Measuring Interethnic Marriage in Africa *

Daniel N. Posner
Department of Political Science
University of California, Los Angeles
dposner@polisci.ucla.edu

February 3, 2026

Abstract

Interethnic marriage is commonly employed in the African politics literature as an indicator of social cohesion. However, using intermarriage rates for this purpose may be misleading because intermarriages are a reflection of both preferences and opportunities. If we are to interpret intermarriage rates as indicators of people's willingness to cross group boundaries (and hence as a measure of social cohesion), we must find a way of controlling for exposure to out-group members in local marriage markets. In this Note, I exploit census data from Zambia to demonstrate how this can be done. The findings, which reveal significant differences across estimates that do and do not control for local exposure to out-group members, underscore a significant weakness in common approaches. The findings also point to important substantive implications for understanding changes in social cohesion in Zambia—and likely other African societies—over time.

*The author thanks Riya Kadam, Galen Murray, and especially Branden Bohrnsen for his indispensable input and data analysis. Elliott Green, Noah Nathan, Volha Charnysh, and two anonymous reviewers provided helpful comments.

Introduction

Interethnic marriage is commonly employed in the African politics literature as an indicator of social cohesion (Bratton and Kimenyi 2008; Crespin-Boucaud 2020; Bandyopadhyay and Green 2021; Kim and Horowitz 2022; Kyei 2022; Pengl, Roessler and Rueda 2022) as well as an explanatory variable in analyzes of ethnic voting (Dulani et al. 2022), conflict (Demarest and Haer 2022), and trust in elected local officials (LeBas 2020), among other outcomes.[1] Because marriage is an intimate relationship, the decision about whether to select a partner from a different ethnic group constitutes perhaps the quintessential behavioral indicator of inter-group social distance (Hechter 1978). In settings where ethnic communities have historically spoken different languages, followed different cultural practices, and exhibited strong patterns of endogamy—as in much of Africa—increases in intermarriage rates over time are interpreted as indicators of weakening ethnic identification and the loosening of communal cleavages in society (Bandyopadhyay and Green 2021; Pengl, Roessler and Rueda 2022).

However, taking intermarriage rates at face value may be misleading. Intermarriage rates are a reflection of both *preferences* about who is acceptable to marry (providing insight into inter-group social distance) and also *opportunities* to interact with people from other groups, and thus be in a position to consider them as suitable marriage partners (Kalmijn 1998).[2] If we are to interpret intermarriage rates as capturing the former—and thus as an indicator of social cohesion—we must find a way of controlling for the latter. A low rate of interethnic marriage has quite different implications in an ethnically homogeneous setting, where it would be challenging to even find a marriage partner from another ethnic group, and in an ethnically diverse environment, where it likely reflects conscious decisions by community members not to cross group boundaries. Controlling for opportunity is especially important for interpreting changes in intermarriage rates over time, as ongoing processes of urbanization and internal migration mean that changes in exposure to out-group members are occurring alongside the changes in preferences that researchers are hoping to capture.

Sophisticated analyses of inter-ethnic marriage in Africa recognize this issue.[3] For example, Crespin-Boucaud (2020) presents interethnic marriage rates in 15 African countries alongside measures of

[1] This Africa-focused work is joined by an even larger body of research exploring the implications of interethnic or interracial marriage on social integration in the US (Fryer 2007; Qian and Lichter 2007; Furtado 2012; Bohra-Mishra and Massey 2015; Goldman, Gracie and Porter 2024) and other settings (Monden and Smits 2005; Kalmijn and van Tubergen 2006; Smits 2010; Ahuja and Ostermann 2016; Bazzi et al. 2019).

[2] Kalmijn (1998) also emphasizes a third factor, the interference of "third parties" in the selection process. In the African setting, this might include religious proscriptions (in instances where ethnic and religious memberships overlap), norms about whether families are involved in the process of selecting marriage partners (Takyi et al. 2003), and the role of traditional elders in formalizing marriages (Epstein 1958; Ault 1983).

[3] For an early example, see Mitchell (1957). The issue is also well recognized in the broader literature on assortative mating (Lichter and Qian 2019).

country-level ethnic diversity to benchmark the rates we would have expected to observe under random matching. While this is a step in the right direction, country-level diversity is not a good proxy for the exposure that individuals have to potential out-group spouses in the local marriage markets in which they actually select their partners.[4] Other approaches, such as controlling for education, urban location, or country or group population size, are also imperfect strategies for conditioning on exposure.

In this Note, I introduce a new approach for estimating intermarriage rates controlling for exposure to out-group members in local marriage markets, leveraging data from Zambia, one of the handful of African countries with publicly available census data that includes ethnic identifiers for household members. I define local marriage markets at the level of the country's 150 electoral constituencies, the smallest administrative unit into which census respondents can be grouped. The large sample size that makes such disaggregation possible (N=151,517 couples after censoring the sample as described below) represents a significant improvement over prior studies, which estimate interethnic marriage rates in African countries using data from Demographic and Health Surveys (DHS), whose country-level samples, even when stacked across multiple surveys, are much smaller, ranging from an average of 2,840 per country in Bandyopadhyay and Green (2021) to an average of 6,474 per country in Crespin-Boucaud (2020).[5] The findings reveal significant differences across estimates that do and do not control for exposure to out-group members and, among estimates that do control for exposure, across those that characterize the marriage market at the national versus local level. The findings also point to important substantive implications for understanding changes in African societies over time.

Why Controlling for Exposure is Crucial

To illustrate the importance of controlling for exposure, consider the observed marriage pairings in the three hypothetical communities depicted in Table 1. All three communities contain six married couples and, in all three, 2 of the 6 couples are married to members of other ethnic groups, generating an identical observed intermarriage rate of 33 percent (see Table 2). However, community 1 is more diverse than communities 2 and 3, implying that the average individual seeking to marry within his/her own ethnic group in community 1 would need to look harder to find a co-ethnic partner than the average individual in the two other communities. If couples were paired at random, the likelihood of being

[4] Accounting for community diversity also does not account for *group-level* variation in the likelihood of randomly matching with a non-coethnic partner, as discussed below.

[5] The Zambia samples in Bandyopadhyay and Green (2021) and Crespin-Boucaud (2020) are 7,198 and 10,711, respectively.

matched to a partner from a different group would be 72 percent in community 1. In communities 2 and 3, it would be just 50 percent—roughly a third lower.

Table 1: Patterns of Interethnic Marriage in Three Hypothetical Communities

		Community 1					Community 2					Community 3				
		wife's group					wife's group					wife's group				
		A	B	C	D	total	A	B	C	D	total	A	B	C	D	total
husband's group	A	2	0	0	0	2	3	0	1	0	4	4	0	0	0	4
	B	0	1	1	0	2	0	1	0	0	1	0	0	1	0	1
	C	0	1	0	0	1	1	0	0	0	1	0	1	0	0	1
	D	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
	total	2	2	1	1		4	1	1	0		4	1	1	0	

A straightforward way to account for these differences is to control for each community's diversity when calculating the intermarriage rate. I do this in row 3 of Table 2 by dividing the observed intermarriage rate by the community's ethnic heterogeneity (ELF), as shown here:

$$I_{community\ diversity} = \frac{r_c}{1 - \sum_g p_{gc}^2} \tag{1}$$

where r_c is the observed intermarriage rate in community c and p_{gc} is the proportion of group g in community c.

Table 2: Intermarriage Rates and Ethnic Diversity in the Three Communities

	Community 1	Community 2	Community 3
Observed intermarriage rate	0.33	0.33	0.33
Ethnic diversity	0.72	0.50	0.50
Controlling for exposure (community diversity)	0.46	0.67	0.67
Controlling for exposure (individual's own group share)	0.45	0.70	0.40

By adjusting the intermarriage rate to account for opportunities to marry outside of one's own ethnic group, this approach better captures the component of intermarriage that reflects intergroup social distance, the outcome that most researchers who invoke intermarriage rates are aiming to capture. In this particular example, the adjustment has the effect of up-weighting the observed intermarriage rate in the communities 2 and 3 to reflect the fact that, given their greater homogeneity (and counterfactual lower intermarriage rate under random matching), achieving the same 33 percent observed rate of out-marriage implies a greater acceptance of out-group unions than in community 1.

Controlling for exposure in this way may still be misleading, however. To see why, consider the differences between communities 2 and 3. These two communities have identical levels of ethnic diversity but group-level patterns of intermarriage vary. In community 2, members of the largest

group, group A, out-marry 25 percent of the time, slightly lower than the 33 percent rate we would have expected to observe if they had chosen marriage partners at random. By contrast, in community 3, members of group A do not out-marry at all, suggesting a conscious reluctance to select partners from other ethnic groups. This reluctance, combined with group A's large size in the community, is reflected in community 3's lower intermarriage rate, as reported in row 4 of Table 2, where the estimated intermarriage rates control for exposure at the *group* level by taking the sum of the ratio of observed and expected out-marriage rates of each group in the community, weighted by the group's size:

$$I_{group\ share} = \sum_{g} \left(\frac{r_{gc}}{(1 - p_{gc})} \cdot p_{gc} \right) \tag{2}$$

where r_{gc} is the observed intermarriage rate for ethnic group g in community c and p_{gc} is the proportion of individuals from group g in community c. If we seek to use intermarriage rates to make inferences about people's willingness to marry across group lines independent of their exposure to potential out-group marriage partners—and if we have the data to make it possible to use this group-level approach—this is the approach we should use.

Data and Analysis

I illustrate this approach using pooled data from the publicly available 10% samples of the 2000 and 2010 Zambian censuses (Ruggles et al. 2024).[6] I begin by identifying households containing married couples, and then code each couple for whether the husband and wife are from different ethnic groups.[7] In the analysis that follows, I focus on the husband, estimating counterfactual out-marriage rates based on the size of the husband's ethnic group in the local marriage market.

In recognition of the multi-dimensional nature of ethnic identity in Africa (Scarritt and Mozaffar 1999; Posner 2004) and in keeping with other analyses of ethnic intermarriage in the region (Bandyopadhyay and Green 2021), I code ethnic group membership in two different ways: in terms of the ethnic categories used in the Zambian census (N=61) and in terms of the broader ethnic blocks into which these ethnic census categories nest (N=5).[8] The former captures something close to what Pos-

[6]The 1990 Zambian census sample is also available, but it cannot be used for the present analysis because it does not identify the constituency in which respondents are enumerated, which is necessary for estimating constituency-level intermarriage rates.

[7]I drop the 0.6% of marriages that are polygamous because intermarriage decisions regarding first and second wives may not be independent of one another. See Table A1 in the Online Appendix for a replication of the main results that includes the marriages between husbands and first-order wives in polygamous households.

[8]The mapping of census categories onto ethnic blocks is provided in Table A4 in the Online Appendix.

ner (2005), in his account of ethnic politics in Zambia, terms "tribal" identities, whereas the latter captures membership in one of the country's five main politically relevant linguistic/regional groups. Intermarriage across ethnic census categories is mechanically higher than across ethnic blocks because many intermarriages of the former kind are defined as within-group marriages under the latter categorization. More important, intermarriage across ethnic blocks indicates a much greater crossing of social boundaries, and thus carries more weight as an indicator of social cohesion.[9] Social cleavages defined by ethnic blocks have also been found to be more salient for outcomes like local public goods provision than cleavages defined by narrower ethnic census categories (Gershman and Rivera 2018). For these reasons, most studies in Africa that employ intermarriage rates as indicators of the rigidity of ethnic boundaries (Crespin-Boucaud 2020; Demarest and Haer 2022; Kyei 2022) aggregate ethnic groups into broader categories before calculating rates of interethnic marriage.[10]

As in other studies, I exploit the data to document variation in interethnic marriage across space and time. Estimating intermarriage rates at the subnational level, whether across regions or across urban and rural areas, is important because countrywide interethnic marriage rates mask substantial local variation. The countrywide intermarriage statistic that is claimed to summarize the character of intergroup relations may not be particularly informative about the rigidity of ethnic boundaries in some areas of the country, and these may be precisely the areas that are responsible for the outcome—conflict, ethnic voting, etc.—that the interethnic marriage rate is invoked to explain.[11]

For the analyses of change over time, I calculate average intermarriage rates by decade based on the wife's year of first marriage, taking advantage of a census question that collects this information.[12] Coding survey respondents as urban or rural is complicated by the fact that urban/rural designations were only recorded in the 2000 census. I therefore code as urban all 2010 respondents who were

[9]For example, in Zambia, a marriage between a Chewa and an Nsenga—two groups that are both part of the broader Nyanja language community and that share many cultural similarities—would constitute a much weaker instance of cultural boundary-crossing than a marriage between an Chewa and a Lozi—two groups from opposite sides of the country with few historical or cultural ties. Viewed in terms of ethnic census categories, both of these couples would be coded as intermarried, but only the Chewa-Lozi couple would be coded as intermarried when defined in terms of ethnic blocks.

[10]Pengl, Roessler and Rueda (2022), and also Crespin-Boucaud (2020), do something analogous by presenting intermarriage rates that account for the linguistic distances between the ethnic groups to which intermarried couples belong—the idea being that marriages across groups at greater linguistic distances are more meaningful indicators of the breakdown of social boundaries.

[11]Recognizing this, Demarest and Haer (2022) calculate intermarriage rates at the level of subnational regions in their analysis of the relationship between intermarriage rates and conflict incidence.

[12]Note that this approach assumes that the observed marriage is the woman's first, which could lead to an undercounting of intermarriages if such unions are more likely to collapse and be replaced by endogamous marriages. In Table A2 in the Online Appendix, I replicate the main results, limiting the sample to couples where the woman was ≤ 29 years old at the time of the census. This has the effect of limiting the analysis to marriages that took place within ten years of the time when most women first marry, thus significantly reducing the likelihood that the marriages analyzed are not first marriages. In the handful of cases where the wife's year of first marriage is not reported, I date the year of marriage for each couple by subtracting 19 from the wife's age and then subtracting that number from the census year, taking advantage of the fact that the age of first marriage for women in Zambia is tightly grouped at 19 years (SD=3.12).

enumerated in a constituency whose population was more than 50 percent urban in the 2000 census, with all others coded as rural.[13] Creating aggregations by province is complicated by the fact that an additional province, Muchinga, was added between 2000 and 2010. Results for Muchinga are calculated for the 2000 census sample by building the new province from its component districts and subtracting those districts from the provinces from which they were shifted.

A significant departure from prior analyses of intermarriage in Africa is the calculation of intermarriage rates in local marriage markets, exploiting information about the demography of the marriage market (in the present analysis, the constituency) in which individuals find their marriage partners.[14] Because each individual has a counterfactual expectation for outmarriage based on their group's population share in the constituency, all estimates reported in this Note are built up from individual-level estimates, weighted by the constituency's proportion of the unit in question (the country as a whole, urban areas, the population of marriageable-aged people in the 1970s, etc.).

While constituencies are in some cases imperfect proxies for actual local marriage markets, they offer a much better approximation of the local environment in which people choose their partners than might be suggested by the demographics of the country as a whole (the approach adopted in Crespin-Boucaud (2020)). It is well established that marriage likelihoods are inversely related to the distance between the residences of would-be partners (Peach and Mitchell 1988). Reducing the size of the unit that defines the marriage market therefore increases the likelihood that it captures the ethnic demographics of the actual population of prospective partners.

Consider, for example, that while ethnic fractionalization in Zambia is 0.28 countrywide when measured in terms of ethnic blocks, it ranges from 0 in the most homogeneous quintile of constituencies to 0.76 in the most heterogeneous quintile.[15] This huge cross-constituency variation in ethnic diversity—which drives the "opportunity" component of interethnic marriage—is ignored if we take no account of the local conditions in which people select their marriage partners. Even if electoral constituencies do not perfectly map onto marriage markets, controlling for local demography defined at this level almost certainly improves our estimates of the component of marriage choice that reflects preferences about who is acceptable to marry rather than simply opportunities for marrying across group lines.[16]

[13] This approach is made possible by the fact that constituency boundaries were unchanged between 2000 and 2010. Results broken down by decile of urbanization are reported in Figures A2 and A1 in the Online Appendix.

[14] Although novel in Africa, research estimating intermarriage rates in local marriage markets is relatively common in the United States. For example, Lichter, LeClere and McLaughlin (1991) estimate intermarriage rates at the level of U.S. Census-designated labor market areas; Qian and Lichter (2018) use U.S. metropolitan areas; and Lichter, Price and Swigert (2020) use census-defined public use microdata areas (PUMAs).

[15] The analogous figures when fractionalization is measured in terms of ethnic census categories are 0.56 countrywide and 0.06 and 0.91 in the most homogeneous and heterogeneous quintiles of constituencies, respectively.

[16] The smallest quintile of constituencies has a median area of 328 sq km, whereas the largest has a median area of just over 11,000 sq km. While constituencies toward the latter end of the spectrum are no doubt too large to constitute

A challenge with using census data to estimate intermarriage rates in local marriage markets is that the location at which couples are enumerated may be different from the location in which they met one another and were married. To address this issue, I limit the analysis to couples whose estimated year of marriage occurred after they were first resident in their current location. Adopting this rule reduces the sample size by roughly half, but it raises confidence that the marriage market employed in the analysis reflects the marriage market that census respondents inhabited when they chose their spouses.[17]

Intermarriage in Zambia

Estimates of intermarriage rates in Zambia based on the procedures just described are presented in Table 3. I report intermarriage rates for the country as a whole and broken down by decade, urban/rural location, and province. For each sub-sample, I present both the observed and exposure-adjusted intermarriage rates, with intermarriages across both ethnic census categories and ethnic blocks. As expected, intermarriage rates are increasing over time and are higher in urban areas and predominantly urban provinces like Lusaka and Copperbelt. Observed intermarriage rates are also higher, as anticipated, when groups are defined in terms of ethnic census categories than in terms of ethnic blocks. Although the exposure-adjusted intermarriage rates are consistently higher than the observed intermarriage rates, this difference is not meaningful in and of itself: it is a mechanical artifact of dividing the observed rate by a value < 1. This said, the exposure-adjusted rates, taken on their own, are highly informative, as they convey how much the observed intermarriage rate diverges from the expected intermarriage rate under random matching.[18] An exposure-adjusted estimate of 1 would indicate perfectly neutral preferences across marrying within versus outside of one's group, whereas values increasingly less than 1 suggest greater reluctance to marry across group lines.[19]

From a substantive standpoint, the real payoff from controlling for exposure comes when we explore the variation in interethnic marriage rates across time and space. The usual story about intermarriage in Africa is that it is rising over time and that this indicates an increasing porosity of ethnic boundaries (Bandyopadhyay and Green 2021; Pengl, Roessler and Rueda 2022). However, the patterns in Figure 1, which plots changes from the 1950s to the 2000s in observed and expected intermarriage rates across

actual marriage markets, constituencies of that size are almost entirely located in homogeneous rural areas, where the exact boundaries of the marriage market matters less for exposure to non-coethnics.

[17] Estimates based on the full sample are provided for reference in Table A3 in the Online Appendix.

[18] The expected intermarriage rate is calculated for each male of marriageable age (i.e., aged 18-65 years) as the proportion of marriageable-aged women in the constituency and decade belonging to other ethnic blocks (i.e., $1 - p_{gc}$, the denominator in equation 2.)

[19] Values greater than 1 would suggest positive preferences for outmarriage.

Table 3: Intermarriage Rates in Zambia

	By ethnic census category		By ethnic block		
	Observed rate	Controlling for exposure	Observed rate	Controlling for exposure	*N* (couples)
	(1)	(2)	(3)	(4)	(5)
Countrywide	32.7	45.4	19.2	50.3	151,517
1951-60	14.7	21.1	7.3	24.3	3,602
1961-70	18.6	27.2	9.4	32.0	7,706
1971-80	24.2	33.4	12.0	31.4	14,829
1981-90	29.0	40.4	15.6	38.6	29,655
1991-00	34.2	48.6	19.4	51.6	64,590
2001-10	36.9	50.4	23.4	60.3	31,135
Urban areas	57.5	65.5	37.3	63.1	44,287
Rural areas	21.0	35.9	10.6	44.1	107,230
Central	40.9	48.4	27.6	49.2	11,805
Copperbelt	55.2	63.2	33.2	62.7	19,622
Eastern	19.7	41.8	4.2	49.7	22,375
Luapula	14.2	20.8	1.3	41.5	14,086
Lusaka	63.2	68.9	45.0	65.2	22,644
Muchinga	13.5	31.5	6.0	38.3	9,820
North-Western	22.7	39.5	7.4	50.7	8,789
Northern	13.0	31.2	2.0	42.2	14,199
Southern	17.2	36.9	15.5	41.6	17,384
Western	26.8	36.1	16.0	38.3	10,793

Note: Sample is restricted to couples where both the husband and wife were resident in their current location prior to their marriage. Intermarriage rates in columns 2 and 4 are calculated as described in equation 2. Sample sizes are from the ethnic census category analyses. Ethnic block analyses occasionally have slightly smaller sample sizes due to missingness of potential marriage partners in out-group blocks.

ethnic blocks, suggest a more nuanced account.[20] In urban areas (right panel), the observed and expected intermarriage rates increase roughly in parallel through the 1970s, suggesting that changes in intermarriage patterns during this period were driven primarily by changes in exposure: as more Zambians migrated from rural to urban areas, urban marriage markets offered greater opportunities for individuals to meet potential partners from other ethnic groups. However, beginning in the 1980s, the expected intermarriage rate flattens out while the observed rate begins climbing more steeply, indicating that something beyond increasing exposure to out-group members—changing *preferences*—is driving the increase in intermarriage. This shift in preferences is what most scholars take the increase in intermarriage rates to imply. But Figure 1 suggests that the change in preferences only began to emerge in urban areas in the 1980s.

[20]Parallel figures showing changes in expected and observed intermarriage rates by ethnic census categories are provided in the Online Appendix.

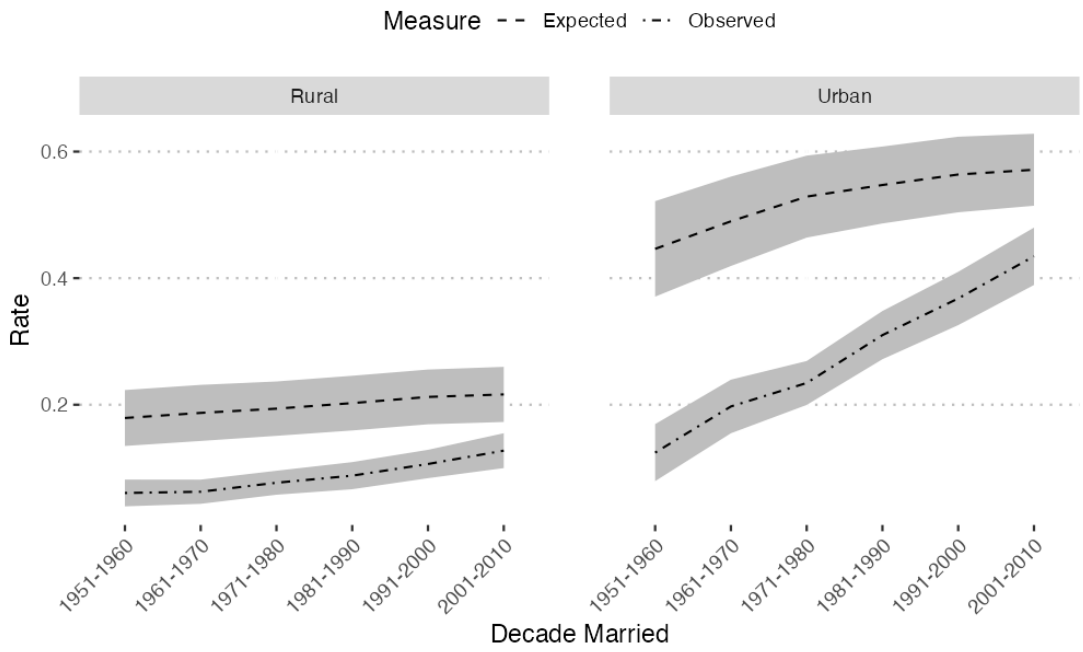

Figure 1: Changes in Observed and Expected Intermarriage Rates, by Urban/Rural Location (Ethnic Block)

When we look at rural Zambia (left panel), we see something different. The observed and expected intermarriage rates move closely in parallel through the entire time series, with the gap between the lines narrowing only slightly beginning in the 1990s. These patterns suggest that the (more modest) increase in intermarriage rates in rural Zambia has been driven mainly by exposure rather than by increasing openness to marrying outside of one's group, with (modest) evidence for changing preferences appearing a full decade after it emerges in urban areas.

Figure 2 plots the change over time in the exposure-adjusted intermarriage rates by ethnic block, as introduced in this Note.[21] I calculate these rates both in terms of the diversity of the marriage market in which couples were located at the time they were married (as in equation 1, solid line) and in terms of the share of the marriage market belonging to ethnic groups other than the individual's own (as in equation 2, dashed line). The trends capture changes in the component of intermarriage rooted in preferences about marrying outside of one's group—theoretically, the net of the two curves presented in Figure 1. The steeper inflection points in the rural panel reflects the fact that marrying outside of one's group is more challenging in rural than urban areas, so even small changes in outmarriage rates generate big effects once we control for exposure. The narrower gap between the two lines in the urban panel likely reflects the greater willingness of members of larger groups to cross ethnic boundaries in urban marriage markets (as illustrated in the difference between communities 2 and 3 in the example

[21] Parallel figures for ethnic census categories are provided in the Online Appendix.

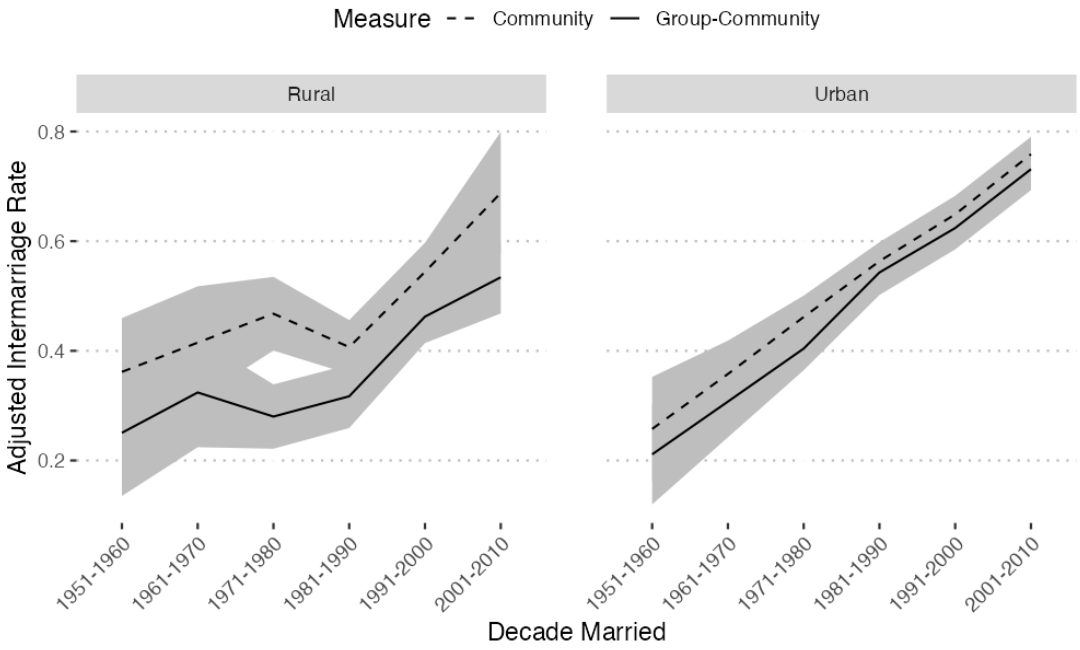

Figure 2: Changes in Exposure-Adjusted Intermarriage Rates, by Urban/Rural Location (Ethnic Block)

provided earlier)—something that the group-level exposure measure captures but the community-level measure does not.[22]

Conclusion

Intermarriage is a product of both preferences and opportunity, but it is often treated as indicating only the former. In this Note, I have introduced and applied a new methodology for calculating intermarriage rates that, by controlling for exposure to non-coethnics in local marriage markets, allows for the disentangling of these two drivers of interethnic marriage. The approach makes three distinct advances over prior efforts to measure interethnic marriage in African countries: it controls for exposure to outgroup members in local marriage markets rather than in the country as a whole; it estimates exposure at the group level, calculating the expected counterfactual outmarriage rate for each group in each marriage market; and it employs much richer data than prior studies, which makes the first two innovations possible. The contribution made by each of these advances is illustrated in Table 4, which presents intermarriage rates at the block level, estimated using different data sources and approaches.

[22]Note that the ability to capture such group-level variation in out-marriage rates is related to the diversity of the marriage market. When the marriage market is diverse (as in urban areas), the largest groups are small vis-a-vis the population more generally, so the weight they carry if their members out-marry at lower rates than members of smaller groups is less. Hence, increasing marriage market diversity is associated with a reduction in the difference between the two adjustment calculations (as illustrated in Appendix Figure **??**). This is why the two lines in the right (urban, heterogeneous) panel of Figure 2 are closer together than the lines in the left (rural, less heterogeneous) panel.

Table 4: Intermarriage Rates in Zambia, by Data Source and Approach (Ethnic Block)

	DHS	Census	Census	Census	Census
	(1)	(2)	(3)	(4)	(5)
Zambia	26.5	16.5	22.7	59.5	50.3
Urban Zambia	35.8	36.8	54.7	65.9	63.1
Rural Zambia	21.1	8.2	10.9	56.5	44.1
Marriage market	country	country	country	constituency	constituency
Controlling for exposure	no	no	ELF	ELF	group share
N (Zambia)	8,799	150,222	150,222	152,000	150,793

Note: Intermarriage rates presented in column 1 are calculated from the the pooled DHS samples from 1996, 2001, and 2013 using the same procedures as employed elsewhere in this Note. Figures in columns 2-4 are from the the pooled 10% census samples from 2000 and 2010. N refers to the number of married couples included in the countrywide estimates presented in the first row.

A comparison between columns 1 and 2 demonstrates the benefits of using stacked 10% census samples rather than data from DHS surveys. In addition to the nearly 20x larger sample size, the census data is truly representative, whereas DHS data, while designed to be nationally representative, is not necessarily representative of all ethnic groups or the demography of sub-national units. The shift to the more comprehensive, census-based data source is associated with quite different intermarriage estimates, especially in rural areas.

Column 3 repeats the analysis from column 2, but controls for outgroup exposure at the country level by dividing the observed intermarriage rate by country-level ethnic diversity, akin to the strategy employed in Crespin-Boucaud (2020). Column 4 does the same thing but controls for exposure at a level that more closely approximates local marriage markets, estimating exposure-adjusted intermarriage rates at the constituency-level by dividing observed intermarriage rates in each constituency by constituency-level diversity before aggregating upward to the desired unit (as in equation 1). The impact of defining the marriage market more locally is immediately apparent: the estimates in column 4 are quite different from those in column 3. The larger estimates in the analyses using the constituency-level exposure adjustments reflect the fact that exposure to outgroup members in the actual locations where people live is significantly lower than might be suggested by the demography of the country as a whole, resulting in higher adjusted intermarriage rates once one accounts for this lower exposure.[23]

Column 5 then recalculates the intermarriage rates presented in column 4, leveraging group-level rather than marriage market-level counterfactuals for outmarriage (as in equation 2). Once again, we see that changing how we control for exposure is associated with changes in the estimates we generate,

[23]This is shown clearly in Online Appendix Figure A5, which plots exposure levels at the national and constituency levels, alongside observed and adjusted intermarriage rates. The difference in the observed and exposure-adjusted estimates in the national versus constituency panels is largely driven by rural marriage markets, which tend to be much more ethnically homogeneous than rural Zambia taken as a whole. Hence the very large differences between the rural estimates (and also the countrywide estimates, which include rural areas) in columns 3 and 4 of Table 4.

with the slightly lower values in column 5 relative to column 4 reflecting the tendency for members of larger groups to outmarry less frequently than members of smaller groups.

Implementing the approach described in this Note requires more data than are often available to researchers. But where the appropriate data are available—for example, in countries with publicly accessible census data that includes information about ethnic group membership—employing the methodology outlined here will provide a superior measure of social cohesion than prior measures that rely on observed intermarriage rates, opening the door to new insights about how intergroup relations vary across space and change over time.

References

Ahuja, Amit and Susan L. Ostermann. 2016. "Crossing caste boundaries in the modern Indian marriage market." *Studies in Comparative International Development* 51(3):365–387.

Ault, James. 1983. "Making 'modern' marriage 'traditional'." *Theory and Society* 12(2):187–210.

Bandyopadhyay, Sanghamitra and Elliott Green. 2021. "Explaining inter-ethnic marriage in Sub-Saharan Africa." *Journal of International Development* 33(4):627–643.

Bazzi, Samuel, Arya Gadueh, Alexander D. Rothenberg and Maisy Wong. 2019. "Unity in diversity? How intergroup contact can foster nation building." *American Economic Review* 109(11):3978–4025.

Bohra-Mishra, Pratikshya and Douglas S. Massey. 2015. "Intermarriage among new immigrants in the USA." *Ethnic and Racial Studies* 38(5):734–758.

Bratton, Michael and Mwangi S. Kimenyi. 2008. "Voting in Kenya: Putting ethnicity in perspective." *Journal of Eastern African Studies* 2(2):272–289.

Crespin-Boucaud, Juliette. 2020. "Interethnic and interfaith marriages in sub-Saharan Africa." *World Development* 125:1–19.

Demarest, Leila and Roos Haer. 2022. "A perfect match? The dampening effect of interethnic marriage on armed conflict in Africa." *Conflict Management and Peace Science* 39(6):686–705.

Dulani, Boniface, Adam S. Harris, Jeremy Horowitz and Happy Kayuni. 2022. "Electoral preferences among multiethnic voters in Africa." *Comparative Political Studies* 54(2):280–311.

Epstein, A.L. 1958. *Politics in an urban African community.* Manchester University Press.

Fryer, Ronald G. 2007. "Guess who's been coming to dinner? Trends in interracial marriage over the 20th century." *Journal of Economic Perspectives* 21(2):71–90.

Furtado, Delia. 2012. "Human capital and interethnic marriage decisions." *Economic Inquiry* 50(1):82–93.

Gershman, Boris and Diego Rivera. 2018. "Subnational diversity in Sub-Saharan Africa: Insights from a new dataset." *Journal of Development Economics* 133:231–263.

Goldman, Benjamin, Jamie Gracie and Sonya R. Porter. 2024. "Who marries whom? The role of segregation by race and class." Unpublished paper.

Hechter, Michael. 1978. "Group formation and the cultural division of labor." *American Journal of Sociology* 84(2):293–318.

Kalmijn, Matthijs. 1998. "Intermarriage and homogamy: Causes, patterns, trends." *Annual Review of Sociology* 24:395–421.

Kalmijn, Matthijs and Frank van Tubergen. 2006. "Ethnic intermarriage in the Netherlands: Confirmations and refutations of accepted insights." *European Journal of Population* 22(4):371–397.

Kim, Hye-Sung and Jeremy Horowitz. 2022. "Rejecting ethnic pandering in urban Africa: A survey experiment on voter preferences in Nairobi, Kenya." *Political Research Quarterly* 75(4):1240–1254.

Kyei, Pearl S. 2022. "Patterns of interethnic marriage in Ghana: 2000-2010." *Ghana Journal of Geography* 14(1):122–137.

LeBas, Adrienne. 2020. "Who trusts? Ethnicity, integration, and attitudes toward elected officials in urban Nigeria." *Comparative Political Studies* 53(10-11):1738–1766.

Lichter, Daniel T., Felicia B. LeClere and Diane K. McLaughlin. 1991. "Local marriage markets and the marital behavior of Black and White women." *American Journal of Sociology* 96(4):843–867.

Lichter, Daniel T., Joseph P. Price and Jeferey M. Swigert. 2020. "Mismatches in the marriage market." *Journal of Marriage and Family* 82(2):796–809.

Lichter, Daniel T. and Zhenchao Qian. 2019. The study of assortative mating: Theory, data, and analysis. In *Analytical Family Demography*, ed. Robert Schoen. Springer pp. 303–337.

Mitchell, J. Clyde. 1957. "Aspects of African Marriage on the Copperbelt of Northern Rhodesia." *Rhodes-Livingstone Journal* 22:1–30.

Monden, Christianan W.S. and Jeroen Smits. 2005. "Ethnic intermarriage in times of social change: The case of Latvia." *Demography* 42(2):323–345.

Peach, C. and J. Clyde Mitchell. 1988. "Marriage distance and ethnicity". In *Human mating patterns*, ed. C. Mascie-Taylor and A. Boyce. Cambridge University Press pp. 31–46.

Pengl, Yannick I., Philip Roessler and Valeria Rueda. 2022. "Cash crops, print technologies, and the politicization of ethnicity in Africa." *American Political Science Review* 116(1):181–199.

Posner, Daniel N. 2004. "Measuring ethnic fractionalization in Africa." *American Journal of Political Science* 48(4):849–863.

Posner, Daniel N. 2005. *Institutions and ethnic politics in Africa.* New York: Cambridge University Press.

Qian, Zhenchao and Daniel T. Lichter. 2007. "Social boundaries and marital assimilation: Interpreting trends in racial and ethnic intermarriage." *American Sociological Review* 72(1):68–94.

Qian, Zhenchao and Daniel T. Lichter. 2018. "Marriage markets and intermarriage: Exchange in first marriages and remarriages." *Demography* 55:849–975.

Ruggles, Steven, Lara Cleveland, Rodrigo Lovaton, Sula Surkar, Matthew Sobek, Derek Burk, Dan Rhrlich, Quinn Heimann and Jane Lee. 2024. "Integrated Public Use Microdata Series, International Version 7.5 [dataset].".

Scarritt, James R. and Shaheen Mozaffar. 1999. "The specification of ethnic cleavages and ethnopolitical groups for the analysis of democratic competition in contemporary Africa." *Nationalism and Ethnic Politics* 5(1):82–117.

Smits, Jeroen. 2010. "Ethnic intermarriage and social cohesion: What can we learn from Yugoslavia?" *Social Indicators Research* 96(4):417–432.

Takyi, Baffour K., Nancy B. Miller, Gay C. Kitson and Yaw Oheneba-Sakyi. 2003. "Marital choice in Sub-Saharan Africa: Comparing structural and cultural influences in contemporary Ghana." *Comparative Sociology* 2(2):375–391.

Online Appendix

The online appendix presents additional analyses and robustness tests, as referred to in the main text.

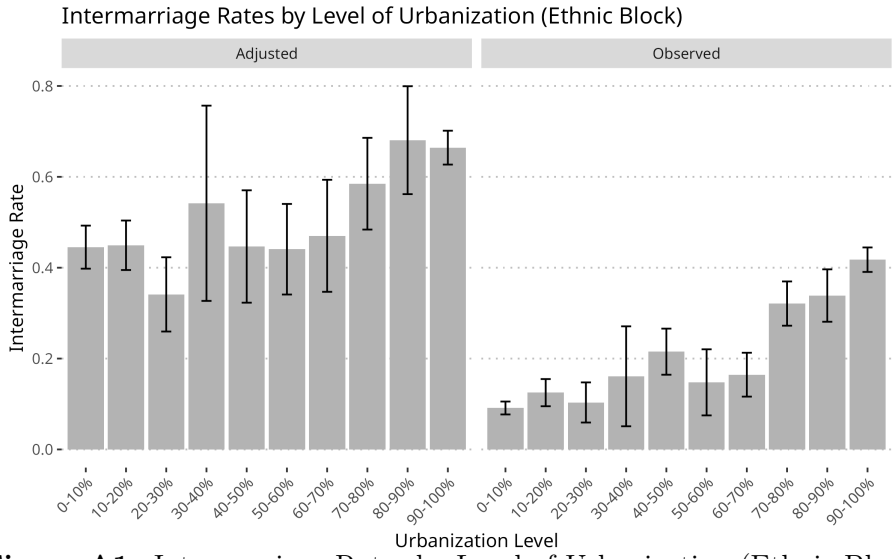

Figure A1: Intermarriage Rates by Level of Urbanization (Ethnic Block)

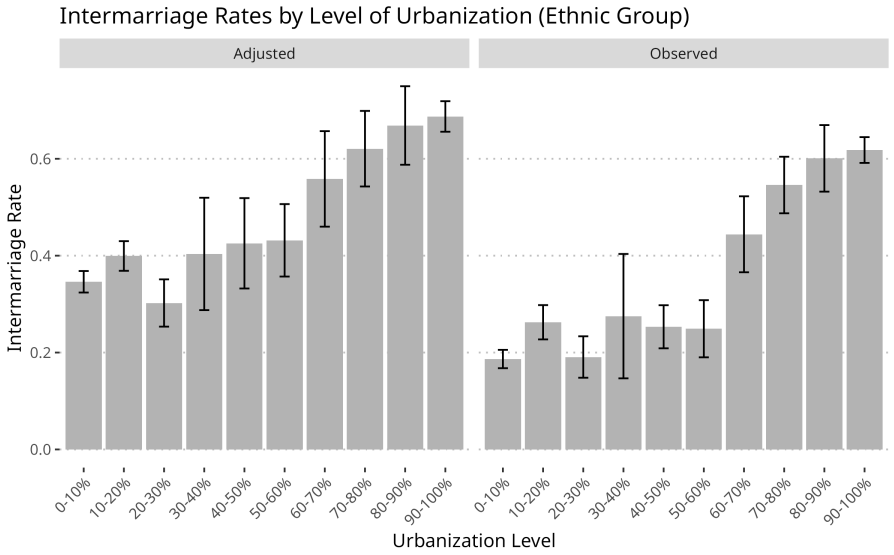

Figure A2: Intermarriage Rates by Level of Urbanization (Ethnic Census Category)

Table A1: Intermarriage Rates in Zambia, including Polygamous Marriages

	By ethnic census category		By ethnic block		
	Observed rate	Controlling for exposure	Observed rate	Controlling for exposure	N (couples)
	(1)	(2)	(3)	(4)	(5)
Countrywide	32.7	45.5	19.2	50.5	152,739
1951-60	14.8	21.2	7.3	24.3	3,613
1961-70	18.5	27.1	9.4	31.9	7,734
1971-80	24.2	33.4	12.0	31.4	14,887
1981-90	28.9	40.3	15.6	38.6	29,786
1991-00	34.2	48.5	19.4	51.6	64,969
2001-10	36.9	50.7	23.4	60.8	31,750
Urban areas	57.4	65.5	37.3	63.1	44,596
Rural areas	21.0	36.1	10.6	44.5	108,143

Note: The main results presented in Table 3 exclude polygamous marriages. This table presents intermarriage rates with this exclusion removed.

Table A2: Intermarriage Rates in Zambia, First Marriages Only

	By ethnic census category		By ethnic block		
	Observed rate	Controlling for exposure	Observed rate	Controlling for exposure	N (couples)
	(1)	(2)	(3)	(4)	(5)
Countrywide	32.6	45.2	19.3	50.2	149,520
1951-60	14.7	21.1	7.3	24.3	3,601
1961-70	18.6	27.2	9.4	32.1	7,690
1971-80	24.2	33.4	12.0	31.5	14,774
1981-90	29.0	40.4	15.6	38.2	29,476
1991-00	34.2	48.5	19.3	51.6	64,147
2001-10	36.8	50.0	23.5	60.3	29,832
Urban areas	57.4	65.4	37.4	63.4	43,486
Rural areas	21.0	35.7	10.6	43.9	106,034

Note: The main analysis presented in Table 3 includes all intact marriages at the time of the census. This may include both first and second (and later) marriages, which may have different likelihoods of being between non-coethnics. This table replicates the results from Table 3 but limits the sample to marriages in which the spouse married at ≤ 29 years old. This increases the likelihood that the analysis is limited to first marriages.

Table A3: Intermarriage Rates in Zambia, Including Couples Married Before Moving to Present Location

	By ethnic census category		**By ethnic block**		
	Observed rate	Controlling for exposure	Observed rate	Controlling for exposure	N (couples)
	(1)	(2)	(3)	(4)	(5)
Countrywide	32.5	48.2	18.8	58.1	283,047
1951-60	15.1	26.3	6.8	33.8	6,558
1961-70	20.2	36.2	9.3	47.3	15,649
1971-80	24.8	40.1	12.1	53.7	31,089
1981-90	28.6	44.1	15.1	52.5	57,873
1991-00	33.7	49.8	19.0	57.0	105,520
2001-10	36.9	52.6	22.9	63.8	66,358
Urban areas	55.4	64.6	35.7	62.6	93,179
Rural areas	21.7	40.4	10.8	55.9	189,868

Note: To ensure that the demography of the local marriage market is indicative of the options couples faced when they chose to marry, the main results presented in Table 3 exclude couples who were married prior to moving to the location in which they were enumerated in the census. This table removes this exclusion.

Table A4: Ethnic Census Categories and Ethnic Blocks

	Ethnic Census Category	Ethnic Block	2000 Percent	2010 Percent
1	Bemba	Bemba	0.177	0.208
2	Lunda (Luapula)	Bemba	0.014	0.010
3	Lala	Bemba	0.032	0.031
4	Bisa	Bemba	0.018	0.016
5	Ushi	Bemba	0.023	0.019
6	Chishinga	Bemba	0.009	0.005
7	Ngumbo	Bemba	0.009	0.007
8	Lamba	Bemba	0.022	0.020
9	Kabende	Bemba	0.005	0.004
10	Tabwa	Bemba	0.008	0.007
11	Swaka	Bemba	0.004	0.004
12	Mukulu	Bemba	0.001	0.000
13	Ambo	Bemba	0.000	0.000
14	Lima	Bemba	0.000	0.000
15	Shila	Bemba	0.002	0.002
16	Unga	Bemba	0.002	0.002
17	Bwile	Bemba	0.004	0.005
18	Luano	Bemba	0.000	0.000
19	Tonga	Tonga	0.127	0.135
20	Lenje	Tonga	0.016	0.016
21	Soli	Tonga	0.008	0.007
22	Ila	Tonga	0.008	0.008
23	Toka-Leya	Tonga	0.005	0.004
24	Sala	Tonga	0.001	0.001
25	Gowa	Nyanja	0.002	0.002
26	Luvale	Northwestern	0.021	0.021
27	Lunda (North-western)	Northwestern	0.028	0.027
28	Mbunda	Northwestern	0.015	0.012
29	Luchazi	Northwestern	0.005	0.004

30	Ndembu	Northwestern	0.001	0.001
31	Mbowe	Northwestern	0.000	0.000
32	Chokwe	Northwestern	0.006	0.005
33	Kaonde subgroup	Northwestern	0.030	0.029
34	Luyana subgroup	Northwestern	0.001	0.000
35	Kwangwa	Lozi	0.004	0.002
36	Kwandi	Lozi	0.001	0.000
37	Koma	Lozi	0.002	0.001
38	Nyengo	Lozi	0.002	0.001
39	Simaa	Lozi	0.001	0.000
40	Mwenyi	Lozi	0.001	0.000
41	Imilangu	Lozi	0.000	0.000
42	Mashi	Lozi	0.004	0.004
43	Lozi	Lozi	0.057	0.059
44	Totela	Lozi	0.001	0.001
45	Subiya	Lozi	0.001	0.000
46	Nkoya	Lozi	0.006	0.005
47	Mashasha	Lozi	0.000	0.000
48	Chewa	Nyanja	0.071	0.074
49	Nsenga	Nyanja	0.054	0.052
50	Ngoni	Nyanja	0.039	0.040
51	Nyanja	Nyanja	0.005	0.004
52	Kunda	Nyanja	0.007	0.007
53	Chikunda	Nyanja	0.003	0.002
54	Lungu	Bemba	0.009	0.008
55	Mambwe	Bemba	0.023	0.025
56	Namwanga	Bemba	0.027	0.027
57	Wina	Lozi	0.000	0.000
58	Tambo	Bemba	0.000	0.000
59	Tumbuka	Nyanja	0.042	0.044
60	Senga	Bemba	0.009	0.009
61	Yombe	Nyanja	0.000	0.000

64	African	Dropped	0.019	0.006
65	American	Dropped	0.000	0.000
66	Asian	Dropped	0.001	0.001
67	European	Dropped	0.001	0.000
68	Other	Dropped	0.003	0.006
99	Unknown	Dropped	0.000	0.008

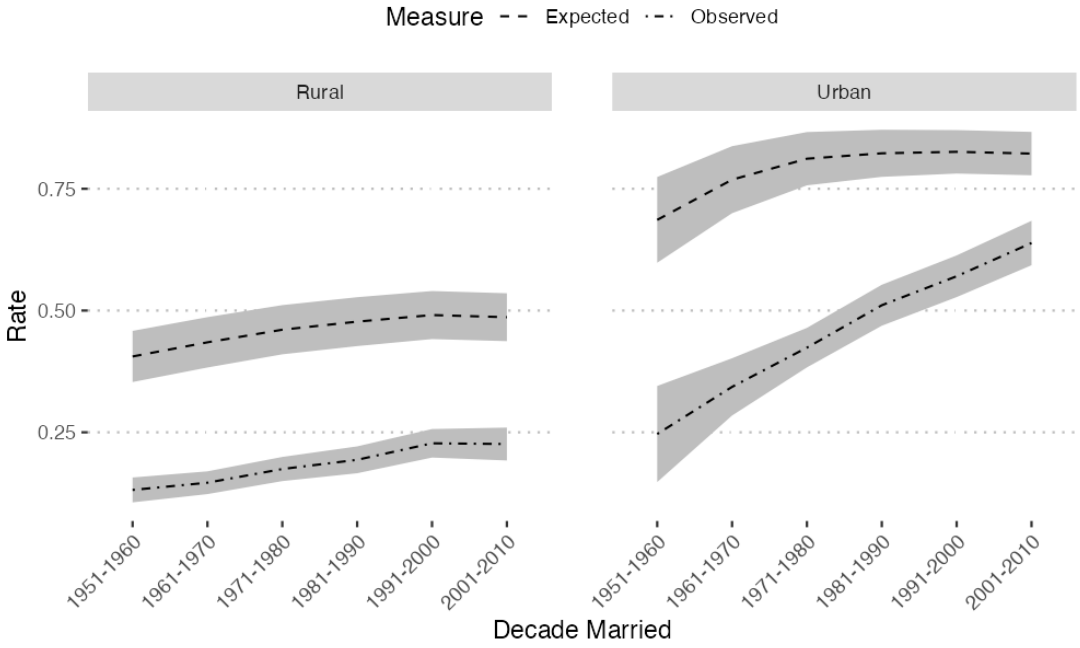

Figure A3: Changes in Observed and Expected Intermarriage Rates, by Urban/Rural Location (Ethnic Census Category)

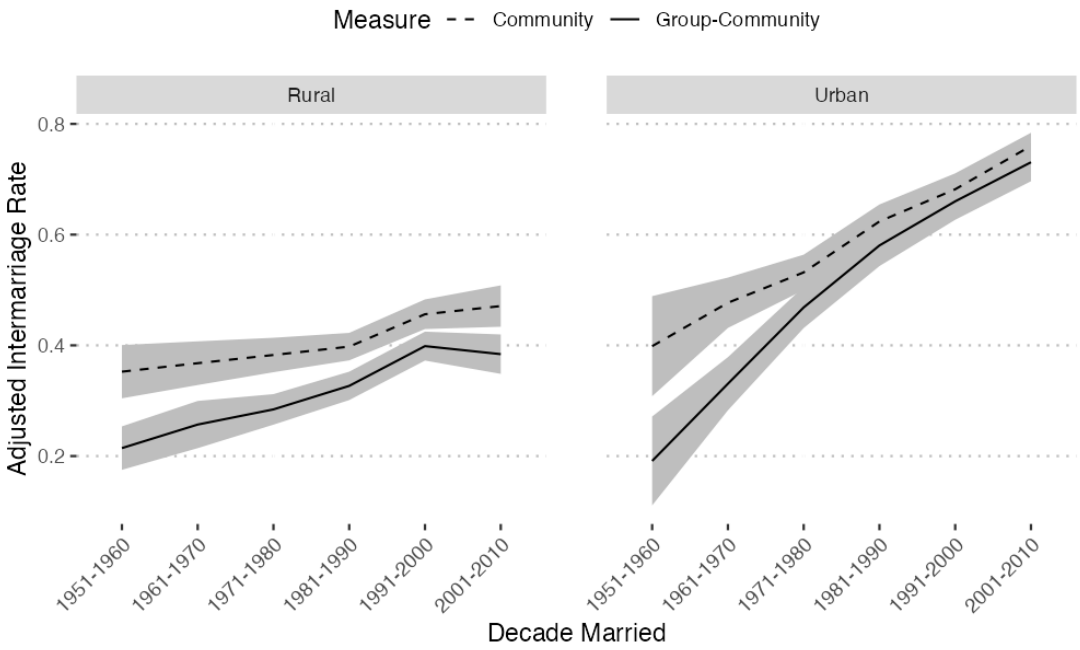

Figure A4: Changes in Exposure-Adjusted Intermarriage Rates, by Urban/Rural Location (Ethnic Census Category)

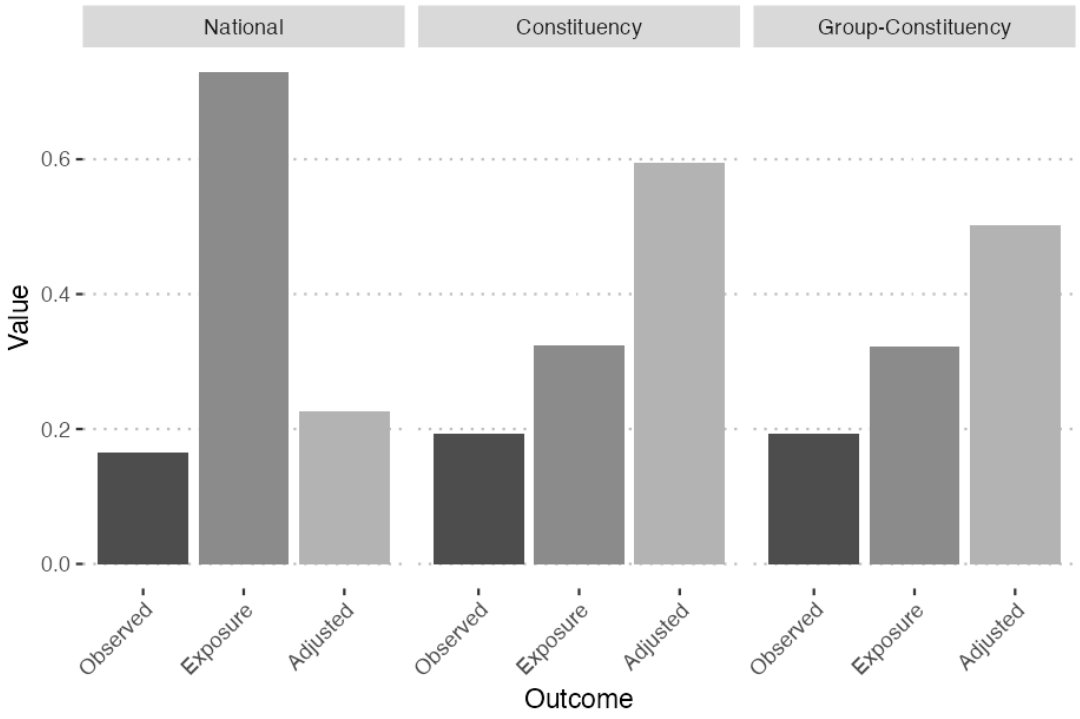

Figure A5: Results by method of calculating marriage market