 0, \qquad \phi(1) = F(-(x-2z)) \ge 0.$$

Thus a solution $p_L^* \in [0,1]$ exists by the Intermediate Value Theorem.

Finally, full revelation cannot be an equilibrium because the lowest lower-caste type would reveal only if

$$\underline{y} \ge -x + 2z,$$

which is ruled out by assumption. Therefore the only equilibria are the two described above. □

Proof. (Remark 1) Under *Hidden*, all mixed-caste pairs end in mutual concealment, so

$$\mathbb{E}[Q_{Hidden}] = q_{HH}.$$

Under *First Name Only*,

$$\mathbb{E}[Q_{FN}] = s\Big[p_L\,\mathbb{E}[q(y_i) \mid y_i \ge y^*] + (1-p_L)q_{RH}\Big] + (1-s)q_{HH}.$$

Therefore

$$\begin{aligned}\mathbb{E}[Q_{FN}] - \mathbb{E}[Q_{Hidden}] &= s\Big[p_L\,\mathbb{E}[q(y_i) \mid y_i \ge y^*] + (1-p_L)q_{RH} - q_{HH}\Big] \\ &= s\Big[p_L\big(\mathbb{E}[q(y_i) \mid y_i \ge y^*] - q_{HH}\big) - (1-p_L)(q_{HH} - q_{RH})\Big].\end{aligned}$$

Hence $\mathbb{E}[Q_{FN}] \ge \mathbb{E}[Q_{Hidden}]$ whenever

$$p_L\big(\mathbb{E}[q(y_i) \mid y_i \ge y^*] - q_{HH}\big) \ge (1-p_L)(q_{HH} - q_{RH}).$$

For the comparison with *Revealed*,

$$\mathbb{E}[Q_{Revealed}] = \mathbb{E}[q(y_i)].$$

So

$$\mathbb{E}[Q_{FN}] - \mathbb{E}[Q_{Revealed}] = (1-s)\big(q_{HH} - \mathbb{E}[q(y_i)]\big) + s(1-p_L)\Big(q_{RH} - \mathbb{E}[q(y_i) \mid y_i < y^*]\Big),$$

whose sign is in general ambiguous as it depends on s. □

B Supplementary Tables

Table A1: Table: Baseline summary statistics for the experimental sample.

	N	Mean	St. Dev.	Min.	Pctl. 25	Median	Pctl. 75	Max.
(a) Age	1198	23.30	4.34	18.00	20.00	22.00	26.00	40.00
(b) Number of siblings	1198	2.86	1.14	0.00	2.00	3.00	4.00	8.00
(c) Household size	1198	5.91	1.71	1.00	5.00	6.00	7.00	20.00
(d) Married	1198	0.27	0.45	0.00	0.00	0.00	1.00	1.00
(e) Number of children (if married)	329	1.38	0.95	0.00	1.00	1.00	2.00	9.00
(f) Household income (midpoint, INR)	1198	10659	8185	2500	2500	7500	17500	55000
(g) Completed 12th standard	1198	0.73	0.44	0.00	0.00	1.00	1.00	1.00
(h) Completed college	1198	0.20	0.40	0.00	0.00	0.00	0.00	1.00
(i) Identity ranking: Caste	1198	3.70	0.99	1.00	3.00	4.00	4.00	5.00
(j) Identity ranking: Religion	1198	3.69	1.11	1.00	3.00	4.00	5.00	5.00
(k) Identity ranking: Indian	1198	1.31	0.83	1.00	1.00	1.00	1.00	5.00
(l) Identity ranking: From U.P.	1198	3.33	1.36	1.00	2.00	3.00	5.00	5.00
(m) Identity ranking Father/Son/Brother	1198	2.97	1.16	1.00	2.00	3.00	4.00	5.00
(n) Overall well-being score (% of Max Possible)	1198	45.22	11.73	33.33	40.00	40.00	40.00	93.33

Notes: This table presents summary statistics for variables collected at baseline. Household income (row (f)) is coded as the midpoint of the income bin, or the minimum value for the top-coded bin. Identity rankings (rows (i)–(m)) are measured on a 1 (most) to 5 (least) ranking scale. Well-being score in row (n) is a composite index normalized to the 0–100 range. All statistics are based on the experimental sample.

Table A2: Baseline differences in means (GEN vs. SC)

	General Caste	Scheduled Caste	Diff. (GEN−SC)	*N*
(a) Age	23.07	23.59	−0.52*	1,064
	(4.24)	(4.46)	(0.27)	
(b) Number of siblings	2.69	3.11	−0.42***	1,064
	(1.10)	(1.16)	(0.07)	
(c) Household size	5.95	5.93	0.03	1,064
	(1.88)	(1.62)	(0.11)	
(d) Married	0.24	0.31	−0.07**	1,064
	(0.43)	(0.46)	(0.03)	
(e) Number of children	1.46	1.29	0.16	296
	(0.77)	(1.08)	(0.11)	
(f) Household income (midpoint, INR)	12,307.33	9,064.85	3,242.48***	1,064
	(9,325.06)	(6,437.23)	(491.27)	
(g) Completed 12th standard	0.77	0.67	0.10***	1,064
	(0.42)	(0.47)	(0.03)	
(h) Completed college	0.24	0.17	0.07***	1,064
	(0.43)	(0.38)	(0.02)	
(i) Identity ranking: Caste	3.88	3.51	0.37***	1,064
	(0.97)	(0.96)	(0.06)	
(j) Identity ranking: Religion	3.92	3.46	0.46***	1,064
	(1.04)	(1.10)	(0.07)	
(k) Identity ranking: Indian	1.19	1.45	−0.26***	1,064
	(0.66)	(0.98)	(0.05)	
(l) Identity ranking: From U.P.	2.91	3.75	−0.83***	1,064
	(1.27)	(1.33)	(0.08)	
(m) Identity ranking: Father/Son/Brother	3.10	2.83	0.27***	1,064
	(1.00)	(1.29)	(0.07)	
(n) Overall well-being score (% of max.)	44.50	46.14	−1.64**	1,064
	(11.58)	(12.08)	(0.73)	

Notes: This table reports baseline differences in means between General and Scheduled caste respondents. Columns (1) and (2) report group means with standard deviations in parentheses. Column (3) reports the difference in means (GEN minus SC); standard errors are shown in parentheses below. Household income is coded as the midpoint of the income bin (minimum for the top-coded bin). Identity rankings are measured on a scale from 1 (most important) to 5 (least important). The overall well-being index is normalized to range from 0 to 100.
Significance levels: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A3: Sample Selection

Sample	Baseline only (1)		Experimental (2)		Diff (2–1)		N	
Variable	Mean	SE	Mean	SE	Mean	SE	1	2
(a) Age	24.21	0.24	23.30	0.13	–0.919***	0.252	448	1198
(b) Number of siblings	2.99	0.06	2.86	0.03	–0.130**	0.065	447	1198
(c) Household size	5.90	0.09	5.91	0.05	0.012	0.098	432	1198
(d) Married	0.29	0.02	0.27	0.01	–0.018	0.024	479	1198
(e) Number of children	1.42	0.08	1.38	0.05	–0.041	0.094	140	329
(f) HH income (INR '000)	10.21	0.35	10.66	0.24	0.451	0.448	431	1198
(g) Completed 12th standard	0.74	0.02	0.73	0.01	–0.012	0.025	431	1198
(h) Completed college	0.25	0.02	0.20	0.01	–0.044*	0.023	431	1198
(i) Identity ranking: Caste	3.70	0.05	3.70	0.03	–0.001	0.056	421	1198
(j) Identity ranking: Religion	3.46	0.05	3.69	0.03	0.224***	0.063	421	1198
(k) Identity ranking: Indian	1.24	0.04	1.31	0.02	0.068	0.046	421	1198
(l) Identity ranking: From U.P.	3.42	0.07	3.33	0.04	–0.088	0.077	421	1198
(m) Identity ranking: Father/Son/Brother	3.18	0.06	2.97	0.03	–0.203***	0.066	421	1198
(n) Overall well-being score (% of Max Possible)	45.95	0.66	45.22	0.34	–0.726	0.726	342	1198

Notes: Means and standard errors (SE) by group. The first group represents the sample of individuals who participated in the baseline survey, but did not take part in the final experiment, whereas the second group represents the latter. Household income (row (f)) is coded as the midpoint of the income bin, or the minimum value for the top-coded bin and is shown in thousands of INR. Identity rankings (rows (i)–(m)) are measured on a 1 (most) to 5 (least) ranking scale. Well-being score in row (n) is a composite index normalized to the 0–100 range. Stars: $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A4: Randomization Checks

	Control (First Name)		Revealed vs Control		*Hidden* vs Control		p: Revealed = *Hidden*
	Mean	SD	β	p-value	β	p-value	p-value
(a) Age	23.53	4.73	−0.30	0.30	−0.36	0.25	0.82
(b) Number of siblings	2.94	1.13	−0.07	0.31	−0.07	0.37	0.95
(c) Household size	5.92	1.79	0.02	0.85	0.02	0.89	0.97
(d) Married	0.30	0.46	−0.03	0.35	−0.04	0.25	0.77
(e) Number of children	1.46	1.15	−0.16	0.26	−0.15	0.35	0.98
(f) Household income (midpoint, INR)	11033	8585	−538	0.30	−562	0.31	0.96
(g) Completed 12th standard	0.72	0.45	0.01	0.74	−0.01	0.81	0.59
(h) Completed college	0.21	0.41	−0.02	0.58	−0.01	0.68	0.89
(i) Identity ranking: Caste	3.69	0.97	0.03	0.64	−0.03	0.64	0.35
(j) Identity ranking: Religion	3.70	1.11	−0.02	0.80	−0.05	0.55	0.71
(k) Identity ranking: Indian	1.34	0.87	−0.06	0.27	−0.02	0.70	0.49
(l) Identity ranking: From U.P.	3.34	1.41	0.00	0.96	0.02	0.85	0.88
(m) Identity ranking: Father/Son/Brother	2.94	1.14	0.04	0.55	0.09	0.29	0.59
(n) Overall well-being score (% of Max Possible)	45.90	12.45	−0.82	0.30	−1.37	0.09	0.45

Notes: Each row reports a separate regression of the covariate on treatment indicators with date fixed effects (absorbed) and standard errors clustered at the work-booth × date level. The control group is "First Name" only group. Columns report control mean and standard deviation, treatment–control differences (β) for "Revealed" and "*Hidden*," and the p-value for equality of treatment effects. Household income (row (f)) is coded as the midpoint of the income bin, or the minimum value for the top-coded bin. Identity rankings (rows (i)–(m)) are measured on a 1 (most) to 5 (least) ranking scale. Well-being score in row (n) is a composite index normalized to the 0–100 range.

Table A5: Heterogeneity by Caste: GEN vs SC

	(1) Identity fusion	(2) Meet up again	(3) Trust sent
General caste (GEN)			
First name only	-0.365^{***}	-0.204^{***}	-4.020^{***}
	(0.104)	(0.051)	(1.311)
Hiding	-0.544^{***}	-0.213^{***}	-4.633^{***}
	(0.107)	(0.051)	(1.353)
Mean dep. var. (Revealed)	2.777	0.741	27.349
Scheduled Caste (SC)			
First name only	-0.199^{**}	-0.101^{*}	-3.140^{**}
	(0.099)	(0.052)	(1.375)
Hiding	-0.386^{***}	-0.186^{***}	-2.928^{**}
	(0.103)	(0.051)	(1.331)
Mean dep. var. (Revealed)	2.783	0.711	26.145
p: GEN = SC, First name only	0.089	0.040	0.489
p: GEN = SC, Hiding	0.105	0.618	0.182
N	1,014	1,014	1,014

Notes: Omitted treatment: Revealed (GEN and SC). Standard errors are clustered at the pair-level. Each column comes from a single regression interacting treatment status with caste identity; the two coefficients shown under each caste panel heading are the interactions of that caste dummy with the First Name Only and Hiding treatment indicators respectively. All specifications control for age and absorb randomisation strata and education fixed effects via `reghdfe`. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A6: Social Relations: Homogeneous vs Heterogeneous Revealed Teams

	(1) Trust sent	(2) Identity fusion	(3) Meet up again
Same-caste pair (=1 if HH or LL, =0 if HL)	3.157*	0.266*	0.001
	(1.913)	(0.138)	(0.072)
Mean dep. var. (HL)	26.778	2.781	0.723
Observations	425	425	425

Standard errors clustered at the pair-level in parentheses. Sample: HH-Revealed, LL-Revealed, and HL-Revealed arms. Homog = 1 if same-caste pair (HH or LL), 0 if cross-caste (HL). Mean dep. var. is for the heterogeneous (HL) group. All specifications include randomization strata fixed effects (date × caste category) and pre-specified controls for baseline education and age. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table A7: Caste Confidence as Mechanism

	(1) Trust sent	(2) Identity fusion	(3) Meet up again	(4) True self
Panel A: Treatment effects only				
First name only	−3.711***	−0.286***	−0.152***	−0.713***
	(1.221)	(0.092)	(0.046)	(0.246)
Hiding	−3.920***	−0.468***	−0.198***	−0.846***
	(1.217)	(0.097)	(0.045)	(0.237)
Mean (Revealed)	26.747	2.780	0.726	7.503
First-name only vs. Hiding p-value	0.844	0.066	0.322	0.571
N	1,014	1,014	1,014	1,014
Panel B: Controlling for caste confidence and beliefs				
First name only	−1.558	−0.091	−0.068	0.005
	(1.138)	(0.089)	(0.047)	(0.207)
Hiding	−0.009	−0.126	−0.067	0.426*
	(1.348)	(0.101)	(0.048)	(0.218)
Caste certainty				
Caste confidence (0–1)	19.002***	1.574***	0.505***	5.933***
	(2.620)	(0.212)	(0.093)	(0.488)
Guessed caste × own caste				
Guessed GEN × own GEN	0.128	−0.022	−0.146*	0.097
	(1.569)	(0.201)	(0.083)	(0.401)
Guessed GEN × own SC	−3.526**	−0.077	0.190***	−0.424
	(1.566)	(0.146)	(0.064)	(0.300)
Mean (Revealed)	26.747	2.780	0.726	7.503
First-name only vs. Hiding p-value	0.167	0.716	0.970	0.051
N	1,014	1,014	1,014	1,014

Notes: Omitted treatment: Revealed. All specifications control for age and absorb randomization strata and education fixed effects. Panel B adds caste confidence (rescaled 0–1) and guessed partner caste interacted with own caste. SEs clustered at the pair-level. Guessed SC × own GEN and Guessed SC × own SC are omitted in Panel B, because of collinearity (strata fixed effects include caste). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

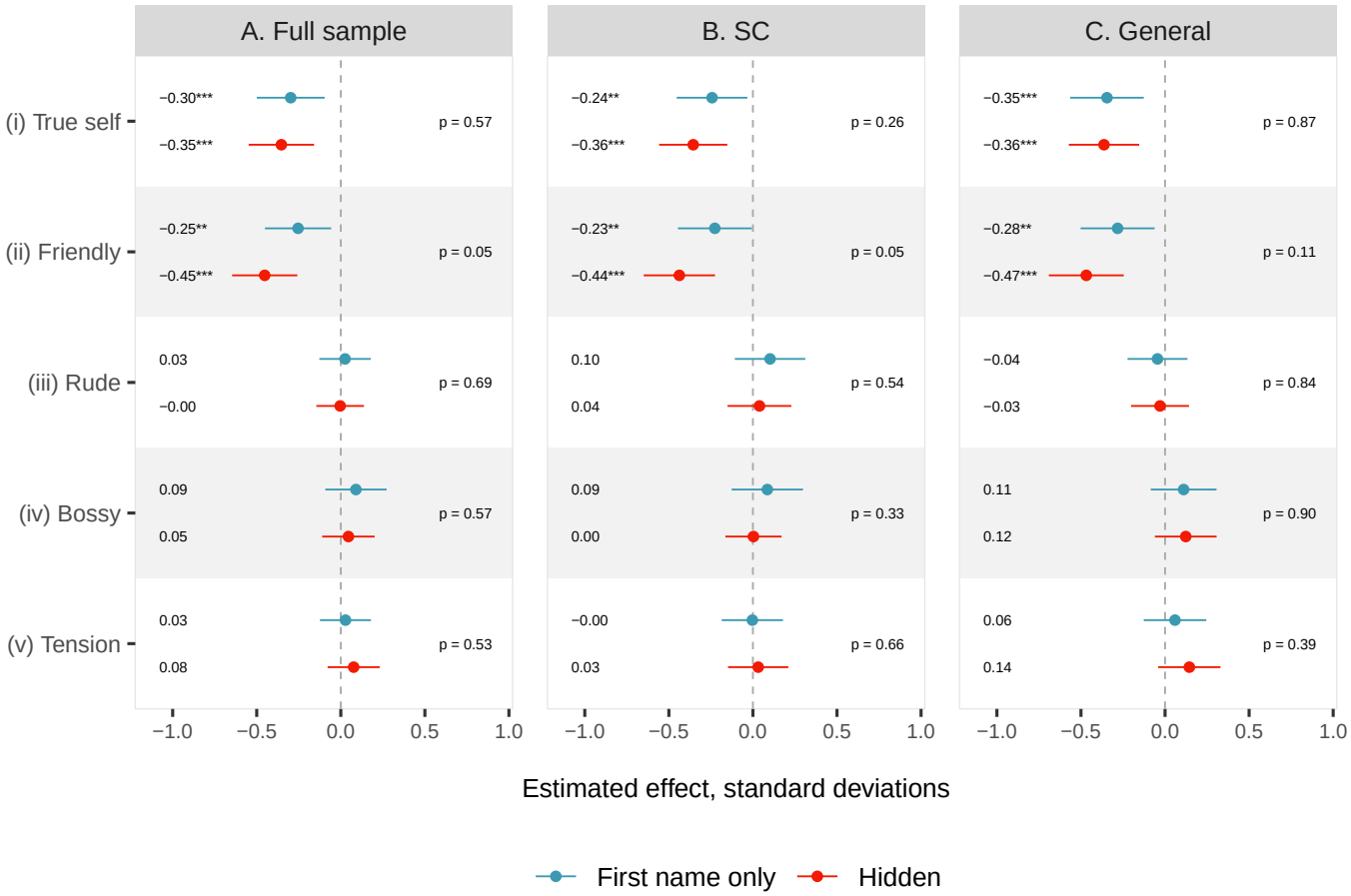
Table A8: Treatment effect on guessing study purpose

	(1)
First-name only	0.006
	(0.014)
Hidden	0.021
	(0.016)
First-name only vs. Hiding p-value	0.39
Mean Dep. Var (Revealed)	0.029
Observations	1198
R-squared	0.131

Notes: Dependent variable is a dummy coded 1 if the subject guessed the purpose of the study correctly. Standard errors clustered at pair-level in parentheses. All specifications include randomization strata fixed effects (date $\times$ caste category) and pre-specified controls for baseline education and age. $*p < 0.1$; $**p < 0.05$; $***p < 0.01$.

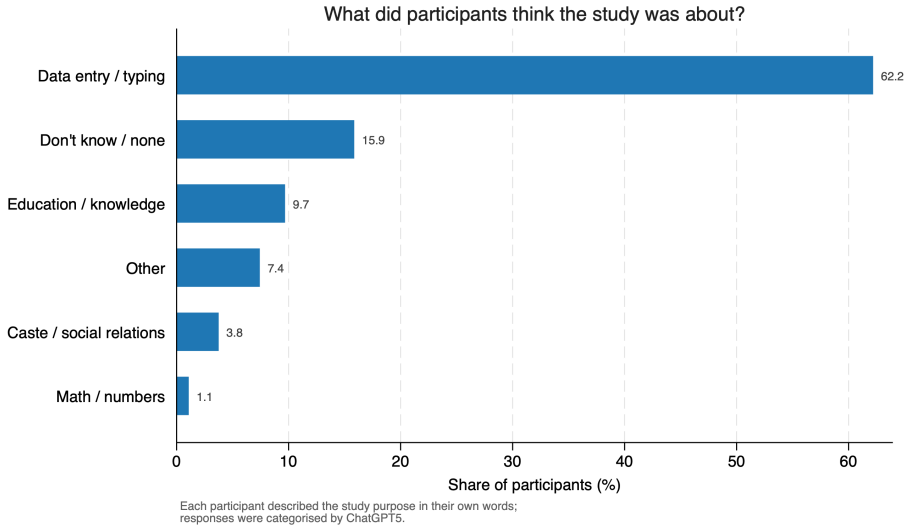
C Supplementary Figures

Figure A1: Estimated treatment effects on partner relations.



Notes: Each facet panel reports estimates from an OLS regression of an outcome on indicators for assignment to the *First name only* and *Hidden* treatments, with *Revealed* as the omitted category. Outcomes are self-reported measures of interaction quality: (i) ability to show one's true self (1 = not at all, 10 = completely); (ii) partner friendliness (1 = very unfriendly, 10 = very friendly); (iii) partner rudeness (1 = not at all rude, 10 = very rude); (iv) partner bossiness (1 = not at all bossy, 10 = very bossy); and (v) experienced tension (1 = no tension, 10 = a lot of tension). Treatment effects are standardized. In Panels A, standard errors are clustered at the pair level; in Panels B and C (the SC and General subsamples), standard errors are heteroskedasticity-robust. All specifications include randomization strata fixed effects and pre-specified controls for participants' education level and age. $*p < 0.1$; $**p < 0.05$; $***p < 0.01$.

Figure A2: The purpose of the study, according to endline respondents.



Notes: This figure summarizes endline subjects' reports of the purpose of the study. Subjects were asked at endline, "What do you think was the purpose of this research study? What was it trying to understand?" Subjects' qualitative responses were recorded by enumerators and categorized using ChatGPT5.

D Supplementary text

D.1 Selecting caste salient last names for the experiment

To validate that the last names used in our experiment are informative about caste identity, we conducted a supplementary beliefs survey with 460 residents of Lucknow (172 General caste, 194 OBC, 94 SC). Enumerators read respondents a list of 18 or 19 last names drawn from the names used in our experimental sample and asked them to guess the caste category of each name and report their confidence in that guess; across respondents, this covered 56 distinct names in total. Figure A3 plots, for each name, the share of respondents who assigned it to the modal caste category, coloured by that category and sorted from most to least informative. The results confirm that caste is highly legible from last names in this context: the majority of names are assigned to a single caste category by more than 90% of respondents, and the names that appear most frequently in our experiment — including GAUTAM, SINGH, RAWAT, MISHRA, TIWARI, SHUKLA, PANDEY, and AVASTHI (shown in bold) — all have modal shares above 87%. Thus last names displayed in the *Revealed* treatment arm provided participants with meaningful and reliable information about their partner's caste identity.

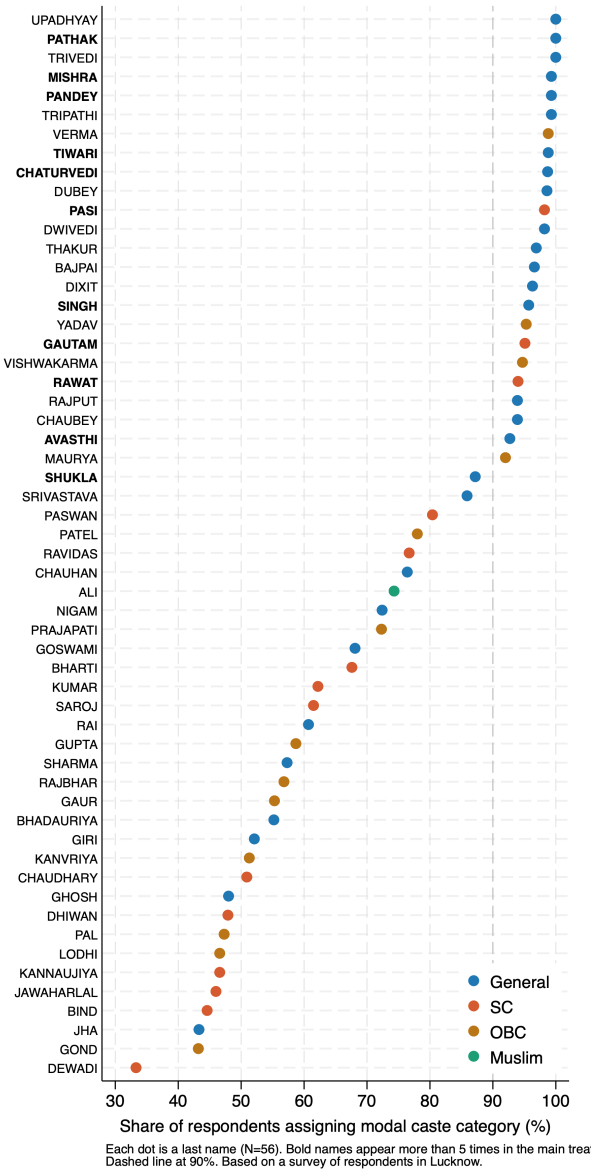


Figure A3: Caste informativeness of last names

Notes: Each dot represents one of 56 last names. The x-axis shows the share of survey respondents who assigned the name to the modal caste category, coloured by that category (blue = General, coral = SC, amber = OBC, teal = Muslim). Names are sorted from most to least informative. Bold names appear more than 5 times in the main experiment sample. Dashed vertical line at 90%. Based on a beliefs survey of 460 residents of Lucknow.

D.2 Deviations from pre-analysis plan

We followed the pre-registration in both outcome construction and empirical analysis. The deviations are minor and listed below.

- **Meal-sharing outcome.** The pre-registration included a meal-sharing measure as part of the social ties outcomes. Due to an implementation error, enumerators did not distribute the registration slips during the first two days of the study. We therefore cannot include this outcome in the analysis.

- **Well-being outcomes.** Ideally, we would have controlled for baseline well-being when estimating treatment effects on well-being measured at endline. However, we discovered an implementation error: the baseline

well-being questions were not translated into Hindi. We therefore do not use the baseline well-being measure, and omit the baseline outcome control in these regressions.

- **Productivity outcomes.** The pre-registration specified productivity outcomes in levels. In the paper, we use log productivity measures instead of raw counts. This does not materially affect the results, but reduces the influence of outliers.

D.3 Ethical considerations

The study protocols were separately reviewed and approved by four Institutional Review Boards, at the University of California San Diego, University of California Davis, Duke University and University of British Columbia.

D.3.1 Informed consent and experimental design

We obtained informed consent from all adult participants before participation. The study was conducted in a supervised workplace setting, and participants were adequately compensated for both attendance and performance. Because the object of interest is how identity visibility shapes interaction, participants were not told in advance that caste was the focus of the study. This was necessary to limit experimenter demand effects and strategic behavior that would have altered the interactions we aimed to study. The intervention itself was deliberately light-touch, operating through small differences in how partners were introduced and, in one condition, whether they were instructed not to reveal caste-linked information.

D.3.2 Risks to participants

The main ethical concern is that revealing or concealing identity may affect participants' comfort, trust, or sense of authenticity during the interaction. We reduced these risks by conducting the study in a controlled environment and avoiding any explicit discussion of caste by the research team during the work session. Participants were not asked to disclose sensitive information beyond what arose naturally in the course of interaction. The structured setting likely reduced the scope for more extreme forms of discriminatory behavior. Consistent with this, we do not find evidence for large wellbeing differences across the treatment arms.

D.3.3 Interpretation of results

Our findings should not be interpreted as a general normative claim that identity concealment is harmful or that disclosure is always beneficial. The experiment identifies the effects of identity legibility on a specific margin: social and workplace interactions in a formal setting. In other contexts, limiting the visibility of identity may protect individuals from more severe forms of discrimination, exclusion, or violence, which our design is not intended to measure.

More broadly, our results should not be read as placing the burden of better social relations on members of marginalized groups. If reduced identity uncertainty improves trust and collaboration, the relevant implication is not that disadvantaged individuals ought to reveal more, but that institutions should be made safer and more inclusive so that identity does not carry such acute social risk in the first place.