



Munich Personal RePEc Archive

Choosing Your Ethnicity: A Longitudinal Analysis of Ethnic Identity Choice and Intra-Individual Ethnicity Change

Rademakers, Robbert and van Hoorn, Andre

Radboud University, Radboud University

20 March 2020

Online at <https://mpra.ub.uni-muenchen.de/99184/>

MPRA Paper No. 99184, posted 23 Mar 2020 08:48 UTC

Choosing Your Ethnicity: A Longitudinal Analysis of Ethnic Identity Choice and Intra-Individual Ethnicity Change

Robbert Rademakers^a & André van Hoorn^b

^a *Corresponding author: Radboud University, Institute for Management Research, PO Box 9108, 6500 HK, Nijmegen, the Netherlands. Tel: +31 243 612 005, Email: R.H.*

Rademakers@fm.ru.nl.

^b *Radboud University, Institute for Management Research, PO Box 9108, 6500 HK, Nijmegen, the Netherlands. Tel: +31 243 611 663, Email: A.vanHoorn@fm.ru.nl.*

Choosing Your Ethnicity: A Longitudinal Analysis of Ethnic Identity Choice and Intra-Individual Ethnicity Change

Abstract

This paper studies individuals' possible choice to forgo their ancestral ethnicity and adopt a specific new ethnicity. We first use individual-level panel data for Indonesia as well as other countries (e.g., the U.S.) to document the pervasiveness of intra-individual ethnicity change and its coincidence with major life events, particularly, interethnic marriage. Next, we focus on individuals who have intermarried and exploit variation in deep-rooted community-level norms on matrilocality (co-residence with the wife's family) to identify how differences in expected costs and benefits of ethnicity change causally affect newlyweds' choice to adopt a specific ethnicity (i.e., their spouses' ethnicity) or not. Results obtained using a three-wave panel comprising more than 13,000 Indonesians confirm the expected effect of matrilocality, as newly intermarried men (women) are significantly more (less) likely to adopt their spouses' ethnicity when the couple lives in a matrilocal community compared to a non-matrilocal one. Because ethnicity change is a means to fit in, important implication of our findings is that in many countries key statistics on ethnic fractionalization and segregation are severely inflated.

JEL classifications: J12, J15, Z13

Keywords: Intra-individual ethnic fluidity; ethnic boundaries; ethnic options; intermarriage; post-marital residence norms; racial identity change

I. INTRODUCTION

This paper studies individuals' possible choice to forgo their original or ancestral ethnic identity and adopt a specific new ethnicity. Ethnic group membership has long been considered an important basis for discrimination and is linked to systematic group differences in key socioeconomic outcomes such as employment, wages, education, health, and credit access, but also in residential area, occupation, romantic partner choice and food preferences (Alesina et al. 2016; Atkin et al. 2019; Bisin and Verdier 2000; Chiswick 1988; Giuliano and Ransom 2013; Goldsmith et al. 2006; Hanna and Linden 2012; Hellerstein and Neumark 2008; Mays et al. 2007; Petersen et al. 2019; Rubinstein and Brenner 2013). Underlying the large literature on ethnic segregation and discrimination is the conceptualization of individuals' ethnicity as something unmalleable, a fixed personal trait that one inherits from one's parents and that is often marked by phenotypical characteristics such as skin tone or hair color (Brubaker 2006; Wimmer 2013). Recent empirical analyses of longitudinal census data, however, provide strong evidence that people's ethnic identity is not fixed and can change over the course of their lives. Although observed intra-individual variation in ethnic identity is at least partly explained by measurement error (Alba et al. 2016; Davenport 2020; Kramer et al. 2016), it appears that individuals are flexible and can identify as one ethnicity, say Black or African American, first and as another ethnicity, say White American, later on (Dahis et al. 2019; Liebler et al. 2017; Saperstein and Penner 2012). However, these analyses remain silent on individuals' reasons for changing their ethnicity.

We seek to understand intra-individual ethnicity change as the outcome of purposeful decision making on the part of the individual. Theoretically, we follow Akerlof and Kranton (2000) and, particularly, the long-standing socio-psychological literature on ethnic boundaries and acculturation (e.g., Barth 1969; Berry 1997; Waters 1990) to propose that changing one's ethnic identity involves, among others, weighing the net benefits of retaining

one's current ethnicity against potential net benefits of converting to another ethnicity. Empirically, we use longitudinal data to consider (i) individual factors predicting intra-individual ethnic identity change, specifically so-called major life events (Cohen et al. 2019; Kanner et al. 1981), and (ii) plausibly exogenous contextual factors affecting the net benefits that individuals can expect to receive from adopting a specific ethnicity, specifically newly intermarried individuals conversion to the ethnicity of their spouse. We thereby follow the dominant understanding of ethnic identity in the social sciences, which is that individuals' ethnicity involves a subjective feeling of groupness based on common ancestry that is the result of shared phenotypical traits, religious beliefs, region of origin and, particularly, shared language (Chandra 2006; Weber 1921; Wimmer 2013).¹

The main research context for this study is Indonesia, although we also present evidence showing that chief empirical observations of our study hold for the U.S. (intra-individual changes in racial identity) and India (intra-individual changes in caste identity) as well. In Indonesia, intra-individual ethnicity change is not a strictly dichotomous phenomenon (you are either White or you are Black) but involves choosing among a rich set of alternative ethnicities. We can therefore consider two distinct forms of ethnicity change, making Indonesia the ideal research context for a longitudinal analysis of individuals' ethnic identity. The first form is so-called generic or *undirected* ethnic identity change, which

¹ The literature correspondingly speaks of ethnoracial groups, ethnoreligious groups, ethnoregional groups and ethnolinguistic groups (Wimmer 2013). Race is thus a form of ethnicity and a famous example of an ethnoracial group division is Blacks vs. Whites in the U.S. Well-known ethnoreligious groups include the division of Shia's, Sunni's, Maronites and Druzes in Lebanon. Worldwide, however, the most prominent ethnic divisions involve ethnolinguistic differences, particularly in Asia and Africa (Alesina et al. 2003; Easterly and Levine 1997).

involves forgoing one's old ethnicity and adopting an unspecified new ethnic identity. The second is so-called *directed* ethnic identity change, which involves the choice of converting to a specific alternative ethnicity. For our empirical analyses, we follow the official practice of statistical agencies worldwide (Government of India 2011; Roth 2010; Statistics Indonesia 2010; U.S. Bureau of the Census 2010) and measure ethnic identity mainly on the basis of self-reports. Obvious drawbacks of self-reported data notwithstanding, this ensures that our results match certified government statistics on ethnicity that, in turn, are grounded in the standard understanding of ethnicity as a subjective feeling of belonging (Weber 1921; Wimmer 2013). Still, evidence indicates that all the different indicators of individuals' ethnicity that we consider—self-reported ethnicity, ethnic ancestry or descent, and ethnic language spoken—are associated with the prevalence of certain ethnicity-specific behaviors such as the eating of beef or the owning of pigs (see Tables A2.1-A2.3 in Appendix A). What is more, we find that changes in self-reported ethnicity are associated with significant behavioral adjustments, particularly speaking or learning the language of the individual's new ethnic group.

We divide our empirical analysis in two parts. In the first part, we present descriptive statistics and use regression analysis with an extensive set of control variables to document the pervasiveness of intra-individual ethnicity change in Indonesia. Longitudinal data from the Indonesian Family Life Survey or IFLS (Strauss et al. 2004, 2009, 2016) provides a three-wave panel of 13,659 Indonesians followed over 14 years and enables us to make two main empirical observations (EOs). The first is that intra-individual ethnicity change is common in Indonesia and not limited to, say, specific socioeconomic classes, age cohorts or ethnic groups (EO1). On average, about 6.86% of all individuals in our sample change their ethnic identity from one wave to the next and churning occurs between all ethnicities: Javanese become Balinese, Balinese become Javanese, Sundanese become Batak, Batak become

Javanese et cetera. The second is that ethnic identity change exhibits structural patterns where intra-individual ethnicity change is more likely to occur coincidental with major life events, specifically interprovincial migration and interethnic marriage (EO2). Both EO1 and EO2 are robust to controlling for individual fixed effects, among others, and consistent with qualitative studies of ethnic group membership in Indonesia (Eilenberg and Reed, 2009; Hoon 2006; Sillander and Alexander 2016).

In the second part of the empirical analysis we move from undirected ethnicity change to analyzing directed ethnicity change. We thereby focus on interethnic marriages, which often involve one spouse forgoing his/her ethnicity and adopting the ethnicity of the other spouse (cf. EO2). The dependent variable that we consider is a dummy variable for directed ethnicity change, specifically whether individuals who have intermarried have also converted to the ethnicity of their spouse. We call this phenomenon *spousal ethnic adaptation* and in our estimations we control for ethnic group and religious group fixed effects and a variety of community-level and time-varying individual-level variables. However, to identify how differences in expected costs and benefits causally affect newly intermarried individuals' decision to convert to the ethnicity of their spouse we exploit variation in deep-rooted, community-specific cultural norms that affect how attractive it is for wives to adopt the ethnicity of their husbands and vice versa.

Historically, some communities in Indonesia are matrilineal, which means that local cultural norms dictate that, after their marriage, couples co-reside with the brides' (extended) family (Guilmoto 2015; Holden et al. 2003; Rammohan and Robertson 2012). We therefore expect that comparable individuals make different ethnicity choices because local post-marital residence norms affect the relative attractiveness of newly intermarried wives retaining or newly intermarried husbands adopting a specific ethnicity, namely the ethnicity of the wife's family. Empirical analysis renders strong support for this expected effect of

plausibly exogenous variation in the costs and benefits that newly intermarried men and women can expect from retaining or changing their ethnicity. Specifically, we find that newly intermarried men living in matrilineal communities are about 13 percentage points more likely to adopt the ethnicity of their wives than men living in non-matrilineal communities are (45.1% vs. 31.4%). Similarly, we find that newly intermarried women living in matrilineal communities are about 13 percentage points less likely to adopt the ethnicity of their husbands than women living in non-matrilineal communities are (41.2% vs. 55.1%).² We conclude that intra-individual ethnicity change is not only systematic but also directional and causally affected by the expected costs and benefits of retaining one's ethnicity vs. adopting a specific alternative ethnicity.

This paper contributes to two bodies of research. The first is the emerging literature on ethnic identity choice in economics. Following Akerlof and Kranton's (2000) conceptual work on the costs and benefits of social identity and group membership, economists have recently started to study possible ethnicity choice empirically, notably in relation to government affirmative action. However, extant papers focus on dichotomous ethnicity choices of a limited number of ethnic groups, notably bi-ethnic individuals choosing between one of their ancestral ethnicities, and consider aggregate-level indicators of ethnic group membership based on repeated cross-sections. Antman and Duncan (2015), for example, consider the effect of a U.S. state-level ban on affirmative action (e.g., in government hiring), and find that the introduction of this ban reduced the proportion of individuals with mixed-black and non-black ancestry that self-identified as Black as opposed to White by almost one-third. Cassan (2015) focuses on the relative size of Indian castes and considers the opposite effect, finding that agricultural castes in India set to benefit from a 1901 land reform grew much faster in size compared to other castes that were not set to benefit from this reform.

² To be sure, in Indonesia, all marriages are between a woman and a man.

Finally, Jia and Persson (2019) consider the effect of China's one-child policy, which gives privileges to ethnic minorities (relative to the dominant ethnic Han), on the ethnicity chosen for children from Han/non-Han (majority/minority) couples. Their results indicate that the implementation of this policy led to a one-fifth increase in the percentage of mixed couples choosing a non-Han, minority ethnicity for their child. We thus advance the empirical literature on ethnic identity choice in three ways. The first is that our use of individual-level longitudinal data means that we consider actual intra-individual ethnicity change. The second is that in our analysis ethnicity change is not a dichotomous phenomenon but involves a much richer set of identity choices whereby individuals choose between a variety of alternative ethnicities. Finally, we expand the analysis of possible intra-individual ethnicity change to consider all individuals regardless of their ancestral ethnic background and not just individuals with mixed ethnic traits or multi-ethnic backgrounds.

The second contribution is to the large literature on ethnic fractionalization and, particularly, ethnicity-based discrimination and segregation (Alesina et al. 2003, 2016; Arrow 1973; Becker 1995; Bertrand et al. 2005; Ewens et al. 2014). Faced with persistent and sizeable ethnic segregation, many governments worldwide use ethnic group membership as a basis for directing policy making, for example, implementing affirmative action policies (Card and Krueger 2005; Cassan 2015; Guan 2005; Lim 1985). Moreover, with few exceptions, official statistics on various forms of ethnic segregation and integration are constructed from population censuses that measure ethnicity on the basis of individuals' self-reports (Government of India 2011; Roth 2010; Statistics Indonesia 2010; U.S. Bureau of the Census. 2013). The evidence on systematic intra-individual ethnicity change presented in this paper, however, raises questions about what we know (or what we think we know) about ethnic fractionalization and the size of various ethnic groups in society, the representation of these groups in different social positions, and the degree to which these groups are

intermingling or not. For instance, our results imply that marital homogamy in Indonesia is rather overstated. Taking cross-sectional data at face value, only 14.5% of all marriages in Indonesia are interethnic marriages. Correcting for intermarried individuals' tendency to convert to their spouses' ethnicity, in contrast, we find that this number understates intermarriage by almost one-third and that the actual prevalence of interethnic marriage is closer to 19.2%. More generally, the evidence on the permeability of ethnic group boundaries and the pervasiveness of intra-individual ethnicity change presented in this paper means that governments need to be careful when using self-reported ethnicity as a basis for public policy and serves as a warning against the reification of ethnic identity.

The next section of this paper provides relevant background information, including details on some of the data and measures that we use, and makes two empirical observations on intra-individual ethnicity change. The section after that elaborates our empirical strategy for identifying causal drivers of individuals' ethnic identity choice and presents the paper's main empirical results. We conclude with a discussion of our findings and the importance of these findings, particularly for our understanding of ethnic segregation.

II. RESEARCH CONTEXT, DATA, AND BACKGROUND

II.A. Ethnic Identities in Indonesia

Indonesia is known for its ethnic patchwork of approximately 300 officially recognized ethnic groups. Individuals' ethnicity is thereby primarily based on fluency in the ethnic language and not on physical appearance (Musgrave 2014). Although most ethnic languages in Indonesia have some degree of resemblance with other Austronesian languages, the former differ significantly in terms of script, vocabulary and pronunciation, mostly because of historical variation in the degree of contact with outside groups (e.g., Arabs, Indians, Dutch, Japanese, Chinese and Malaysians). Some of the most prominent ethnicities in Indonesia are

Javanese, Sundanese, Batak and Madurese, accounting for 40.2%, 15.5%, 3.6% and 3.2% of the total population (Statistics Indonesia 2010).

The importance of language as the defining feature of ethnic group membership means that, in Indonesian society, individuals' choice for adopting a certain ethnicity is not heavily constrained by physical appearance, for instance, by being light-skinned vs. dark-skinned. Moreover, the Indonesian national government is rather accommodative of intra-individual ethnicity change. Individuals' ethnicity is not determined at birth / by their parents but by the individual him-/herself: "The ethnicity of a household member is whatever he thinks it is" (Statistics Indonesia 2010). Members of ethnic groups tend to be geographically concentrated in Indonesia's 34 provinces, however. The Sundanese, for instance, mostly live in West Java and Central Java, whereas the Balinese mostly live in Bali. With some exceptions, ethnic groups in Indonesia do not have a rank-ordered social position and are considered of equal status. Because of regional ethnic concentration, Javanese are the dominant ethnic group on Java but not on Bali, where the Balinese are the dominant ethnic group. Vice versa, the Balinese are the dominant ethnic group on Bali but not on Java. Hence, an ethnic group that is underprivileged and suffers socioeconomic exclusion in one province or geographic area in Indonesia may be the dominant ethnicity in another area. Overall, the lack of a universal rank-order for Indonesian ethnicities means that a change in ethnic self-identification in Indonesia is not necessarily a matter of crossing group boundaries to seek acceptance by a high-status group.

Finally, it is not uncommon for Indonesians to leave their region of birth and migrate to a different province or district within Indonesia. Sukamdi and Mujahid (2015), for example, report that between 2005 and 2010, 2.4% of the population of five years and older migrated between one of the 34 provinces of Indonesia. Lottum and Marks (2011) similarly find that more than 10% of Indonesians have migrated between provinces in Indonesia at one

point in their lives. Wajdi et al. (2017) estimate a gravity model of interregional migration in Indonesia finding that people mostly migrate towards developed regions, whereas geographic distance has a negative effect on migration flows between regions in Indonesia.

II.B. Data and Measures

Data for our analysis come from the Indonesian Family Life Survey or IFLS. Availability of specific variables means that we use the last three waves of this panel with data collected in 2000 (Wave 3), 2007 (Wave 4) and 2014 (Wave 5) (Strauss et al. 2004, 2009, 2016). The IFLS contains an extensive set of variables concerning ethnicity, not only of respondents themselves but also of their parents and spouse, if applicable. In addition, much information is available on respondents' communities. The IFLS data that we use comprise up to 13,659 unique individuals and 27,008 observations overall, depending on whether we consider directed or undirected intra-individual ethnicity change. Throughout this paper, we only consider individuals included in all three waves.

II.B.1. Measures of Ethnic Identity and Undirected Intra-Individual Ethnicity Change

The key variable in our analysis concerns individuals' ethnicity or ethnic identity. The main measure of ethnicity that we consider derives from the IFLS questionnaire item asking respondents to which ethnic group they belong ("What is your ethnicity?"). The accompanying answer scale distinguishes 29 ethnicities (see Table A1 in Appendix A for details). The survey allows individuals to select more than one ethnicity but very few respondents (<0.5%) do. We have removed these respondents from the sample for the main analysis but we obtain similar results when we do include these individuals. Across waves, the categories of possible answers to the ethnicity item have changed to cover different ethnic groups. Hence, we also exclude individuals reporting ethnicities that were not available as

possible answers in all three waves of the IFLS that we consider. Excluded answer categories include “Other,” “Cirebon” and “Banten.” Using the ethnicity item, we measure (undirected) intra-individual ethnicity change simply as a dummy variable that equals 1 if the individual’s self-reported ethnicity at time t is not the same as his/her ethnicity at time $t-1$ and equals 0 if the individual’s self-reported ethnicity at time t is the same as his/her ethnicity at time $t-1$.

Further information on individuals’ ethnicity comes from the item asking respondents to indicate the ethnicity of their parents: “What is your father’s/mother’s ethnicity?” As before, the accompanying answer scale distinguishes 29 ethnicities (although this scale has also changed slightly across waves). Respondents can indicate different ethnicities for their mother and their father. We measure (undirected) intra-individual ethnicity change as a dummy variable that equals 1 if individuals’ self-reported ethnicity at time t is not the same as the reported ethnicity of both their father and their mother at time t and equals 0 otherwise.

II.B.2. Other Measures and Control Variables

For our purpose, two other variables included in the IFLS are particularly relevant. These variables concern the community in which individuals are living (a categorical variable) and the ethnicity of individuals’ spouses, if applicable (also a categorical variable). The latter variable derives from the item asking the respondents’ spouse “What is your ethnicity?,” although this question can also be answered by the respondent in case the spouse is absent during the interview. When estimating regression models using the IFLS data, we typically control for a large set of possible confounders. First, we include various time-varying individual-level control variables such as age, age squared, educational degree, employment status and an individual’s ability to speak the official national language “Bahasa Indonesia” (1=yes). Second, we control for community fixed effects and, if necessary, community-wave fixed effects and individual fixed effects. To make sure that our results are general and not

driven by systematic (dyadic) changes in ethnic identity (e.g., Balinese becoming Javanese), we further control for both ethnic group fixed effects and religious group fixed effects. Finally, because measurement error may be a factor (Alba et al. 2016; Davenport 2020; Kramer et al. 2016), we control for interviewers' perceived accuracy of respondents' answers and for interviewers' own ethnicity and their prior experience interviewing for the IFLS. Table A1 in Appendix A presents details on these variables and their measurement.

II.C. Theoretical Background

Understanding intra-individual ethnicity change as the outcome of purposeful decision making on the part of the individual requires a framework that considers possible motives and corresponding benefits and costs of maintaining one's current ethnic identity vs. adopting another ethnicity. Akerlof and Kranton (2000) consider individuals' social identity, which includes their ethnic or racial identity, as an argument in their utility function. We extend Akerlof and Kranton's (2000) framework to incorporate the long-standing socio-psychological literature on ethnic boundaries and ethnic identity formation (Barth 1969; Waters 1990), particularly Berry's (1997) work on acculturation (see, also, e.g., Bourhis et al. 1997). This refers to individuals undergoing social and psychological change to adapt or adjust to a particular socio-cultural environment. We generalize this idea of acculturation, which is most significant among international migrants, asylum seekers, and sojourners (Ward et al. 2005), to propose that individuals may use their social identity, particularly their ethnicity, as an instrument and change this identity as a means to fit in with a particular social environment. A concrete example involves forgoing one's birth language to adopt the ethnic language spoken by most others with whom one meets or interacts regularly. Language is a key constituent of individuals' ethnic identity (Brubaker 2006; Kertzer and Arel 2002; Weber 1921; Wimmer 2013). However, language also has important economic ramifications for

individuals, as evidenced by many studies showing that adopting the language of the host country increases migrants' labor market earnings (see Chiswick and Miller 2015 for a survey). In terms of Akerlof and Kranton's (2000) framework, we expect that individuals not only maximize their utility conditional on their ethnic identity but also choose the ethnic identity that maximizes their utility (cf. Shayo 2009).

Individuals' ethnic identity mostly derives from their cultural upbringing in a particular ethnic tradition, which is the result of their parents' socialization efforts and ethnic ancestry (Erikson 1968; Pahl and Way 2006). Adopting another ethnicity therefore comes at the cost of renouncing one's (genetic) ancestry and breaking with one's childhood rearing. In addition, switching ethnic identities means leaving the comfort of one's current ethnic group in favor of an ethnic group with which one tends to be less familiar. Being a member of a social group has generic psychological benefits (Abrams and Hogg 1988; Balliet et al. 2014; Bergami and Bagozzi 2000; Reynolds et al 2001), but also involves being part of a social network and being at the receiving end of ingroup favoritism. At the same time, as emphasized by the acculturation literature, individuals may change their ethnic identity precisely because being or becoming a member of a specific social group fosters social acceptance in a particular social environment (Hogg and Terry 2000; Hornsey 2008; Phinney 2003; Sam and Berry 2010). In fact, when their social environment changes, individuals are likely to experience pressure from this new environment to assimilate (Bauer et al. 2000; Berry 1991; Rumbaut 1994), for instance, by changing their ethnic identity. Ethnicities are delineated by boundaries that are defined by, among others, shared language, religion and culture but also by phenotypical traits such as skin tone or hair color (Brubaker 2006; Wimmer 2013). Some of these ethnic group boundaries can be crossed or permeated more easily than others. In general, however, switching ethnicities takes time and effort and is a stressful experience (Berry 2003). Moreover, even when pressured by a particular ethnic

group, the degree to which the individual ends up being accepted by this group is uncertain. Still, when it comes to fitting into a particular socio-cultural environment, individuals' best course of action will often be to adapt their ethnic identity.

Given these expected costs and benefits of ethnicity change, we predict that, over time, a small percentage of all individuals may forgo their original ethnicity and adopt a new ethnic identity. Of course, the prevalence of intra-individual ethnicity change is strongly dependent on local circumstances, including the presence of other ethnic groups. Moreover, we expect that intra-individual ethnicity change is more common among individuals who have experienced a so-called major life event (e.g., interethnic marriage or internal migration) and a corresponding shock to their social environment. The underlying idea is that major life events are likely to prompt individuals to seek acceptance by a new social group, which is to say that the expected benefits of ethnicity change will be higher in case a major life event has occurred. An interethnic marriage, for instance, exposes the individual to a family-in-law that may be reluctant to accept him/her, not least because he/she is unacquainted with the family's ethnic customs and traditions. Two specific expectations are that intra-individual ethnicity change is more likely to occur coincidental with (i) internal migration and with (ii) an interethnic marriage.³

II.D. Two Empirical Observations on Intra-Individual Ethnicity Change

The IFLS data enable us to make two chief empirical observations (EO1 and EO2) that fit the above argument about the likelihood of intra-individual ethnicity change in general and coincidental with major life events in particular.

³ Other major life events exist, including the birth of a child or the death of a spouse. However, we expect that, on average, intermarriage and internal migration bring about much larger changes in individuals' social environment.

II.D.1. Intra-Individual Ethnicity Change is Common (EO1)

The first empirical observation (EO1) is that intra-individual ethnic identity is not uncommon and occurs universally among different ethnic and demographic groups in Indonesia. Based on raw descriptive statistics for the IFLS sample, 6.92% of Indonesians have changed their self-reported ethnicity sometime between 2000 and 2014 (Table A4 in Appendix A). However, in the IFLS sample, certain (ethnic) subgroups in Indonesia are overrepresented and others underrepresented. Hence, we have also estimated the marginal probability of intra-individual ethnicity change while controlling for various individual-level, ethnicity-level and community-level factors and year/wave fixed effects. After these corrections, the marginal probability of intra-individual ethnicity change is slightly lower, 6.86% (95%CI: 6.53–7.19%). The prevalence of intra-individual ethnicity change is further lower among individuals with only primary education or less but higher among individuals with at least some tertiary education. Ethnic identity change is also less common among older people and more common among younger people. Figure 1 presents the marginal probability of undirected intra-individual ethnicity change for these various groups. Table B1 in Appendix B presents details on the underlying empirical results.

<Insert Figure 1 about here>

Table 1 reports descriptive results on the prevalence of specific dyadic ethnicity switches in the IFLS sample. Ethnicity change appears particularly common among the Melaya and the Komerling. Only about 27-30% of all members of these two ethnic groups have kept their ethnicity during the period 2000-2014. Similarly, among Javanese that have changed their ethnicity, the most popular switch has been to Sundanese. Although only 3% of

Javanese have changed their ethnicity, 1.2% of all Javanese (or about two-fifths of all Javanese ethnicity switchers) have become Sundanese. Two particularly popular switches involve Melaya becoming Betawi (almost 39% of all Melaya) and Komerling becoming South-Sumatrese (44% of all Komerling). Even though some specific ethnic switches are more prevalent in the IFLS sample than others are, it nevertheless seems that ethnic identity switching occurs in all ethnic groups.

<Insert Table 1 about here>

II.D.2. Intra-Individual Ethnicity Change Occurs Systematically (EO2)

The second empirical observation (EO2) is that intra-individual ethnicity change is more likely to occur coincidental with major life events, specifically interprovincial migration and interethnic marriage (also known as intermarriage or exogamy) (Figure 2). We define an intermarriage as a marriage where an individual's ethnicity at time $t-1$ is different from the ethnicity of the individual's spouse at time t . About 9.81% of the individuals in our sample have engaged in an interethnic marriage between 2000 and 2014. We define interprovincial migrants as individuals who live in one of Indonesia's 34 provinces at time t but lived in another province at time $t-1$. Since different ethnic groups in Indonesia are geographically concentrated, interprovincial migration tends to be associated with becoming exposed to a different locally dominant ethnic group. However, to be complete, we also consider intraprovincial, intercommunity migrants, which we define as individuals who live in one community at time t but in another community in the same province at time $t-1$. Some 21.6% of the individuals in our sample have migrated either from one province to another or from one community to another at some point between 2000 and 2014. As expected, a comparatively small share of this group, though still 4.27% of the individuals in our sample,

have migrated from one Indonesian province to another in the same period.

<Insert Figure 2 about here>

In line with the literature on acculturation and major life events, results indicate that the marginal probability that a newlywed individual that married a spouse with a different ethnicity also changed his/her ethnic identity is about 27.3% (95%CI: 25.2–29.5%) or almost four times higher than the marginal probability of intra-individual ethnicity change for all Indonesians. Moreover, the marginal probability that a newlywed individual that married a spouse with the same ethnicity also changed his/her ethnic identity is only 3.07% (95%CI: 2.70–3.43%) and thus significantly lower than the marginal probability of intra-individual ethnicity change for all Indonesians. Similarly, the marginal probability that an individual that has moved to another province within Indonesia also changed his/her ethnicity is 15.9% (95%CI: 13.7–18.0%) or more than twice as high as the marginal probability of intra-individual ethnicity change for all Indonesians. The marginal probability that an individual that has moved to another community also changed his/her ethnicity is lower, about 9.50% (95%CI: 8.78–10.2%), but also significantly higher than the marginal probability of intra-individual ethnicity change for all Indonesians.

Importantly, we have estimated all the marginal probabilities presented in Figure 2 with both ethnic group fixed effects and religious group fixed effects controlled for (see Table B2 in Appendix B for details). Hence, the apparent relationship between interethnic marriage and (interprovincial) migration on the one hand and intra-individual ethnicity change on the other is not driven by a specific form of ethnic churning, particularly by a generic tendency of all non-Javanese to become Javanese. Results are further robust to controlling for individual fixed effects (Table B3). In principle, it is possible that the

occurrence of a major life event and ethnic identity change are both affected by the same unobserved individual trait. Individuals who are more socially apt, for instance, are probably both more likely to find an other-ethnic spouse and to change their ethnicity. However, since controlling for individual fixed effects does not render substantially different results, it seems that there is a genuine coincidence between major life events and undirected intra-individual ethnicity change. Finally, we obtain qualitatively similar descriptive statistics and results on the prevalence of intra-individual ethnicity change when we measure intra-individual ethnicity change by considering whether individuals self-report an ethnic identity that is different from both the ethnicity of their father and of their mother (Table B4).

Two extensions to these basic empirical observations are as follows. First, additional analyses indicate that changes in self-reported ethnic identity are associated with relevant behavioral adjustments (Table B5). In particular, more than 85% of the Indonesians who have changed their ethnicity between two waves have also learned the language of their new ethnic group in this period. In addition, more than 14% of ethnic identity switchers started learning the language of their new ethnic group in the seven years since they changed their self-reported ethnicity. As before, these results are robust to controlling for individual fixed effects. However, with individual fixed effects controlled for, the probability that an Indonesian that has changed his/her ethnicity has also learned the language of his/her new ethnic group drops to about 38%. Second, we find comparable evidence on intra-individual identity change for other countries, specifically on intra-individual racial identity change in the U.S. (Appendix D) and intra-individual changes in caste identity in India (Appendix E). In the U.S., for instance, the marginal probability of intra-individual racial identity change for newlyweds that have married a spouse with a different racial identity is about 20.2% (95%CI: 15.6–24.7%), which is much higher than the marginal probability for newlyweds that have married a spouse with the same racial identity (-0.19%; 95%CI: -0.59–0.21%) (see Models

D1-D3 in Table D1 in Appendix D). Races in the U.S., but also castes in India, are associated with systematic differences in social class or status. According to the empirical evidence, however, such a hierarchy of social identities does not prevent churning among these identities. We conclude that the idea of intra-individual ethnicity change is not unique to Indonesia and that systematic changes in individuals' chief social identity also occur when group boundaries appear less permeable.

III. EMPIRICAL STRATEGY AND RESULTS

III.A. Contextual Variation in Expected Costs and Benefits of Directed Ethnicity Change

III.A.1. Intermarriage, Matrilocality, and Spousal Ethnic Adaptation

The evidence that intra-individual ethnicity change coincides with major changes in individuals' social environment, as proxied by interprovincial migration and marrying an individual from a different ethnic group, is highly supportive of the idea that ethnicity is, in part, an outcome of deliberate choices that individuals make. At the same time, the evidence underlying EO2 does not rule out that there might be some form of systematic measurement error that drives our results on individuals' generic choice to change their ethnicity (cf. Alba et al. 2016; Davenport 2020; Kramer et al. 2016). In particular, it is possible that experiencing a major life event makes individuals' ethnic identities more salient, which, in turn, affects their subjective feeling of belonging to particular ethnic groups or not. Hence, for the purpose of making definitive claims that ethnicity is, in part, a choice based on cost/benefit considerations, in this section we assess how external, contextual factors systematically affect directed ethnicity change (i.e., an individual's choice to convert to a specific ethnicity or not), because they influence the individual's expected costs and benefits of ethnicity change.

We focus this part of the analysis on individuals who have married a spouse from a different ethnic group and their choice whether to convert to the ethnicity of their spouse, what we call *spousal ethnic adaptation*. Interethnic marriages often involve one spouse forgoing his/her ethnicity and adopting the ethnicity of the other spouse, which is a form of directed ethnicity change. The rationale is that adopting the ethnicity of one's spouse provides a chief means for getting accepted by the spouse's circle of friends and social network, not least the spouse's family. Of course, intra-individual ethnicity change following intermarriage means that one of the spouses retains his/her ethnicity, while the other spouse forgoes his/her original ethnicity in favor of the former's ethnicity. The contextual factor that we consider concerns local, community-specific norms on matrilocality (i.e., married couples' co-residence with the bride's family). Variation in these deep-rooted norms, specifically whether a community is matrilocal or non-matrilocal, is a plausibly exogenous source of variation in the net benefits that individuals can expect to receive from converting to a specific ethnicity or retaining their original ethnicity. Specifically, matrilocality norms affect how much a man or woman can expect to benefit or lose from giving up his/her ethnicity or how much one's husband or one's wife would benefit or lose from forgoing his/her ethnicity and convert to one's own ethnicity.⁴

There are three main ways in which matrilocality residence norms (or local post-marital residence norms more generally) affect the tradeoff that otherwise identical intermarried men and women face between retaining their original ethnicity vs. adopting a specific alternative ethnic identity, i.e., the ethnicity of their spouse. First, for newly intermarried men,

⁴ The literature distinguishes two other norms on post-marital residence, patrilocality and neolocality (Gruijters and Ermisch 2019; Holden et al. 2003). Patrilocality norms prescribe that, after their marriage, couples will live at the (parents of the) grooms' place. Neolocality norms prescribe that newlyweds reside separately from both the wife's and the husband's family.

matrilocal means co-residing with in-laws from another ethnic group. Living together with certain people intensifies an individual's contact and interactions with these people. Moreover, spending most time with in-laws that have a different ethno-cultural tradition, speak a different language and/or adhere to a different religion increases the salience of the husband's otherness, causing him to feel alienated and socially excluded (Tajfel and Turner 2004). As a result, a husband is likely to want to fit in with and be accepted by his new family. He can do this by acculturating into his wife's ethnic group. Second, matrilineal post-marital residence norms increase the bargaining power of women in interethnic households (Bowie 2008; Fortunato 2012; Hyde 1999). Interethnic couples often emphasize one ethnic background within the household, not least when deciding on the ethnic tradition in which to raise their children. At the same time, both the wife and the husband typically have a strong preference for retaining their own, ancestral ethnic identity (and for socializing their children in the same tradition) (Berry 1997; Jia and Persson 2019). The bargaining power that matrilineal residence norms afford women are therefore likely to increase the probability of a husband conforming to his wife's ethnic preferences and identity rather than the other way around. Third, for an intermarried woman who lives with her parents there is likely a great deal of pressure not to abandon her family's ethnic identity. Living with her family, a newly intermarried woman is less willing to risk alienating herself from this family by switching to another ethnic group. Finally, there is an indirect effect of matrilocality on spousal ethnic adaptation involving the approximate co-occurrence of matrilineal residence norms and so-called matrilineality, which refers to the cultural norm of giving women priority in succession whereas men may be excluded from inheriting wealth or a social position belonging to the family (Fortunato 2012; Gneezy et al. 2009). In many matrilineal communities there is therefore an added reason why a wife does not want to alienate herself from her family, namely the risk of losing (part of) her inheritance. In fact, even a husband may encourage his

wife not to abandon her family's ethnicity in order to maximize the chance of their joint household receiving a substantial inheritance from his wife's family. Overall, we expect that deep-rooted community-specific norms on matrilocal co-residence are a powerful influence guiding newly intermarried women's and men's choice between forgoing their original or ancestral ethnicity and adopting their spouses' ethnicity or not. Hence, we propose the following set of twin hypotheses on spousal ethnic adaptation and intra-individual ethnicity change coincidental with marrying a spouse with a different ethnicity:

Hypothesis 1a: *Newly intermarried men living in matrilocal communities are more likely to adopt their wives' ethnicity than newly intermarried men living in communities with other post-marital residence norms are.*

Hypothesis 1b: *Newly intermarried women living in matrilocal communities are less likely to adopt their husbands' ethnicity than newly intermarried women living in communities with other post-marital residence norms are.*

We test these hypotheses later on in this section. First, however, we provide details on our measurement of spousal ethnic adaptation and community-level cultural norms on matrilocality.

III.A.2. Measures of Spousal Ethnic Adaptation as Directed Ethnicity Change

Focusing on newly intermarried individuals, we define spousal ethnic adaptation as the choice to forgo one's original ethnicity and convert to the ethnicity of one's spouse. The IFLS item on the (self-reported) ethnicity of an individual's spouse is akin to the item asking individuals about their own ethnicity. The item simply asks a respondent's spouse to indicate

the ethnic group to which he/she belongs, should a spouse be present (see Table A1 in Appendix A for details on this item). Alternatively, if the spouse is not home at the time of the interview, the respondent him-/herself indicates the ethnic group to which the spouse belongs. We combine this measure of spousal ethnicity with various measures of an individual's own ethnicity (see Section II) to provide different operationalizations of spousal ethnic adaptation.

First, we use the item on an individual's own, self-reported ethnicity to construct a dummy variable that has value 1 if the individual's self-reported ethnicity at time t is unequal to the individual's self-reported ethnicity at time $t-1$ but is equal to the ethnicity of the individual's spouse at time t . This variable thus has value 0 if the individual's self-reported ethnicity at time t is equal to the individual's self-reported ethnicity at time $t-1$ or if the individual's self-reported ethnicity at time t is not equal to the ethnicity of the individual's spouse at time t . In similar fashion, we consider the item asking individuals about their parents' ethnicity to create a second dummy variable on spousal ethnic adaptation. This dummy variable has value 1 if an individual's self-reported ethnicity at time t is unequal to the individual's ancestral ethnicity at time $t-1$ but is equal to the ethnicity of the individual's spouse at time t . This variable thus has value 0 if the individual's self-reported ethnicity at time t is equal to his/her ancestral ethnicity at time $t-1$ or if the individual's self-reported ethnicity at time t is not equal to the ethnicity of the individual's spouse at time t .

Spousal ethnic adaptation is quite common in our sample of individuals who have intermarried ($N = 722$). Depending on the specific measure of spousal ethnic adaptation considered, 42.2% (95%CI: 42.1–42.4%) of all individuals who have married a spouse with a different ethnicity convert to the ethnicity of their spouse, meaning that they forgo their original ethnicity. Table A1 in Appendix A presents details on the two measures of spousal ethnic adaptation that we consider as well as on the IFLS items that we use to construct these

measures.

III.A.3. Measure of Post-Marital Residence Norms

To measure the presence of a matrilineal post-marital residence norm in a particular community we draw on the “Adat” (traditional lifestyle) module included in Wave 4 of the IFLS, which was conducted in 2007. This module asks the mayor of a particular community (or a similar type of high-ranked government official) about specific cultural features of the local community. One of the questions concerns the residence of newlyweds: “Putting aside economic constraints, where does the newly married couple live after their marriage?” Possible answers to this item are that couples reside at the (parents of the) bride’s place, at the (parents of the) groom’s place or neither, which includes the norm that newlyweds reside at their own place. We create a dummy variable of matrilineality that has a value of 1 if the norm in the community is that married couples reside at the (parents of the) brides’ place and a value of 0 otherwise. The sample of newly intermarried men and women that we can consider comprises 162 different communities of which 90 are matrilineal. Moreover, 57.4% of the individuals in the sample of newly intermarried individuals live in a community classified as matrilineal. Given the geographic areas covered by the IFLS sample (e.g., Sumatera, Java), these numbers match the prevalence of matrilineal residence norms in Indonesia reported in the literature (Guilmoto 2015; Rammohan and Robertson 2012). The Adat module has only been included in Wave 4 of the IFLS. However, because cultural differences are highly stable over time (Beugelsdijk et al. 2015; Hofstede 2001), we can use the same measure of local marital residence norms for observations from Wave 5 of the IFLS. Table A5 in Appendix A presents a list of how the communities in our sample are spread across provinces. Figure A1 illustrates the prevalence of matrilineal communities for the Indonesian provinces included in the analysis.

III.A.4. Empirical Model and Control Variables

Following Hypotheses 1a and 1b, for the empirical analysis, we estimate differences in the likelihood of spousal ethnic adaptation of newly intermarried men and newly intermarried women living in matrilocal vs. non-matrilocal communities. Our sample comprises individuals nested in communities that are measured in different waves/years. Hence, we have cross-sections of individuals that are repeated at the community level and the basic empirical model that we estimate is:

$$E_{ict} = \beta_0 + \beta_1 F_i + \beta_2 G_i M_c + \beta_3 Z_{ict} + \beta_4 S_{it} + \beta_5 I_{it} + R_i + D_c + \delta_t + \gamma_{ct} + u_{ict}. \quad (1)$$

In this model, E_{ict} denotes spousal ethnic adaptation of newly intermarried individual i living in community c at time t , G_i denotes gender (male or female), M_c denotes whether the community is matrilocal, $G_i M_c$ denotes the interaction between gender and matrilocal residence norms in one's community, Z_{ict} denotes possible individual-level and/or (time-varying) community-level control variables. The specific control variables that we include are mostly the same control variables that we considered in Section II. Hence, we control for a variety of individual-level factors such as educational degree or ability to speak the official national language and also for interviewer-level variables that speak to possible sources of measurement error. In addition, we control for ethnic minority status, as there is a possibility that individuals use intermarriage as a means to escape their minority status and become a member of the ethnic majority in a particular area. Because ethnic spousal adaptation involves both the individual and his/her spouse as decision makers we further consider the education level of the spouse as a proxy for the spouse's bargaining power (Spierings et al. 2010).

As before, including ethnic group and religious group fixed effects (R_i) helps ensure that our results do not reflect specific patterns in intra-individual ethnicity change, for instance, that all non-Javanese intermarried individuals adopt the Javanese ethnicity of their spouse. Because community-level cultural norms do not vary over time, for most analyses we prefer estimating a model that includes community fixed effects (D_c), year fixed effects (δ_t) and community-year interactions (γ_{ct}). We include year fixed effects and/or year fixed effects interacted with community fixed effects to rule out that our results are driven by time-invariant or time-varying community-level factors. A specific example is that we take into account any biases that might occur when the incidence of ethnic conflicts or tensions in a region (e.g., in North-Sumatera, Bali or Sulawesi), which likely affects the permeability of ethnic boundaries (Bertrand 2003), is correlated with the presence of non-matrilocal residence norms in the region (see Table A5 and Figure A1 in Appendix A). Of course, including community fixed effects means that we do not consider the direct effect of matrilocality (M_c) on spousal ethnic adaptation.

As a robustness check, we extend the baseline results that we obtain from estimating Equation 1 by considering several additional moderating effects. First, in some Indonesian regions, ethnicity is not a particularly salient issue (Bazzi et al. 2019), which, in turn, may affect the likelihood of spousal ethnic adaptation. We explore this issue using the ability of an individual to speak Bahasa Indonesia, the national language. We expect that ethnicity is less salient among couples that use the national language in their daily lives instead the ethnic language of one of the spouses. Hence, as an added control we include the interaction between matrilocality marital residence norms in one's community and the ability of an individual to speak Bahasa Indonesia. In addition, we correct for the possibility that matrilocality has an indirect influence on the decision to adopt the ethnicity of one's spouse through individual-level factors that correlate with matrilocality. Specifically, we take into

account that women, on average, may be (relatively) more highly educated in matrilocal communities than in non-matrilocal communities (Lowes 2020), and that, therefore, they are less likely to adopt the ethnicity of their husbands. Because both women's absolute level of education and their level of education relative to their husbands may be affected, we include two further interaction terms: matrilocality interacted with the educational degree of the individual and matrilocality interacted with a measure of the difference in education level between spouses.

Finally, to overcome the heteroscedasticity issues that result from considering a nominal dependent variable, we estimate all models in terms of log odds (logistic models). Similarly, to be conservative, we correct estimated standard errors for clustering at the level immediately above the community level. Tables A1 and A4 in Appendix A present a brief description and summary statistics for the main variables in the analyses, including the three additional interaction terms.

III.B. Results: Directed Ethnicity Change under Different Post-Marital Residence Norms

III.B.1. Baseline Results

Figure 3 presents the estimated marginal probabilities of newly intermarried men and women adopting the ethnicity of their spouse when living in matrilocal communities vs. when living non- matrilocal communities. These probabilities differ (statistically) highly significantly: 45.1% (95%CI: 41.4–48.7%) vs. 31.4% (95%CI: 23.2–39.7%) for men and 41.2% (95%CI: 38.3–44.2%) vs. 55.1% (95%: 46.0–64.3%) for women. Specifically, our results indicate that the probability of newly intermarried men adopting the ethnicity of their wives is 13.7 percentage points (95%CI: 1.75–25.5%) higher among couples living in matrilocal communities compared to couples living in non-matrilocal communities, which supports

Hypothesis 1a. Similarly, results indicate that the probability of newly intermarried women adopting the ethnicity of their husbands is 13.9 percentage points (95%CI: 1.76–26.0%) lower among couples living in matrilocal communities compared to couples living in non-matrilocal communities, which supports Hypothesis 1b.

<Insert Figure 3 about here>

Table C1 in Appendix C presents the complete results underlying the marginal probabilities presented in Figure 3. Moreover, to be complete, Table C2 repeats the baseline analyses presented in Table C1 but considers a measure of spousal ethnic adaptation based on the ethnicity of individuals' parents. Results are highly similar although estimates are less precise, particularly for men.

III.B.2. Robustness Checks with Added Moderators

Digging deeper, Table C3 in Appendix C reports results of analyses that consider several additional community-level and individual-level moderators. First, our results are robust to controlling for individuals' ability to speak the national language as a factor that moderates the influence of matrilocal post-marital residence norms on spousal ethnic adaptation (Model C7). Second, results are robust to controlling for the possibility that matrilocality has an indirect moderating effect involving a possible influence of matrilocality on individuals' own level of education (Model C8) or on the difference in education levels between spouses (Model C9). We obtain similar results when considering all three additional moderators simultaneously (Model C10). Hence, we deem it unlikely that the apparent moderating effect of matrilocality (Figure 3, Tables C1 and C2) is spurious and does not reflect the influence of variation in expected costs and benefits of ethnicity change on newly intermarried men's and

women's choice to convert to the ethnicity of their spouse or not.

We draw two conclusions. The first is that men's and women's likelihood of spousal ethnic adaptation is systematically affected by deep-rooted post-marital residence norms. The second conclusion, however, is more general, namely that ethnic identity and intra-individual ethnic identity change is a choice that is causally affected by differences in the expected costs and benefits of retaining one's original ethnicity vs. adopting a specific other ethnicity.

IV. DISCUSSION AND CONCLUSION

IV.A. Do People Choose their Ethnicities?

Individuals' ethnic identities have implications for many aspects of their lives, including how they are treated by other people, ranging from (prospective) employers to potential romantic partners and from banks to online buyers. At the same time, there is a growing consensus among social scientists that individuals' ethnicity is malleable and not a fixed trait defined by their (genetic) ancestry. In line with the emerging consensus, this paper has presented extensive empirical evidence indicating that intra-individual changes in (self-reported) ethnicity are not uncommon and associated with important individual behavioral adjustments, particularly the learning of a new language. More importantly, further empirical results strongly suggest that individuals choose between specific alternative ethnicities and that this choice is deliberate, guided by a weighing of expected costs and benefits of either retaining one's original ethnicity or forgoing this ethnicity in favor of a particular new ethnicity. These results are robust to the inclusion of a variety of control variables, including controlling for individual fixed effects.

Moreover, focusing on intra-individual ethnicity change among newly intermarried men and women, we have used variation in the form of deep-rooted post-marital residence norms to identify how differences in expected costs and benefits causally affect individuals'

decisions to convert to the ethnicity of their spouses (and of their spouses' family). A local cultural norm on matrilocality—newlyweds are expected to co-reside with the wife's family—bolsters plausibly exogenous variation in the costs and benefits that newly intermarried men and women can expect from retaining their ethnicity or adopting a new one, specifically wives forgoing their families' ethnicity or husbands adopting the ethnicity of their wives' family. Empirical results indicate that newly intermarried women living in matrilocal communities are far less likely to forgo their original or ancestral ethnicity than newly intermarried women living in non-matrilocal communities are. Vice versa, newly intermarried men living in matrilocal communities are far more likely to adopt their wives' ethnicity than newly intermarried men living in non-matrilocal communities are. Hence, there is convincing evidence that one's ethnic identity is, at least partly, a choice that individuals make after careful consideration of the costs and benefits of retaining their current ethnicity vs. adopting a specific alternative ethnicity.

Our findings show the usefulness of a broadened approach to the economics of social identity in which individuals' identity is endogenous and not merely a parameter in their utility function as in Akerlof and Kranton's (2000) seminal work. More generally, our study adds an important cost/benefit perspective to the vast literature on social identity formation, particularly the formation of ethnic and racial identities (Berry 1997; Erickson 1968; Helms 1990; Wimmer 2013). Even when individuals' ethnicity is strongly rooted in their (genetic) ancestry and family-level ethnic socialization, individuals seem both willing and able to change their ethnicity if they expect certain benefits from becoming a member of another ethnic group.

At the same time, we should note that some of our empirical findings could be partly specific to the Indonesian context and its language-based ethnic boundaries. Although we present comparable empirical observations on intra-individual identity change in India and

the U.S., the permeability of ethnic group boundaries in Indonesia is probably higher than in a number of other countries, particularly countries in which ethnic group boundaries are based on less malleable, phenotypical traits. As a result, in these other countries, the ability of individuals to switch ethnic identities is likely lower than it is in Indonesia. Notwithstanding that language is the most common ethnic group boundary (Alesina et al. 2003; Easterly and Levine 1997), we should thus be careful not to overstate the ability to switch ethnicities in different ethnic contexts. In general, changing one's ethnicity is not easy but stressful and not something that individuals are likely to do lackadaisically.

IV.B. Implications of Ethnicity as Choice

Implications of ethnicity or ethnic group membership as a choice are far-reaching and concern scholarly research and government policies and public policy making alike. Most fundamentally, the idea of ethnicity as choice raise questions about what we know (or what we think we know) about ethnic cleavages and inequalities in society as well as about ethnic fractionalization and the size of various ethnic groups in society. There is much work in the social sciences documenting a general lack of social interactions between members of different ethnic groups, exemplified by infrequent interethnic contact and few interethnic friendships as well as residential, marital and school segregation. Similarly, many studies have reported systematic differences in socioeconomic outcomes and status between members of different ethnic groups. Moreover, extant evidence on the extent of ethnic segregation and discrimination has motivated many governments worldwide to implement ethnicity-based policies, for instance, affirmative action plans or deliberate geographic spreading of ethnic group members.

There is a large literature discussing and empirically investigating potential challenges that limit the effectiveness of using ethnicity or some other social categorization as a basis for

directing policy making, particularly affirmative action (Bagde et al. 2016; Fryer and Loury 2005, 2016; Holzer and Neumark 2000, 2006). Moreover, there is some work discussing the usefulness and legitimacy of using ethnic or racial identity as an eligibility criterion for affirmative action compared to using a direct measure of eligibility based on actual socioeconomic disadvantage (Cancian 1998; Darity et al. 2011; Kahlenberg 1996; Malamud 1995). However, these literatures do not yet consider how ethnicity as choice affects both the evidence on ethnic segregation and discrimination as well as the usefulness of ethnicity-based policies. Concerning the former in particular we find that directed ethnicity change undermines received wisdom on the degree to which ethnic groups in society are divorced or intermingling. Changing one's ethnicity is a chief means for individuals to fit into a particular social environment. Neglecting the potential for intra-individual ethnicity change therefore means overlooking a significant part of real-world interactions among and inclusion of different ethnic groups in society. The use of longitudinal population or census data, in contrast, would make it possible to correct for systematic changes in individuals' ethnicity and thus address this measurement problem.

<Insert Table 2 about here>

As an example, consider marital homogamy and intermarriage. Intergroup friendship, intergroup relationships and, particularly, intermarriage (or co-habitation) are widely taken as signs of decreasing social distance between groups in society and social acceptance of members of one group by members of another group. The cross-sectional data for Indonesia suggest that 14.5% (95%CI: 13.9–15.0%) of marriages in Indonesia is interethnic (Table 2, Row 1). Correcting for systematic intra-individual ethnicity change, in contrast, renders a share that is significantly larger, 19.2% (95%CI: 18.6–19.8%). Results for the U.S. and India

are even more pronounced (Table 2, Rows 2-4). In the U.S., the prevalence of interracial marriage nearly doubles, from 5.87% to 8.60%. In India, intercaste marriages used to be officially illegal but amount to 17.7% of all marriages when taking into account individuals' tendency to change their caste identity coincidental with marrying a spouse who belongs to a different caste. It thus seems that taking social identities such as ethnicity as fixed indeed results in a sizeable overstatement of actual levels of ethnic, racial or caste-based segregation in society. Moreover, this sort of overstatement will bias assessments of the effectiveness of policies aimed at fostering desegregation and social inclusion. Policy interventions for decreasing ethnic segregation in schools or housing, for instance, are difficult to evaluate accurately when individuals deliberately change their ethnic self-perception to match the ethnic group that is most salient in their new school or new neighborhood. More generally, any analysis of ethnic or racial fractionalization in society would do well not to look at raw numbers on group sizes only but to take into account individuals' crossing of ethnic and racial boundaries. Nation building, for instance, would not only be fostered by decreased ethnic fractionalization but also by increased permeability of ethnic group boundaries. In such a case, society can still comprise a plethora of ethnic groups but polarization would nevertheless be low, as continuous churning among these groups boosts social cohesion. Overall, a chief lesson from this paper is that properly studying and assessing social segregation requires detailed analysis of the degree to which boundaries between groups in society are permeable and enable individuals to switch their chief social identities.

Appendix A: Description of Indonesian Data (IFLS) and Construct Validity of Measured Ethnic Identity

<Insert Tables A1, A2.1, A2.2, A2.3, A3, A4 and A5 and Figure A1 here>

**Appendix B: Regression Analyses Underlying Empirical Observations 1 and 2 (EO1
and EO2)**

<Insert Tables B1, B2, B3, B4 and B5 here>

Appendix C: Robustness Checks of Main Analyses

<Insert Tables C1, C2 and C3 here>

**Appendix D: Intra-Individual Racial Identity Change and Racial Intermarriage in the
U.S.**

<Insert Tables D1 and D2 here>

**Appendix E: Intra-Individual Change in Caste Identity and Intercaste Marriage in
India**

<Insert Tables E1 and E2 here>

REFERENCES

- Abrams, Dominic, and Michael A. Hogg. "Comments on the Motivational Status of Self-Esteem in Social Identity and Intergroup Discrimination." *European Journal of Social Psychology* 18, no. 4 (August 1988): 317–34. <https://doi.org/10.1002/ejsp.2420180403>.
- Akerlof, George A., and Rachel E. Kranton. "Economics and Identity." *Quarterly Journal of Economics* 115, no. 3 (August 2000): 715–53. <https://doi.org/10.1162/003355300554881>.
- Alba, Richard D., Noura E. Insolera, and Scarlett Lindeman. "Is Race Really So Fluid? Revisiting Saperstein and Penner's Empirical Claims." *American Journal of Sociology* 122, no. 1 (July 2016): 247–62. <https://doi.org/10.1086/687375>.
- Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat, and Romain Wacziarg. "Fractionalization." *Journal of Economic Growth* 8, no. 2 (2003): 155–194.
- Alesina, Alberto, Stelios Michalopoulos, and Elias Papaioannou. "Ethnic Inequality." *Journal of Political Economy* 124, no. 2 (April 2016): 428–88. <https://doi.org/10.1086/685300>.
- Antman, Francisca, and Brian Duncan. "Incentives to Identify: Racial Identity in the Age of Affirmative Action." *Review of Economics and Statistics* 97, no. 3 (July 2015): 710–13. https://doi.org/10.1162/REST_a_00527.
- Arrow, Kenneth J. "The Theory of Discrimination." In *In Discrimination in Labor Markets*, 3–33. Princeton University Press, 1973.
- Atkin, David, Eve Colson-Sihra, and Moses Shayo. "How Do We Choose Our Identity? A Revealed Preference Approach Using Food Consumption." Cambridge, MA: National Bureau of Economic Research, March 2019. <https://doi.org/10.3386/w25693>.
- Balliet, Daniel, Junhui Wu, and Carsten K. W. De Dreu. "Ingroup Favoritism in Cooperation: A Meta-Analysis." *Psychological Bulletin* 140, no. 6 (2014): 1556–81. <https://doi.org/10.1037/a0037737>.
- Barth, Fredrik, ed. *Ethnic Groups and Boundaries: The Social Organization of Culture Difference*. 2. repr. Oslo: Universitetsforlaget, 1969.
- Bauer, Thomas K., Magnus Lofstrom, and Klaus F. Zimmermann. "Immigration Policy, Assimilation of Immigrants and Natives' Sentiments Towards Immigrants: Evidence

- from 12 OECD-Countries.” IZA Discussion Paper, no. 187 (2000).
- Bazzi, Samuel, Arya Gaduh, Alexander D. Rothenberg, and Maisy Wong. “Unity in Diversity? How Intergroup Contact Can Foster Nation Building.” *American Economic Review* 109, no. 11 (November 1, 2019): 3978–4025. <https://doi.org/10.1257/aer.20180174>.
- Becker, Gary S. *The Economics of Discrimination*. 2. ed., 6. Economics Research Studies of the Economics Research Center of the University of Chicago. Chicago: Univ. of Chicago Press, 1995.
- Bergami, Massimo, and Richard P. Bagozzi. “Self-Categorization, Affective Commitment and Group Self-Esteem as Distinct Aspects of Social Identity in the Organization.” *British Journal of Social Psychology* 39, no. 4 (December 2000): 555–77. <https://doi.org/10.1348/014466600164633>.
- Berry, John W. “Conceptual Approaches to Acculturation.” In *Acculturation: Advances in Theory, Measurement, and Applied Research*, edited by Kevin M. Chun, Pamela Balls Organista, and Gerardo Marín, 17–37. Washington: American Psychological Association, 2003. <https://doi.org/10.1037/10472-004>.
- . “Immigration, Acculturation, and Adaptation.” *Applied Psychology* 46, no. 1 (January 1997): 5–34. <https://doi.org/10.1111/j.1464-0597.1997.tb01087.x>.
- Bertrand, Jacques. *Nationalism and Ethnic Conflict in Indonesia*. Cambridge: Cambridge University Press, 2003. <http://ebooks.cambridge.org/ref/id/CBO9780511559341>.
- Bertrand, Marianne, Dolly Chugh, and Sendhil Mullainathan. “Implicit Discrimination.” *American Economic Review* 95, no. 2 (April 2005): 94–98. <https://doi.org/10.1257/000282805774670365>.
- Beugelsdijk, Sjoerd, Robbert Maseland, and André van Hoorn. “Are Scores on Hofstede’s Dimensions of National Culture Stable over Time? A Cohort Analysis: Are Hofstede’s Culture Scores Stable over Time?” *Global Strategy Journal* 5, no. 3 (August 2015): 223–40. <https://doi.org/10.1002/gsj.1098>.
- Bisin, Alberto, and Thierry Verdier. “‘Beyond the Melting Pot’: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits.” *Quarterly Journal of Economics* 115, no. 3 (August 2000): 955–88. <https://doi.org/10.1162/003355300554953>.
- Chiswick, Barry R., and Paul W. Miller. “International migration and the economics of language.” *Handbook of the economics of international migration*. Vol. 1. North-Holland, 2015. 211–269.

- Bourhis, Richard Y., Lena Celine Moise, Stephane Perreault, and Sacha Senecal. "Towards an Interactive Acculturation Model: A Social Psychological Approach." *International Journal of Psychology* 32, no. 6 (December 1997): 369–86. <https://doi.org/10.1080/002075997400629>.
- Bowie, Katherine. "Standing in the Shadows: Of Matrilocality and the Role of Women in a Village Election in Northern Thailand." *American Ethnologist* 35, no. 1 (February 2008): 136–53. <https://doi.org/10.1111/j.1548-1425.2008.00010.x>.
- Brubaker, Rogers. *Ethnicity without Groups*. Harvard Univ. Press paperback ed. Cambridge, Mass.: Harvard Univ. Press, 2006.
- Card, David, and Alan Krueger. "Would the Elimination of Affirmative Action Affect Highly Qualified Minority Applicants? Evidence from California and Texas." Cambridge, MA: National Bureau of Economic Research, March 2004. <https://doi.org/10.3386/w10366>.
- Cassan, Guilhem. "Identity-Based Policies and Identity Manipulation: Evidence from Colonial Punjab." *American Economic Journal: Economic Policy* 7, no. 4 (November 2015): 103–31. <https://doi.org/10.1257/pol.20130290>.
- Chandra, Kanchan. "What Is Ethnicity and Does It Matter?" *Annual Review of Political Science* 9, no. 1 (June 2006): 397–424. <https://doi.org/10.1146/annurev.polisci.9.062404.170715>.
- Chiswick, Barry R. "Differences in Education and Earnings Across Racial and Ethnic Groups: Tastes, Discrimination, and Investments in Child Quality." *The Quarterly Journal of Economics* 103, no. 3 (August 1988): 571. <https://doi.org/10.2307/1885546>.
- Cohen, Sheldon, Michael L.M. Murphy, and Aric A. Prather. "Ten Surprising Facts About Stressful Life Events and Disease Risk." *Annual Review of Psychology* 70, no. 1 (January 4, 2019): 577–97. <https://doi.org/10.1146/annurev-psych-010418-102857>.
- Dahis, Ricardo, Emily Nix, and Nancy Qian. "Choosing Racial Identity in the United States, 1880-1940." Cambridge, MA: National Bureau of Economic Research, November 2019. <https://doi.org/10.3386/w26465>.
- Davenport, Lauren. "The Fluidity of Racial Classifications." *Annual Review of Political Science* 23, no. 1 (May 11, 2020). <https://doi.org/10.1146/annurev-polisci-060418-042801>.
- Deai, Sonalde, Reeve, Vanneman and National Council of Applied Economic Research, "India Human Development Survey (IHDS)", New Delhi ICPSR22626-v8. Ann Arbor, MI: Inter-university Consortium for Political and Social Research (2005).

- . “India Human Development Survey-II (IHDS-II)”, 2011-12. ICPSR36151-v2. Ann Arbor, MI: Inter-university Consortium for Political and Social Research (2012). <http://doi.org/10.3886/ICPSR36151.v2>
- Easterly W., and R. Levine. “Africa’s Growth Tragedy: Policies and Ethnic Divisions” *The Quarterly Journal of Economics* 112, no. 4 (November 1, 1997): 1203–50 <https://doi.org/10.1162/003355300555466>.
- Eilenberg, Michael, and Reed L. Wadley. “Borderland Livelihood Strategies: The Socio-Economic Significance of Ethnicity in Cross-Border Labour Migration, West Kalimantan, Indonesia.” *Asia Pacific Viewpoint* 50, no. 1 (April 2009): 58–73. <https://doi.org/10.1111/j.1467-8373.2009.01381.x>.
- Erikson, Erik H. *Identity, Youth and Crisis*. New York: Norton, 1994.
- Ewens, Michael, Bryan Tomlin, and Liang Choon Wang. “Statistical Discrimination or Prejudice? A Large Sample Field Experiment.” *Review of Economics and Statistics* 96, no. 1 (March 2014): 119–34. https://doi.org/10.1162/REST_a_00365.
- Fortunato, Laura. “The Evolution of Matrilineal Kinship Organization.” *Proceedings of the Royal Society B: Biological Sciences* 279, no. 1749 (December 22, 2012): 4939–45. <https://doi.org/10.1098/rspb.2012.1926>.
- Gneezy, Uri, Kenneth Leonard, and John List. “Gender Differences in Competition: Evidence from a Matrilineal and a Patriarchal Society.” *Econometrica* 77, no. 5 (2009): 1637–64. <https://doi.org/10.3982/ECTA6690>.
- Giuliano, Laura, and Michael R Ransom. “Manager Ethnicity and Employment Segregation.” *ILR Review* 66, no. 2 (April 2013): 346–79. <https://doi.org/10.1177/001979391306600203>.
- Goldsmith, Arthur H, Darrick Hamilton, and William Darity. “Shades of Discrimination: Skin Tone and Wages.” *American Economic Review* 96, no. 2 (April 2006): 242–45. <https://doi.org/10.1257/000282806777212152>.
- Government of India. “Census of India 2011 ‘Towards a Bright Future,’” 2011.
- Grujters, Rob J., and John Ermisch. “Patrilocal, Matrilocal, or Neolocal? Intergenerational Proximity of Married Couples in China.” *Journal of Marriage and Family* 81, no. 3 (June 2019): 549–66. <https://doi.org/10.1111/jomf.12538>.
- Guan, Lee Hock. “Affirmative Action in Malaysia.” *Southeast Asian Affairs*, 2005, 211–28.
- Guilmoto, Christophe Z. “Mapping the Diversity of Gender Preferences and Sex Imbalances in Indonesia in 2010.” *Population Studies* 69, no. 3 (September 2, 2015): 299–315. <https://doi.org/10.1080/00324728.2015.1091603>.

- Hanna, Rema N, and Leigh L Linden. "Discrimination in Grading." *American Economic Journal: Economic Policy* 4, no. 4 (November 2012): 146–68. <https://doi.org/10.1257/pol.4.4.146>.
- Hellerstein, Judith K., and David Neumark. "Workplace Segregation in the United States: Race, Ethnicity, and Skill." *Review of Economics and Statistics* 90, no. 3 (August 2008): 459–77. <https://doi.org/10.1162/rest.90.3.459>.
- Helms, Janet E., ed. *Black and White Racial Identity: Theory, Research, and Practice*. Contributions in Afro-American and African Studies 129. Westport, Conn.: Praeger, 1993.
- Hofstede, Geert. *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations across Nations*. 2. ed., [Nachdr.]. Thousand Oaks, Calif.: Sage Publ, 2001.
- Hogg, Michael A., and Deborah J. Terry. "Social Identity and Self-Categorization Processes in Organizational Contexts." *The Academy of Management Review* 25, no. 1 (January 2000): 121. <https://doi.org/10.2307/259266>.
- Holden, C. "Matriliney as Daughter-Biased Investment." *Evolution and Human Behavior* 24, no. 2 (March 2003): 99–112. [https://doi.org/10.1016/S1090-5138\(02\)00122-8](https://doi.org/10.1016/S1090-5138(02)00122-8).
- Holden, Clare Janaki, and Ruth Mace. "Spread of Cattle Led to the Loss of Matrilineal Descent in Africa: A Coevolutionary Analysis." *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270, no. 1532 (December 7, 2003): 2425–33. <https://doi.org/10.1098/rspb.2003.2535>.
- Hoon, Chang-Yau. "Assimilation, Multiculturalism, Hybridity: The Dilemmas of the Ethnic Chinese in Post-Suharto Indonesia 1." *Asian Ethnicity* 7, no. 2 (June 2006): 149–66. <https://doi.org/10.1080/14631360600734400>.
- Hornsey, Matthew J. "Social Identity Theory and Self-Categorization Theory: A Historical Review." *Social and Personality Psychology Compass* 2, no. 1 (January 2008): 204–22. <https://doi.org/10.1111/j.1751-9004.2007.00066.x>.
- Hyde, Abbey. "Matrilocality and Female Power." *Women's Studies International Forum* 22, no. 6 (November 1999): 597–605. [https://doi.org/10.1016/S0277-5395\(99\)00070-9](https://doi.org/10.1016/S0277-5395(99)00070-9).
- Jia, Ruixue, and Torsten Persson. "Individual vs. Social Motives in Identity Choice: Theory and Evidence from China." Cambridge, MA: National Bureau of Economic Research, June 2019. <https://doi.org/10.3386/w26008>.
- Kanner, Allen D., James C. Coyne, Catherine Schaefer, and Richard S. Lazarus. "Comparison of Two Modes of Stress Measurement: Daily Hassles and Uplifts versus

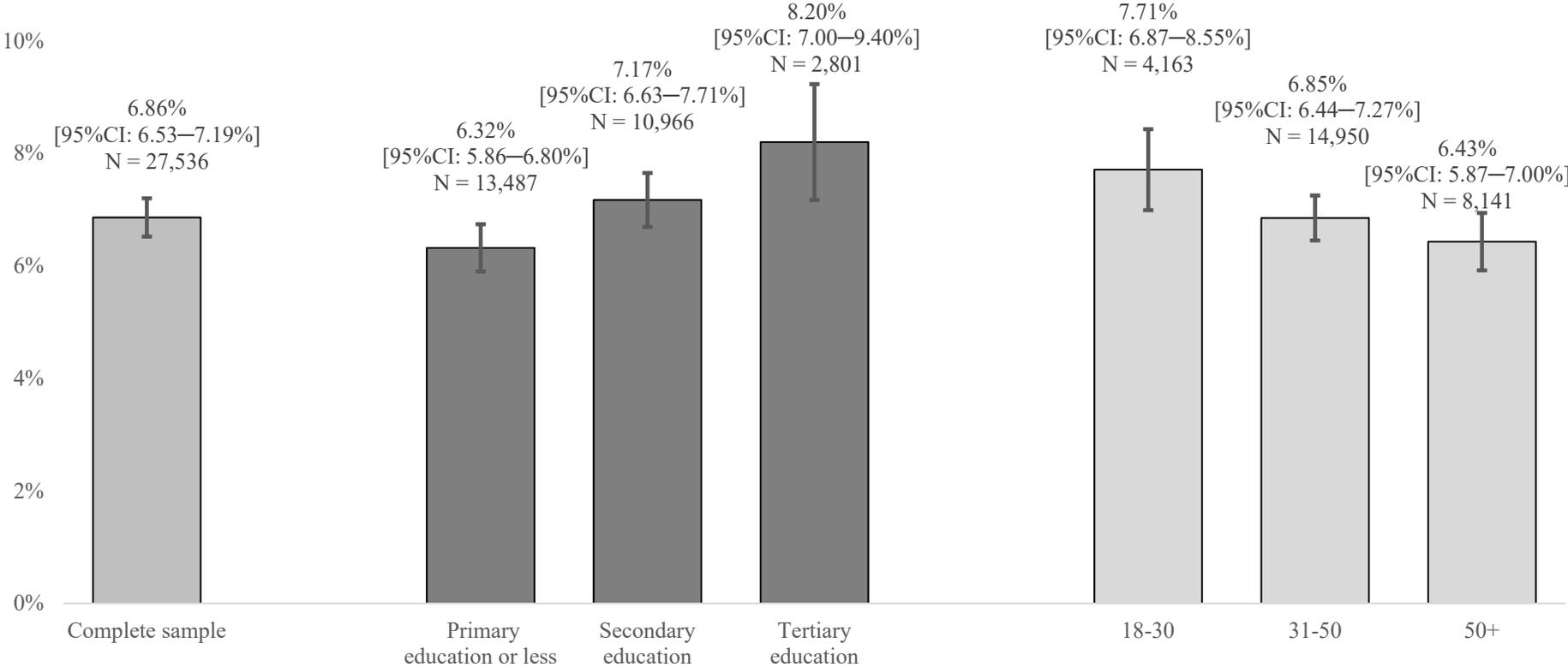
- Major Life Events.” *Journal of Behavioral Medicine* 4, no. 1 (March 1981): 1–39.
<https://doi.org/10.1007/BF00844845>.
- Kertzner, David I., and Dominique Arel, eds. *Census and Identity: The Politics of Race, Ethnicity, and Language in National Censuses*. 1st ed. Cambridge University Press, 2001. <https://doi.org/10.1017/CBO9780511606045>.
- Kramer, Rory, Robert DeFina, and Lance Hannon. “Racial Rigidity in the United States: Comment on Saperstein and Penner.” *American Journal of Sociology* 122, no. 1 (July 2016): 233–46. <https://doi.org/10.1086/687374>.
- Liebler, Carolyn A., Sonya R. Porter, Leticia E. Fernandez, James M. Noon, and Sharon R. Ennis. “America’s Churning Races: Race and Ethnicity Response Changes Between Census 2000 and the 2010 Census.” *Demography* 54, no. 1 (February 2017): 259–84. <https://doi.org/10.1007/s13524-016-0544-0>.
- Lim, Mah Hui. “Affirmative Action, Ethnicity and Integration: The Case of Malaysia*.” *Ethnic and Racial Studies* 8, no. 2 (April 1985): 250–76. <https://doi.org/10.1080/01419870.1985.9993485>.
- Lottum, Jelle van, and Daan Marks. “The Determinants of Internal Migration in a Developing Country: Quantitative Evidence for Indonesia, 1930–2000.” *Applied Economics* 44, no. 34 (December 2012): 4485–94. <https://doi.org/10.1080/00036846.2011.591735>.
- Lowes, Sara. “Kinship Structure & Women: Evidence from Economics.” *Daedalus* 149.1 (2020): 119-133.
- Mays, Vickie M., Susan D. Cochran, and Namdi W. Barnes. “Race, Race-Based Discrimination, and Health Outcomes Among African Americans.” *Annual Review of Psychology* 58, no. 1 (January 2007): 201–25. <https://doi.org/10.1146/annurev.psych.57.102904.190212>.
- Musgrave, Simon. “Language Shift and Language Maintenance in Indonesia.” In *Language, Education and Nation-Building: Assimilation and Shift in Southeast Asia*, edited by Peter Sercombe and Ruanni Tupas, 87–105. London: Palgrave Macmillan UK, 2014. https://doi.org/10.1057/9781137455536_5.
- Pahl, Kerstin, and Niobe Way. “Longitudinal Trajectories of Ethnic Identity Among Urban Black and Latino Adolescents.” *Child Development* 77, no. 5 (September 2006): 1403–15. <https://doi.org/10.1111/j.1467-8624.2006.00943.x>.
- Petersen, Emily E., Nicole L. Davis, David Goodman, Shanna Cox, Carla Syverson, Kristi Seed, Carrie Shapiro-Mendoza, William M. Callaghan, and Wanda Barfield. “Racial/Ethnic Disparities in Pregnancy-Related Deaths — United States, 2007–

- 2016.” *MMWR. Morbidity and Mortality Weekly Report* 68, no. 35 (September 6, 2019): 762–65. <https://doi.org/10.15585/mmwr.mm6835a3>.
- Phinney, Jean S. “Ethnic Identity and Acculturation.” In *Acculturation: Advances in Theory, Measurement, and Applied Research*, 63–81. Washington, DC: American Psychological Association, 2003.
- Rammohan, Anu, and Peter Robertson. “Do Kinship Norms Influence Female Education? Evidence from Indonesia.” *Oxford Development Studies* 40, no. 3 (September 2012): 283–304. <https://doi.org/10.1080/13600818.2012.711303>.
- Reynolds, Katherine J., John C. Turner, S. Alexander Haslam, and Michelle K. Ryan. “The Role of Personality and Group Factors in Explaining Prejudice.” *Journal of Experimental Social Psychology* 37, no. 5 (September 2001): 427–34. <https://doi.org/10.1006/jesp.2000.1473>.
- Roth, Wendy D. “Racial Mismatch: The Divergence Between Form and Function in Data for Monitoring Racial Discrimination of Hispanics*: Racial Mismatch.” *Social Science Quarterly* 91, no. 5 (December 2010): 1288–1311. <https://doi.org/10.1111/j.1540-6237.2010.00732.x>.
- Rubinstein, Yona, and Dror Brenner. “Pride and Prejudice: Using Ethnic-Sounding Names and Inter-Ethnic Marriages to Identify Labour Market Discrimination.” *The Review of Economic Studies* 81, no. 1 (January 1, 2014): 389–425. <https://doi.org/10.1093/restud/rdt031>.
- Ruggles, Steven, Flood, Sarah, Goeken, Ronald, Grover, Josiah Meyer, Erin, Pacas, Jose, and Sobek, Matthew. IPUMS USA: Version 9.0 [CPS2005-2012] Minneapolis, MN: IPUMS, 2019. <https://doi.org/10.18128/D010.V9.0>
- Rumbaut, Ruben G. “The Crucible Within: Ethnic Identity, Self-Esteem, and Segmented Assimilation among Children of Immigrants.” *International Migration Review* 28, no. 4 (December 1994): 748–94. <https://doi.org/10.1177/019791839402800407>.
- Sam, David L., and John W. Berry. “Acculturation: When Individuals and Groups of Different Cultural Backgrounds Meet.” *Perspectives on Psychological Science* 5, no. 4 (July 2010): 472–81. <https://doi.org/10.1177/1745691610373075>.
- Saperstein, Aliya, and Andrew M. Penner. “Racial Fluidity and Inequality in the United States.” *American Journal of Sociology* 118, no. 3 (November 2012): 676–727. <https://doi.org/10.1086/667722>.
- Shayo, Moses. “A Model of Social Identity with an Application to Political Economy: Nation, Class, and Redistribution.” *American Political Science Review* 103, no. 2

- (May 2009): 147–74. <https://doi.org/10.1017/S0003055409090194>.
- Sillander, Kenneth, and Jennifer Alexander. “Belonging in Borneo: Refiguring Dayak Ethnicity in Indonesia.” *The Asia Pacific Journal of Anthropology* 17, no. 2 (March 14, 2016): 95–101. <https://doi.org/10.1080/14442213.2016.1152882>.
- Spierings, Niels, Jeroen Smits, and Mieke Verloo. “Micro- and Macrolevel Determinants of Women’s Employment in Six Arab Countries.” *Journal of Marriage and Family* 72, no. 5 (September 29, 2010): 1391–1407. <https://doi.org/10.1111/j.1741-3737.2010.00772.x>
- Statistics Indonesia. “2010 Population Census,” 2010.
- Strauss, John, Firman Witoelar, and Bondan Sikoki. “The Fifth Wave of the Indonesia Family Life Survey: Overview and Field Report: Volume 1.” *RAND Corporation*, no. Volume 1 (2016).
- . “The Fourth Wave of the Indonesia Family Life Survey: Overview and Field Report: Volume 1.” *RAND Corporation*, no. Volume 1 (2009).
- . “The Third Wave of the Indonesia Family Life Survey: Overview and Field Report: Volume 1.” *RAND Corporation*, no. Volume 1 (2004).
- Sukamdi, and Mujahid. “UNFPA Indonesia Monograph Series: No.3, Internal Migration in Indonesia.” UNFPA Indonesia Monograph Series (2015).
- Tajfel, Henri, and John C. Turner. “The Social Identity Theory of Intergroup Behavior.” In *Political Psychology*, edited by John T. Jost and Jim Sidanius, 0 ed., 276–93. Psychology Press, (2004). <https://doi.org/10.4324/9780203505984-16>.
- U.S. Bureau of the Census. US Census Bureau, the American Community Survey Design and Methodology Report”, US Census Bureau, Washington, DC, USA, (2013)
- Wajdi, Nashrul, Clara H. Mulder, and Sri M. Adioetomo. “Inter-Regional Migration in Indonesia: A Micro Approach.” *Journal of Population Research* 34, no. 3 (September 2017): 253–77. <https://doi.org/10.1007/s12546-017-9191-6>.
- Ward, Christopher, and Styles, Irene. “Culturing Settlement Using Pre- and Post-Migration Strategies.” *Journal of Psychiatric and Mental Health Nursing* 12, no. 4 (August 2005): 423–30. <https://doi.org/10.1111/j.1365-2850.2005.00853.x>.
- Waters, Mary C. *Ethnic Options: Choosing Identities in America*. Berkeley: University of California Press, 1990.
- Weber, Max. *Economy and Society: An Outline of Interpretive Sociology*. Vol. 2. Nachdr. Berkeley: Univ. of California Press, 1921 [1978].
- Wimmer, Andreas. *Ethnic Boundary Making Institutions, Power, Networks*. New York;

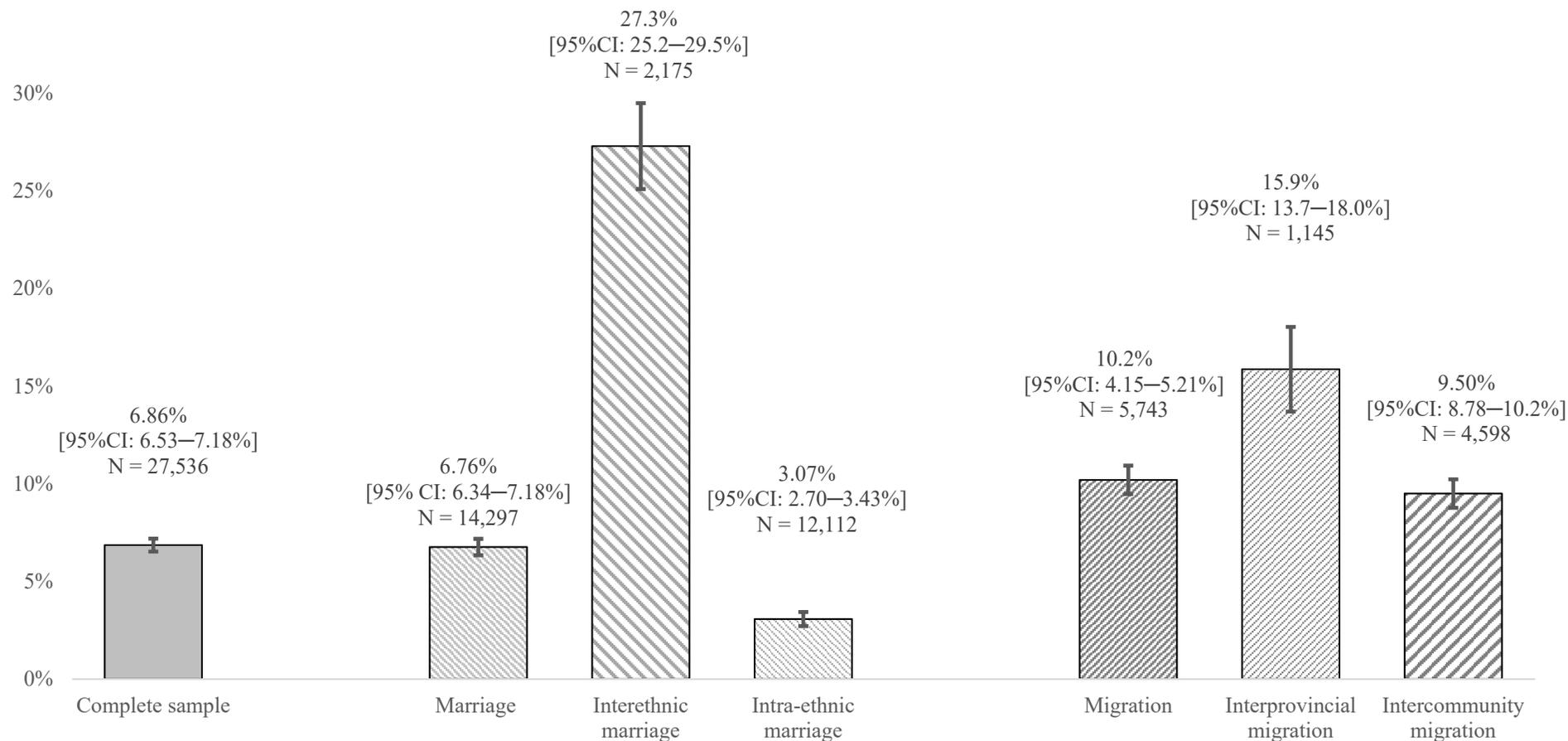
Oxford: Oxford University Press, 2013.

Figure 1. Marginal probability of intra-individual ethnicity change among different age cohorts and educational groups (EO1)



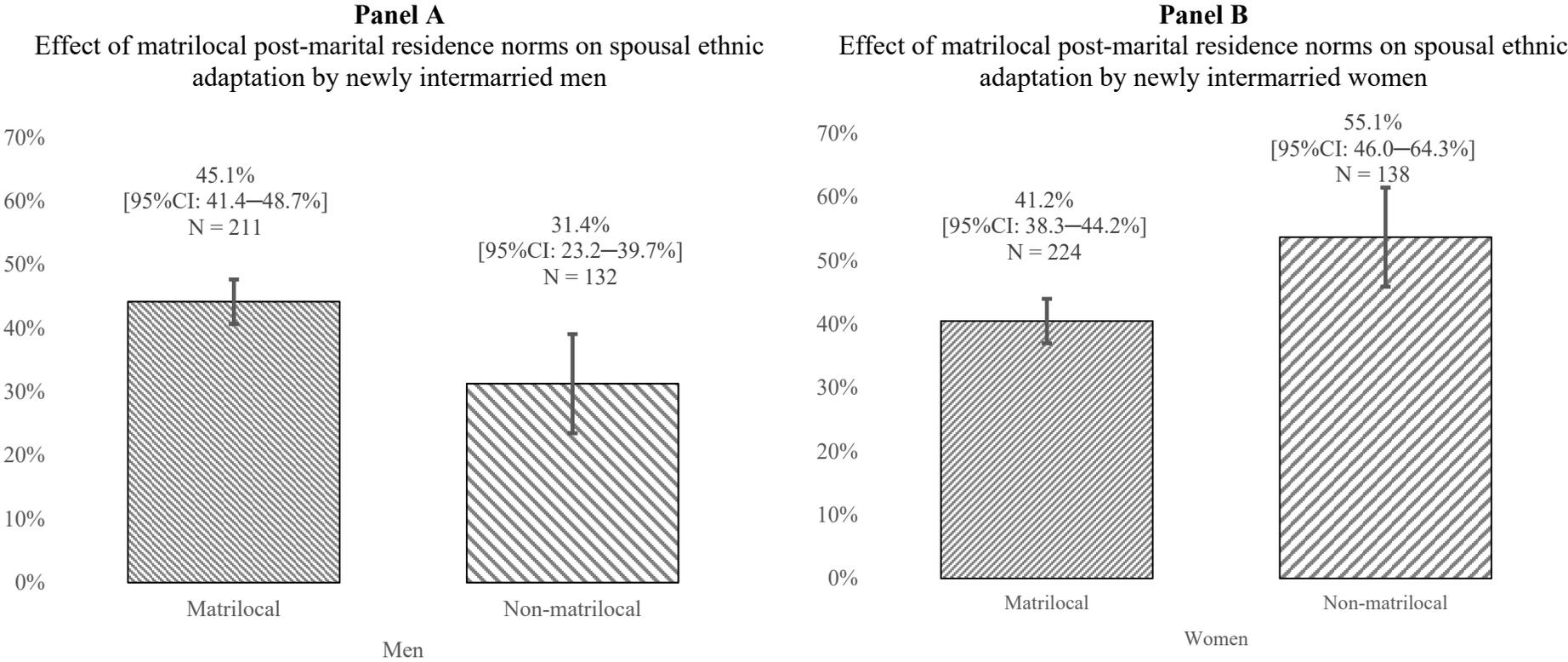
Notes: Figure presents the marginal probability of undirected intra-individual ethnicity change among different groups in society. These marginal probabilities derive from the estimation of Model B1 in Table B1 in Appendix B and are estimated at means. 95% confidence intervals (CIs) are calculated with robust standard errors.

Figure 2. Marginal probability of intra-individual ethnicity change coincidental with different major life events (EO2)



Notes: Figure presents the marginal probability of undirected intra-individual ethnicity change coincidental with different major life events, specifically marriage and internal migration. To estimate the marginal effects of interethnic and intra-ethnic marriage, we restrict the sample to married individuals. To estimate the marginal effects of interprovincial and intercommunity migration, we restrict the sample to internal migrants. The reported marginal probabilities derive from the estimation of the models in Table B2 of Appendix B and are estimated at means. 95% confidence intervals (CIs) are calculated using robust standard errors.

Figure 3. Marginal probabilities of spousal ethnic adaptation among newly intermarried men and women in matrilocal vs. non-matrilocal communities



Notes: Figure presents marginal probabilities of spousal ethnic adaptation, which is a form of directed intra-individual ethnicity change. These marginal probabilities derive from the estimation of Models C2 (newly intermarried men) and C3 (newly intermarried women) in Table C1 of Appendix C. 95% confidence intervals (CIs) are calculated with cluster robust standard errors.

Table 1. Churning and ethnic identity switching between major ethnic groups in Indonesia (in %)

Previous ethnicity→	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
New ethnicity↓																					
(1) Javanese [n=12,345]	97	5.1	0.9	1.6	0.7	5.2	9.9	0.2	1.2	2.6	0.4	1.1	0.0	9.7	0.8	0.0	8.2	3.4	11	4.6	7.3
(2) Sundanese [n=3,456]	1.2	91	0.2	0.1	0.0	2.1	0.1	0.0	0.6	0.5	0.0	0.3	0.0	4.4	0.8	0.0	9.5	8.4	0.0	0.8	2.4
(3) Balinese [n=1,389]	0.1	0.1	96	0.1	0.0	0.0	0.5	0.1	0.0	0.0	0.0	0.0	0.0	0.9	0.8	0.8	0.1	0.0	0.0	0.0	0.0
(4) Batak [n=925]	0.1	0.1	0.3	95	0.1	1	0.6	0.0	0.3	0.2	0.0	0.0	2.1	0.9	0.0	0.8	0.2	5.6	0.0	0.0	0.0
(5) Buginese [n=1,036]	0.1	0.0	1.3	0.1	96	3.1	0.1	0.4	0.0	0.7	0.2	7.0	0.0	0.0	0.8	0.8	0.3	0.0	0.0	0.0	0.0
(6) Tionghao [n=115]	0.1	0.0	0.0	0.1	0.4	72	0.4	0.0	1.1	0.0	0.0	0.0	2.1	1.8	0.0	0.0	0.8	0.0	0.0	0.6	0.0
(7) Madurese [n=860]	0.7	0.1	0.2	0.4	0.0	1.0	88	0.0	0.1	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.6	0.0
(8)Sasak [n=1,215]	0.0	0.0	0.4	0.0	0.4	1.0	0.0	95	0.0	0.0	0.7	0.3	0.0	0.0	0.8	0.0	0.0	0.0	0.0	0.0	0.0
(9) Minang [n=1,416]	0.1	0.1	0.0	0.7	0.0	5.2	0.0	0.0	94	0.0	0.0	0.0	4.2	1.8	0.0	0.0	0.4	12	1.8	1.0	2.4
(10) Banjarese [n=990]	0.2	0.1	0.0	0.0	0.4	0.0	0.2	0.0	0.0	95	0.0	0.3	0.0	0.0	0.0	0.0	0.1	1.1	0.0	0.0	0.0
(11) Bina- Dompu [n=532]	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.2	0.0	0.0	98	0.3	0.0	0.0	0.8	0.0	0.0	0.0	0.0	0.0	0.0
(12) Makassarese [n=372]	0.0	0.0	0.0	0.2	2.2	1.0	0.0	0.2	0.1	0.1	0.6	90	2.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0

Table 1, ctd.

Previous ethnicity→	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)
New ethnicity↓																					
(13) Nias [n=41]	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.3	90	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
(14) Palembang [n=118]	0.1	0.1	0.0	0.0	0.0	2.1	0.0	0.0	0.1	0.0	0.0	0.0	0.0	49	0.8	0.0	0.2	0.0	5.5	6.4	0.0
(15) Sumbawanese [n=174]	0.0	0.0	0.1	0.1	0.1	0.0	0.0	3.9	0.0	0.0	0.0	0.0	0.0	0.8	91	0.0	0.1	0.0	1.8	0.2	0.0
(16) Toraja [n=132]	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.8	98	0.1	0.0	0.0	0.0	0.0
(17) Betawi [n=973]	0.4	2.0	0.0	0.1	0.0	3.1	0.3	0.1	0.1	0.0	0.0	0.3	0.0	0.9	0.0	0.0	74	39	9.1	0.2	2.4
(18) Melaya [n=214]	0.3	0.7	0.0	0.8	0.0	0.0	0.0	0.0	1.9	0.2	0.0	0.0	0.0	4.4	1.6	0.0	4.8	30	0.0	1.8	4.9
(19) Komerling [n=43]	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	2.6	0.0	0.0	0.2	0.0	27	3.3	0.0
(20) South-Sumatran [n=616]	0.1	0.1	0.1	0.1	0.0	1.0	0.1	0.0	0.3	0.0	0.0	0.0	0.0	23	0.8	0.0	0.3	0.6	44	81	4.9
(21) other [n=46]	0.0	0.1	0.1	0.2	0.0	2.1	0.0	0.0	0.1	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.6	0.0	0.2	76
Total (sum of Rows 1 to 21)	100%																				

Notes: Table presents the raw percentages of all individuals in the sample switching to / retaining a specific ethnicity, say, Javanese. To improve readability, we have collapsed all ethnicities with fewer than 20 individual observations in a single group called “other.” This group includes “Ambonese,” “Ache,” “Manado” and “Dyak.”

Table 2. Understatement of intermarriage in Indonesia, the United States and India

	Percentage of marriages that is an intermarriage		Percentage difference corrected vs. uncorrected	Percentage point difference corrected vs. uncorrected
	Raw, not corrected for spousal social identity adaptation	Corrected		
Indonesia (interethnic marriage)	14.5% [95%CI: 13.9–15.0%]	19.2% [95%CI: 18.6–19.8%]	+32.4%	4.72 (0.002) [p=0.000]
United States (interracial marriage)	5.87% [95%CI: 5.56–6.18%]	8.60% [95% CI: 8.24–8.97%]	+46.5%	2.73 (0.001) [p=0.000]
India, indirect measure based on self-reported caste at time t-1 and time t (intercaste marriage)	0 ^s	17.7% [95%CI: 17.4–17.9%]	x	17.7 (0.001) [p=0.000]
India, direct measure based on item on spouse's caste (intercaste marriage)	4.92% [95%CI: 4.65–5.18%]	17.7% [95%CI: 17.4–17.9%]	+259%	12.8 (0.003) [p=0.000]

Notes: Results concern only newly married individuals, meaning individuals who have married or re-married between one wave/year and the next. The difference between the raw and corrected measures for interethnic/interracial/intercaste marriage is that the former uses the post-marriage ethnic self-identification of newly married individuals to calculate the prevalence of intermarriages whereas the latter uses these individuals' ethnic self-identification before their marriage. We consider a marriage to be an intra-ethnic/intraracial/intra-caste marriage if the self-reported identity of a newly married individual is the same as the self-reported identity of his/her spouse. We consider a marriage to be an interethnic/interracial/intercaste marriage if the self-reported identity of a newly married individual is not the same as the self-reported racial identity of his/her spouse. Data for Indonesia are from the IFLS, Waves 3-5 (2000, 2007 & 2014). The prevalence of interethnic marriages in Indonesia is based on comparing the self-reported ethnicity of an individual with the self-reported ethnicity of the individual's spouse (see Table A1 and the main text). Data for the U.S. are from the 2005-2012 Current Population Survey (Ruggles et al. 2019). The unadjusted measure of the prevalence of interracial marriage is based on individuals' (and spouses') self-reported racial identity, which is measured with the questionnaire item asking "What is your Race?" (see Appendix D for details). Data for India are from the India Human Development Survey or IHDS (Deai et al. 2005, 2012). We use two measure to estimate the prevalence of intercaste marriage in India. The first measure is based on individuals' self-reported caste identity. This measure of caste identity derives from the questionnaire item asking "Which caste do you belong to?" (see Appendix E for details). The second measure is based on the item asking "Is your husband's family the same caste as your natal family?" If this question is answered positively, we consider the marriage to be an intra-caste marriage (=0). If this question is answered negatively, we consider the

marriage to be an intercaste marriage (=1).

Table A1. Overview and description of key variables

Variable	Figure	Table	Definition/description	Level of measurement
Intra-individual ethnicity change (based on individual's self-reported ethnicity)	1,2	B1, B2, B3	Dummy variable that indicates whether an individual changes (=1) or retains (=0) his/her current self-reported ethnicity in comparison to the individual's self-reported ethnicity in the previous wave. The underlying measure of ethnic group membership is based on the item asking respondents "What is your Ethnicity?" Respondents can select (at least) one of the following ethnicities: "Jawa," "Sunda," "Bali," "Batak," "Bugis," "Tionghoa," "Madura," "Sasak," "Minang," "Banjar," "Bima-Dompou," "Makassar," "Nias," "Palembang," "Sumbawa," "Toraja," "Betawi," "Dayak," "Melayu," "Komerling," "Ambon," "Manado," "Aceh," "Other South Sumatera," "Banten," "Cirebon," "Gorontalo," "Kutai" and "Other." This dummy variable captures generic, undirected intra-individual ethnicity change.	Individual
Intra-individual ethnicity change (based on parental / ancestral ethnicity)	-	B4	Dummy variable that indicates whether an individual changes (=1) or retains (=0) his/her current self-reported ethnicity in comparison to the individual's self-reported ethnicity of his/her parents. The underlying measure of ethnic group membership is based on the items asking respondents "What is your ethnicity", "What is your father's ethnicity?" and "What is your mother's ethnicity?" For each question the respondent can select (at least) one of the following ethnicities: "Jawa," "Sunda," "Bali," "Batak," "Bugis," "Tionghoa," "Madura," "Sasak," "Minang," "Banjar," "Bima-Dompou," "Makassar," "Nias," "Palembang," "Sumbawa," "Toraja," "Betawi," "Dayak," "Melayu," "Komerling," "Ambon," "Manado," "Aceh," "Other South Sumatera," "Banten," "Cirebon," "Gorontalo," "Kutai" and "Other." As above, this dummy variable captures generic, undirected intra-individual ethnicity change.	Individual
Intra-individual ethnicity change (based on ethnic language spoken)	-	B5	Dummy variable that indicates whether individual reports speaking the ethnic language of his/her ethnic group after a change in one's self-reported ethnic identity (=1) or whether the individual reports not currently speaking the ethnic language of his/her new ethnic group after a change in self-reported ethnic identity (=0). The underlying measure of ethnic language speaking is based on the item "What languages do you speak in your daily life at home?" on which respondents can answer by selecting (at least) one of the following ethnic languages: "Indonesian," "Javanese," "Sundanese," "Maduranese," "Balinese," "Sasak," "Minang," "Batak," "Bugis," "Chinese," "Banjar," "Bima,"	Individual

Variable	Figure	Table	Definition/description	Level of measurement
			“Makassar,” “Nias,” “Palembang,” “Sumbawa,” “Toraja,” “Lahat,” “Other South Sumatra,” “Betawi,” “Lampung,” or “Other.” As above, this dummy variable captures generic, undirected intra-individual ethnicity change.	
Spousal ethnic adaptation (based on individual’s self-reported ethnicity)	3	C1, C3	Dummy variable that indicates whether a newly intermarried individual changes his/her self-reported ethnic identity and adopts the self-reported ethnicity of his/her spouse (=1) or retains the same ethnicity as reported in the previous wave (=0). The underlying measure of ethnic group membership is based on the item asking respondents “What is your ethnicity?” for both the individual and his/her spouse. This dummy variable captures directed intra-individual ethnicity change.	Individual
Spousal ethnic adaptation (based on parental / ancestral ethnicity)	-	C2	Dummy variable that indicates whether a newly intermarried individual changes his/her self-reported ethnic identity and adopts the self-reported ethnicity of his/her spouse (=1) or retains the same ethnicity as reported in the previous wave (=0). The underlying measure of ethnic group membership is based on the items asking respondents “What is your ethnicity?”, “What is your father’s ethnicity?” and “What is your mother’s ethnicity?” For each question the respondent can select one of the following categories: “Jawa,” “Sunda,” “Bali,” “Batak,” “Bugis,” “Tionghoa,” “Madura,” “Sasak,” “Minang,” “Banjar,” “Bima-Dompu,” “Makassar,” “Nias,” “Palembang,” “Sumbawa,” “Toraja,” “Betawi,” “Dayak,” “Melayu,” “Komerling,” “Ambon,” “Manado,” “Aceh,” “Other South Sumatera,” “Banten,” “Cirebon,” “Gorontalo,” “Kutai,” and “Other.” As above, this dummy variable captures directed intra-individual ethnicity change.	Individual
Interethnic marriage	2	B2, B3 B4	Dummy variable that indicates whether an individual’s self-reported ethnicity before his/her marriage is different from the current self-reported ethnicity of his/her spouse (=1) or not (=0). The ethnic categories to determine the match between the ethnicity of the individual and his/her spouse are the same as the categories mentioned in the description of the measure of intra-individual ethnicity change.	Individual
Intra-ethnic marriage	2	B2, B3 B4	Dummy variable that indicates whether an individual’s self-reported ethnicity before his/her marriage is the same as the current self-reported ethnicity of his/her spouse (=1) or not (=0). The ethnic categories to determine the match between the ethnicity of the individual and his/her spouse are the same as the categories mentioned in the description of the measure of intra-individual ethnicity change.	Individual

Variable	Figure	Table	Definition/description	Level of measurement
Marriage	2	B1, B2, B3, B4	Dummy variable that indicates whether an individual has gotten married (=1) or not (=0) between waves. This measure is constructed based on the availability of a unique personal identifier for a spouse. In the case the spouse is identified in the current wave but not in the previous, we consider the individual to have gotten married.	Individual
Migration	2	B1, B2, B3, B4	Dummy variable that indicates whether an individual has migrated to a different community between waves (=1) or otherwise (=0). The variable is based on the interviewer's recording of the community in which the interviewed household lives.	Individual
Interprovincial migration	2	B1, B2, B3, B4	Dummy variable that indicates whether an individual has migrated to a different province between waves (=1) or otherwise (=0). The variable is based on the interviewer's recording of the community in which the interviewed household lives.	Individual
Intercommunity migration	2	B1, B2, B3, B4	Dummy variable that indicates whether an individual has migrated to a different community within the same province between waves (=1) or otherwise (=0). The variable is based on the interviewer's recording of the community in which the interviewed household lives.	Individual
Age	2, 3	B1, B2, B3, B4, B5, C1, C2, C3	Interval variable running from 18-104.	Individual
Employment status	1,2, 3	B1, B2, B3, B4, B5, C1, C2, C3	Set of dummies indicating the employment status of the individual. The dummies derive from the item asking "Which category best describes the work that you do" with the following possible answers: 0 = No work, 1 = Self-employed (reference), 2 = Waged Work, 3= Unpaid Family Work, 4 = Flexible agricultural work.	Individual
Education	2, 3	B1, B2, B3, B4, B5, C1, C2, C3	Set of dummies indicating the highest level of education attained by the individual. The dummies derive from the item asking "What is the highest education level attended" with the following possible answers: 1 = None or Primary (reference), 2= Lower-Secondary, 3= Higher-Secondary, 4= Tertiary, 5 = Adult Education.	Individual
Religious denomination	1,2, 3	B1, B2, B3, B4, B5, C1,	Set of dummies to indicate the religious denomination of individuals. The dummies derive from the item asking "What is your religion?" with the following possible answers: 1 = Islam, 2 = Catholic, 3 = Protestant, 4 = Hindu, 5 = Buddha, 6 = Konghucu,	Individual

Variable	Figure	Table	Definition/description	Level of measurement
		C2, C3	7 = Not applicable.	
Member of local ethnic minority	3	C1, C2, C3	Dummy variable that indicates whether the individual belongs to the largest ethnic group in one's community (=0) or not (=1). The underlying measure for the individual's ethnicity is based on the question: "What is your ethnicity" and for the dominant community ethnicity we use the question: "How is the ethnic composition in this village/township [with regard to the largest ethnic group]?" that is asked to one of the local community leaders.	Individual
Double ethnicity	1,2	B1, B2, B3, B4, B5	Dummy variable that indicates whether the individual self-identifies with multiple ethnic backgrounds (=1) or one ethnicity (=0). The variable is based on the item asking "What is your Ethnicity?," which allows individuals to indicate more than one ethnicity (see above).	Individual
Educational difference spouse	3	C1, C2, C3	Set of dummies that indicates whether the individual has the same, more, or less education than his/her spouse does. The underlying measure for the variable is based on the question asked to the individual and one's spouse: "What is the highest education level attended" with the following possible items for the individual and one's spouse 0 = None, 1= Primary (reference), 2= Lower-Secondary, 3= Higher-Secondary, 4= Tertiary.	Individual
Perceived economic well-being	1,2, 3	B1, B2, B3, B4, B5, C1, C2, C3	Ordinal variable that indicates the subjective economic well-being from 0 (=poorest) until 6 (richest). The underlying measure for the variable is based on the question: "Please imagine a six-step ladder where on the bottom (the first step), stand the poorest people, and on the highest step (the sixth step), stand the richest people. On which step are you today?"	Individual
Speaks Bahasa national language	1,2, 3	B1, B2, B3, B4, B5, C1, C2, C3	Dummy variable that indicates whether the individual is able to speak the national language Bahasa Indonesia (=1) or not (=0). The underlying measure for the variable is based on the question: "What languages do you speak in your daily life at home?". If the individual answers "Bahasa Indonesia", we consider her able to speak this language.	Individual
Perceived accuracy of respondent's interview answers	1,2, 3	B1, B2, B3, B4, B5, C1, C2, C3	Ordinal variable that indicates the interviewer's perception on the accuracy of the interviewee's answers from 1 (=not so good) until 4 (Excellent). The underlying measure for the variable is based on the question: "What is your evaluation of the accuracy of the respondent's answers?"	Individual

Variable	Figure	Table	Definition/description	Level of measurement
Interviewer has conducted IFLS surveys before	1,2, 3	B1, B2, B3, B4, B5, C1, C2, C3	Dummy variable that indicates whether the interviewer has conducted surveys before or whether it is the time that the interviewer participates in a large-scale survey(=0). The underlying measure for the variable is based on the question: “Before IFLS (4 or 5) have you ever been involved in any other survey?”	Interviewer
Ethnicity of interviewer	1,2, 3	B1, B2, B3, B4, B5, C1, C2, C3	Set of dummies to indicate the self-reported ethnicity of the interviewer. The underlying measure for the variable is based on the question: ”What is your ethnicity”	Interviewer
Matrilocal / matrilocality community	3	C1, C2, C3	Dummy variable that indicates whether a community has matrilocality residence norms (=1) or not (i.e., non-matrilocality residence norms) (=0). The variable is based on the questionnaire item asking a community’s mayor, where newly married couples live after their wedding? If the mayor answered “The bride’s own place” or “The bride’s parents place,” we consider the community matrilocality (=1). Other possible answers, for instance, “Wherever they want”, “The groom’s own place” or “Groom’s parent’s house” are coded as non-matrilocality residence norms.	Community

Notes: Details on the wording of the various questionnaire items used can be found in the questionnaire books of the Indonesia Family Life Survey (Strauss et al. 2004, 2009, 2016).

Table A2.1. Self-reported ethnic identity and ethnic behavior (in %)

Self-reported ethnicity	Ethnic language spoken	Parental / ancestral ethnicity	Religious denomination							Customs		
			Islam			Catholic	Protestant	Hindu	Other religious denomination	Own dog	Own pig	Eat beef
			Overall	Traditional	Modern							
Javanese [N=12,310]	89.4	99.6	96.0	83.6	16.4	2.0	1.7	0.2	0.1	0.39	0	13.2
Sundanese [N=3,373]	83.8	99.0	99.7	94.1	5.85	0.12	0.09	0.06	0	0.30	0	14.1
Balinese [N=1,403]	96.0	99.6	1.93	100	0	0.80	0.20	97.0	0.1	5.41	6.12	5.0
Batak [N=872]	71.4	99.8	34.9	86.9	13.1	5.05	59.8	0.34	0	5.71	0.91	14.2
Buginese [N=993]	81.0	98.7	85.2	88.1	11.9	0.10	0.10	14.5	0.10	2.79	0	4.44
Tionghao [N=102]	10.8	99.0	25.5	73.3	26.7	19.6	29.4	1.0	24.5	1.85	0	35.3
Madurese [N=837]	93.0	98.9	99.4	96.8	3.29	0	0.48	0	0.12	0	0	26.2
Sasak [N=630]	96.8	100	100	92.3	7.67	0	0	0	0	0.48	0	35.2
Minang [N=1,360]	90.8	99.7	99.9	54.4	45.6	0.07	0	0	0	1.32	0	29.3
Banjarese [N=940]	98.2	99.2	100	94.7	5.3	0.	0.	0	0	0	0	9.90
Bina-Dompu [N=530]	98.2	99.6	100	75.2	24.8	0	0	0	0	0	0	18.3
Makassarese [N=345]	90.4	99.7	100	73.2	26.8	0	0	0	0	0	0	2.03
Nias [N=37]	81.1	100	8.10	50.0	50.0	24.3	67.6	0	0	0	0	27.0

Table A2.1, ctd.

Self-reported ethnicity	Ethnic language spoken	Parental / ancestral ethnicity	Religious denomination						Customs			
			Islam			Catholic	Protestant	Hindu	Other religious denomination	Own dog	Own pig	Eat beef
			Overall	Traditional	Modern							
Palembangnese [N=90]	64.4	98.9	95.9	78.6	21.4	0	0	0	1.11	0	0	17.8
Sumbawane [N=166]	92.8	100	100	84.9	15.1	0	0	0	0	0	0	20.5
Toraja [N=128]	90.6	100	15.8	47.1	52.9	15.0	69.3	0	0	4.76	7.94	5.47
Betawi [N=795]	-	94.0	99.4	95.8	0	0.50	0	0	0.13	0	0	14.8
Dayak [N=13]	-	84.6	53.9	66.7	33.3	15.4	30.8	0	0	0	0	15.4
Melaya [N=186]	-	98.4	100	94.6	5.4	0	0	0	0	0.78	0	8.1
Komering [N=37]	-	97.3	100	80.0	20.0	0	0	0	0	0	0	16.2
Ambonese [N =5]	-	60	40	0	100	0	40	20	0	0	0	0
Acheh [N=10]	-	100	100	100	0	0	0	0	0	0	0	10.0
South-Sumatran [N=581]	89.0	98.5	99.1	87.3	22.7	0.5	0.2	0	0.2	0	0	11.5

Notes: Source is data from the IFLS, Wave 4 (2007) and Wave 5 (2014), and own calculations. The number of individuals who self-report belonging to the ethnic group is in square brackets.

Table A2.2. Ethnic language spoken and ethnic behavior (in %)

Ethnic language spoken	Self-reported ethnicity	Parental / ancestral ethnicity	Religious denomination							Customs		
			Islam			Catholic	Protes-tant	Hindu	Other religious denomination	Own dog	Own pig	Eat beef
			Overall	Traditional	Modern							
Javanese [n=11,261]	97.7	98.4	95.7	83.3	16.7	2.21	1.94	0.11	0.08	0.34	0.02	12.8
Sundanese [n=3,011]	93.9	99.8	99.6	94.3	5.7	0.13	0.07	0.03	0.0	0.33	0	13.0
Balinese [n=1,401]	96.2	96.9	3.36	90.5	4.5	0.50	0.21	95.9	0.07	5.50	5.92	4.72
Batak [n=644]	96.7	97.4	23.6	89.2	10.8	5.43	70.7	0	0.31	7.12	1.24	13.3
Buginese [n=834]	96.4	97.7	82.9	86.2	13.8	0	0.12	17.0	0	3.23	0	3.60
Tionghao [n=16]	68.8	81.3	18.8	66.7	33.3	0	56.3	0	25.0	0	0	31.3
Madurese [n=875]	88.9	93.8	99.8	96.8	3.2	0	0	0.11	0.11	0.23	0	25.3
Sasak [n=635]	96.1	97.3	99.7	91.8	8.2	0	0.16	0.16	0	0.49	0	35.3
Minang [n=1,306]	94.6	95.3	99.2	53.5	46.5	0.54	0.15	0	0.08	1.36	0	30.1
Banjarese [n=1,032]	89.4	92.3	99.1	94.8	5.3	0.29	0.58	0	0	0.	0	10.4
Bina-Dompu [n=526]	98.9	98.9	100	74.7	25.3	0	0	0	0	0	0	18.3
Makassarese [n=348]	89.7	92.2	98.6	74.3	25.7	0	1.44	0	0	0	0	1.44
Nias [n=32]	93.8	93.8	3.13	-	-	31.3	65.6	0	0	0	0	18.8

Table A2.2, ctd.

Ethnic language spoken	Self-reported ethnicity	Parental / ancestral ethnicity	Religious denomination							Customs		
			Islam			Catholic	Protestant	Hindu	Other religious denomination	Own dog	Own pig	Eat beef
			Overall	Traditional	Modern							
Palembangnese [n=209]	27.8	45.0	98.7	80.0	20.0	0	0	0.66	0.66	0	0	23.0
Sumbawane [n=159]	96.9	97.5	99.4	85.6	14.4	0	0.63	0	0	0	1.59	20.8
Toraja [n=121]	95.9	97.5	16.7	53.3	46.7	16.7	66.7	0	0	4.92	6.56	4.96
South-Sumatran [n=682]	75.8	81.1	98.5	86.7	13.3	0.50	0.99	0	0.25	0.84	0	11.0

Notes: Source is data from the IFLS, Wave 4 (2007) and Wave 5 (2014), and own calculations. The number of individuals who self-report speaking the ethnic language is in square brackets.

Table A2.3. Ethnic ancestry and ethnic behaviors (in %)

Parental / ancestral ethnicity	Self- reported ethnicity	Ethnic language spoken	Religious denomination							Customs		
			Islam			Catholic	Protes- tant	Hindu	Other religious denomination	Own dog	Own pig	Eat beef
			Overall	Traditional	Modern							
Javanese [n=12,331]	99.5	89.2	95.7	83.6	16.4	2.2	1.86	0.16	0.08	0.39	0	13.2
Sundanese [n=3,380]	98.7	83.3	99.8	94.1	5.88	0.09	0.09	0.06	0	0.29	0	13.9
Balinese [n=1,402]	99.7	95.9	1.93	100	0	0.79	0.21	97.0	0.07	5.41	6.13	4.93
Batak [n=871]	99.9	71.5	34.9	86.9	13.1	5.05	59.8	0.23	0	5.72	0.92	14.2
Buginese [n=982]	99.8	81.6	85.1	88.2	11.9	0.10	0.10	14.6	0.10	2.82	0	4.50
Tionghao [n=103]	98.1	12.6	25.2	75.0	25.0	19.4	29.1	1.94	24.2	3.70	0	35.0
Sasak [n=631]	99.8	96.7	100	92.4	7.65	0	0	0	0	0.48	0	35.2
Minang [n=1,363]	99.6	90.6	100	53.9	46.1	0	0	0	0	1.17	0	23.4
Banjarese [n=940]	99.2	97.9	100	94.8	5.3	0	0	0	0	0	0	10.1
Bina- Dompu [n=530]	99.8	97.9	100	74.9	25.1	0	0	0	0	0	0	18.5
Makassarese [n=351]	98.0	90.0	100	73.7	26.3	0	0	0	0	0	0	1.71
Nias [n=37]	100	81.1	8.11	50.0	50.0	34.3	67.6	0	0.1	0	0	27.0

Table A2.3, ctd.

Parental / ancestral ethnicity	Self- reported ethnicity	Ethnic language spoken	Religious denomination						Customs			
			Islam		Catholic	Protes- tant	Hindu	Other religious denomination	Own dog	Own pig	Eat beef	
			Overall	Traditional								Modern
Palembang- nese [n=85]	98.8	65.5	98.8	79.3	20.8	0	0.0	0	1.18	0	0	18.8
Sumbawa- nese [n=167]	99.4	92.2	100	84.9	15.1	0	0	0	0	0	0	20.4
Toraja [n=130]	98.5	90.0	17.1	47.1	52.9	14.7	68.2	0	0	4.69	7.81	5.38
Betawi [n=755]	98.9	-	99.3	96.1	3.91	0	0.53	0	0.13	0	0	14.9
Dyak [n =12]	91.7	-	50	50	50	16.7	33.3	0	0	0	0	8.33
Melaya [n=186]	99.4	-	100	94.6	5.41	0	0	0	0	0.77	0	8.06
Komering [n=37]	97.3	-	100	80	20	0	0	0	0	0	0	16.2
Ambonese [n=3]	100	-	33.3	0	100	0	66.7	0	0	0	0	0
Aceh [n=10]	100	-	100	100	0	0	0	0	0	0	0	10.0
South- Sumatran [n=580]	98.6	88.8	99.1	87.1	15.4	0.52	0.17	0	0.17	0	0	11.6

Notes: Source is data from the IFLS, Wave 4 (2007) and Wave 5 (2014), and own calculations. The number of individuals who self-report having parents belonging to the ethnic group is in square brackets.

Table A3. Cramer's V for measures of ethnic identity

	Self-reported ethnicity	Parental ethnicity
Parental ethnicity	0.975 [p=0.000]	-
Ethnic language	0.908 [p=0.000]	0.905 [p=0.000]

Notes: Cramer's V indicates to what degree categorical variables coincide in meaning, for instance, that 97.5% of the Indonesians who self-identify as member of a specific ethnic group have parents that are members of the same ethnic group. Table A1 presents details on the different measures of ethnic identity considered.

Table A4. Descriptive statistics for key variables

Variables	Generic, undirected intra-individual ethnicity change [n=27,536]	Spousal ethnic adaptation (directed intra-individual ethnicity change) [n=705]
Main dependent variables		
Change self-reported ethnicity	6.86% (25.3%)	-
Spousal ethnic adaptation	-	43.0% (49.5%)
Major life events		
Marriage	51.8% (50.0%)	-
Interethnic marriage	7.89% (27.0%)	-
Intra-ethnic marriage	43.9% (49.6%)	-
Migration	21.6 (41.2%)	-
Interprovincial migration	4.27% (20.2%)	-
Intercommunity migration	18.2% (38.5%)	-
Individual characteristics		
Age (in years)	44.0 (13.0)	38.5 (11.1)
Education		
Primary education or less	49.5% (50.0%)	40.2% (49.1%)
Secondary education	40.2% (49.0%)	49.5% (50.0%)
Tertiary education	10.3% (30.4%)	10.3% (30.4%)
Employment status		
Unemployed	21.8% (41.3%)	23.7% (42.5%)
Self-employed	34.9% (47.6%)	28.9% (45.4%)
Waged work	25.3% (43.5%)	31.3% (46.4%)
Unpaid family work	9.90% (29.9%)	7.66% (26.6%)
Agricultural flex work	8.10% (27.3%)	8.37% (27.7%)
Perceived economic well-being (1 poorest - 6 richest)	2.95 (0.957)	-
Ability to speak Bahasa national language (1=yes)	26.4% (44.1%)	61.3% (48.7%)
Education difference between individual and spouse		
Same education	-	51.8% (50.0%)
Spouse > individual	-	23.7% (42.5%)
Individual > spouse	-	24.5% (43.1%)
Perceived accuracy of respondent's interview answers (1 lowest - 5 highest)	4.10 (0.532)	4.13 (0.536)
Member ethnic minority (1=yes)	-	36.3% (48.1%)
Individual has double ethnicity (1=yes)	0.03% (5.55%)	-
Interviewer characteristics		
Interviewer has conducted IFLS interviews before (1=yes)	71.6% (45.1%)	70.2% (45.8%)
Community characteristics		
Matrilocal (1=yes)	-	61.7% (48.6%)

Notes: This table presents descriptive statistics for key variables in the analysis of undirected intra-individual ethnicity change (Column 1) and of spousal ethnic adaptation (Column 2).

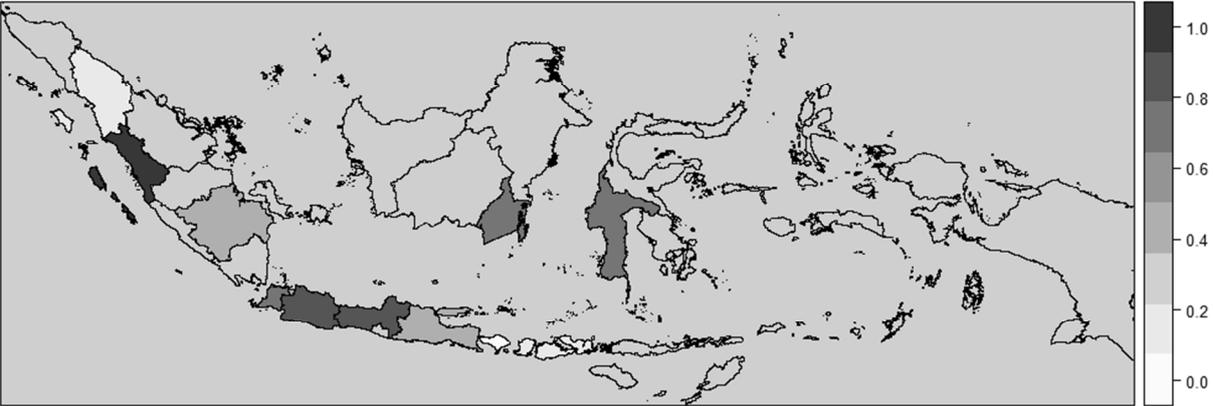
Standard deviations are in parentheses and the number of observations in square brackets. Source of the data is the IFLS, Waves 3-5 (Strauss et al., 2004, 2009, 2016). See the main text and Table A1 for details on these variables.

Table A5. Matrilocal and non-matrilocal communities

Island	Province	Number of communities	Percentage of communities in province that is matrilocal
Sumatera	North-Sumatera	15	20%
	West-Sumatera	5	100%
	South-Sumatera	2	50%
	Lampung	11	27.3%
Java	DKI Jakarta	28	46.4%
	West-Java	36	86.1%
	Central-Java	5	80%
	East-Java	18	50%
	Banten	4	75%
Bali	Bali	6	0%
Nusa Tenggara	West-Nusa Tenggara	10	20%
Kalimantan	South-Kalimantan	12	75%
Sulawesi	South-Sulawesi	9	66.7%
	West-Sulawesi	1	100%

Notes: Matrilocal vs. non-matrilocal post-marital residence norms are measured using the “Adat” (traditional lifestyle) module included in Wave 4 of the IFLS (see the main text and Table A1 for details).

Figure A1. The prevalence of matrilocal communities per province



Notes: The colors indicate the percentage of communities that have matrilocal (1; dark grey) vs. non-matrilocal (0; light grey / white) co residence norms for the Indonesian provinces in in the sample, Matrilocal vs. non-matrilocal post-marital residence norms are measured using the “Adat” (traditional lifestyle) module included in Wave 4 of the IFLS (see the main text and Table A1 for details).

Table B1. Individual random effects estimation of education and age as factors predicting the likelihood of intra-individual ethnicity change

Dependent = Intra-individual ethnicity change (1=yes)	Model B1
Educational groups	
Primary education or less (reference)	0
Secondary education	0.008 (0.004) [p=0.024]
Tertiary education	0.019 (0.007) [p=0.006]
Age cohorts	
18-30 (reference)	0
31-50	-0.009 (0.004) [p=0.055]
50+	-0.0123 (0.005) [p=0.019]
Individual-level control variables	
Gender (1=female)	0.001 (0.004) [p=0.753]
Double ethnicity (1=yes)	0.355 (0.058) [p=0.000]
Ability to speak Bahasa national language (1=yes)	0.032 (0.004) [p=0.000]
Dummies for perceived accuracy of respondent's interview answers	Yes
Dummies for subjective economic well-being	Yes
Dummies for employment status	Yes
Interviewer-level control variables	
Interviewer has conducted IFLS surveys before (1=yes)	-0.007 (0.003) [p=0.027]
Dummies for ethnicity of interviewer	Yes
Ethnic group fixed effects	Yes
Religious group fixed effects	Yes
Year/wave fixed effects	Yes
Community fixed effects	Yes
No. of observations	27,254
No. of individuals	13,734
Within R ²	0.051
Between R ²	0.212
Overall R ²	0.170

Notes: Table reports estimated linear probabilities. Standard errors (in parentheses) are robust standard errors and clustered at the individual level. Robust p-values are in square brackets. The estimated model includes baseline fixed effects to control for time-invariant ethnic group, religious group, community and year-wave factors. To make the estimates of our coefficients more interpretable, we report the outcomes of the linear probability models only. However, logistic models render similar results (available on request).

Table B2. Individual random effects estimation of factors predicting the likelihood of intra-individual ethnicity change considering the role of major-life events

Dependent = Intra-individual ethnicity change (1=yes)	Model B2	Newly married individuals only		Model B5	Migrants only	
		Model B3	Model B4		Model B6	Model B7
Marriage (1=yes)	-0.002 (0.003) [p=0.401]	-	-		-	-
Interethnic marriage (1=yes)	-	0.243 (0.012) [p=0.000]		-		-
Intra-ethnic marriage (1=yes)	-	-	-0.243 (0.012) [p=0.000]	-	-	-
Migration (1=yes)	-	-		0.011 (0.004) [p=0.006]	-	-
Interprovincial migration (1=yes)	-	-		-	0.070 (0.012) [p=0.000]	-
Intercommunity migration (1=yes)	-	-		-	-	-0.070 (0.012) [p=0.000]
Individual-level control variables						
Gender (1=female)	0.002 (0.004) [p=0.681]	0.001 (0.005) [p=0.823]	0.001 (0.005) [p=0.823]	0.001 (0.004) [p=0.717]	-0.003 (0.008) [p=0.679]	-0.003 (0.008) [p=0.679]
Double ethnicity (1=yes)	0.355 (0.058) [p=0.000]	0.349 (0.081) [p=0.000]	0.349 (0.081) [p=0.000]	0.356 (0.058) [p=0.000]	0.108 (0.129) [p=0.403]	0.108 (0.129) [p=0.403]
Age & age squared	Yes	Yes	Yes	Yes	Yes	Yes
Ability to speak Bahasa national language (1=yes)	0.031 (0.004) [p=0.000]	-0.001 (0.006) [p=0.835]	-0.001 (0.006) [p=0.835]	0.030 (0.004) [p=0.000]	0.021 (0.009) [p=0.024]	0.021 (0.009) [p=0.024]
Dummies for perceived accuracy of respondent's interview answers	Yes	Yes	Yes	Yes	Yes	Yes
Dummies for subjective economic well-being	Yes	Yes	Yes	Yes	Yes	Yes
Dummies for employment categories	Yes	Yes	Yes	Yes	Yes	Yes

Table B2, ctd.

Dependent = Intra-individual ethnicity change (1=yes)	Model B2	Newly married individuals only		Model B5	Migrants only	
		Model B3	Model B4		Model B6	Model B7
Dummies for educational categories	Yes	Yes	Yes	Yes	Yes	Yes
Interviewer-level control variables						
Interviewer has conducted IFLS surveys before (1=yes)	-0.007 (0.003) [p=0.017]	-0.001 (0.004) [p=0.862]	-0.001 (0.004) [p=0.862]	-0.007 (0.003) [p=0.019]	-0.007 (0.009) [p=0.438]	-0.007 (0.009) [p=0.438]
Dummies for ethnicity of interviewer	Yes	Yes	Yes	Yes	Yes	Yes
Ethnic group fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Religious group fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year/wave fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Community fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	27,536	14,295	14,295	27,536	5,953	5,953
No. of individuals	13,776	9,881	9,881	13,776	4,913	4,913
Within R ²	0.049	0.128	0.128	0.049	0.070	0.070
Between R ²	0.212	0.291	0.291	0.213	0.308	0.308
Overall R ²	0.169	0.267	0.267	0.169	0.275	0.275

Notes: Table reports estimated linear probabilities. Models B2 and B5 estimate the relationships between marriage and migration and the probability of individuals changing their ethnicity. Models B3 and B4 estimate the relationships between interethnic and intra-ethnic marriage and the probability of individuals changing their ethnicity restricted to consider only newly married individuals. Models B6 and B7 estimate the relationships between interprovincial and intercommunity migration and the probability of individuals changing their ethnicity restricted to a sample of migrants. Standard errors (in parentheses) are robust standard errors and clustered at the individual level. Robust p-values are in square brackets. All estimated models include baseline fixed effects to control for time-invariant ethnic group, religious group, community and year-wave factors. To make the estimates of our coefficients more interpretable, we report the outcomes of the linear probability models only. However, logistic models render similar results (available on request).

Table B3. Individual fixed effects estimation of factors predicting the likelihood of intra-individual ethnicity change

Dependent = Intra-individual ethnic identity change (1=yes)	Model B8	Model B9	Model B10	Model B11
Marriage (1=yes)	0.001 (0.003) [p=0.835]	-	-	-
Interethnic marriage (1=yes)	-	0.097 (0.007) [p=0.000]	-	-
Intra-ethnic marriage (1=yes)	-	-0.017 (0.004) [p=0.000]	-	-
Migration (1=yes)	-	-	0.010 (0.004) [p=0.029]	-
Interprovincial migration (1=yes)	-	-	-	0.058 (0.014) [p=0.000]
Intercommunity migration (1=yes)	-	-	-	0.000 (0.004) [p=0.924]
Individual-level control variables				
Ability to speak Bahasa national language (1=yes)	-0.001 (0.005) [p=0.814]	-0.002 (0.005) [p=0.620]	-0.001 (0.005) [p=0.788]	-0.001 (0.005) [p=0.824]
Age and age squared	Yes	Yes	Yes	Yes
Dummies for perceived accuracy of respondent's interview answers	Yes	Yes	Yes	Yes
Dummies for subjective economic well-being	Yes	Yes	Yes	Yes
Dummies for employment status	Yes	Yes	Yes	Yes
Interviewer-level control variables				
Dummies for ethnicity of interviewer	Yes	Yes	Yes	Yes
Interviewer has conducted IFLS survey before (1=yes)	-0.012 (0.004) [p=0.001]	-0.013 (0.003) [p=0.000]	-0.012 (0.004) [p=0.001]	-0.011 (0.004) [p=0.002]
Individual fixed effects	Yes	Yes	Yes	Yes
Ethnic group fixed effects	Yes	Yes	Yes	Yes
Year/wave fixed effects	Yes	Yes	Yes	Yes
Community fixed effects	Yes	Yes	Yes	Yes

Table B3, ctd.

Dependent = Intra-individual ethnic identity change (1=yes)	Model B8	Model B9	Model B10	Model B11
No. of observations	27,536	27,536	27,536	27,536
No. of individuals	13,776	13,776	13,776	13,776
Within R ²	0.099	0.115	0.099	0.102
Between R ²	0.001	0.009	0.001	0.002
Overall R ²	0.005	0.015	0.005	0.006

Notes: Table reports estimated linear probabilities. Models B8 and B9 estimate the relationship between (intra-ethnic and interethnic) marriage and the likelihood of individuals changing their ethnicity. Models B10 and B11 estimate the relationship between (interprovincial and intercommunity) migration and the likelihood of individuals changing their ethnicity. Standard errors (in parentheses) are robust standard errors and clustered at the individual level. Robust p-values are in square brackets. All estimated models include baseline fixed effects to control for time-invariant ethnic group, community and year-wave factors. Because there are no intra-individual religious conversions, estimated models do not include religious group fixed effects. To make the estimates of our coefficients more interpretable, we report the outcomes of the linear probability models only. However, logistic models render similar results (available on request).

Table B4. Individual random effects estimation of factors predicting the likelihood of intra-individual ethnicity change with parental definition of ethnicity

Dependent = Intra-individual ethnic identity change (1=yes)	Model B12	Model B13	Model B14	Model B15
Marriage (1=yes)	-0.001 (0.002) [p=0.614]	-	-	-
Interethnic marriage (1=yes)	-	0.083 (0.009) [p=0.000]	-	-
Intra-ethnic marriage (1=yes)	-	-0.020 (0.003) [p=0.000]	-	-
Migration (1=yes)	-	-	-0.001 (0.003) [p=0.875]	-
Interprovincial migration (1=yes)	-	-	-	0.030 (0.008) [p=0.000]
Intercommunity migration (1=yes)	-	-	-	-0.008 (0.003) [p=0.023]
Individual-level control variables				
Gender (1=female)	-0.001 (0.003) [p=0.706]	-0.003 (0.003) [p=0.444]	-0.001 (0.003) [p=0.666]	-0.001 (0.003) [p=0.705]
Double ethnicity (1=yes)	0.396 (0.068) [p=0.000]	0.390 (0.068) [p=0.000]	0.396 (0.068) [p=0.000]	0.397 (0.067) [p=0.000]
Age & age squared	Yes	Yes	Yes	Yes
Ability to speak Bahasa national language (1=yes)	0.010 (0.004) [p=0.008]	0.003 (0.004) [p=0.400]	0.010 (0.004) [p=0.008]	0.009 (0.004) [p=0.012]
Dummies for perceived accuracy of respondent's interview answers	Yes	Yes	Yes	Yes
Dummies for subjective economic well-being	Yes	Yes	Yes	Yes
Dummies for educational categories	Yes	Yes	Yes	Yes
Dummies for employment categories	Yes	Yes	Yes	Yes

Table B4, ctd.

Dependent = Intra-individual ethnic identity change (1=yes)	Model B12	Model B13	Model B14	Model B15
Interviewer-level control variables				
Dummies for ethnicity of interviewer	Yes	Yes	Yes	Yes
Interviewer has conducted IFLS surveys before (1=yes)	-0.002 (0.003) [p=0.445]	-0.002 (0.003) [p=0.440]	-0.002 (0.003) [p=0.442]	-0.002 (0.003) [p=0.512]
Year/wave fixed effects	Yes	Yes	Yes	Yes
Community fixed effects	Yes	Yes	Yes	Yes
No. of observations	26,306	26,306	26,306	26,306
No. of individuals	13,161	13,161	13,161	13,161
Within R ²	0.188	0.177	0.188	0.186
Between R ²	0.193	0.231	0.193	0.196
Overall R ²	0.191	0.219	0.192	0.194

Notes: Table reports estimated linear probabilities. Models B12 and B13 estimate the relationship between (intra-ethnic and interethnic) marriage and the likelihood of individuals changing their ethnicity. Models B14 and B15 estimate the relationship between (inter-provincial and inter-community) migration and the likelihood of individuals changing their ethnicity. Standard errors (in parentheses) are robust standard errors and clustered at the individual level. Robust p-values are in square brackets. All estimated models include baseline fixed effects to control for time-invariant ethnic group, community and year-wave factors. Because there are no intra-individual religious conversions, estimated models do not include religious group fixed effects (see, also, Table B3). To make the estimates of our coefficients more interpretable, we report the outcomes of the linear probability models only. However, logistic models render similar results (available on request).

Table B5. Individual linear probability models estimating whether individuals learn the language of their new ethnicity

Dependent = Learning language of new ethnic identity (1=yes)	Model B16 Cross-sectional	Model B17 Cross-sectional	Model B18 Panel
Intra-individual ethnicity change (1=yes)	0.734 (0.028) [p=0.000]	0.855 (0.018) [p=0.000]	0.380 (0.112) [p=0.002]
Intra-individual ethnicity change in the previous wave (1=yes)	-	0.143 (0.028) [p=0.000]	-
Individual-level control variables			
Gender (1=female)	0.001 (0.002) [p=0.480]	0.003 (0.004) [p=0.427]	-
Double ethnicity (1=yes)	0.106 (0.097) [p=0.274]	0.145 (0.109) [p=0.185]	-
Age and age squared	Yes	Yes	Yes
Ability to speak Bahasa national language (1=yes)	0.024 (0.007) [p=0.001]	0.034 (0.011) [p=0.003]	0.818 (0.258) [p=0.002]
Dummies for perceived accuracy of respondent's interview answers	Yes	Yes	Yes
Dummies for subjective economic well-being	Yes	Yes	Yes
Dummies for employment status	Yes	Yes	Yes
Dummies for educational categories	Yes	Yes	Yes
Interviewer-level control variables			
Interviewer has conducted IFLS surveys before (1=yes)	0.001 (0.002) [p=0.539]	-0.004 (0.004) [p=0.301]	0.001 (0.001) [p=0.508]
Dummies for ethnicity of interviewer	Yes	Yes	Yes
Individual fixed effects	No	No	Yes
Ethnic group fixed effects	Yes	Yes	Yes
Religious group fixed effects	Yes	Yes	No
Year/wave fixed effects	Yes	Yes	Yes
Community fixed effects	Yes	Yes	Yes

Table B5, ctd.

Dependent = Learning language of new ethnic identity (1=yes)	Model B16 Cross-sectional	Model B17 Cross-sectional	Model B18 Panel
No. of observations	17,933	8,931	17,933
No. of individuals	10,287	8,931	10,287
Within R ²	x	x	0.179
Between R ²	x	x	0.020
Overall R ²	0.313	0.384	0.025

Notes: Table reports estimated linear probabilities. Model B16 estimates the probability of individuals learning the ethnic language of their new ethnic group after a change in their ethnicity. Model B17 estimates the probability of individuals learning the ethnic language of their new ethnic group with a lag of maximum 7 years after a change in their ethnicity. Model B18 estimates the probability of individuals learning the ethnic language of their new ethnicity after a change in their ethnicity with individual fixed effects controlled for. Standard errors (in parentheses) are robust standard errors and clustered at the individual level. Robust p-values are in square brackets. All regressions include baseline fixed effects to control for time-invariant ethnic group, religious group, community and year-wave factors. To make the estimates of our coefficients more interpretable, we report the outcomes of the linear probability models only. However, logistic models render similar results (available on request).

Table C1. Individual logistic analysis estimating the log odds of an individual adopting the ethnic identity of his/her spouse in matrilocal vs. non-matrilocal communities

	Model C1 Men as reference category (Gender = 1 for women, 0 for men)	Model C2 Women as reference category (Gender = 1 for men, 0 for women)	Model C3 Men as reference category (Gender = 1 for women, 0 for men)
Dependent = Spousal ethnic adaptation (1=yes)			
Gender * matrilocal residence norm	-	0.993 (0.382) [p=0.009]	-0.993 (0.382) [p=0.009]
Gender	0.393 (0.172) [p=0.022]	-0.910 (0.403) [p=0.024]	0.910 (0.403) [p=0.024]
Individual-level control variables			
Ability to speak Bahasa national language (1=yes)	-0.399 (0.303) [p=0.189]	-0.572 (0.308) [p=0.063]	-0.572 (0.308) [p=0.063]
Educational difference with spouse			
Spouse has same level of education (=reference)	0	0	0
Spouse has less education	-0.112 (0.284) [p=0.694]	-0.109 (0.294) [p=0.710]	-0.109 (0.294) [p=0.710]
Spouse has more education	-0.589 (0.226) [p=0.009]	-0.539 (0.224) [p=0.016]	-0.539 (0.224) [p=0.016]
Individual belongs to community's ethnic minority before marriage (1=yes)	0.905 (0.684) [p=0.186]	0.896 (0.680) [p=0.188]	0.896 (0.680) [p=0.188]
Dummies for educational categories	Yes	Yes	Yes
Dummies for perceived accuracy of respondent's interview answers	Yes	Yes	Yes
Interviewer-level control variables			
Interviewer has conducted IFLS surveys before (1=yes)	Yes	Yes	Yes
Dummies for ethnicity interviewer	Yes	Yes	Yes
Ethnic group fixed effects	Yes	Yes	Yes
Religious group fixed effects	Yes	Yes	Yes
Year/wave fixed effects	Yes	Yes	Yes
Community fixed effects	Yes	Yes	Yes
Year-community fixed effects	Yes	Yes	Yes

Table C1, ctd.

	Model C1 Women as reference category (Gender = 1 for women, 0 for men)	Model C2 Men as reference category (Gender = 1 for men, 0 for women)	Model C3 Women as reference category (Gender = 1 for women, 0 for men)
Dependent = Spousal ethnic adaptation (1=yes)			
No. of observations	705	705	705
No. of clusters (provinces)	14	14	14
Pseudo R ²	0.31	0.314	0.314

Notes: Sample is restricted to newly intermarried men and women. Table reports estimated log odds. Model C1 estimates the effect of gender (1=female) on the likelihood of spousal ethnic adaptation. Models C2 and C3 estimate the effect of the interaction between matrilineal residence norms and gender on the likelihood of spousal ethnic adaptation using either men (Model C2) or women (Model C3) as the reference category. Standard errors (in parentheses) are cluster-robust standard errors. Robust p-values are in square brackets. The estimated model includes baseline fixed effects to control for time-invariant ethnic group, religious group, community and year-wave factors. Linear probability models render similar results (available on request).

Table C2. Individual logistic analysis estimating the log odds of an individual adopting the ethnic identity of his/her spouse measured using parental definition of ethnic adaptation

	Model C4 Men as reference category (Gender = 1 for women, 0 for men)	Model C5 Women as reference category (Gender = 1 for men, 0 for women)	Model C6 Men as reference category (Gender = 1 for women, 0 for men)
Dependent = Spousal ethnic adaptation (1=yes)			
Gender * matrilocal residence norm	-	1.15 (0.439) [p=0.009]	-1.15 (0.439) [p=0.009]
Gender	0.322 (0.198) [p=0.104]	-1.25 (0.510) [p=0.014]	1.25 (0.510) [p=0.014]
Individual-level control variables			
Ability to speak Bahasa national language (1=yes)	Yes	Yes	Yes
Educational difference spouse			
Spouse has same level of education (reference)	Yes	Yes	Yes
Spouse has less education	Yes	Yes	Yes
Spouse has more education	Yes	Yes	Yes
Individual belongs to community's ethnic minority before marriage (1=yes)	Yes	Yes	Yes
Dummies for educational categories	Yes	Yes	Yes
Dummies for perceived accuracy of respondent's interview answers	Yes	Yes	Yes
Interviewer-level control variables			
Interviewer has conducted IFLS surveys before (1=yes)	Yes	Yes	Yes
Dummies for ethnicity interviewer	Yes	Yes	Yes
Ethnic group fixed effects	Yes	Yes	Yes
Religious group fixed effects	Yes	Yes	Yes
Year/wave fixed effects	Yes	Yes	Yes
Community fixed effects	Yes	Yes	Yes
Year-community fixed effects	Yes	Yes	Yes
No. of observations	584	584	584
No. of clusters (provinces)	14	14	14
Pseudo R ²	0.324	0.332	0.332

Notes: See Table C1. Sample is restricted to newly intermarried men and women. Table reports estimated log odds. Standard errors (in parentheses) are cluster-robust standard errors. Robust p-values are in square brackets. The estimated model includes baseline fixed effects to control for time-invariant ethnic group, religious group, community and year-wave factors. Linear probability models render similar results (available on request).

Table C3. Individual logistic analysis estimating the log odds of an individual adopting the ethnicity of his/her spouse with added moderating effects

Dependent = Spousal ethnic adaptation (1=yes)	Model C7 Bahasa Indonesia as added moderator	Model C8 Education as added moderator	Model C9 Educational differences between spouses as added moderators	Model C10 All moderating effects
Gender * matrilocal residence norm	-1.00 (0.394) [p=0.034]	-0.836 (0.412) [p=0.042]	-0.918 (0.405) [p=0.023]	-0.854 (0.401) [p=0.033]
Gender (1=women)	1.06 (0.383) [p=0.006]	0.974 (0.390) [p=0.013]	0.996 (0.378) [p=0.008]	0.981 (0.382) [p=0.010]
Additional moderating effects				
Ability to speak Bahasa national language * matrilocal residence norm	Yes	No	No	Yes
Educational categories * matrilocal residence norm	No	Yes	No	Yes
Educational difference spouse * matrilocal residence norm	No	No	Yes	Yes
Individual-level control variables				
Ability to speak Bahasa national language (1=yes)	Yes	Yes	Yes	Yes
Educational difference spouse	Yes	Yes	Yes	Yes
Individual belongs to community's ethnic minority before marriage (1=yes)	Yes	Yes	Yes	Yes
Dummies for educational categories	Yes	Yes	Yes	Yes
Dummies for perceived accuracy of respondent's interview answers	Yes	Yes	Yes	Yes
Interviewer-level control variables				
Interviewer has conducted IFLS surveys before (1=yes)	Yes	Yes	Yes	Yes
Dummies for ethnicity interviewer	Yes	Yes	Yes	Yes
Ethnic group fixed effects	Yes	Yes	Yes	Yes
Religious group fixed effects	Yes	Yes	Yes	Yes
Year/wave fixed effects	Yes	Yes	Yes	Yes

Table C3, ctd.

Dependent = Spousal ethnic adaptation (1=yes)	Model C7 Bahasa Indonesia as added moderator	Model C8 Education as added moderator	Model C9 Educational differences between spouses as added moderators	Model C10 All moderating effects
Community fixed effects	Yes	Yes	Yes	Yes
Year-community fixed effects	Yes	Yes	Yes	Yes
No. of observations	703	705	705	703
No. of clusters (provinces)	14	14	14	14
Pseudo R ²	0.315	0.319	0.314	0.323

Notes: See Table C1. Sample is restricted to newly intermarried men and women. Table reports estimated log odds. Standard errors (in parentheses) are cluster-robust standard errors. Robust p-values are in square brackets. The estimated model includes baseline fixed effects to control for time-invariant ethnic group, religious group, community and year-wave factors. Linear probability models render similar results (available on request).

Table D1. Individual random and fixed effects estimation of factors predicting the likelihood of intra-individual racial identity change coincidental with (interracial) marriage

	Married individuals only				
	Model D1 Generic, undirected intra-individual racial identity change	Model D2 Generic, undirected intra-individual racial identity change	Model D3 Spousal racial adaptation (directed intra-individual racial identity change)	Model D4 Generic, undirected intra-individual racial identity change	Model D5 Spousal racial adaptation (directed intra-individual racial identity change)
Interracial marriage (1=yes)	0.204 (0.024) [p=0.000]	-	0.189 (0.022) [p=0.000]	0.135 (0.018) [p=0.000]	0.141(0.017) [p=0.000]
Intraracial marriage (1=yes)	-	-0.204 (0.024) [p=0.000]	-	0.000 (0.000) [p=0.573]	-
Dummies for family income	Yes	Yes	Yes	Yes	Yes
Dummies for educational categories	Yes	Yes	Yes	Yes	Yes
Dummies for household income	Yes	Yes	Yes	Yes	Yes
Age and age squared	Yes	Yes	Yes	Yes	Yes
Individual fixed effects	No	No	No	Yes	Yes
Community fixed effects	Yes	Yes	Yes	Yes	Yes
Racial group fixed effects	Yes	Yes	Yes	Yes	Yes
Month-wave fixed effects	Yes	Yes	Yes	Yes	Yes
No. of observations	5,351	5,351	5,336	9,171,406	9,164,433
No. of individuals	4,684	4,684	4,670	2,300,088	2,299,667
Within R ²	0.149	0.149	0.153	0.019	0.006
Between R ²	0.265	0.265	0.275	0.004	0.000
Overall R ²	0.239	0.239	0.248	0.004	0.000

Notes: Table reports estimated linear probabilities. Models D1, D2, and D4 estimate the relationship between interracial/intraracial marriage and the probability of individuals changing their racial identity. Models D3 and D4 estimate the relationship between interracial/intraracial marriage and the probability of individuals adopting their spouse's racial identity. Models D1-D3 estimate the relationship between interracial/intraracial marriage for a restricted sample of newly married individuals whereas models D4 and D5 estimate these effects for the whole sample. All regressions include baseline fixed effects to control for time-invariant individual (if possible), racial group, community and year-wave factors. Standard errors (in parentheses) are robust standard errors and clustered at the individual level. Robust p-values are in square brackets. To make

the estimates of our coefficients more interpretable, we report the outcomes of linear probability model only. However, logistic models render similar results (available on request). The dependent variables measure whether the individual has changed his/her self-reported racial identity (undirected intra-individual racial identity change) or whether the individual has adopted the racial identity of his/her spouse (directed intra-individual racial identity change). Undirected intra-individual racial identity change is a dummy variable that indicates whether an individual changed (=1) or retained (=0) his/her current self-reported race or racial identity in comparison to the individual's self-reported racial identity in the previous wave. The underlying measure of racial group membership is based on the item asking respondents "I am going to read you a list of five race categories. You may choose one or more races. For this survey, Hispanic origin is not a race. (Are/Is) (NAME/you) White; Black or African American; Pacific Islander, Asian, Indian or Alaska Native; or Native Hawaiian?" Spousal racial adaptation is a dummy variable that indicates whether an individual adapted his/her self-reported racial identity to have the same self-reported racial identity as his/her spouse (=1) or not (=0). The measure underlying this variable is similar to the dummy variable on undirected intra-individual racial identity change. Interracial marriage is a dummy variable that measures whether the individual had a different self-reported racial identity before his/her marriage than the self-reported racial identity of his/her spouse after their marriage. Data for the U.S. are from the 2005-2012 Current Population Surveys (Ruggles et al. 2019). Based on the results for Model D1, the marginal probability of intra-individual racial identity change among individuals that have married a spouse with another racial identity is 20.2% (95%CI: 15.6–24.7%). Based on the results for Model D2, the marginal probability of intra-individual racial identity change among individuals that have married a spouse with the same racial identity is -0.19% (95%CI: -0.59–0.21%). Based on the results for Models D1-D2, the marginal probability of intra-individual racial identity change among all newlyweds is 1.65% (95%CI: 1.10–2.21%).

Table D2. Churning and racial identity switching between major racial groups in the U.S. (in %)

Former self-reported race→	(1) White	(2) Black or African American	(3) Mixed	(4) Pacific Islander	(5) Asian	(6) Indian or Alaska Native	(7) Native Hawaiian
(1) White [n=10,615,509]	99.9	0.35	0.63	4.4	0.39	0.64	0.57
(2) Black or African American [n=1,346,862]	0.05	99.6	0.26	0.58	0.08	0.06	0.10
(3) Mixed [n=75,745]	0.01	0.03	99.0	0.04	0.00	0.02	0.0
(4) Pacific Islander [n=5,149]	0.00	0.00	0.00	65.5	0.00	0.00	0.0
(5) Asian [628,878]	0.03	0.03	0.04	27.05	99.5	0.01	0.28
(6) Indian or Alaska Native [n=88,150]	0.01	0.00	0.01	0.02	0.00	99.3	0.01
(7) Native Hawaiian [n=56,783]	0.00	0.00	0.01	2.4	0.02	0.00	99.0
Total (sum of Rows 1 to 7)	100%						

Notes: Table presents the raw percentages of all individuals in the sample switching to / retaining a specific racial identity, say, White. Calculations are based on data from the U.S. Current Population Survey (2008-2018). Individuals' racial identity is based on the following questionnaire item: "I am going to read you a list of five race categories. You may choose one or more races. For this survey, Hispanic origin is not a race. (Are/Is) (NAME/you) White; Black or African American; Pacific Islander, Asian, Indian or Alaska Native; or Native Hawaiian?" The number of individuals who self-report belonging to a certain racial group is reported in square brackets.

Table E1. Individual random effects estimation of factors predicting the likelihood of intra-individual caste identity change coincidental with (intercaste) marriage

	Newly married individuals only		Model E3 Generic, undirected intra-individual caste identity change	Model E4 Spousal adaptation (directed intra-individual caste identity change)
	Model E1 Generic, undirected intra- individual caste identity change	Model E2 Generic, undirected intra- individual caste identity change		
Intercaste marriage (1=yes)	0.800 (0.039) [p=0.000]	-	0.619 (0.040) [p=0.000]	0.643 (0.040) [p=0.000]
Intra-caste marriage (1=yes)	-	-0.800 (0.039) [p=0.000]	-0.155 (0.006) [p=0.000]	-
Individual-level control variables				
Household below poverty line	Yes	Yes	Yes	Yes
Years of educations	Yes	Yes	Yes	Yes
Age and age squared	Yes	Yes	Yes	Yes
Community fixed effects	Yes	Yes	Yes	Yes
Caste fixed effects	Yes	Yes	Yes	Yes
Year/wave fixed effects	Yes	Yes	Yes	Yes
No. of observations	5,018	5,018	150,851	68,305
No. of individuals	5,018	5,018	150,851	68,305
Adjusted R ²	0.742	0.742	0.091	0.113

Notes: Table reports estimated linear probabilities. Models E1, E2, and E3 estimate the relationship between intercaste/intra-caste marriage and the probability that an individual changes his/her caste identity. Model E4 estimates the relationship between an intercaste/intra-caste marriage and the probability that an individual adopts the caste identity of his/her spouse or not. Models E1 and E2 estimate this relationship for a restricted sample of newly married individuals whereas models E3 and E4 estimate this relationship for the whole sample. All regressions include baseline fixed effects to control for time-invariant caste, community and year/wave factors. Standard errors (in parentheses) are robust standard errors. Robust p-values are in square brackets. To make the estimates of our coefficients more interpretable, we report the outcomes of linear probability model only. However, logistic models render similar results (available on request). The dependent variables measure whether the individual has changed his/her self-reported caste identity (undirected intra-individual caste identity change) or whether the individual has adopted the caste identity of his/her spouse (directed intra-individual caste identity change). Undirected intra-individual caste identity change (Models E1, E2, and E3) is a dummy variable that indicates whether an individual has changed (=1) or retains (=0) his/her self-reported caste or caste identity compared to the individual's self-reported ethnicity in the previous wave. The underlying measure of caste membership is based on

the item asking respondents “Which caste do you belong to?” Spousal adaptation is a dummy variable that indicates whether an individual has adapted his/her self-reported caste identity to have the same self-reported caste identity as his/her spouse (=1) or not (=0). Intercaste marriage is a dummy variables that indicates whether the individual had a different self-reported caste identity before his/her marriage than the self-reported caste identity of his/her spouse after the marriage or whether the individual self-reports to have had an intercaste marriage measured with the questionnaire item: “Is your husband’s family the same caste as your natal family?” Data are from the 2005 and 2012 India Human Development Survey (IHDS 2005, 2012). Based on the results for Model E1, the marginal probability of intra-individual caste identity change coincidental with an intercaste marriage is 77.8% (95%CI: 70.0–85.7%). Based on the results for Model E2, the marginal probability of intra-individual caste identity change among individuals that have married a spouse from the same caste is only 1.34% (95%CI: -0.82 –3.50%) and not statistically significant at usual levels. Based on the results for Models E1-E2, the marginal probability of intra-individual caste change among all individuals is 22.2% (95%CI: 21.8-22.7%).

Table E2. Churning and caste identity switching between main castes in India (in %)

Former self-reported caste→	(1) Brahmin	(2) Forward caste	(3) Other backward caste	(4) Dalit	(5) Adivasi	(6) Muslim	(7) Christian, Sikh, and Jain
Current self-reported caste↓							
(1) Brahmin [n=9,299]	77.7	14.4	5.09	2.18	0.39	0.03	0.23
(2) Forward caste [n=132,107]	3.48	69.2	20.7	3.34	1.45	0.56	1.15
(3) Other backward caste [n=60,575]	1.00	10.9	81.9	3.89	1.59	0.29	0.44
(4) Dalit [n=35,732]	0.38	2.43	6.42	87.0	2.23	0.93	0.57
(5) Adivasi [n=14,131]	0.70	1.99	5.15	5.68	85.9	0.42	0.18
(6) Muslim [n=19,635]	0.03	0.22	0.35	1.03	0.71	97.6	0.12
(7) Christian, Sikh, and Jain [n=5,386]	0.41	7.52	6.92	8.11	1.42	0.43	75.2
Total (sum of Rows 1 to 7)				100%			

Notes: Table presents the raw percentages of all individuals in the sample switching to / retaining a specific caste, say, Brahmin. Calculations are based on data from the India Human Development Survey (2005 & 2012). Individuals' caste identity is based on the item asking "Which caste do you belong to?" The number of individuals who self-report belonging to a certain caste is reported in square brackets.