

INTERNATIONAL
CENTER FOR
PUBLIC POLICY

—
WITHIN THE
PUBLIC FINANCE
RESEARCH CLUSTER

Putting a Face on the Waiting List: Recipient Photographs and Organ- Donation Registration in Peru

Alberto Chong
Daniela Valdivia

October 2026



ANDREW YOUNG SCHOOL
OF POLICY STUDIES

Tel: 404-413-0233

Website: pfrc.gsu.edu

Address:
55 Park Place NE
7th Floor
Atlanta, GA 30303

Mail:
Public Finance Research Cluster
P.O. Box 3992
Atlanta, GA 30302-3992

International Center for Public Policy Andrew Young School of Policy Studies

The Andrew Young School of Policy Studies was established at Georgia State University with the objective of promoting excellence in the design, implementation, and evaluation of public policy. In addition to four academic departments, including economics and public administration, the Andrew Young School houses nine leading research centers, including the International Center for Public Policy.

The mission of the International Center for Public Policy (ICePP) is to provide academic and professional training, applied research, and technical assistance in support of sound public policy and sustainable economic growth in developing and transitional economies.

ICePP is recognized worldwide for its efforts in support of economic and public policy reforms through technical assistance and training around the world. This reputation has been built serving a diverse client base, including the World Bank, the U.S. Agency for International Development (USAID), the United Nations Development Programme (UNDP), finance ministries, government organizations, legislative bodies, and private sector institutions.

The success of ICePP reflects the breadth and depth of its in-house technical expertise. The Andrew Young School's faculty are leading experts in economics and public policy and have authored books, published in major academic and technical journals, and have extensive experience in designing and implementing technical assistance and training programs. Andrew Young School faculty have been active in policy reform in over 70 countries around the world. Our technical assistance strategy is not merely to provide technical prescriptions for policy reform, but to engage in a collaborative effort with host governments and donor agencies to identify and analyze the issues at hand, arrive at policy solutions, and implement reforms.

ICePP specializes in four broad policy areas:

- Fiscal policy (e.g., tax reforms, public expenditure reviews)
- Fiscal decentralization (e.g., reform, intergovernmental transfer systems, urban finance)
- Budgeting and fiscal management (e.g., local, performance-based, capital, and multi-year budgeting)
- Economic analysis and revenue forecasting (e.g., micro-simulation, time series forecasting)

For more information about our technical assistance activities and training programs, please visit our website at icepp.gsu.edu or contact us at paulbenson@gsu.edu.

Putting a Face on the Waiting List: Recipient Photographs and Organ-Donation Registration in Peru

Alberto Chong and Daniela Valdivia*

October 2026

Abstract

Deceased-donor registries are chronically undersupplied. This is particularly true in developing countries. We ask whether putting a face on potential donors changes behavior. In a laboratory experiment in Peru, sessions were randomly assigned to a text-only registration request or to one accompanied by a photograph of a prospective recipient. Since treatment is assigned by session, we take the session as the unit of inference. When registering is costly, a photograph raises registration by 7.0 percentage points on a control mean of 36.2 percent; the effect is concentrated in the initial costly block and is not detected in the subsequent costless block. Given that block order is confounded with cost, this is consistent with a cost-dependent treatment effect. Every subject also ranked photographs of two ethnicities; subjects register more often in the rounds showing the photograph they themselves ranked more favorably, while the same variable constructed for control subjects is null.

Keywords: organ donation, photograph, donor registration, laboratory experiment, Peru

JEL Codes: C91, C92, D64, D91, I11, I18

* Chong (achong6@gsu.edu): Department of Economics, Georgia State University, 55 Park Place, Room 786, Atlanta and Department of Economics, Universidad del Pacífico, Av. Salaverry 2020, Lima. Valdivia (danielav@iadb.org) Inter-American Development Bank, Calle Dean Valdivia 148, Piso 10, Centro Empresarial Platinum Plaza, San Isidro, Lima.

We are very grateful to Rexvi Rivera who co-wrote previous versions of this research and participated very actively in the design and application of the experimental sessions. We are thankful to Marco Chong, Young-Yun Choi, Virgilio Galdo, Erica Mtenga, Pierre Ngumkeu, Anna Serrichio, and Luisa Zanforlin and to seminar participants at Georgia State University, Queen's University, Michigan State University, Trinity College, Universidad del Pacífico, and University of Regina, for comments and suggestions. This research received IRB approval from Universidad del Pacífico. The standard disclaimer applies

1. Introduction

Every deceased donor can save several lives, registering as one requires little more than a signature, and yet in most countries a minority of adults are registered. This is particularly glaring in developing countries. The economics of this gap has been studied almost exclusively as a problem of institutional design. In the laboratory, allocation priority roughly doubles registration (Kessler and Roth 2012); an opt-out default recovers about eighty percent of that gain (Li et al. 2013); eliminating the costs of registering recovers about eighty percent of the gain from an outright payment (Hawley et al. 2018); and a persuasive message about the benefit to recipients raises registration, though only temporarily (Li 2016). The common structure of these studies is that they manipulate the rules of the game. The prospective recipient is in an anonymous position on a waiting list. However, outside the laboratory the recipient is rarely presented anonymously. Donor-recruitment campaigns are built around faces

Does adding a photograph of a prospective recipient to the registration request change the decision to register?¹In the canonical paradigm a statistical victim is contrasted with a particular person whose name, age and circumstances are given (Schelling 1968; Jenni and Loewenstein 1997; Small and Loewenstein 2003). Our subjects see a photograph and nothing else: no name, no history, and no assurance that this individual will receive their organ. The treatment is therefore a recipient-photograph cue. It may operate through visual attention, affective response to faces, humanization of an abstract problem, or experimenter demand; the design does not distinguish among these channels.

We run the experiment in Peru, a country that reported two deceased donors per million population at the end of 2018, among the lowest rates in the world (Gonzales Moreno et al. 2020). This, in the context of a country where ethnic stratification along an indigenous–mestizo–white gradient is a central feature of the society and discrimination against indigenous-looking individuals has been broadly demonstrated (e.g., Moreno et al. 2012; Galarza and Yamada 2014; Planas et al. 2015). Sessions were randomly assigned to one of three conditions: a text-only registration request, a request accompanied by a photograph of a white-looking prospective recipient, or a request accompanied by a photograph of an indigenous-looking prospective recipient. Within every

¹ The question is narrow by design. We are explicit at the outset about what it is not. In particular, it is not a test of the identifiable-victim effect.

session the sex of the depicted person alternates across rounds, so that four photographs appear in total. Rounds one through eight impose the registration and donation costs of Li et al. (2013); rounds nine through sixteen remove them. We take the session as the unit of inference, report the unadjusted difference as the primary causal estimate, and give four procedures in parallel, of which randomization inference over the nineteen sessions is our preferred benchmark. We designate the costly first block as the primary sample and treat the second as complementary.²

The central result is that adding a photograph raises registration in the costly block. The unadjusted estimate is 7.0 percentage points on a control mean of 36.2 percent, with a 95 percent confidence interval of 1.7 to 12.3 points, a session-clustered p-value of 0.012 and a randomization-inference p-value of 0.063. Adjusting for pre-treatment covariates raises the estimate to 8.8 points and the randomization p-value falls to 0.018; the equally weighted session-means estimator on nineteen observations gives 6.1 points (s.e. 3.2, $p = 0.072$). That estimator weights a five-subject session the same as a fourteen-subject one; weighting sessions by their number of subjects gives 7.0 points (s.e. 3.2, $p = 0.045$), the same as the individual-level estimate. The experiment provides evidence of a positive photograph effect of roughly six to nine percentage points. The effect is of the order of a fifth of the control registration rate, comparable to the informational message of Li (2016) and roughly a third of the priority rule of Kessler and Roth (2012).

Given external constraints, we are only able to employ two photographs per ethnic category, so the ethnicity a photograph conveys is confounded with the identity of the photograph itself. The difference between the two photographic conditions is essentially zero: -0.6 points unadjusted, with a confidence interval running from -10.1 to $+8.9$. However, this is not evidence against ethnic discrimination because the two conditions differ in the photographs used as much as in the ethnicity those photographs convey. The variation in estimated effects across the two photographs within an ethnic category is of the same order as, and in the white pair larger than, the 0.6-point variation across categories, and none of these differences is distinguishable from zero. No amount of statistical precision separates a category from the single stimulus used to represent it (Clark 1973; Judd et al. 2012). Having said this, the null is informative. Subjects clearly perceive the photographs along an ethnic dimension: asked to rank all four, 86 percent of control-arm subjects

² To our knowledge, the vast majority of the experimental organ-donation literature has been conducted in the United States; Herr and Normann (2016) is a prominent exception.

placed a white recipient first and fewer than five percent the indigenous woman, an ordering that correlates with a measured discrimination index. This juxtaposition would be merely descriptive if the ranking told us nothing about behavior, but it tells us something narrower and better identified than the arm comparison does. Every subject in a treated session saw two of the four photographs, one in odd rounds and one in even rounds, and ranked both. Holding the subject and the round fixed, registration is higher in the rounds showing the photograph that subject personally ranked more favorably: -0.019 per rank position under one reading of the stored ranking ($p = 0.055$) and -0.014 under the other ($p = 0.149$), both signed as the mechanism predicts. The identical variable constructed for control subjects, who ranked the photographs but saw none of them while deciding, is null. This test asks nothing about ethnicity and requires no generalization across photographs, which is why we regard it as the most direct evidence the design provides about stimulus-specific response: they are not interchangeable stimuli, and subjects respond to the particular one in front of them.

Notice that the photograph effect is concentrated in the costly block. Between the first block and the second, registration rises by twenty-four points in all three conditions, and the photograph effect is no longer detected. Because block order is confounded with the cost change, that contrast is consistent with a cost-dependent effect but does not identify one. A threshold model in which a subject registers when the perceived benefit from the depicted recipient exceeds her cost rationalizes the pattern, since a treatment that shifts perceived benefit can only move subjects near the margin. The blocks differ in order, experience and learning as well as in cost, and treatment decay of the kind Li (2016) documents would produce the same pattern.

The paper makes two contributions. First, to our knowledge, it provides experimental evidence within the organ-donation paradigm that the presentation of the recipient, and not only the rules governing the queue, is a margin along which registration responds. Second, it adds rare experimental evidence from a developing-country setting. The paper is organized as follows. Section 2 presents the current literature. Section 3 describes the experimental design. Section 4 describes the empirical strategy. Section 5 presents the results. Section 6 presents robustness and inference diagnostics, while Section 7 discusses mechanisms, external validity, and the implications of the findings. Heterogeneity analyses, the recipient-sex decomposition, and the full balance table are presented in the appendix.

2. Related Literature

The paradigm originates with Kessler and Roth (2012), who model registration as a decision in a repeated game: subjects hold one non-transplantable and two transplantable organs, face period-by-period failure risks, and may join a waiting list, with registration costly to stand in for the psychological cost of contemplating one's own death. They find that granting registered donors' priority on the waiting list raises registration from about 36 to about 74 percent, and that a cash rebate calibrated to the expected value of priority reproduces the effect only once subjects have experience with the game (Lavee et al. 2010).

Li et al. (2013) reframe the instructions in terms of organs, show that this contextual framing alone raises registration by roughly five points, and add an opt-out default, a lever whose power in registration data Johnson and Goldstein (2003) had documented. Their two-by-two design establishes that opt-out and priority each raise registration substantially, that opt-out delivers about eighty percent of the priority gain, and that the two together dominate either alone. Kessler and Roth (2014) show that a loophole allowing non-donors to obtain priority destroys the incentive entirely. Herr and Normann (2016) replicate the priority result in a German subject pool, let subjects vote on the rule, and find that two-thirds endorse it after experiencing the dilemma; Herr and Normann (2019) show that a medium-sized priority bonus performs as well as absolute priority. Hawley et al. (2018) vary the level of donation costs and establish that halving costs does nothing while eliminating them raises registration from 21.8 to 63.5 percent, capturing about eighty percent of the gain from an outright payment. More recent work has extended the paradigm to blood-type compatibility constraints (Li et al. 2023), to the welfare consequences of adverse selection under priority and the value of information provision (Fang et al. 2026), to the regulation of procurement organizations (Chan and Roth 2024), and to dynamic social norms (Zang et al. 2025).

Closest to our research is Li (2016), who supplements the registration request with an informational message stating how much a saved subject earned on average. Registration rises by about ten points over the first fifteen rounds and then decays to nothing, and the effect is concentrated among subjects who are not registered donors in real life. Two features of that paper shape ours. The first is the finding that behavior in the laboratory tracks real-world donor status, which we replicate: registered donors in our sample register 11.6 points more often in the game.

The second is the paper’s closing agenda, which identifies message framing and the identifiable-versus-statistical-victim contrast as the obvious next question. Ours is the visual analogue of that question. As explained above, a bare photograph is a weaker manipulation than the canonical identifiable-victim design and we do not claim otherwise; nor do we claim that varying the ethnicity a photograph conveys identifies an effect of recipient ethnicity.

Our treatment resembles several literatures that are worth keeping apart. The identifiable-victim literature contrasts a statistical victim with a particular person whose identity and circumstances are given. A second strand concerns the salience of the presentation itself: an image draws attention and processing effort that text does not, and elicits affective responses that arise pre-attentively, neither of which requires the depicted person to be identified as anyone in particular. Because our subjects see a photograph and nothing else—no name, no age, no history—our manipulation sits squarely in the second strand and only loosely in the first. We discuss the first because it motivated the question and because Li (2016) poses it explicitly, not as evidence that ours is a study in it. Schelling (1968) drew the distinction that organizes this literature: harm to a particular person invokes sentiment and responsibility, while statistical harm invokes little. Jenni and Loewenstein (1997) formalized the resulting identifiable-victim effect, and Small and Loewenstein (2003) demonstrated it in a modified dictator game in which the only difference between conditions was whether the recipient’s identity had already been determined. Kogut and Ritov (2005) established the related singularity effect, and Small, et al. (2007) showed that prompting deliberation reduces giving to identified victims without raising giving to statistical ones. Cryder, et al. (2013) generalize the phenomenon to tangible detail about the intervention. Photographs in particular have been shown to raise giving relative to silhouettes, with the shift mediated by affective arousal (Genevsky, et al. 2013), and to act principally on the extensive margin of whether to give at all (Bergh and Reinstein 2021).

We take two cautions from this literature. The meta-analysis of Lee and Feeley (2016) puts the pooled effect at a modest $r = 0.05$ and finds it reliable mainly for a single photographed victim who bears little responsibility for the need, and Maier et al. (2023) fail to replicate Small et al. (2007) in a pre-registered study. Our setting has the features the meta-analysis identifies as favorable, and, unlike much of that literature, the decision is incentivized and repeated. We treat the size of our estimate as informative about this setting rather than as a general parameter.

That giving might depend on the recipient's group is among the most studied propositions in experimental economics. From Tajfel et al. (1971) through Akerlof and Kranton (2000) and Chen and Li (2009), the standard finding is that group membership shifts allocations toward the in-group. Meta-analytic and survey treatments (Balliet, et al. 2014; Lane 2016; Li 2020) report in-group favoritism across a wide range of games, though Lane's meta-analysis notes that discrimination is stronger when identity is artificially induced than when it is natural. The conceptual distinction between taste-based discrimination (Becker 1957) and statistical discrimination (Arrow 1973; Phelps 1972) is difficult to draw in dictator-type games, which is one reason our design uses a setting with no strategic component: the depicted recipient cannot reciprocate and cannot be inferred to have any productivity-relevant characteristic.

The evidence is less uniform than the headline suggests. Berge et al. (2020) find no ethnic bias in dictator and public goods games among a large Nairobi sample even in the run-up to a contested election, and a growing set of studies reports that economics games often fail to detect the bias that audit studies find in the same populations. Grimm, et al. (2017) show that favoritism across multiple out-groups is driven substantially by beliefs about how others behave, which is inoperative in a design like ours where the recipient is a photograph. Our contribution to this literature is not a clean test of ethnic favoritism, which this design cannot deliver. What it offers is the observation that in a domain with unusually high stakes and unusually low strategic content, in a population where discrimination is demonstrable on other margins, a stark stated ordering over recipients coexists with no behavioral difference between them, while the individual response to a particular depicted recipient still tracks how that individual ranked it.

Discrimination along this gradient has been demonstrated experimentally. Galarza and Yamada (2014) sent 4,820 fictitious CVs to 1,205 vacancies in Lima and found that Indigenous applicants must send about eighty percent more applications than observationally equivalent white applicants to obtain the same callback rate, with the gap largest in professional occupations. Moreno et al. (2012) reach similar conclusions using monitored intermediation services. Castillo, et al, (2007), running laboratory experiments on group sorting in urban Lima, find that looking Indigenous makes a subject a less desirable group member and looking white makes one more desirable. Most directly relevant, Planas et al. (2015) sent simulated patients enacting Indigenous and mestizo profiles to 351 public health establishments in Metropolitan Lima in a randomized crossover trial and documented differential quality of family-planning care. Peru has among the lowest deceased-

donation rates in the Americas. Gonzales Moreno et al. (2020), analyzing records of the national procurement authority over 2013–2018, report roughly two donors per million population at the end of 2018 against 48 in Spain and 33 in the United States, and attribute the shortfall to cultural and religious concerns, misinformation, weak clinical approach protocols, and ignorance of the deceased’s wishes. Regional comparisons place Peru at the bottom of the Latin American distribution alongside Paraguay, at under one effective donor per million in some years, while Uruguay, Argentina and Brazil have converged toward European levels. Family refusal is unusually high. The margin that our treatment addresses—what the registration prompt says and shows—is precisely the margin a low-capacity registry can act on without legislative change, which is why the comparison to opt-out and priority reform is a policy-relevant one here.

Ethnic categories in Peru are graded. The mestizaje ideology of the twentieth century recast racial difference as cultural difference, producing a society in which the great majority self-identify as mestizo while phenotype, ancestry and self-identification each independently predict socioeconomic outcomes (de la Cadena 2000; Paredes 2018). Indigenous ancestry and darker skin are associated with lower educational attainment and fewer household assets, and the share of the population classified as Indigenous ranges from about nineteen to about seventy-five percent depending on whether language, birthplace or self-identification is used (Galarza and Yamada 2014).

3. Experimental Design

Our findings are organized by a threshold condition, not by a model of any particular psychological mechanism. That is deliberate: the design identifies who moves, not why. Let subject i register in a given round when

$$B(i, j) - C(i) > 0 \quad (1)$$

where $C(i)$ is the subject’s cost of registering, comprising the monetary registration and donation costs imposed in the first block together with whatever non-monetary cost she attaches to the decision, and $B(i, j)$ is the benefit she perceives from an organ reaching a recipient of type j . The control condition presents j as an anonymous queue position, the photograph conditions present j as a visible person. Write the treatment as a shift ΔB in perceived benefit. In this context we test two hypotheses.

Hypothesis 1. Adding a recipient photograph to the registration request raises registration, by an amount equal to the mass of subjects whose net benefit lies within ΔB of zero.

Hypothesis 2. The effect is concentrated among subjects near the registration margin and is therefore larger where $C(i)$ is larger, so removing the monetary costs shrinks it toward zero as the constraint stops binding.

The underlying game follows Li et al. (2013), which is itself the organ-framed adaptation of Kessler and Roth (2012); the instructions are stated in terms of organs. Each subject begins every round with one A organ and two B organs. In each period the A organ fails with probability 0.10 and both B organs fail together with probability 0.20; the two events are mutually exclusive. A subject whose A organ fails ceases to participate in that round, and if the subject had registered, the two B organs become available to subjects on the waiting list. A subject whose B organs fail joins the waiting list for at most five periods, earning nothing while waiting, and ceases to participate if no organ arrives. Subjects earn a fixed amount in every period in which they hold a functioning A organ and at least one functioning B organ; mean round earnings in our data are 7.65 experimental currency units. Waiting-list position is determined by time waited. Subjects observe their own health status, per-period and cumulative earnings, and, once on the list, their rank and waiting time. 18.7 percent of subject-rounds are preceded by a round in which the subject received an organ. The experiment was implemented in oTree (Chen et al. 2016); the decision is taken at the start of each round, before health status is realized, so registration is never conditioned on being about to need or supply an organ. The default is non-donor throughout: we do not vary the default, and we do not implement a priority rule. Holding the allocation institution fixed at the Peruvian status quo—opt-in, no priority—is deliberate, since our object of interest is the content of the request, not the rules of the queue.

Sessions were assigned to one of three arms. In the control arm the registration request is text only, as in the previous literature. In the two photograph conditions the request is accompanied by an image of a potential recipient: an indigenous-looking person in one arm and a white-looking person in the other. Within every session, including the control, the sex of the depicted recipient alternates deterministically with round parity, so that odd rounds present a man and even rounds a woman. This yields a three-by-two structure in which ethnicity varies between sessions and sex varies within subject. Six sessions and 65 subjects were assigned to the control arm, six sessions and 65

subjects to the Indigenous arm, and seven sessions and 65 subjects to the white arm. Session size ranges from five to fourteen subjects.³

Rounds one through eight impose the two-part cost structure of Li et al., (2013): a registration cost charged whenever a subject registers, and a larger donation cost charged only if a realized donation takes place. Rounds nine through sixteen remove both. Subjects were stopped after round eight and informed of the change, as in the source papers.⁴ One hundred and ninety-five undergraduates at a Lima university participated in nineteen sessions, each making sixteen registration decisions, for 3,120 subject-round observations in a perfectly balanced panel. Sixty-seven percent are women and the mean age is 22.0 years. Ninety percent identify as white or mestizo and ten percent as members of an ethnic minority, with only nine percent identifying as white. Mestizo identification in Peru spans a wide range of ancestry, phenotype and status (de la Cadena 2000; Paredes 2018), and the ten percent minority share is far too small to support an in-group test. That is why below we use region of birth instead. Seventy-five percent report a religious affiliation and 65 percent attend services. Thirty-seven percent are registered organ donors on their national identity card and eighty percent state a willingness to donate, but only seventeen percent know the registration procedure. After the sixteen rounds, subjects ranked all four recipient photographs; no subject, including those in the control arm, saw any photograph before the experiment began.⁵

4. Empirical Strategy

Let $y(i,s,t)$ be an indicator for subject i in session s registering as a donor in round t . The principal specification is the linear probability model:

$$y(i,s,t) = \beta_1 \cdot \text{Indigenous}(s) + \beta_2 \cdot \text{White}(s) + \delta(t) + X(i)' \theta + \varepsilon(i,s,t) \quad (2)$$

³ All participants were junior or senior undergraduate students at Universidad del Pacífico, Lima. The laboratory experiment was performed during the Spring of 2024.

⁴ Two features of this arrangement differ from the current literature. First, the block structure was not counterbalanced: every session received costs first and no costs second. This means that the block difference confounds cost removal with order, experience and the restart effect that this literature documents at exactly such a break. Kessler and Roth (2012), Li et al. (2013), Hawley et al. (2018) and Herr and Normann (2016) all counterbalanced, and all report that order matters. Second, the eight-plus-eight structure with a mid-experiment break reproduces the restart effect documented throughout this literature.

⁵ A post-experiment questionnaire elicited demographics, attitudes toward donation, and four multi-item indices of perceived discrimination in Peruvian society—general, racial, socioeconomic and gender—together with an overall index. Each item is a five-point Likert agreement scale, and each index is the normalized mean of its items. See the Appendix.

where Indigenous and White are session-level arm indicators with the control arm omitted, $\delta(t)$ is a full set of round fixed effects, and $X(i)$ is a vector of subject-level controls. We report the linear model as the primary specification because the parameter of interest is a difference in probabilities across randomly assigned groups. The linear probability model also makes treatment contrasts and their interactions directly interpretable as percentage-point differences; nonlinear models are reported as robustness checks (Ai and Norton, 2003; Kessler and Roth, 2012).⁶

The photograph treatment does not vary within session. The effective number of independent units for the treatment contrast is nineteen, as the appropriate level of clustering is the level at which treatment was assigned (Cameron and Miller 2015; MacKinnon et al. 2023). With few clusters, cluster-robust t-tests over-reject, so we report four procedures in parallel. First, CR1 cluster-robust standard errors referred to a t distribution with $G-1 = 18$ degrees of freedom. Second, the restricted wild cluster bootstrap of Cameron et al. (2008) with Rademacher weights and 9,999 replications, imposing the null, the procedure MacKinnon and Webb (2018) and Canay et al. (2021) recommend when clusters are few. Third, randomization inference that permutes the arm label across the nineteen sessions holding the six-six-seven allocation fixed, recomputes the estimate under 5,000 draws, and reports the share of placebo estimates at least as large in absolute value. This is design-based and requires no asymptotics in the number of clusters (Young, 2019; MacKinnon and Webb, 2020). Fourth, an equally weighted session-means regression on nineteen observations in the spirit of Donald and Lang (2007). Of the four, we treat randomization inference as the primary benchmark: it is design-based, requires no asymptotics in the number of clusters, and answers exactly the question nineteen randomly assigned sessions can answer. Because the treatment is randomly assigned, the unadjusted difference is the primary causal estimate, and covariate adjustment is a precision device; we report the unadjusted estimate first everywhere and adjustment alongside it. We give confidence intervals and do not lean on p-values. It should be mentioned that the cost block is a deterministic function of the round and was not counterbalanced. The second block is the costless one. Registration rises by 24.0 percentage points from the first block to the second, and by 34.4 points between rounds eight and nine alone. That quantity

⁶ We include round fixed effects because the depicted recipient's sex is perfectly collinear with round parity, a specification without them loads the common parity profile onto the photograph dummies asymmetrically. Since the cost block is a deterministic function of the round, they also absorb the block. A second specification decomposes each arm by the sex of the depicted recipient, in which case four image indicators replace the two arm indicators, and the comparison remains against the control arm in the same round.

confounds the removal of costs with order, accumulated experience, and the restart effect (Kessler and Roth, 2012; Li, 2016; and Herr and Normann, 2016). It is not a cost effect. What remains identified is the interaction of the photograph treatment with the block, since the block is common to all arms, and the within-block arm contrasts. We designate the first block as the primary sample, as it is the block in which registration is costly, which is the empirically relevant case, and it is uncontaminated by prior experience or the restart.⁷

The design uses four photographs, two male and two female, spanning two ethnic categories, Indigenous and white. Every subject in a treated session saw two of the four photographs, one male and one female, and freely ranked both. The contrast between the Indigenous and white conditions is a contrast between particular photographs, which differ in age, expression, apparent warmth, dress, and photographic composition as well as in ethnicity. Given this stimulus-sampling issue, we write “photograph” rather than “face” throughout: the stimulus carries clothing, expression, composition, and image quality as well as facial features, so the confound is broader than facial identity alone (Clark 1973; Wells and Windschitl 1999; Judd et al. 2012). We therefore describe the Indigenous-versus-white comparison as a difference between photographic treatments, not as an ethnic effect. The recipient sex alternates within session, allowing us to estimate the effect of each photograph separately against the control condition. In the costly block, the two white-looking photographs differ from each other by 6.2 percentage points and the two Indigenous-looking photographs by 1.9 points, compared with a 0.6-point difference between the ethnic categories; none of these differences is statistically distinguishable from zero. Finally, with one photograph per category-by-sex cell, ethnicity and photograph identity are collinear by construction.⁸ Every subject, including every control subject, was shown all four photographs and asked to rank them. This yields a subject-specific measure of how each individual perceives each particular photograph, which a design with only between-photograph variation does not possess.

⁷ Two diagnostics indicate how this matters. The intraclass correlation of registration at the session level is only 0.023, against 0.244 at the subject level, so session clustering inflates standard errors less than one might expect. The between-session standard deviation of the registration rate within an arm is 0.067, implying minimum detectable effects at eighty percent power of 12.1 points for the Indigenous-versus-control contrast, 11.5 for white-versus-control and 9.9 for any photograph versus control. The design is therefore underpowered to reliably detect effects smaller than these thresholds, including effects of the size estimated here.

⁸ The sex of the depicted recipient cannot be separated from the round, since no session shows a man and a woman in the same round. The sex decomposition is therefore a set of contrasts against control, not a within-photograph comparison.

It permits a within-subject test of whether the perceived standing of the specific photograph shown predicts the response to it, holding the subject and round fixed.⁹

5. Results

Table 1 reports means by arm and the p-value of a session-clustered joint test that both arm indicators are zero. Five of twenty-four characteristics fail at the ten percent level, against roughly 2.4 expected by chance: age, mother's education, the general discrimination index, and—more consequentially—two pre-treatment attitudes toward donation. Subjects in the Indigenous arm are more likely to state that they are willing to donate (87.7 versus 73.8 percent, $p = 0.020$) and report a more favorable perception of donation ($p = 0.025$). The variable most directly comparable to the outcome, registration as a donor on the national identity card, is balanced (41.5, 35.4 and 35.4 percent, $p = 0.76$), as are sex, ethnicity, religiosity, knowledge and the racial and overall discrimination indices. This pattern is what nineteen-cluster randomization produces, but it is not innocuous, since the two imbalanced attitudes favor the arm with the largest raw effect. Every estimate is therefore reported both unadjusted and adjusted for baseline attitudes. Adjustment leaves the Indigenous coefficient at 0.092 in the full sample and raises it from 0.067 to 0.081 in the primary block, so the imbalance does not generate the result, though the unadjusted estimate remains the causal one.

⁹ With nineteen sessions, imbalance on some baseline characteristic is likely. Section 7 shows that it occurred, and reports estimates with and without adjustment.

Table 1. Sample Characteristics and Randomization Balance

	Control	Indigenous	White	p (joint)
Age	22.031 (3.122)	21.308 (1.960)	22.708 (4.457)	0.099
Female	0.662 (0.477)	0.662 (0.477)	0.692 (0.465)	0.833
Semester at university	6.831 (2.434)	6.938 (2.609)	7.969 (2.651)	0.222
Number of siblings	1.478 (0.863)	1.538 (0.920)	1.523 (0.954)	0.851
Father's education	3.785 (1.053)	3.769 (1.101)	3.908 (1.155)	0.742
Mother's education	3.554 (1.090)	3.892 (0.970)	3.600 (0.997)	0.080
Health insurance	0.815 (0.391)	0.800 (0.403)	0.877 (0.331)	0.203
Economics major	0.431 (0.499)	0.508 (0.504)	0.385 (0.490)	0.204
White or mestizo	0.938 (0.242)	0.877 (0.331)	0.892 (0.312)	0.531
Attends religious services	0.631 (0.486)	0.631 (0.486)	0.677 (0.471)	0.754
Registered donor on ID card	0.415 (0.497)	0.354 (0.482)	0.354 (0.482)	0.763
Willing to donate	0.738 (0.443)	0.877 (0.331)	0.769 (0.425)	0.020
Knowledge about donation	0.672 (0.159)	0.669 (0.143)	0.682 (0.186)	0.904
Closeness to donation	0.215 (0.306)	0.154 (0.278)	0.138 (0.287)	0.210
Perception of donation	0.646 (0.242)	0.692 (0.198)	0.631 (0.205)	0.025
Discrimination index: racial	0.481 (0.164)	0.470 (0.160)	0.447 (0.176)	0.605
Discrimination index: overall	0.497 (0.093)	0.508 (0.089)	0.503 (0.137)	0.557
Sessions	6	6	7	
Subjects	65	65	65	

Notes: Subject-level means with standard deviations in parentheses. The final column reports the p-value of a joint test that both arm indicators are zero in a regression of the characteristic on the arm dummies with standard errors clustered on session (19 clusters). Five of twenty-four characteristics reject at the ten percent level. The full set of twenty-four is reported in Appendix Table A1.

Table 2 reports registration rates. Pooling all sixteen rounds, 48.8 percent of control decisions, 58.0 percent of Indigenous-arm decisions and 51.7 percent of white-arm decisions are registrations. The control rate is high relative to the American studies—25.3 percent in Li et al. (2013) and 21.8 percent in the baseline of Hawley et al. (2018)—and close to the 40 percent of Herr and Normann (2016). Part of this is composition: our sample states unusually high willingness to donate. Part is the block structure, since half of our rounds are costless.

Table 2. Registration Rates by Treatment Arm and Cost Block

	Control	Indigenous	White	All
All 16 rounds	0.4875	0.5798	0.5173	0.5282
Block 1: rounds 1–8 (costly)	0.3615	0.4288	0.4346	0.4083
Block 2: rounds 9–16 (costless)	0.6135	0.7308	0.6000	0.6481
Round 1 only	0.1692	0.2154	0.1692	0.1846
Sessions	6	6	7	19
Subjects	65	65	65	195
Decisions	1,040	1,040	1,040	3,120

Notes: Cell entries are the share of decisions in which the subject registered as a donor. The block-2 minus block-1 difference of 24.0 points pools the removal of registration and donation costs with order, experience and the restart effect and is not interpretable as a cost effect (Section 6).

Figure 1 shows the round-by-round profile. Three features stand out. Registration is low in round one in every arm, between 17 and 22 percent, and rises sharply thereafter, the standard learning pattern. Within the first block all three series decline from round three onward, with the two photograph arms above control from round five. At the block break registration jumps by more than thirty points in every arm simultaneously—the confounded quantity discussed above. Within the second block the Indigenous arm remains above control throughout while the white arm falls to or below it.

Figure 1. Donor Registration by Round and Recipient Photograph

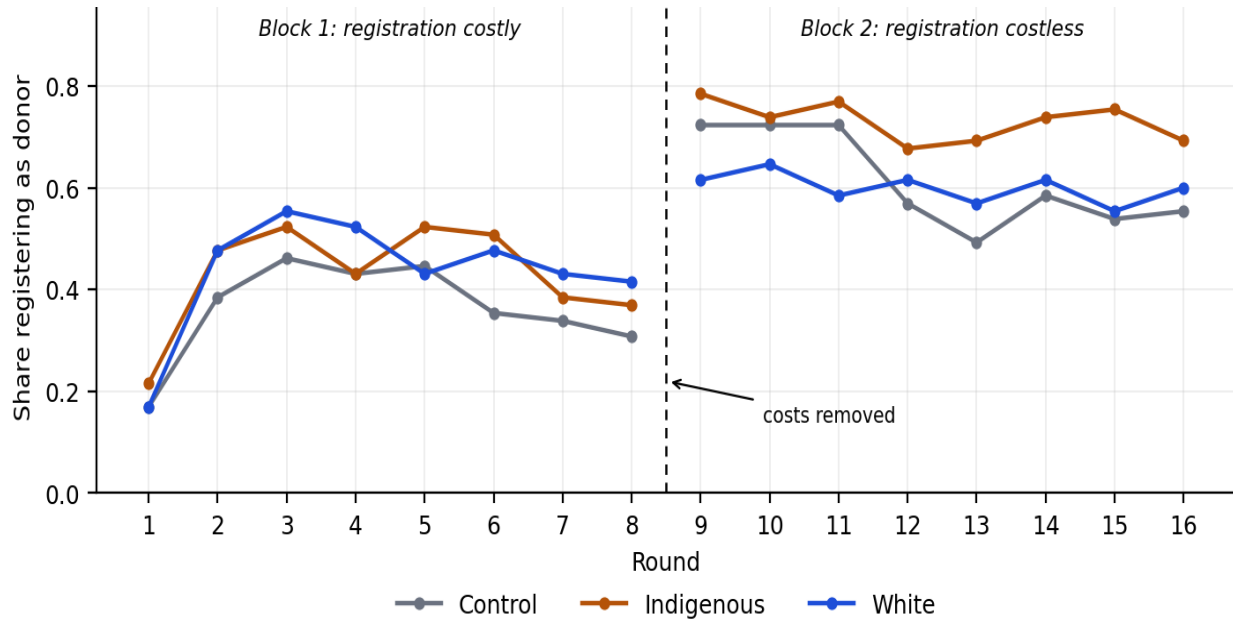


Figure 1. Share of subjects registering as a donor, by round and treatment arm. The vertical line marks the removal of registration and donation costs after round eight. Each series aggregates six or seven sessions.

Table 3 reports the primary results, estimated on the costly first block. Column 1 is the unadjusted difference, which is the estimate the randomization licenses: adding a photograph of a prospective recipient raises registration by 7.0 percentage points, with a session-clustered standard error of 0.025 and a 95 percent confidence interval running from 1.7 to 12.3 points. Against the control mean of 36.2 percent in Table 2 this is an increase of about a fifth. (Table 4 reports 0.363, the session-weighted mean; the regressions are observation-weighted.) The wild cluster bootstrap gives $p = 0.021$ and randomization inference $p = 0.063$. Column 2 adds demographic and baseline-attitude controls. The estimate rises to 8.8 points, and the randomization p -value falls to 0.018. This is a precision device because with randomly assigned treatment adjustment cannot be required for consistency. Column 3 goes the other way: collapsing to the nineteen session means, the unit at which treatment was actually assigned, gives 6.1 points (s.e. 3.2, $p = 0.072$), with a 95 percent confidence interval from -0.6 to 12.8. That estimator weights a five-subject session the same as a fourteen-subject one, which matters here because the white arm contains the three smallest sessions; weighting sessions by their number of subjects gives 7.0 points (s.e. 3.2, $p = 0.045$), the same as the individual-level estimate in column 1. The point estimates cluster between six and nine points across every procedure, the sign is consistent. Our preferred inferential benchmark, randomization inference over the nineteen sessions, returns $p = 0.063$ unadjusted.

Table 3. Effect of a Recipient Photograph on Registration, Block 1 (Rounds 1–8)

	Unadjusted (1)	Adjusted (2)	Session means (3)
<i>Panel A. Any recipient photograph</i>			
Estimate	0.070** (0.025)	0.088*** (0.027)	0.061* (0.032)
95% confidence interval	[0.017, 0.123]	[0.032, 0.145]	[−0.006, 0.128]
Wild bootstrap p	0.021	0.004	—
Randomization p	0.063	0.018	—
<i>Panel B. By photograph type</i>			
Indigenous-looking	0.067* (0.033)	0.081** (0.029)	0.062 (0.039)
95% CI	[−0.003, 0.138]	[0.020, 0.142]	[−0.019, 0.144]
Randomization p	0.152	0.082	—
White-looking	0.073** (0.034)	0.096** (0.036)	0.060 (0.038)
95% CI	[0.001, 0.145]	[0.020, 0.172]	[−0.018, 0.138]
Randomization p	0.100	0.026	—
<i>Panel C. Difference between the two photographs</i>			
Indigenous – White	−0.006	−0.015	0.002
95% CI	[−0.101, 0.089]	[−0.095, 0.065]	[−0.076, 0.081]
Randomization p	0.906	0.758	—
Round FE / covariates	Yes / No	Yes / Yes	— / No
Unit of observation	subject-round	subject-round	session
Observations	1,560	1,560	19

Notes: Column 1 is the primary causal estimate: treatment is randomly assigned, so no adjustment is needed for consistency. Column 2 adds demographics and baseline attitudes as a precision device. Column 3 collapses the data to the nineteen session means, the unit at which treatment was assigned, and is the most conservative estimator. Standard errors are CR1 clustered on session in columns 1–2 and conventional in column 3; intervals use $t(18)$ and $t(16)$ respectively. Column 3 weights every session equally; weighting by the number of subjects, which the unequal session sizes make relevant, gives 0.070 (s.e. 0.032, $p = 0.045$), the same as column 1. Randomization-inference p-values come from 5,000 permutations of the arm label across the nineteen sessions. The Panel C contrast is a difference between two photographic treatments and is not an estimate of an ethnic effect; see Section 6.3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 4 shows the nineteen numbers on which all of this rests. The control sessions run from 0.323 to 0.391, the indigenous-photograph sessions from 0.350 to 0.573, and the white-photograph sessions from 0.292 to 0.527. The treated distributions sit above the control distribution but overlap it, and the Indigenous mean is pulled up appreciably by one session. Figure 2, which plots these

units directly, is in our view a closer representation of the experiment than any regression table, and Figure 3 shows the estimate against 5,000 placebo re-assignments of the nineteen sessions. The magnitude bears comparison with the institutional levers. Li (2016) obtains about ten points from an informational message over the first fifteen rounds. Li et al. (2013) obtain twenty-five points from an opt-out default and thirty-two to thirty-seven from a priority rule. Our six to nine points is smaller than the institutional levers and comparable to the informational one, which is the right reference class: like a message and unlike a priority rule, a photograph is a change to the content of a request, not to the rules of allocation. Panel B of Table 3 separates the two photograph conditions. Unadjusted, the Indigenous-looking photograph raises registration by 6.7 points and the white-looking photograph by 7.3 points. The difference between them is -0.6 points with a 95 percent confidence interval from -10.1 to $+8.9$ and a randomization p-value of 0.906. Adjusted, the difference is -1.5 points with an interval from -9.5 to $+6.5$. The equally weighted session-means estimator puts it at 0.2 points. On every estimator the two photographs perform alike, and the interval is wide enough to accommodate differences of eight or nine points in either direction. We deliberately do not describe this as evidence about ethnicity. The two conditions differ in the photographs used as much as in the ethnicity those photographs convey, and with two photographs per category the two cannot be separated. The quantity Panel C estimates is the difference between these particular photographic treatments in this particular sample. That quantity is well defined, and close to zero. It is not the average effect of recipient ethnicity.

Table 4. Full-sample Estimates and Specification Checks

	(1)	(2)	(3)	(4)
Indigenous-looking photograph	0.092** (0.040)	0.110** (0.039)	0.092** (0.042)	0.092* (0.046)
White-looking photograph	0.030 (0.037)	0.040 (0.041)	0.044 (0.044)	0.045 (0.049)
Difference	0.062	0.070	0.048	0.047
randomization p	0.365	—	0.365	—
Round fixed effects	Yes	Yes	Yes	Yes
Demographics	No	Yes	Yes	Yes
Baseline attitudes	No	No	Yes	Yes
Lagged outcomes	No	No	No	Yes
Observations	3,120	3,120	3,120	2,925
Sessions	19	19	19	19

Notes: All sixteen rounds, which pools the costly and costless blocks and is therefore not our primary sample. CR1 standard errors clustered on session in parentheses; stars from the restricted wild cluster bootstrap. An equally weighted session-means regression gives 0.086 (s.e. 0.041, $p = 0.055$) for the Indigenous-looking photograph and 0.035 (s.e. 0.040, $p = 0.393$) for the white-looking one. The decomposition by the sex of the depicted person is reported in Appendix Table A2; it is not identified as a male-versus-female contrast because recipient sex is collinear with round parity. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure 2. Session-level Registration Rates (each dot = one session)

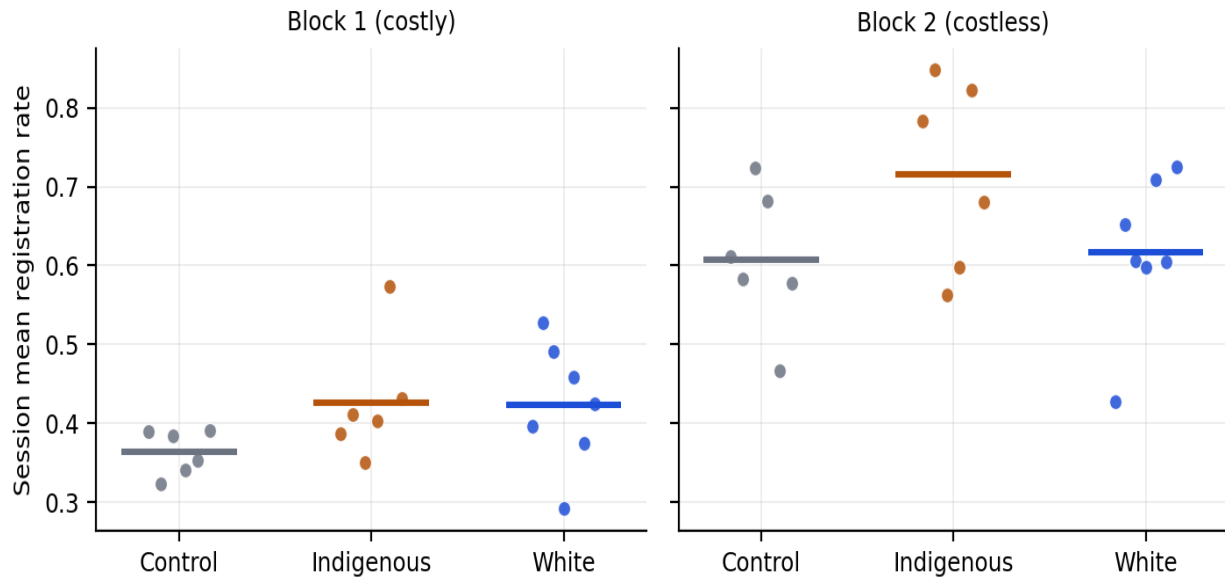


Figure 2. Session-level registration rates by arm and block. Each dot is one session; horizontal bars are arm means. The left panel is the primary sample. The figure shows the unit of variation that inference must respect.

Figure 3. Randomization Inference, Block 1 (5,000 Re-assignments of the 19 Sessions)

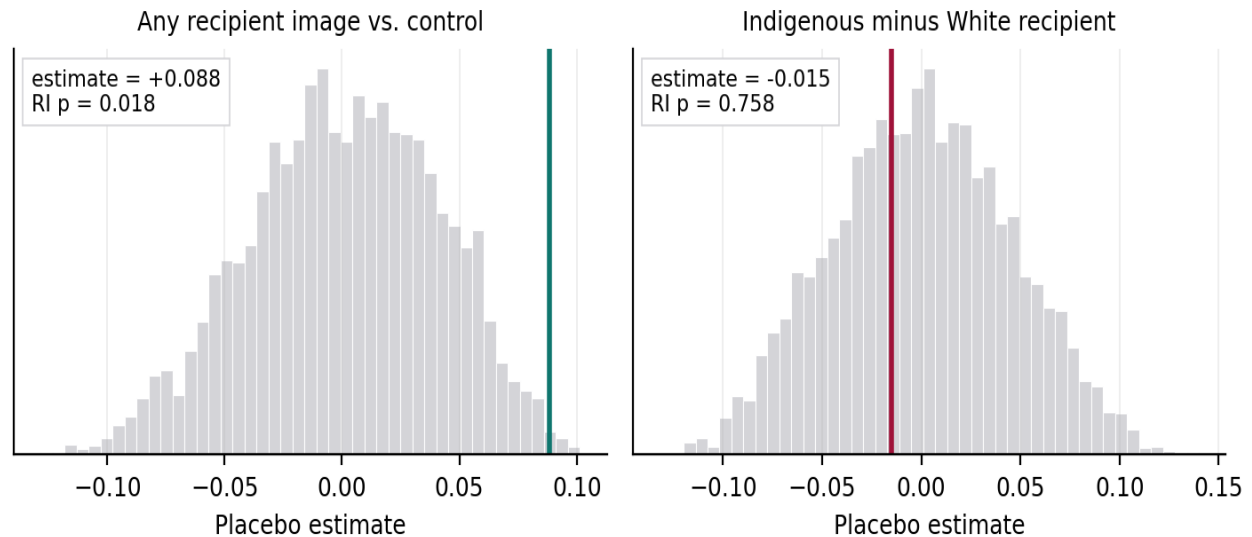


Figure 3. Randomization inference in the primary block. Histograms show estimates from 5,000 random re-assignments of the arm label across the nineteen sessions holding the six-six-seven allocation fixed; the vertical line is the estimate in the realized assignment.

Table 4 extends to all sixteen rounds, which pools the costly and costless blocks and is therefore not our primary sample. There the Indigenous coefficient is stable between 0.092 and 0.110 and individually distinguishable from zero, while the white coefficient sits between 0.030 and 0.045 and is not. The difference between them is nevertheless never significant, with a randomization p-value of 0.365.

Along the same lines Table 5 splits the sample by block. In the costly block both photographs raise registration and are statistically indistinguishable from one another. In the costless block neither is individually significant: the indigenous coefficient is 0.102 with a randomization p-value of 0.193, and the white coefficient is -0.008 with $p = 0.908$. Estimated jointly with an interaction, the photograph-by-block terms are 0.050 for the Indigenous condition and -0.087 for the white condition, neither significant. The pooled photograph effect falls from 0.088 in the costly block to 0.047 in the costless block.

Table 5. Treatment Effects by Cost Block

	Block 1 (costly)	Block 2 (costless)	Pooled, with Interaction
	(1)	(2)	(3)
Any recipient photograph	0.088*** (0.027)	0.047 (0.055)	
randomization p	{0.018}	{0.476}	
Indigenous-looking photograph	0.081** (0.029)	0.102 (0.064)	0.067** (0.028)
randomization p	{0.082}	{0.193}	
White-looking photograph	0.096** (0.036)	-0.008 (0.062)	0.087** (0.034)
randomization p	{0.026}	{0.908}	
Indigenous $\times$ costless			0.050 (0.055)
White $\times$ costless			-0.087 (0.052)
Indigenous – White	-0.015	0.110	
randomization p	{0.758}	{0.138}	
Observations	1,560	1,560	3,120
Sessions	19	19	19

Notes: All specifications include round fixed effects, demographics and baseline attitudes. Columns 1 and 2 are separate regressions estimated on each block. Column 3 pools both blocks and adds the interaction terms; in that

column the photograph rows are the Block 1 effects and the effect in Block 2 is the photograph row plus its corresponding $\times$ costless row, giving 0.117 for the Indigenous-looking photograph and 0.001 for the white-looking one. These implied values need not equal the column 2 estimates of 0.102 and -0.008 , because columns 1 and 2 fit the round fixed effects and covariates separately within each block while column 3 fits them once across all sixteen rounds. The block main effect is absorbed by the round fixed effects in column 3, so the interaction terms are difference-in-differences that net out the common block shift, including the restart effect. Stars from the restricted wild cluster bootstrap. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

The framework of Section 4 offers a reading: a treatment that shifts perceived benefit can only move subjects near the registration margin, and removing the monetary cost moves the margin. We state the inferential position plainly. Because block order is perfectly confounded with the cost change, the pattern across blocks is consistent with a cost-dependent treatment effect but does not identify one. The effect is concentrated in the costly block and not detected in the costless block. The second block differs from the first in the removal of costs, but also in order, accumulated experience, fatigue, the restart effect, the higher base rate against which any increase must register, and simple treatment decay of the kind Li (2016) documents for her message over exactly this horizon. Any of these produces the observed attenuation. The block split is descriptive and for the same reason we record but do not interpret the fact that the largest gap between the two photographs anywhere in our data—eleven points, in the costless block—is also not significant, at a randomization p -value of 0.138. It is a hypothesis for a counterbalanced replication.

The absence of an average difference conceals two responses of opposite sign. Twenty-five percent of our subjects were born outside Lima. In Peru provincial origin is among the strongest available correlates of Indigenous heritage, and it is observed without relying on self-identification, which only ten percent of the sample endorses. It is balanced across arms (24.6, 27.7 and 21.5 percent; $p = 0.67$) and, unlike the treatment, varies within session, so the interaction is better identified than the main effect. Table 6 reports the split. Among Lima-born subjects the Indigenous-looking photograph raises registration by 15.8 points ($p = 0.002$) and the white-looking photograph by 12.4 points ($p = 0.033$). Among subjects born outside Lima the Indigenous-looking photograph lowers registration by 6.7 points while the white-looking one does nothing. The interaction with provincial origin is -0.307 (wild bootstrap $p = 0.002$); the white interaction is -0.084 and insignificant. Dropping each session in turn moves the interaction only within -0.361 to -0.280 , and it is significant at five percent in all nineteen leave-one-out samples. A composite index adding public schooling, subsidized insurance and low tuition scale reproduces the pattern (-0.539 , $p < 0.001$), while none of the socioeconomic components does alone, which suggests that the heterogeneity is

more closely associated with region of birth than with the individual socioeconomic components examined here; the mechanism is not identified.

Table 6. Photograph Treatment Interacted with the Subject's Region of Birth (Block 1)

	Interaction model	Lima-born	Provincial
Indigenous-looking photograph	0.160*** (0.036)	0.158*** (0.043)	−0.067 (0.058)
White-looking photograph	0.116* (0.056)	0.124** (0.052)	0.032 (0.069)
Indigenous × provincial	−0.307*** (0.057)		
wild bootstrap p	[0.002]		
White × provincial	−0.084 (0.116)		
wild bootstrap p	[0.604]		
Subjects	195	147	48
Observations	1,560	1,176	384
Sessions	19	19	17

Notes: Block 1 only. All columns include round fixed effects, demographics and baseline attitudes; column 1 also includes the provincial main effect. Session-clustered standard errors in parentheses, restricted wild cluster bootstrap p-values in brackets. Leave-one-session-out estimates of the Indigenous interaction range from −0.361 to −0.280 and are significant at five percent in 19 of 19 samples. Provincial origin is balanced across arms ($p = 0.67$) and varies within session. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Figure 4 and the raw means make the pattern stark. In the control arm, provincial-origin subjects register more often than Lima-born subjects, 47.7 against 32.4 percent. Shown a white recipient the two groups are indistinguishable; shown an Indigenous recipient they diverge in opposite directions, Lima-born rising to 48.9 percent and provincial falling to 27.1 percent. The average effect over the whole sample is the sum of these two offsetting movements. This is exploratory and cannot offer a confident mechanism. It is the opposite of in-group favoritism. One reading is identity distancing: our subjects are students at an elite Lima university, and the mestizaje literature describes upward mobility in Peru as historically entailing distance from indigeneity (de la Cadena

2000; Paredes 2018). A second is that provincial-origin subjects hold different beliefs about who is on the waiting list. The instrument also elicited a complete ranking of the four photographs from every subject, including control subjects, who saw no photograph at any point during play. The four ranking variables form a permutation of one to four for every subject, and twenty-two of the twenty-four possible orderings appear. The stored ranking admits two conventions, since the value may be the rank assigned to each photograph or the photograph occupying each rank; we report only conclusions invariant to which is correct, giving both figures throughout. The ranking was elicited after the sixteen rounds, and no subject saw any photograph beforehand, so the control condition is a clean no-photograph baseline. Because the ranking is measured after treatment, the differences across arms in Table 7 are themselves a downstream effect of exposure: both treated arms place the Indigenous man higher than control subjects do. Interestingly, the ranking is not random. In the control arm a white recipient is placed first by 86.2 percent of subjects under the first convention and 69.2 percent under the second, against 50 percent if ranking were random over the two white and two Indigenous photographs; both reject at conventional levels (binomial $p = 2 \times 10^{-9}$ and $p = 0.003$), and a chi-square test of independence between photograph and rank gives 105.9 on nine degrees of freedom. The ordering is monotone in both ethnicity and sex, from the white man to the Indigenous woman, who is placed first by fewer than five percent of subjects. Figure 4 shows the pattern.

Table 7. Stated Ranking of Photographs and Behavior

	Convention A	Convention B	
<i>Panel A. Share ranking each photograph first, control arm (n = 65)</i>			
White man	0.585	0.585	
White woman	0.277	0.108	
Indigenous man	0.092	0.200	
Indigenous woman	0.046	0.108	
Any white photograph first	0.862	0.692	$p < 0.01$ vs 0.50
<i>Panel B. Correlates of ranking an Indigenous photograph first</i>			
Overall discrimination index (z)	+0.79		$p < 0.001$
Racial discrimination index (z)	+0.57		$p = 0.001$
Attends religious services	+0.21		$p = 0.012$
Born outside Lima	-0.09		$p = 0.22$

	Convention A	Convention B	
<i>Panel C. Ranking × photograph treatment on registration (Block 1)</i>			
Indigenous × ranks white first	+0.064	+0.031	ns
White × ranks white first	+0.009	+0.029	ns
Ranks white first, main effect	−0.057	+0.013	ns

Notes: Convention A treats the stored value as the rank assigned to each photograph; convention B treats it as the photograph occupying each rank. The ranking was administered after the 16 rounds; no subject saw any photograph before the experiment. Panel A restricts to the control arm, whose subjects saw no photograph at any point during play. Panel B reports differences in means between subjects ranking an Indigenous and a white photograph first. Panel C reports interaction coefficients from the Block 1 specification of Table 3 with session-clustered standard errors; none is significant at 10 percent.

Figure 4. Stated Ranking of the Four Photographs and the Behavioral Response

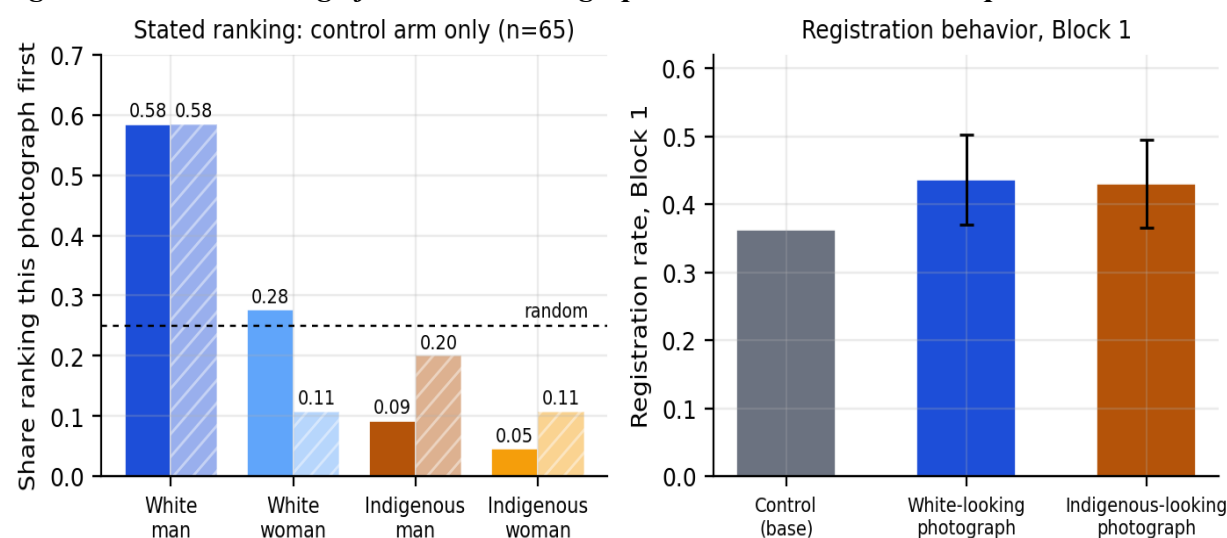


Figure 4. Left: share of control-arm subjects ranking each photograph first, under both storage conventions (solid and hatched); the dashed line is the twenty-five percent that random ranking would produce. Right: registration in the costly block by arm, with 95 percent session-clustered intervals.

Two features are worth mentioning. First, the ranking measures something that appears to be discrimination-related: subjects who place an Indigenous recipient first score 0.79 standard deviations higher on the overall discrimination index than those who place a white recipient first ($p < 0.001$), and 0.57 higher on the racial index ($p = 0.001$). The direction is itself informative about what those indices capture. They behave like measures of the perceived prevalence of discrimination in Peruvian society, not of the respondent's own endorsement of it: subjects more alert to discrimination rank the disadvantaged recipient first and register less often overall. Second, the ranking coheres with the region-of-birth split: subjects born outside Lima place a white recipient first more often (85.4 against 76.2 percent), the same direction as their negative

behavioral response. And yet the stated hierarchy does not move the average difference between the two arms. Interacting an indicator for ranking a white recipient first with the photograph treatment yields interaction terms of 0.064 and 0.009 under the first convention and 0.031 and 0.029 under the second, none remotely significant, with the main effect of the ranking also indistinguishable from zero. Subjects who put the Indigenous woman last do not, on average, register less often in the arm that shows her. This is an attitude–behavior gap in a domain where the behavior is incentivized, and the stakes are framed as survival. Two qualifications belong with it. The ranking is over the same four photographs as the treatment and so inherits the stimulus-sampling limitation mentioned above. And the region-of-birth split shows the behavioral average is itself masking opposed responses, so the absence of an average difference is a statement about the mean, not about every subject.

As mentioned above, every subject in a treated session saw two of the four photographs, one in odd rounds and one in even rounds, and freely ranked both. We can therefore ask whether a subject registers more often in the rounds showing the photograph that subject personally ranked more favorably, holding the subject and the round fixed. This asks nothing about ethnicity and requires no generalization across photographs. The specification includes subject and round fixed effects, so identification comes from within-subject variation across rounds interacted with the subject's own ordering of the pair shown. Because the two photographs in a pair differ in sex, and sex is collinear with round parity, the round fixed effects absorb the parity profile common to everyone. Table 8 reports the result. Registration falls by 1.9 percentage points for each position lower the shown photograph occupies in the subject's own ranking under the second storage convention ($p = 0.055$), and by 1.4 points under the first ($p = 0.149$). Both are signed as the mechanism predicts, and the implied difference between showing a subject their best-ranked and their worst-ranked photograph is about 5.7 points, of the same order as the seven-point average effect of showing a photograph at all. The indicator version, for whether the round shows the preferred member of the pair, gives 2.8 points ($p = 0.097$) and 1.8 points ($p = 0.348$). The placebo is the reason we report this at all. Control subjects ranked the same four photographs and saw none of them at the moment of decision. Constructing the identical variable for them—the rank they gave to the photograph they would have seen in that round had they been assigned to a treated session—yields -0.6 points ($p = 0.82$) and -1.8 points ($p = 0.46$), both indistinguishable from zero. The ranking predicts behavior only among subjects for whom the ranked object was actually on the screen. Two things

limit what this can bear. The estimate clears five percent under one reading of an ambiguous variable and not the other. More importantly, the ranking is measured post-treatment: a subject who registered readily when shown a photograph may have come to rank it higher, rather than the ranking having driven the registration. The placebo shows the association is not mechanical and the round fixed effects rule out a common time profile, but neither establishes the direction of causation, so we read the result as a consistency check. What it establishes is narrower than a mechanism: the photographs are not interchangeable stimuli, and the dimension along which they differ is one that subjects register and that predicts their own behavior.

Table 8. Within-subject Test: Subject's Own Ranking of Photograph Shown

	Convention A	Convention B	Placebo (control)
	(1)	(2)	(3)
<i>Panel A. Rank of the photograph shown (1 = best)</i>			
Estimate	-0.014	-0.019*	-0.018
	(0.009)	(0.009)	(0.023)
95% CI	[-0.034, 0.006]	[-0.039, 0.001]	[-0.077, 0.040]
p	0.149	0.055	0.458
<i>Panel B. Shows the preferred member of the pair</i>			
Estimate	0.018	0.028*	-0.006
	(0.018)	(0.015)	(0.024)
95% CI	[-0.022, 0.057]	[-0.006, 0.061]	[-0.068, 0.056]
p	0.348	0.097	0.820
Subject fixed effects	Yes	Yes	Yes
Round fixed effects	Yes	Yes	Yes
Sample	treated arms	treated arms	control arm
Observations	2,080	2,080	1,040

Notes: Linear probability models with subject and round fixed effects; standard errors clustered on session. Conventions A and B are the two readings of the stored ranking described in the text accompanying Table 8. Column 3 builds the identical regressor for control subjects using the photograph they would have been shown in that round under a treated assignment; they ranked the photographs but saw none of them while deciding. Panel B replaces the rank with an indicator for whether the round shows the higher-ranked member of the subject's assigned pair. The ranking is measured after the rounds and is therefore post-treatment; see the discussion in the text. * $p < 0.10$.

6. Robustness

Table 9 collects the robustness and inference diagnostics; Figure 2 shows the session-level variation these diagnostics quantify. Dropping each session in turn moves the full-sample Indigenous coefficient within a range of 0.073 to 0.115, and it remains significant at five percent in ten of nineteen leave-one-out samples—evidence that the estimate is not driven by a single session but also that it sits close to the boundary of what nineteen clusters can resolve. Session size ranges from five to fourteen subjects, and because the availability of organs depends on how many other subjects are playing, thin sessions differ economically from full ones. The white arm contains the three smallest sessions. Session size is weakly positively associated with registration ($r = 0.16$ across sessions; coefficient 0.010 per subject, $p = 0.097$ in the primary block), so this works against the white arm. Controlling for it strengthens both estimates slightly, to 0.084 for Indigenous and 0.101 for white in the primary block and does not change any conclusion.

Table 9. Robustness and Inference Diagnostics

Check	Result	Assessment
Leave-one-session-out (Indigenous, all rounds)	0.073 to 0.115	Stable; 10/19 significant at 5%
Session-means regression, all rounds (Indigenous)	0.086 (s.e. 0.041, $p = 0.055$)	Conservative, same sign and size
Controlling for session size	0.084 / 0.101	Slightly stronger
Population-averaged logit, session-clustered	0.42 / 0.41 (Indigenous)	Same picture as LPM
Session vs subject clustering, s.e. ratio	0.96	Variance is not the binding issue
Intraclass correlation, session	0.023	Low; subject-level ICC is 0.244
Between-session s.d. of registration	0.067	Basis for the MDE below
MDE, Indigenous vs control (80% power)	0.121	Underpowered for moderate gaps
MDE, White vs control (80% power)	0.115	Underpowered for moderate gaps
MDE, any image vs control (80% power)	0.099	Adequate for the effect found

Notes: Minimum detectable effects computed from the between-session standard deviation of the registration rate within arm, at five percent size and eighty percent power, with the realized numbers of sessions per arm.

A binary-response model gives the same picture. Estimating a population-averaged logit with the four photograph indicators and the covariate set used in the original analysis of these data, the coefficients on the Indigenous man and woman are 0.44 and 0.43 with subject-clustered standard errors and 0.42 and 0.41 with session-clustered ones; the average ratio of session- to subject-clustered standard errors across the four image terms is 0.96. This is the diagnostic that matters for how the present results relate to a conventional analysis of these data: the correction is not mainly a variance correction. Moving from a specification without round fixed effects and with normal critical values to one with round fixed effects and design-based p-values shifts the Indigenous result from p below 0.01 to p of about 0.08.

7. Discussion and Conclusion

The cleanest thing this experiment establishes is that a photograph of a prospective recipient, attached to a registration request that is otherwise identical, raises registration by something in the range of six to nine percentage points when registering is costly. Within the organ-donation paradigm this is new. The literature from Kessler and Roth (2012) onward has varied the rules of allocation, the default, the price and the accompanying message, and has left the recipient an anonymous position in a queue. Our result says that the presentation of the recipient is also a margin. We collected no measure of perceived recipient need, emotional response, empathy, perceived similarity, deservingness, attention, or beliefs about what the experimenter wanted. Identifiability, visual salience, affective response to a human face, humanization of an abstract problem and experimenter demand all predict the sign we observe. The framework of Section 4 is deliberately agnostic among them: each is a story about why perceived benefit shifts, and the experiment measures the shift. Distinguishing these channels requires a design that varies the information attached to a fixed photograph—name, age, history—which is the natural next experiment and would also settle whether anything here deserves the label identifiable victim.

On policy we want to be more disciplined than motivation usually is. The marginal printing cost of a photograph is negligible, but it does not follow that photographs would raise actual registration. The subjects are 195 students at a selective Lima university, and the decision is a repeated laboratory game, not a once-in-a-lifetime encounter with a registry form. The depicted person is not the actual recipient of anything, and the costs are monetary, while Kessler and Roth (2012) identify the psychological costs as the real barrier. The effect is confined to the costly block,

and the closest antecedent, the message treatment of Li (2016), decayed to nothing within fifteen rounds.

The photographic contrast is a different matter. We did not find a difference between the two conditions, the interval is wide, and we showed that even a precise estimate would not identify an ethnic effect with two photographs per category. Furthermore, it is unlikely to be a story about subjects failing to register the ethnic dimension: they rank the photographs by it almost unanimously, the response splits by region of birth, and on the one margin the design identifies cleanly, subjects register more in the rounds showing the photograph they themselves ranked higher. Assigning treatment at the session level and analyzing it at the decision level is natural in a repeated game and inflates apparent precision invisibly. Here the design effect from session-level clustering is modest in variance terms, since the session-level intraclass correlation is only 0.023; the substantive change in inference instead comes from three other choices: including round fixed effects, using the unadjusted estimate as primary, and testing the difference between the two treated arms.

In the future, several extensions may follow. Varying the information attached to a fixed photograph—a name, an age, a sentence of history—would separate identification from visual presentation. Multiple independently randomized photographs within each ethnic category, pre-tested for comparability and supported by many more sessions, would deliver the ethnic test this design cannot. Counterbalancing the cost blocks would identify the cost effect and distinguish it from decay. And stratifying on region of birth would turn the region-of-birth heterogeneity from an exploratory pattern into a test. Given that Peru registers roughly two deceased donors per million population, knowing which prompt works is particularly relevant.

References

- Ai, C., and E. C. Norton. 2003. "Interaction Terms in Logit and Probit Models." *Economics Letters* 80(1): 123–129.
- Akerlof, G. A., and R. E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics* 115(3): 715–753.
- Arrow, K. J. 1973. "The Theory of Discrimination." In *Discrimination in Labor Markets*, edited by O. Ashenfelter and A. Rees. Princeton: Princeton University Press.
- Balliet, D., J. Wu, and C. K. W. De Dreu. 2014. "Ingroup Favoritism in Cooperation: A Meta-Analysis." *Psychological Bulletin* 140(6): 1556–1581.
- Becker, G. S. 1957. *The Economics of Discrimination*. Chicago: University of Chicago Press.
- Berge, L. I. O., K. Bjorvatn, S. Galle, E. Miguel, D. N. Posner, B. Tungodden, and K. Zhang. 2020. "Ethnically Biased? Experimental Evidence from Kenya." *Journal of the European Economic Association* 18(1): 134–164.
- Bergh, R., and D. Reinstein. 2021. "Empathic and Numerate Giving: The Joint Effects of Victim Images and Charity Evaluations." *Social Psychological and Personality Science* 12(3): 407–416.
- Clark, H. H. 1973. "The Language-as-Fixed-Effect Fallacy: A Critique of Language Statistics in Psychological Research." *Journal of Verbal Learning and Verbal Behavior* 12(4): 335–359.
- Judd, C. M., J. Westfall, and D. A. Kenny. 2012. "Treating Stimuli as a Random Factor in Social Psychology: A New and Comprehensive Solution to a Pervasive but Largely Ignored Problem." *Journal of Personality and Social Psychology* 103(1): 54–69.
- Wells, G. L., and P. D. Windschitl. 1999. "Stimulus Sampling and Social Psychological Experimentation." *Personality and Social Psychology Bulletin* 25(9): 1115–1125.
- Cameron, A. C., J. B. Gelbach, and D. L. Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* 90(3): 414–427.
- Cameron, A. C., and D. L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50(2): 317–372.
- Canay, I. A., A. Santos, and A. M. Shaikh. 2021. "The Wild Bootstrap with a 'Small' Number of 'Large' Clusters." *Review of Economics and Statistics* 103(2): 346–363.
- Castillo, M., R. Petrie, and M. Torero. 2007. "Social Barriers to Cooperation: Experiments on the Extent and Nature of Discrimination in Peru." *Experimental Economics Center Working Paper 2007-01*, Georgia State University.
- Chan, A., and A. E. Roth. 2024. "Regulation of Organ Transplantation and Procurement: A Market-Design Lab Experiment." *Journal of Political Economy* 132(11): 3827–3866.
- Chen, D. L., M. Schonger, and C. Wickens. 2016. "oTree—An Open-Source Platform for Laboratory, Online, and Field Experiments." *Journal of Behavioral and Experimental Finance* 9: 88–97.
- Chen, Y., and S. X. Li. 2009. "Group Identity and Social Preferences." *American Economic Review* 99(1): 431–457.

- Cryder, C. E., G. Loewenstein, and R. Scheines. 2013. "The Donor Is in the Details." *Organizational Behavior and Human Decision Processes* 120(1): 15–23.
- de la Cadena, M. 2000. *Indigenous Mestizos: The Politics of Race and Culture in Cuzco, Peru, 1919–1991*. Durham: Duke University Press.
- Donald, S. G., and K. Lang. 2007. "Inference with Difference-in-Differences and Other Panel Data." *Review of Economics and Statistics* 89(2): 221–233.
- Fang, M., M. Li, and Y. E. Riyanto. 2026. "Boosting Donation and Welfare: Augmenting Donor-Priority Rule with Information Provision." *European Economic Review* 181: 105181.
- Galarza, F. B., and G. Yamada. 2014. "Labor Market Discrimination in Lima, Peru: Evidence from a Field Experiment." *World Development* 58: 83–94.
- Genevsky, A., D. Västfjäll, P. Slovic, and B. Knutson. 2013. "Neural Underpinnings of the Identifiable Victim Effect: Affect Shifts Preferences for Giving." *Journal of Neuroscience* 33(43): 17188–17196.
- Gonzales Moreno, M., J. Duran, O. Ponce, G. Navarro, and M. Benavides. 2020. "Organ Donation in Perú: Why Are We Failing? Conclusions from the Last 5 Years." *American Journal of Transplantation* 20(suppl. 3), abstract.
- Grimm, V., V. Utikal, and L. Valmasoni. 2017. "In-Group Favoritism and Discrimination among Multiple Out-Groups." *Journal of Economic Behavior and Organization* 143: 254–271.
- Hawley, Z., D. Li, K. Schnier, and N. Turgeon. 2018. "Can We Increase Organ Donation by Reducing the Disincentives? An Experimental Analysis." *Economics and Human Biology* 29: 128–137.
- Herr, A., and H.-T. Normann. 2016. "Organ Donation in the Lab: Preferences and Votes on the Priority Rule." *Journal of Economic Behavior and Organization* 131(B): 139–149.
- Herr, A., and H.-T. Normann. 2019. "How Much Priority Bonus Should Be Given to Registered Organ Donors? An Experimental Analysis." *Journal of Economic Behavior and Organization* 158: 367–378.
- Jenni, K. E., and G. Loewenstein. 1997. "Explaining the Identifiable Victim Effect." *Journal of Risk and Uncertainty* 14(3): 235–257.
- Johnson, E. J., and D. G. Goldstein. 2003. "Do Defaults Save Lives?" *Science* 302(5649): 1338–1339.
- Kessler, J. B., and A. E. Roth. 2012. "Organ Allocation Policy and the Decision to Donate." *American Economic Review* 102(5): 2018–2047.
- Kessler, J. B., and A. E. Roth. 2014. "Loopholes Undermine Donation: An Experiment Motivated by an Organ Donation Priority Loophole in Israel." *Journal of Public Economics* 114: 19–28.
- Kogut, T., and I. Ritov. 2005. "The 'Identified Victim' Effect: An Identified Group, or Just a Single Individual?" *Journal of Behavioral Decision Making* 18(3): 157–167.
- Lane, T. 2016. "Discrimination in the Laboratory: A Meta-Analysis of Economics Experiments." *European Economic Review* 90: 375–402.

- Lavee, J., T. Ashkenazi, G. Gurman, and D. Steinberg. 2010. "A New Law for Allocation of Donor Organs in Israel." *Lancet* 375(9720): 1131–1133.
- Lee, S., and T. H. Feeley. 2016. "The Identifiable Victim Effect: A Meta-Analytic Review." *Social Influence* 11(3): 199–215.
- Li, D. 2016. "Effect of Persuasive Messages on Organ Donation Decisions: An Experimental Test." *Journal of Economic Behavior and Organization* 131(B): 150–159.
- Li, D., Z. Hawley, and K. Schnier. 2013. "Increasing Organ Donation via Changes in the Default Choice or Allocation Rule." *Journal of Health Economics* 32(6): 1117–1129.
- Li, M., Y. E. Riyanto, and M. Xu. 2023. "Prioritized Organ Allocation Rules under Compatibility Constraints." *Games and Economic Behavior* 141: 403–427.
- Li, S. X. 2020. "Group Identity, Ingroup Favoritism, and Discrimination." In *Handbook of Labor, Human Resources and Population Economics*. Cham: Springer.
- MacKinnon, J. G., M. Ø. Nielsen, and M. D. Webb. 2023. "Cluster-Robust Inference: A Guide to Empirical Practice." *Journal of Econometrics* 232(2): 272–299.
- MacKinnon, J. G., and M. D. Webb. 2018. "The Wild Bootstrap for Few (Treated) Clusters." *Econometrics Journal* 21(2): 114–135.
- MacKinnon, J. G., and M. D. Webb. 2020. "Randomization Inference for Difference-in-Differences with Few Treated Clusters." *Journal of Econometrics* 218(2): 435–450.
- Maier, M., et al. 2023. "Revisiting and Rethinking the Identifiable Victim Effect: Replication and Extension of Small, Loewenstein, and Slovic (2007)." *Collabra: Psychology* 9(1): 90203.
- Moreno, M., H. Ñopo, J. Saavedra, and M. Torero. 2012. "Detecting Gender and Racial Discrimination in Hiring through Monitoring Intermediation Services: The Case of Selected Occupations in Metropolitan Lima, Peru." *World Development* 40(2): 315–328.
- Paredes, C. L. 2018. "Multidimensional Ethno-Racial Status in Contexts of Mestizaje: Ethno-Racial Stratification in Contemporary Peru." *Socius* 4: 1–17.
- Phelps, E. S. 1972. "The Statistical Theory of Racism and Sexism." *American Economic Review* 62(4): 659–661.
- Planas, M.-E., P. J. García, M. Bustelo, C. P. Carcamo, S. Martinez, H. Ñopo, J. Rodriguez, M.-F. Merino, and A. Morrison. 2015. "Effects of Ethnic Attributes on the Quality of Family Planning Services in Lima, Peru: A Randomized Crossover Trial." *PLoS ONE* 10(2): e0115274.
- Schelling, T. C. 1968. "The Life You Save May Be Your Own." In *Problems in Public Expenditure Analysis*, edited by S. B. Chase. Washington, DC: Brookings Institution.
- Small, D. A., and G. Loewenstein. 2003. "Helping a Victim or Helping the Victim: Altruism and Identifiability." *Journal of Risk and Uncertainty* 26(1): 5–16.
- Small, D. A., G. Loewenstein, and P. Slovic. 2007. "Sympathy and Callousness: The Impact of Deliberative Thought on Donations to Identifiable and Statistical Victims." *Organizational Behavior and Human Decision Processes* 102(2): 143–153.

- Tajfel, H., M. G. Billig, R. P. Bundy, and C. Flament. 1971. "Social Categorization and Intergroup Behaviour." *European Journal of Social Psychology* 1(2): 149–178.
- Young, A. 2019. "Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results." *Quarterly Journal of Economics* 134(2): 557–598.
- Zang, R., J. Li, and X. Niu. 2025. "Dynamic Norms and Organ Donation." *Economic Inquiry* 63(2).

Appendix

Table A1. Full Randomization Balance

Variable	Control	Indig.	White	p (Indig.)	p (White)
Age	22.031	21.308	22.708	0.315	0.422
Female	0.662	0.662	0.692	1.000	0.670
Semester	6.831	6.938	7.969	0.830	0.093
Fee scale	3.923	4.031	4.015	0.478	0.584
Number of siblings	1.478	1.538	1.523	0.568	0.752
Father's education	3.785	3.769	3.908	0.917	0.502
Mother's education	3.554	3.892	3.600	0.037	0.757
Health insurance	0.815	0.800	0.877	0.786	0.305
Economics major	0.431	0.508	0.385	0.357	0.518
White or mestizo	0.938	0.877	0.892	0.360	0.350
Ethnic minority	0.062	0.123	0.108	0.360	0.350
Religious affiliation	0.769	0.769	0.708	1.000	0.345
Attends services	0.631	0.631	0.677	1.000	0.563
Donor on ID card	0.415	0.354	0.354	0.510	0.500
Willing to donate	0.738	0.877	0.769	0.008	0.697
Knows the procedure	0.200	0.154	0.154	0.300	0.316
Knowledge about donation	0.672	0.669	0.682	0.913	0.727
Closeness to donation	0.215	0.154	0.138	0.142	0.149
Perception of donation	0.646	0.692	0.631	0.060	0.533
Discrimination: general	0.401	0.429	0.431	0.007	0.016
Discrimination: racial	0.481	0.470	0.447	0.647	0.309
Discrimination: socioeconomic	0.569	0.581	0.564	0.435	0.855
Discrimination: gender	0.452	0.465	0.493	0.508	0.092
Discrimination: overall	0.497	0.508	0.503	0.279	0.764

Notes: Subject-level means. p-values from a regression of the characteristic on the two arm indicators with standard errors clustered on session.

Table A2. Decomposition by Sex of Depicted Recipient

Photograph	Block 1, unadj.	Block 1, adj.	All rounds, adj.
White-looking man	0.042 (0.044)	0.065 (0.045)	0.016 (0.047)
White-looking woman	0.104** (0.041)	0.127*** (0.043)	0.072 (0.046)
Indigenous-looking man	0.058 (0.040)	0.071* (0.036)	0.094** (0.041)
Indigenous-looking woman	0.077* (0.044)	0.091** (0.041)	0.090* (0.045)

Notes: Since the sex of the depicted person alternates deterministically with round parity, no session ever shows a man and a woman in the same round, and the four photograph indicators are each identified only against the control arm in the corresponding rounds. They do not identify a male-versus-female recipient effect. We report them because the within-category spread is the input to the stimulus-sampling discussion of Section 6. Session-clustered standard errors in parentheses; all specifications include round fixed effects. In the costly block the two white-looking photographs differ from each other by 6.2 points and the two Indigenous-looking photographs by 1.9, against a gap between the ethnic categories of 0.6 points; none of these three differences is distinguishable from zero ($p = 0.23, 0.71$ and 0.90 respectively). $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A3. Photograph Treatment Interacted with Characteristics

Moderator	Indigenous $\times$ M	White $\times$ M	Sample
Overall discrimination index (z)	-0.043 (0.040)	0.015 (0.038)	All
Racial discrimination index (z)	-0.071 (0.051)	0.003 (0.049)	All
Attends religious services	0.013 (0.086)	0.237** (0.101)	All
Attends religious services	-0.074 (0.116)	0.158 (0.111)	Block 1
Female subject	-0.037 (0.072)	0.026 (0.079)	All
White or mestizo self-ID	0.140 (0.127)	0.226 (0.133)	All
Registered donor on ID card	0.045 (0.068)	0.107 (0.084)	All

Level effect of discrimination index on registration: -0.048** (0.017) per s.d.

Notes: Table reports interactions of the two photograph indicators with six pre-specified moderators. Taken at face value this would suggest that the white-looking photograph activates something specific to observant subjects, however, restricted to the primary block the interaction falls to 0.158 with $p = 0.172$, so the result appears to be driven by the confounded second block. Each row is a separate regression including both image indicators, the moderator, their interactions, round fixed effects, demographics and baseline attitudes. Session-clustered standard errors in parentheses. Twelve interaction terms are tested; none survives a conservative multiplicity correction. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A4. What Subjects Know About Donation in Peru

Item	Share Correct
Who makes the donation decision	0.990
Donation involves no payment	0.954
Donor characteristics	0.851
Registration requirements	0.574
Role of the family in the decision	0.385
Costs borne by the donor family	0.292
Overall knowledge score	0.674

Notes: The knowledge score used throughout is the mean of six true/false items, which we verified reproduces the summary variable. Ninety-nine percent of subjects know who makes the decision and 95 percent know that donation is not paid, but only 38.5 percent answer the item on the role of the family correctly and only 29.2 percent the item on costs borne by the donor family. N = 195 subjects. The six items are the components of the knowledge index.