Ethnicity in Children and Mixed Marriages: Theory and Evidence from China*

Ruixue Jia† and Torsten Persson‡

October 5, 2013

Abstract

This paper provides a framework to link the ethnic choice for children with interethnic marriage. Our model is constructed in such a way to be consistent with four motivating facts for ethnic choices in China, but it also delivers a rich set of auxiliary predictions. The empirical tests on Chinese microdata generally find support for these predictions. In particular, we provide evidence that social norms can crowd in or crowd out material benefits in ethnic choices. We also evaluate how sex ratios affect interethnic marriage patterns and how their effects are strengthened or dampened by ethnic choices for children in mixed marriages.

*We are grateful to the ERC and the Torsten and Ragnar Söderberg Foundations for financial support.

†School of International Relations and Pacific Studies, University of California San Diego, rxjia@ucsd.edu.

‡Institute for International Economic Studies, Stockholm University, torsten.persson@iies.su.se

1 Introduction

How can government policies shape ethnic identification? The answer to this question has important policy implications. For instance, ethnic conflict could be moderated or exacerbated by political institutions that induce ethnic identification (Horowitz, 2000). Broadly, this is the topic of the seminal paper by Bisin and Verdier (2000). Motivated by a large existing sociological literature, they set the task for themselves to theoretically understand why cultural convergence is so slow, even in the US. They model the persistence of ethnic and religious minorities, through a propensity to marry within its own kin and to socialize their children in the same mold.

Yet, there are a number of examples of how groups gradually or suddenly change their identity, often to reap some material benefits. So for example, Botticini and Eckstein (2007) discuss how material incentives have played an important historical role in transitions between Judaism and Christianity. Cassan (2012) shows how groups in Punjab at the turn of the past century adopted a lower caste identity, in order to take advantage of a large land-distribution program. Green (2011) argues that in the process of urbanization, Africans have often assimilated into larger ethnic groups in order to find security and prestige in the difficult urban environment. The Ibo of Nigeria, Jola of Senegal, Duala of Cameroon, Luyia and Mijikenda of Kenya, and Bangala of the DRC can be seen as examples of previously different ethnic groups amalgamating into larger ethnic identities.

Intuitively, such cultural switchovers are likely to involve a tradeoff between extrinsic material benefits and intrinsic costs, where the latter are shaped by existing self-images or social norms. Which way that tradeoff goes is far from clear, however. Indeed, recent theoretical work by Benabou and Tirole (2011) shows that extrinsic incentives to make a certain choice may either be crowded out or crowded in by intrinsic incentives. Most of the existing literature focuses on the impact of government ethnic policies on choices by a single generation. However, it is easy to imagine that ethnic choices also have intergenerational effects. For instance, which ethnicity a mixed couple expect to transmit to their children, is likely to affect the marriage decision.

China is an interesting testing ground for investigating government ethnic policies and family choices. It is a multiethnic society with 55 officially recognized ethnic groups, beyond the dominant Han group. At the same time, China is a relatively homogenous society despite some ethnic tensions,

such as occasional riots in Tibet and Xijiang. The national and provincial governments have instituted some policies, which are similar to "affirmative action" for US minorities. Moreover, mixed ethnic couples are free to choose whichever of their two ethnicities for their children.

A few facts stand out from the data (the censuses 1982, 1990, 2000 and a mini-census 2005 – see Section 4 for precise sources). First, the propensity to choose minority identity for the children is much higher in mixed marriages with a minority man and a Han woman than in those with a Han man and a minority woman: 94% vs. 41% on average. Figure 1 plots the probability of having minority children for the two types of mixed marriages by every five birth years. The figure also shows a second fact: the share of minority children in mixed marriages are clearly increasing in the mixed couples with a Han man, especially after 1980 – the mean of having minority children is 36% in the cohorts born before 1980 whereas it is 45% in the cohorts born after 1980. Differently, there is little change in those with a minority man – from 94% in the cohorts born before 1980 to 93% in the cohorts born after 1980.

[Figure 1 about here]

Third, the frequency of marrying across ethnic lines is much smaller, namely 1.4%, in the dominant Han group than in the average minority group, where it is 11.8%. Moreover, in a cross-sectional comparison across China's prefectures, this wedge is clearly increasing in the Han population share. Panels A and Panels B in Figure 2 plot the correlation between the share of Han population in a prefecture and the probability of mixed marriage for a Han man and a minority man for those marred in the 1980s (the correlations have a similar pattern for other marriage cohorts). Clearly, the share of Han population is negatively associated with mixed marriages for a Han man (the slope of the fitted line is around -0.15) but positively associated with mixed marriages for a minority man (the slope of the fitted line is around 0.52). The opposite effects of relative population size on mixed marriages for a Han versus a minority is the fourth fact we would like to explain.

[Figure 2 about here]

Existing research on the ethnicity in children and mixed marriages in China mainly comes from sociology. On the ethnicity of children, Guo and Li

(2008) document a similar pattern to our first observation – i.e., the propensity to choose minority identity for children is higher in mixed marriages with a minority man and a Han woman – using the 0.095% sample of the 2000 census. On average, the probability of having a minority children is more than one half. They argue that this fact leads to an increase of minority population share over time. On interethnic marriage, Li (2004) uses aggregate-level information from the 2000 census to document three stylized facts. First, for a minority, the probability of marrying a Han dominates that of marrying a spouse of another minority. Related to this fact, we will focus on the comparison between marrying a Han and marrying a minority (regardless of which minority group). Second, the distribution of ethnic population in a region matters. Third, Muslim religious minorities are more likely to marry within the ethnic groups. Thus, it will be important to allow for variations in population sizes and religious differences.

To the best of our knowledge, no existing study has systematically analyzed ethnic choices in China from a rational-choice perspective. Neither do we know of any existing study – on China or other countries – that has linked the choice of ethnicity of children and the decision to marry across ethnic lines. Our paper tries to fill these two gaps.

We do this in two steps. First, we provide a model that links the choices about ethnicity of children and marriage partner. Agents choose how to search for a spouse across ethnicities, as well as the ethnicity of their children if they end up in a mixed ethic marriage. Any observed correlation between ethnic choices for children and interethnic marriages is thus an equilibrium outcome, which is endogenous to government policies and other economic or social determinants. The model is constructed to be consistent with the four facts noted above on the choice of ethnicity for children and the choice of mixed marriages. But the model also delivers several auxiliary predictions.

In a second step, we take these predictions to Chinese microdata. For example, we empirically evaluate the interplay between social norms and incentives on ethnic choices. We are not aware of any existing empirical study on how social norms crowd in or crowd out material incentives, as in Benabou and Tirole (2011). A similar methodology may also apply to other contexts. Another auxiliary prediction relates to sex ratios. Specifically, we examine how sex ratios together with material benefits affect inter-ethnic marriages, which contributes to the existing literature evaluating the consequences of imbalanced sex ratios in China. Wei and Zhang (2011) argue that higher ratios of men to women explain a large part of increased saving rates in

China, whereas Edlund et al. (2013) document that higher sex ratios lead to more crimes. Although these papers do not study the marriage market directly, marriage search may be an important underlying mechanism.

In what follows, Section 2 sets up our model where agents choose how to search in the marriage markets and what ethnicity to pick for their children. Section 3 shows that the model is consistent with the four facts mentioned above, and spells out a number of other predictions. Section 4 discusses which data can be used to test them. Section 5 brings the predictions to the data and presents our econometric results. Section 6 concludes the paper. An Appendix collects the proofs of some theoretical results, and a Web Appendix provides some additional empirical results.

2 The Model

In this section, we set up a model to generate predictions for the determinants of mixed marriages, and the ethnicity choices for the children in such marriages. The model has two connected stages: a marriage stage and a family stage. Given their information, agents at the marriage stage have rational expectations about the outcome at the family stage. Hence, we consider the stages in reverse order. For the family stage, we use a framework similar to the one in Benabou and Tirole (2011) to model the ethnicity choice for children as a choice that involves material payoffs as well as immaterial payoffs (social norms and culture). For the marriage stage, we use a framework with costly directed search similar to the one in Bisin and Verdier (2000) to model search behavior in the marriage markets for different ethnic groups.

The main purpose of the model is to set the stage for our empirical work. Therefore, we include in the model only those prospective determinants of ethnicity choices that we can actually measure with some degree of confidence. These variables include material benefits for minority children, cultural differences across ethnicities, and sex ratios across and within groups. Actually, we can measure most of the observables at the regional (province or prefecture) level and some at the individual level, as further discussed in Section 4. We should thus think about the model as capturing these individual or regional conditions. The model is certainly highly stylized, but it does yield a number of predictions, summarized in Section 3, which we take to the data in Section 5.

2.1 The Family Stage

Consider a region (province or prefecture) with a continuum of households. There are two ethnicities $J \in \{H, M\}$,where H denotes Han and M Minority. Households have children which yield the same basic benefit for everyone v. Each household has a single discrete decision to make: whether to choose minority status for their children, $m = 1$, or not, $m = 0$. In line with the social situation in China, we assume that this choice primarily reflects the husband's preferences. We focus on the decisions by mixed couples (H, M) or (M, H), where the first entry is the ethnicity of the man. Non-mixed couples, which are kept in the background, always choose their joint ethnicity for their children (this is not only plausible theoretically, but true empirically). The framework considers extrinsic incentives (material benefits or costs) as well as intrinsic incentives (social norms or self-image), and – not the least – the interaction between the two.

Han-Minority mixed couples Suppose first that the man is Han and the woman is minority. Then, the preference function of the couple is

$$v + (b - e(H) - \varepsilon)m - \mu E(\widetilde{\varepsilon} \mid m) \ , \tag{1}$$

where b is the net *extrinsic* benefit of having minority children, which is controlled by the regional government. This parameter could differ across regions or time, due to different policies favoring minority children (such as they themselves being allowed to have more children, or advantages in the education system). Further $e(H) + \varepsilon$, is the *intrinsic* cost of not having a Han child. Its first component is common and deterministic, and possibly different across regions; it could also differ across ethnicities depending on "cultural distance". The second component ε instead varies across households. An important source of heterogeneity in the model is ε, which is distributed across couples with mean $E(\varepsilon) = 0$, c.d.f. $G(\varepsilon)$, and continuous, differentiable, single-peaked p.d.f. $g(\varepsilon)$,which is symmetric around zero.

The final term captures the household's social reputation, or self image – how society views the mixed couple, or the couple views itself – given the ethnicity decision that it makes. It is defined over $\widetilde{\varepsilon}$, which is the truncated mean of ε in all households with the household's *peer group*, who make the same choice as the household does. Parameter μ, is the weight on this social reputation. Depending on the strength with which the social norm is held, this parameter could vary across different peer groups. One definition of the

relevant peer group would be the household's region, but there could also be separate within-region peer groups, say households with or without higher education, or households in urban vs. rural areas.

For the analysis to follow, it is useful to define the variable

$$\Delta(\varepsilon) = E(\widetilde{\varepsilon} \mid m = 1) - E(\widetilde{\varepsilon} \mid m = 0) \ . \tag{2}$$

Following the terminology in Benabou and Tirole (2011), the first term on the RHS of (2) can be interpreted as the "stigma" for the Han-man household in a particular peer group of having a child of minority identity, which will be the choice of households with a sufficiently low value of ε. Instead, the second term is the "honor" of having a child of the man's own identity, which will be the choice of households with sufficiently high ε.

Specifically, it follows from (1) and (2) that the mixed couple will have a minority child if

$$\begin{aligned} \varepsilon &< b - e(H) - \mu[E(\widetilde{\varepsilon} \mid m = 1) - E(\widetilde{\varepsilon} \mid m = 0)] \\ &= b - e(H) - \mu\Delta(\varepsilon^*) = \varepsilon^*_H(b, e(H), \mu). \end{aligned} \tag{3}$$

The second equality implicitly defines a cutoff value of ε, below which agents have minority children, as a function of b, e and μ. The properties of the equilibrium and its comparative statics will crucially depend on the derivative Δ_ε. Suppose ε^* goes up so that more Han-minority couples have minority children. Then, both the honor and the stigma terms go up, so the question is which goes up by more. By the results in Jewitt (2004), the single peak of g implies that Δ has a unique interior minimum, so $\Delta_\varepsilon < 0$ for low values of ε^*, when few people have minority kids, and $\Delta_\varepsilon > 0$ for high values of ε^*, when many have minority kids. It follows that ethnicity choices for children are strategic complements when $\Delta_\varepsilon < 0$, while they are strategic substitutes when $\Delta_\varepsilon > 0$. For one of the results below, we also assume that the second derivative has the same sign $\Delta_{\varepsilon\varepsilon} < 0$.

Minority-Han mixed couples In a M, H mixed couple, where the man is Minority rather than Han, the preference function can be written:

$$v + mb - (1 - m)(e(M) + \varepsilon) - \mu E(\widetilde{\varepsilon} \mid m) \ , \tag{4}$$

where $e(M)$ and ε now denote the deterministic and stochastic parts of the intrinsic cost of having a Han child (different from the minority man's own

ethnicity). Similar calculations as before, shows that this mixed couple will have a minority child when

$$-\varepsilon < b + e(M) + \mu\Delta(\varepsilon^*) = \varepsilon^*_M(b, e(M), \mu) \tag{5}$$

By the symmetry of the ε-distribution, the probability of having a minority child is thus $G(\varepsilon^*_M(b, e(M), \mu))$.

2.2 The Marriage Stage

To model the marriage market, we use a model of directed search similar to the one in Bisin and Verdier (2000). There are two *restricted* marriage-matching pools, where only individuals with the same ethnicity can match in marriage. Consistent with our assumption that the ethnicity choices are dominated by the preferences of men, we assume that men are the active agents in the marriage market. Thus, we consider the search behavior only of men. When evaluating the prospects of marriage with women of different ethnicities, a man internalizes the expected utility given by the expected outcomes at the family stage, as derived in the previous subsection.

Basics Let C be a convex function with $C'(0) = 0$. With (directed) search effort $C(\delta\alpha^J)$, a man with ethnicity J enters the restricted marriage pool with probability α^J, where he is always married with a woman of the same ethnicity. The slope parameter δ could capture individual level search difficulties. With probability $1 - \alpha^J$, he instead enters the common pool with all women, who have not been matched in the restricted pools of their own ethnicity. In this common pool, individuals match randomly notwithstanding their ethnicity. Let A^J be the fraction of men of ethnicity J, who search in the restricted pool, a share that must be consistent with the share of women who passively get matched in that pool. In equilibrium, every man with the same ethnicity, in the same peer group, behaves identically and hence we have $\alpha^J = A^J$.

Denote by λ the population share of the Han, and by S^J the (inverted) sex ratio in ethnicity J – the number of women per man – for a constant population share of ethnicity J. Then, the assumptions in the previous paragraph imply that the probability of a Han man to marry a Han woman is:

$$\pi^H = \alpha^H + (1 - \alpha^H)P^H \ , \tag{6}$$

where $P^H = \frac{(1-A^H)\lambda S^H}{(1-A^H)\lambda S^H + (1-A^M)(1-\lambda)S^M}$ is the probability to meet a partner of Han ethnicity in the unrestricted (common) pool. It follows that $\frac{\partial P^H}{\partial A^H} < 0$, $\frac{\partial P^H}{\partial A^M} > 0$, $\frac{\partial P^H}{\partial \lambda} > 0$, $\frac{\partial P^H}{\partial S^H} > 0$, and $\frac{\partial P^H}{\partial S^M} < 0$. The corresponding probabilities π^M and P^M for Minority men are defined accordingly.

The Han man's marriage problem A Han-man chooses α^H to maximize:

$$v + (1 - \pi^H)V^H - C(\delta\alpha^H) = v + (1 - \alpha^H)(1 - P^H)V^H - C(\delta\alpha^H) \ ,$$

where the equality follows from the definition in (6), and where $V^H = G(\varepsilon_H^*)[b - e(H) - \mu\Delta(\varepsilon_H^*)] - \mu E(\widetilde{\varepsilon} \mid m = 0)$. This objective function incorporates the expected outcome from the family stage of the model. Independently of the match, the utility of a child is v. With probability $1 - \pi^H$ the Han man will end up in a mixed marriage. V^H is the continuation value of such marriage and obtained by taking expectations of the expression in (1). The ex-ante probability, not knowing the shock ε, of having a minority child is given by the population probability $G(\varepsilon_H^*)$ derived in the previous subsection. In this event, the man expects additional extrinsic benefits b and expected intrinsic cost $(1 + \mu)e(H) + \mu\Delta(\varepsilon_H^*)$. Thus, the social norms regarding the ethnicity choice for children spill over onto the marriage-search decisions.

Of course, in his individual (and atomistic) decision of choosing α^H, the Han man takes as given the decisions by others in their marriage search and ethnicity choices, although he has rational expectations about their behavior. The first-order condition for this decision becomes (ignoring the constant term in $\mu E(\widetilde{\varepsilon} \mid m = 0)$):

$$\delta C'(\delta\alpha^H) \geq -(1 - P^H)V^H \quad \text{c.s. } \alpha^H \geqslant 0 \ . \tag{7}$$

To get $\alpha^H > 0$, we require that $V^H < 0$. In other words, a Han man searches in the restricted pool only when the (unconditional) expected cost of minority children are higher than the benefits.

The Minority man's marriage problem A minority man's problem is to choose α^M to maximize

$$v + \pi^M b + (1 - \pi^M)V^M - C(\delta\alpha^M) \ ,$$

where π^M is defined in the same way as π^H, and where $V^M = G(\varepsilon_M^*)b - (1 - G(\varepsilon_M^*))[e(M) + \mu\Delta(\varepsilon_M^*)) - \mu E(\widetilde{\varepsilon} \mid m = 1)]$. This continuation payoff, obtained from (4), is somewhat different than the one for a Han man. The Minority man's probability of getting a minority child, and hence benefits b, is given by $\pi^M + (1 - \pi^M)G(\varepsilon_M^*)$. With probability $(1 - \pi^M)(1 - G(\varepsilon_M^*))$ he gets a Han child and suffers an expected cost $e(M) + \mu\Delta(\varepsilon_H^*)$.

Defining $P^M = \frac{(1-A^M)(1-\lambda)S^M}{(1-A^M)(1-\lambda)S^M+(1-A^H)\lambda S^H}$ analogously to P^H, and using this to rewrite π^M in terms of α^M and P^M, we can write the first-order condition to this problem as

$$\delta C'(\delta\alpha^M) \geq (1 - P^M)b - (1 - P^M)V^M \quad \text{c.s. } \alpha^M \geqslant 0 \ . \tag{8}$$

We can rewrite the RHS as $(1 - P^M)(b - V^M) = (1 - P^M)(1 - G(\varepsilon_M^*))[b + e(M) + \mu\Delta(\varepsilon_M^*) - \mu E(\widetilde{\varepsilon} \mid m = 1)] > 0$.

It is clear from this condition that the Minority man always puts in search effort to get access to the restricted own-ethnicity pool, as such access avoids the risk of meeting a Han woman in the unrestricted pool with probability $(1 - P^M)$ and end up with a Han child with probability $1 - G(\varepsilon_M^*)$, which carries intrinsic costs of $e(M) + \mu\Delta(\varepsilon_M^*)$ and foregoes extrinsic benefits of b.

3 Comparative Statics

This simple model delivers a rich set of comparative statics both for the family stage and the marriage stage. These are summarized by the family stage and the marriage stage below. For the predictions on mixed marriages, we relegate the proofs of the results to the Appendix.

3.1 Predictions at the Family Stage

We start by showing that the model predictions are consistent with the two facts noted in the introduction on ethnicity choices for children. Then, we derive three sets of auxiliary predictions.

Comparison across mixed marriages (ε_M^* versus ε_H^*) It follows from (5) and (3) that $\varepsilon_M^*(b, e(M), \mu) > \varepsilon_H^*(b, e(H), \mu)$. This implies that $G(\varepsilon_M^*) > G(\varepsilon_H^*)$ – i.e., we have a straightforward prediction: Minority children are more frequent in mixed marriages when the man is Minority rather than

Han. Clearly, this prediction is consistent with the fact in the introduction about the average status of children in different types of mixed marriages.

The effect of material benefits (b) Let's first look at a Han-minority family. Consider the proportion of minority kids in the population of these couples. This can be written $m(b, e, \mu) = G(\varepsilon^*_H(b, e, \mu))$, as a function of the cutoff value, which itself is a function of the benefits and costs of having minority children. Using the definition $b - (1 + \mu)e - \mu\Delta(\varepsilon^*) = \varepsilon^*$, we can calculate the shift in the proportion of minority kids in response to a higher net benefit:

$$m^H_b(b, e, \mu) = g(\varepsilon^*_H(b, e, \mu))\frac{1}{1 + \mu\Delta_\varepsilon(\varepsilon^*_H(b, e, \mu))} > 0. \tag{9}$$

Similarly, the effect of extrinsic incentives (b) for a minority-Han family is:

$$m^M_b(b, e, \mu) = g(\varepsilon^*_M(b, e, \mu))\frac{1}{1 + \mu\Delta_\varepsilon(\varepsilon^*_M(b, e, \mu))} > 0. \tag{10}$$

Given that $g(\varepsilon^*_M(b, e, \mu))$ is smaller than $g(\varepsilon^*_H(b, e, \mu))$ (minority-Man couples having Han children is more of a tail event) and $\Delta_\varepsilon(\varepsilon^*_M(b, e, \mu)) > \Delta_\varepsilon(\varepsilon^*_H(b, e, \mu))$, we get the following prediction: material benefits raise the probability of having a minority child. This effect is larger for the Han-Minority families.

This prediction, *together with* increasing benefits of having minority children over time (see Section 4 for a discussion of the benefits), gives a prospective explanation for the second fact highlighted in the introduction: a more pronounced trend to have minority kids over time in mixed marriages with Han men than in those with minority men (recall Figure 1).

Having established the link between our model and the facts on ethnicity for children, we turn our interest to the auxiliary predictions implied by the model. These are the ones we will test empirically.

Material benefits (b) and social norms (Δ_ε) From (9), we see that net material benefits are crowded in by social reputation – i.e., the multiplier is larger than 1 – when few people have minority kids and their ethnicity choices are strategic complements. Instead, benefits are crowded out when many people have minority kids. Note that the amount of crowding in or crowding out is also governed by the relative strength of the social norm μ.This implies

a specific empirical prediction across regions (or more generally across peer groups) summarized as:

F1 *If all prefectures (peer groups) in a province (prefecture) experience the same change in benefits, due to a provincial (prefecture) policy, we should see a larger effect on individual behavior in prefectures where the share of minority children in male-Han mixed marriages is smaller.*[1]

Heterogeneity in material effects (b) Another auxiliary prediction is straightforward.

F2 *The minority groups that enjoy smaller material benefits are less likely to choose minority for their children than those who enjoy more material benefits.*

Material benefits (b) and intrinsic costs (e) Shifts in the deterministic intrinsic costs e – say due to a successful socialization campaign – can be analyzed in similar fashion as b. We are interested in the interaction effect of b and e:

$$\frac{\partial m_b^H(b,e,\mu)}{\partial e} = \frac{\partial m_b^H(b,e,\mu)}{\partial \varepsilon_H^*}\frac{\partial \varepsilon_H^*}{\partial e}$$

We know that $\frac{\partial \varepsilon_H^*}{\partial e} < 0$. As for $\frac{\partial m_b^{HM}(b,e,\mu)}{\partial \varepsilon_H^*}$, it has two terms. One is positive, by our assumption that $\Delta_{\varepsilon\varepsilon} < 0$. The sign of the other term depends on the change in the density $g_\varepsilon(\varepsilon_H^*(b,e,\mu))$ which, in turn, is positive before the single peak of g and negative thereafter.

F3 *The interaction effect of b and e on the share of male-Han mixed marriages is negative if the share of mixed marriages is small, but less negative or even positive if this share is large.*

[1] Note that we also get different comparative statics for minority-Han families. As Δ_ε is monotonically increasing from a negative value when the number of minority kids is small, the social multiplier is smaller for mixed household with minority men than for those with Han men. This means that the same increase in net extrinsic benefits produces a smaller effect on the share of minority kids in M, H couples than in H, M couples – with a larger share of couples having minority children, there is more crowding out (or less crowding in) via the social reputation mechanism. (As can be seen from (10), this also requires that the density g is relatively flat across the two equilibrium points.) To test this prediction emprically, however, we need enough variation in ε_M^*. This is difficult, howevere, since $G(\varepsilon_M^*)$ is close to 1 in most cases.

3.2 Predictions at the Marriage Stage

To get clear and specific signs in the comparative statics for the marriage stage, we postulate the following throughout this subsection:

Assumption 1 $(\delta)^2 C''(\delta\alpha^H) < V^H \frac{dP^H}{dA^H}$ and $(\delta)^2 C''(\delta\alpha^M) < -(b - V^M)\frac{dP^M}{dA^M}$.

In words, this says that the convexity of the search costs for group J is low enough to be dominated by the effect on the expected cost of having a child of different ethnicity when a higher share of ethnicity J searches in the restricted marriage pool.

The effect of population shares (λ) We first consider the effect of the majority group's population share on the incidence of mixed marriages. Due to the mechanical reason of a large share of Han population, we get the result that the frequency of marriages across ethnic lines is higher among minority men than among Han men, i.e., $\pi^H > \pi^M$. Note that $1 - \pi$ indicates the mixed marriage probability.

Moreover, we have the following prediction: a higher population share of Han, decreases the proportion of male-Han mixed marriages ($\frac{d\pi^H}{d\lambda} > 0$), but increases the proportion of male-Minority mixed marriages ($\frac{d\pi^M}{d\lambda} < 0$). The proof is presented in the appendix. Intuitively, the main effect of a higher population share for the Han (λ) raises the probability to meet a partner of Han ethnicity in the unrestricted (common) pool (P^H) for a Han man and hence increases π^H. The effect of is the opposite for a Minority man.

This prediction makes the model consistent with the third and fourth facts highlighted in the introduction about the average and cross-sectional mixed-marriage pattern across prefectures (recall Figure 2). Given that our model is consistent with these facts, we examine the auxiliary predictions on sex ratios in the population.

The effect of sex ratios (S) For the Han sex ratio, we have the following results:

M1 *A higher sex ratio (men to women) among the Han, raises the proportion of male-Han mixed marriages, but lowers the proportion of male-Minority mixed marriages. Moreover, the former effect becomes larger, the higher are the material benefits of minority children, while*

the latter effect is dampened by these material benefits. ($\frac{d\pi^H}{dS^H} > 0$, $\frac{d\pi^M}{dS^H} < 0$, $\frac{d^2\pi^H}{dS^H db} > 0$ and $\frac{d^2\pi^M}{dS^H db} > 0$).

When it comes to the Minority sex ratio, we focus on the comparison across different minority groups

M2 *A higher sex ratio (men to women) within a Minority, raises the proportion of male-Minority mixed marriages. Moreover, this effect is dampened by material benefits.* ($\frac{d\pi^M}{dS^M} > 0$ and $\frac{d^2\pi^M}{dS^M db} < 0$).

The proofs of these two predictions are presented in the Appendix. The intuition for the prediction that $\frac{d\pi^H}{dS^H} > 0$, $\frac{d\pi^M}{dS^H} < 0$ goes as follows: a higher sex ratio (lower S^H) makes it more difficult for a Han man to meet a Han woman and hence decreases his effort to search within his group. This decreases π^H and increases interethnic marriage. The effect on a Minority man is the opposite.

The interaction effects ($\frac{d^2\pi^H}{dS^H db}$ and $\frac{d^2\pi^M}{dS^H db}$) are determined by the main effect that we have just discussed, interacted with the continuation values in the family stage. The continuation value for a Han man to marry a minority is increasing in the material benefits (b), whereas the effect of b makes the continuation value of mixed marriages lower than that of within-ethnic marriages for a Minority man. The effect of sex ratios across minorities is similar.

4 Data and Measurement

This section discusses how to measure the relevant variables and parameters in the model. The outcomes and some control variables are measured at the individual level, while the material and intrinsic motivations are measured at the regional or ethnicity levels.

Linking of datasets We draw on two sources of data. The first one are excerpts from three China censuses: the 1% sample of the 1982 census, the 1% sample of the 1990 census and the 0.095% sample of the 2000 census. The second source is the 2005 population survey that covers about 1% of the population, also known as the mini-census. These data provides demographic information and some limited information on social economic status for altogether about 25 million people.

As in the model, we are interested in the husband-wife-children structure of households. The husband or wife data is based on the information of the gender of household head. In some cases, parents or parents-in-law of the household-head or the spouse cohabit with them. We drop this relatively small part of the sample, as the censuses do not distinguish parents from parents-in-law in the censuses in 1982 and 1990.

The administrative unit we focus the most on is the areas defined by the four-digit codes in the census. Considering that some areas change names and codes over time, we unify the boundaries based on the year 2000 information to end up with 346 prefectures or cities. Since over 330 of these are prefectures, we refer to all of them broadly as prefectures.

Measuring outcomes (m and π in the model) Following the family stage of the model, we study the ethnicity of children in mixed marriages. We can identify children in the census 2000 and mini-census 2005. The 1982 and 1990 censuses do not distinguish between children and children-in-law. To identify children in these two censuses, we therefore limit ourselves to unmarried children who still live with their parents. The results we report below are robust to using the census 2000 and the mini-census 2005 only.

In all the waves of data, we know each individual's ethnicity as well as her birth year. This way, we know whether $m = 0$ or $m = 1$ and whether a child is subject to certain state or province policies implemented in his or her birth cohort. As shown in Panel A of Table 41% of the children in Han-minority families are minorities whereas 94% of the children in minority-Han families are minority. This is the first of the facts mentioned in the introduction of the paper.

[Table 1 about here]

To study the inter-ethnic marriage decisions, we follow the model's marriage stage and ask whether a Han man decides to marry a Minority woman (related to probability $1 - \pi^H$) and whether a Minority man marries a Han woman (related to probability $1 - \pi^M$). Because the 2000 census and 2005 mini-census report the marriage year, we know to what extent a man is affected by the state or province policies implemented relevant for different marriage cohorts (Marriage-year information is not available in the 1982 and 1990 censuses). As shown in Panel B of Table 1, based on the censuses 2000 and 2005, the probability of marrying a minority woman for a Han man is

1.4%. One reason for this small number is that the average population share of Han is above 90%. The probability of marrying a Han woman for a minority man is 11.8%. This difference across Han and minority men is the third of the facts highlighted in the introduction.

In the Web Appendix, Tables W1 and W3, we show that these two facts are not only true at the aggregate level, but also at the individual level (which is the domain of the model), even when we control for prefecture fixed effect and cohort fixed effects.

Measuring material benefits (b in the model) We use three ways to measure material benefits.

The first measure is based on cohorts. The People's Republic of China (1949-) has employed different policies to the benefit of ethnic minorities. These policies can be broadly grouped into three groups:

(1) **Family planning**. When the family planning policy started in the 1960s, minorities were exempted from it. Over time, there is also regional variation in the treatment of minorities.

(2) **Entrance to college**. Since the restoration of entrance exams to colleges in 1977, minorities enjoy some extra points in the exams. These benefits too vary by provincial policy.

(3) **Employment**. The national ethnic policy states that minorities should have favorable treatment in employment. However, there is rarely explicit quotas for minority employment. As minorities are often discriminated in employment, it is unclear that this policy would make people tend to choose minority identity.

It is not straightforward to quantify regional variation over time in these policies. However, it is clear that the effect should be larger after 1980. Since the 1980s, family planning was switched more strictly to one-child policy. Hence, the benefits of being exempted from this policy became larger. On top of this, there came the extra benefits of better opportunities for higher education. Hence we use the dummy indicating post 1980 as a measure of material benefits in the baseline estimations.

The increasing benefits to having minority children over time, together with the theoretical result in Section 3 that the effects of benefits are larger in mixed marriages with a Han man, makes the model consistent with the second fact in the introduction as illustrated in Figure 1. Table W2 in the Web Appendix shows that the increasing propensity for such couples to have

minority children also holds up at the individual level, when we control for cohort and prefecture fixed effects.

The second measure of minority benefits exploits the gradual rollout of one-child policy across provinces. The precise timing is based on when the family-planning organization was set up in a province (data is available for 27 provinces, which is used in the working-paper version of Edlund et al. (2013)).[2] The advantage of this measure is that it is staggered across provinces as the organizations are established between the 1970s and the 1980s. The disadvantage is that it does not capture other benefits, such as those in education and employment. Naturally, the measure is correlated with the post-1980 dummy (with a correlational coefficient around 0.8).

A third measure we explore considers heterogeneity in the beneficiaries of pro-minority policies. In particular, most of the preferential policies are limited to minorities with a population smaller than 10 million. As the size of Zhuang minority was over 13 million already in the 1982 census, this group enjoyed many fewer benefits than did other minority groups. Therefore, we will compare the Zhuang minority with other minority groups. As shown in Table 1, the probability of having a Zhuang wife in a Han-man mixed marriage is about 17%.

Measuring social norms (Δ_ε in the model) Following the discussion about crowding out or crowding in the model, we measure social norms primarily by the share of minority children in mixed marriages, separately for male-Han and male-Minority mixed marriages. To make sure that our results are reasonably robust, we define the peer group relevant for the social norms in a few different ways.

(1) **The 1970s cohort in the same prefecture.** We first exploit the variation across prefectures in the birth cohort of the 1970s, i.e., before the dramatic changes in ethnic policies and sex ratios (see below).

(2) **The previous cohort in the same prefecture.** Considering the dramatic economic development in the past few decades, social norms may have changed fairly quickly. A second way to define the peer group relevant

[2] Beijing, Shanghai, Tianjin and Chongqing are not included. We thank Lena Edlund for providing this data. The working-paper version of Edlund et al. (2013), considers three types of family-planning organizations: (i) family-planning science and technology-research institutes, (ii) family-planning education centers, and (iii) family-planning associations. As the timing of these organizations are close, the results do not depend much on which ones are used. Below, we present the results using a measure based on (i).

for the prevailing social norms in a cohort is to use the previous birth cohort from the previous ten years in the same prefecture. That is, the 1970s cohort becomes the peer group for the cohort norm in the 1980s, and so on.

(3) **Same residency and previous cohort in the same prefecture**. Measures (1) and (2) use only birth cohort and prefecture to define a peer group. Conceptually, the effect of social norms might be stronger within a more specific peer group. Hence, we also distinguish urban and rural residency and define the peer group at the prefecture-cohort-residency level. A limitation of this method is that it implies smaller groups at the same time way as rural/urban information is only available in the 2000 and 2005 censuses. Hence, the number of observations in each cell becomes much smaller than for measures (1) and (2).

Figure 3 plots the distribution of having a minority child in the two types of mixed families. It shows a great deal of variation across prefectures for male-Han mixed families. However, for male-Minority mixed families, most prefectures are concentrated at the right end, leaving little variations across prefectures. Hence, we will focus on the effect of social norms on Han-minority families.

[Figure 3 about here]

Figure 4 further maps the spatial distribution across China of the social norms (based on the 1970s cohort) for male-Han mixed families. It indicates that social norms seem to vary quite a bit across prefectures, and that this variation is not strongly clustered geographically. For Han-minority families, the model predicts a strategic complementarity $\Delta_\varepsilon < 0$ for low values of ε^* (when few people have minority kids) and a strategic substitutability $\Delta_\varepsilon > 0$ for high values of ε^* (when many have minority kids). We do not observe the distribution of ε^* and thus cannot measure the cutoff value. Instead, we vary the assumed cutoff value and examine the estimates for different assumed cutoffs.

[Figure 4 about here]

Measuring intrinsic costs (e in the model) A first measure we use is whether the child is a son or a daughter. Consistent with the Confucian values, the intrinsic costs of having sons with different ethnicity are higher. A second measure of intrinsic costs is whether the spouse belongs to a religious

minority group. It is conceivable that it is more costly for a Han man if his child needs to practice religion due to a minority identity. It is worthwhile pointing out that marrying a religious wife can be endogenous. Our interest is how the effect of material benefits on ethnic choice for the children is contigent on having a religious wife rather than the effect of having a religious wife itself.

Table 1 shows that the share of male-Han mixed families with a religious wife is about 17%. We have also experimented with two other potential measures: linguistic distance and genetic distance. However, the former may be less important for China, where Mandarin is the dominant language, and the latter may be less important within a country than between countries.

Measuring population shares (λ in the model) To measure λ in the model, we calculate the share of Han population and minority population by prefecture-birth cohort. We pool all censuses together to increase the sample size. Still, the size of population in some prefecture-birth cohort cells may be small and their ratios may be outliers. To deal with this concerns, we trim both the right and the left 5% tails in our baseline estimates, but include them as a robustness check. Considering that the marriage age for men is around their twenties, we use the population ratios for those born in cohort $t-20$ to measure the population shares faced by a man in marriage cohort t. For example, those married in the 2000s face the population share effect measured by those born in the 1980s.

This information is used to generate the fourth fact in the introduction as illustrated in Figure 2. Table W4 in the Web Appendix, shows that the opposite effects of the Han population ratio on mixed marriages with Han and minority men is true not only at the prefecture level but also at the individual level, also when we control for individual socioeconomic status and cohort fixed effects.

Measuring sex ratios (S in the model) Similar to population size ratios for the Han population, we calculate sex ratios by prefecture-birth cohort. Again, those married in the 2000s face the population share effect measured by those born in the 1980s and so forth.

Panel A of Figure 5 plots the distribution of sex ratios after trimming the 5% tails (to diminish the weight of outliers in the distribution). Compared with the birth cohort of the 1950s, it is clear that the distribution of sex

ratios moves right in the birth cohort of 1990s, reflecting the effect of the one-child policy. This figure also suggests that there is a lot of variation in sex ratios across cohorts within a prefecture. A regression of the sex ratio on prefecture fixed effects yields an R-square of around 0.24. Thus, we can exploit a large portion of unexplained variation within prefectures over time to test the model predictions.

For minorities, we are primarily interested in how sex ratios across ethnic groups affect inter-ethnic marriages. Thus, we calculate sex ratios across province-ethnicity-birth cohort. Similar to the Han sex ratios, we trim 5% tails in the baseline estimates. Panel B of Figure 5 plots the distribution of these sex ratios. It shows that change in the distribution across cohorts is much narrower than the corresponding distribution of the Han sex ratios.

[Figure 5 about here]

Individual socioeconomic status Finally, our model revolves around choices at the individual or family level. As these choices may also reflect socioeconomic conditions, or social norms in a more narrow peer group than a prefecture-wide cohort, we would also like to hold constant individual socioeconomic status. Two important dimensions are rural vs. urban identity and college education. Both dimensions are available and consistently measured in the census 2000 and mini-census 2005. As shown in Panel B of Table 1, among the Han men, 32% have urban identities and 8% have college education. Among the minority men, 20% have urban identities and 6% have college education. In the 1982 and 1990 censuses, however, the information on rural/urban identities is unavailable, while the coding of education is different from that in the latter sources.

Since we focus on the 2000 census and 2005 mini-census in the estimation of the marriage stage, we present the results including individual urban identity and college education in our baseline estimates. For the family stage, we use all censuses as our baseline. Here, we present the results using censuses 2000 and 2005 only with individual controls for urban identity and college education as a robustness check.

5 Empirical Results

This section presents our empirical specifications and estimation results, beginning with those on the ethnic choices of children followed by those on

mixed marriages.

5.1 Ethnic Choice for Children

Material incentives and social norms – Prediction F1 We focus on the ethnic choice for children in mixed marriages between Han men and minority women because almost all mixed marriages between minority men and Han women result in minority children (recall Figure 1).

Prediction F1 about the influence of social norms says that the effect of higher material benefits should be larger in places or groups where the initial share of minority children is smaller, because then the material benefits are crowded in (crowded out less) by prevailing social norms. To test this, we run the specification:

$$MinorityChild_{i,p,t} = \beta_b Post1980_t \times Cutoff_p + pref_p + birth_t + prov \times t + \varepsilon_{i,p,t},$$

where $MinorityChild_{i,p,t}$ is a dummy indicating whether a child i, in prefecture p, and birth cohort t is a minority.

We use a dummy of $Post1980_t$ to measure material benefits. $Cutoff_p$ is a dummy variable which indicates whether the peer group – defined in the three ways discussed in the previous section – has a share of minority children (defined in three ways) smaller than some cutoff value. To be flexible, we use a wide range of cutoff values from 0.3 to 0.7.

To control for the effect of prefecture characteristics that are time-invariant or change slowly over time, we control for prefecture fixed effects ($pref_p$). To control for factors that affect different cohorts, we control for birth cohort fixed effect ($birth_t$, for every ten years). We also include province-specific (linear) trends ($prov \times t$) to control for different evolutions across provinces, such as different growth rates or different provincial policies other than ethnic policies.

[Figure 6 about here]

(1) **Results using the 1970s cohort.** We first define social norms for a prefecture by the share of minority children in the 1970s cohort in the prefecture. To save space, Table 2A presents the results for the cutoff range between 0.45 and 0.65 while Figure 6 visualizes all the results. The estimated effect of material incentives is indeed generally larger when the share is smaller than the cutoff value. On average, the effect of $Post1980_t$ is around 0.08

(presented in column (2) of Table W2 in the Web Appendix). The estimates suggest that the differences on the two sides of the cutoff can be half the average effect of material benefits (represented by $Post1980_t$).

[Table 2A about here]

Table W5 in the Web Appendix present results using the provincial timing of one-child policy instead of $Post1980_t$. The results are similar to those in Table 2A.

(2) **Results using the previous cohort.** The second way defines social norms for a cohort by the share of minority children in the previous 10-year cohort in the same prefecture. The estimates of β_b are presented in Table 2B. They are very similar to those using the 1970s cohort only, although the magnitudes are a bit smaller.

[Table 2B about here]

(3) **Results using the previous cohort plus rural and urban residency.** This measure is similar to that in (2), but further separates those with rural and urban residencies. For this measure, we can only use the 2000 and 2005 census data. Tables 2C and 2D present the results for rural-residency and urban-residency members of the same prefecture-cohort, respectively. These tables deliver a similar message as the results based on the first two measures, but now the estimated values of β_b are generally larger.

[Tables 2C and 2D about here]

In sum, the data is clearly consistent with Prediction F1.

Heterogeneity in material benefits – Prediction F2 Another auxiliary prediction of our model implies that the effect of higher benefits should be smaller for the Zhuang ethnicity than for other minorities, simply because the increase in benefits were smaller for the Zhuang. To test this prediction, we check whether its effect of benefits is smaller if the wife is a Zhuang:

$$\begin{aligned} MinorityChild_{i,p,t} &= \beta_z Post1980_t \times ZhuangWife_i + \gamma ZhuangWife_i \\ &\quad + pref_p + birth_t + prov \times t + \varepsilon_{i,p,t}. \end{aligned}$$

Our model predicts that $\beta_z < 0$.

The estimates are presented in Table 2. Column (1) shows the result controlling for $Post1980_t$ (but not $birth_t$), whereas Column (2) controls for $birth_t$. The results show that having a Zhuang wife decreases the effect of material benefits, as represented by coefficient β_z. When we further control for province-specific time trends, the mitigating effect of β_z is smaller in size but still negative. As a further check for potential time trends, Column (4) restricts the cohorts of children to those born between 1970 and 1990, before and after the main policy changes around 1980. The magnitude is similar to the result in Column (3). Column (5) presents the results using the provincial one-child policy timing instead of $Post1980_t$ (controlling for all fixed effects and trends). Again, having a Zhuang wife significantly decreases the effect of material benefits by around one fourth to one half.

[Table 3 about here]

Together, these econometric estimates are also consistent Prediction F2.

Material benefits and intrinsic costs – Prediction F3 Our final prediction about the family stage concerns the interaction effect of material benefits and cultural distance on the choice of a minority child. Our model predicts that this interaction effect is negative. To measure intrinsic costs, we use a dummy for whether the child a son and another dummy for whether the minority wife is also religious (although we recognize that the selection of such a wife may not be random). Thus, we estimate:

$$MinorityChild_{i,p,t} = \beta_s Post1980_t \times Son_i + \delta Son_i + pref_p + birth_t + prov \times t + \varepsilon_{i,p,t},$$

and

$$\begin{aligned} MinorityChild_{i,p,t} &= \beta_r Post1980_t \times ReligiousWife_i + \delta ReligiousWife_i \\ &\quad + pref_p + birth_t + prov \times t + \varepsilon_{i,p,t}. \end{aligned}$$

The estimates are found in Table 4, where columns (1)-(4) present the effect of having a son and columns (5)-(8) present the results of having a wife belonging to religious minorities. Columns (1) and (2) and columns (4)-(5) show that both having a son and having a religious wife cut the effect of material benefits, consistent with the predicted negative interaction effect. However, as shown in columns (3) and (7), the dampening effect becomes

weaker and loses statistical significance, once we further control for province-specific trends. Column (4) and (8) present the results using the provincial one-child policy timing instead of $Post1980_t$ (controlling for all the fixed effects and trends), with similar results.

[Table 4 about here]

On balance, we find that the estimates square with Prediction F3.

5.2 Inter-Ethnic Marriage

Han sex ratios and mixed marriages – Predictions M1 To study the links between sex ratios and mixed marriages, we first examine the prediction that a higher Han sex ratio should raise the probability that a Han man marries an minority wife. To check this for a Han man's marriage choice in cohort t, we look at the effect of sex ratios in prefecture p and cohort $t-20$:

$$H_{i,p,t} = \beta_h(\frac{H^{man}}{H^{woman}})_{p,t-20} + pref_p + marry_t + X_i + prov \times t + \varepsilon_{i,p,t} \ .$$

where $H_{i,p,t}$ is a dummy indicating marrying a minority or not for Han man i in prefecture p and year i. Since the mean is very small (1.4%), we multiply the dummy with 100 and the results can be interpreted as percentage points changes. As in the children results, we control for prefecture fixed effects and marriage cohort fixed effects ($marry_t$). X_i is a vector indicating whether man i has an urban identity and/or a college education

Our model predicts that $\beta_h > 0$. In addition, it predicts that this effect is strengthened by higher material benefits. That is, we expect that $\theta_h > 0$ in the following specification:

$$\begin{aligned} H_{i,p,t} &= \theta_h(\frac{H^{man}}{H^{woman}})_{p,t-20} \cdot Post1980_t + \beta_h(\frac{H^{man}}{H^{woman}})_{p,t-20} \\ &\quad + pref_p + marry_t + X_i + X_i \times Post1980_t + prov \times t + \varepsilon_{i,p,t}. \end{aligned}$$

The estimates are presented in Table 5. The result in Column (1) implies that if the sex ratio increases by one standard deviation (0.23), the probability of marrying a minority indeed goes up by about 3.4 percentage points, which doubles the average probability of doing so for a Han man. Column (2) shows the results after including urban identity and college education.

Unsurprisingly, having a college education raises the probability of a mixed marriage. Column (3) further includes province trends and finds a similar result.

Columns (4) and (5) present the interaction estimates, with and without controlling for X_i, while Column (6) also includes the interaction $X_i \cdot Post1980_t$ as well as province trends. Column (7) uses one-child policy timing instead of $Post1980_t$ to measure material benefits and shows that the results are robust. These results show that the effect of sex ratios is indeed strengthened by higher material benefits.

[Table 5 about here]

We also look the effect of sex ratios among Han on the inter-ethnic marriage probability for a minority man by replacing the dependent variables above to be $M_{i,p,t}$. Our model predicts that $\beta_h < 0$ and $\theta_h > 0$ for a minority man. The results are presented in Table 6.

As in Table 5, Columns (1)-(3) present the results for the sex ratios alone, whereas Columns (4)-(6) present the results for the interaction effect of sex ratios and material benefits. Consistent with the prediction that $\beta_h < 0$, we find that the effect of Han sex ratios on the marriage choices of a minority man is negative but it is not significant. The sign of the interaction effect is also consistent with our prediction but it is not significant. Using the provincial one-child policy timing instead of $Post1980_t$ in Column (7), the interaction effect is not significant and even changes sign.

Not surprisingly, as seen in Table 5 and Table 6, college education increases the chance of mixed marriages for both Han men and Minority men. Urban minority men are more likely to marry across ethnic lines than their rural counterparts, whereas the effect of urban identity is not significant for Han men.

[Table 6 about here]

In sum, the estimates using $Post1980_t$ to measure benefits have the sign predicted in M2, but are not statistically significant. Thus, the effect of Han sex ratios on the marriage choice of minorities is weak.

Sex ratios across minorities – Predictions M2 Our theory predicts that the effect of sex ratios across minority groups has a positive effect on mixed marriages for a minority man, i.e., $\beta_m > 0$ in:

$$M_{i,p,t} = \beta_m(\frac{M^{man}}{M^{woman}})_{g,t-20} + pref_p + marry_t + X_i + prov \times t + \varepsilon_{i,p,t.}$$

Further, due to the findings in the family stage, our theory implies that state policies have little effect on top of sex ratios, i.e., θ_m should be close to 0 if we run:

$$\begin{aligned} M_{i,p,t} &= \theta_m(\frac{M^{man}}{M^{woman}})_{g,t-20} \times Post1980_t + \beta_m(\frac{M^{man}}{M^{woman}})_{g,t-20} \\ &\quad + pref_p + marry_t + X_i + X_i \times Post1980_t + prov \times t + \varepsilon_{i,p,t}. \end{aligned}$$

The results are presented in Table 7. Consistent with our theory, sex ratios across minority groups have a strong positive effect on the probability of marrying a Han for a minority man. A one standard-deviation increase in the minority sex ratio (0.7) increases the probability of entering a mixed marriage for a minority man by about 8 percentage points, which is about 80% of the mean probability for a minority man. Meanwhile, as predicted, the interaction material benefits and sex ratios has a negative significant effect on the mixed marriage choice for a Minority man. Once again, both urban identity and college education increase the chance of marrying a Han woman for a Minority man.

[Table 7 about here]

Together, these estimates are entirely consistent with Predictions M2.

6 Conclusions

We provide a framework to link the ethnic choice for children with interethnic marriage. Our model is constructed in such a way to be consistent with a set of motivating facts for China. It also delivers a rich set of auxiliary predictions. The empirical tests on Chinese microdata generally find support for these predictions. More generally, our results speak to two rarely studied issues: the interplay between incentives and social norms, and the interplay between sex ratios and interethnic marriage patterns.

References

[1] Benabou, Roland, and Jean Tirole (2011), "Laws and Norms", NBER Working Paper, No 17579.

[2] Bisin, Alberto, and Thierry Verdier (2000), "Beyond the Melting Pot: Cultural Transmission, Marriage, and the Evolution of Ethnic and Religious Traits", *Quarterly Journal of Economics* 115, 955-988.

[3] Botticini, Maristella and Zvi Eckstein (2007), "From Farmers to Merchants, Conversions, and Diaspora: Human Capital and Jewish History," *Journal of the European Economic Association* 5, 885-926.

[4] Cassan, Guilhem (2013), "Identity-Based Policies and Identity Manipulation: Evidence from Colonial Punjab", mimeo, University of Namur.

[5] Edlund, Lena, Hongbin Li, Junjian Yi and Junsen Zhang (2013), "Sex Ratios and Crime: Evidence from China", *Review of Economics and Statistics*, forthcoming.

[6] Green, Elliott (2011), "Endogenous Identity", mimeo, London School of Economics .

[7] Guo, Zhigang and Rui Li (2008), "Cong Renkoupuchashuju kan zujitonghun fufu de hunling, shengyushu jiqi zinu de minzu xuanze" (Marriage Age, Number of Children Ever Born, and the Ethnic Identification of Children of Inter-ethnic Marriage: Evidence from China Population Census in 2000), Shehuixue Jianjiu (*Research on Sociology*) 5, 98-116.

[8] Horowitz, Donald (2000), *Ethnic Groups in Conflict*, University of California Press.

[9] Jewitt, Ian (2004), "Notes on the Shape of Distributions", mimeo, Oxford University.

[10] Li, Xiaoxia (2004), "Zhongguo geminzujian zujihunyin de xianzhuang fenxi" (An Analysis of Interethnic Marriages in China), Renkou yanjiu (*Population Research*) 3, 68-75.

[11] Wei, Shang-jin and Xiaobo Zhang (2011), "The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China," *Journal of Political Economy* 119, 511-564.

Appendix: Proofs

Proof of consistency with the facts in Figure 2 We wish to establish that the model is consistent with the fourth fact in the introduction, i.e.,

$$\frac{d\pi^H}{d\lambda} > 0 \texttt{ and } \frac{d\pi^M}{d\lambda} < 0.$$

Proof. Consider the two FOCs (with interior solutions):

$$\begin{aligned}
\delta C'(\delta\alpha^H) &= -(1-P^H)V^H \\
\delta C'(\delta\alpha^M) &= (1-P^M)b - (1-P^M)V^M,
\end{aligned}$$

where $P^H = \frac{(1-A^H)\lambda S^H}{(1-A^H)\lambda S^H + (1-A^M)(1-\lambda)S^M}$ and $P^M = \frac{(1-A^M)(1-\lambda)S^M}{(1-A^M)(1-\lambda)S^M + (1-A^H)\lambda S^H}$. The derivatives of these probabilities are:

$$\begin{aligned}
\frac{dP^H}{d\lambda} &= \frac{(1-A^H)(1-A^M)S^H S^M}{[(1-A^H)\lambda S^H + (1-A^M)(1-\lambda)S^M]^2}; \\
\frac{dP^M}{d\lambda} &= -\frac{(1-A^H)(1-A^M)S^H S^M}{[(1-A^M)(1-\lambda)S^M + (1-A^H)\lambda S^H]^2}; \\
\frac{dP^H}{dA^H} &= -\frac{\lambda S^H(1-A^M)(1-\lambda)S^M}{[(1-A^H)\lambda S^H + (1-A^M)(1-\lambda)S^M]^2}; \\
\frac{dP^H}{dA^M} &= \frac{(1-A^H)\lambda S^H(1-\lambda)S^M}{[(1-A^H)\lambda S^H + (1-A^M)(1-\lambda)S^M]^2}; \\
\frac{dP^M}{dA^M} &= -\frac{(1-A^H)(1-\lambda)S^M\lambda S^H}{[(1-A^M)(1-\lambda)S^M + (1-A^H)\lambda S^H]^2}; \\
\frac{dP^M}{dA^H} &= \frac{(1-A^M)(1-\lambda)S^M\lambda S^H}{[(1-A^M)(1-\lambda)S^M + (1-A^H)\lambda S^H]^2}.
\end{aligned}$$

Differentiating the FOCs, one gets the comparative statics:

$$\begin{aligned}
(\delta)^2 C''(\delta\alpha^H)\frac{d\alpha^H}{d\lambda} &= V^H\{\frac{dP^H}{d\lambda} + \frac{dP^H}{dA^H}\frac{d\alpha^H}{d\lambda} + \frac{dP^H}{dA^M}\frac{d\alpha^M}{d\lambda}\} \\
(\delta)^2 C''(\delta\alpha^M)\frac{d\alpha^M}{d\lambda} &= -(b-V^M)\{\frac{dP^M}{d\lambda} + \frac{dP^M}{dA^H}\frac{d\alpha^H}{d\lambda} + \frac{dP^M}{dA^M}\frac{d\alpha^M}{d\lambda}\},
\end{aligned}$$

which can be written on matrix form as:

$$\begin{bmatrix} (\delta)^2 C''(\delta\alpha^H) - V^H \frac{dP^H}{dA^H} & -V^H \frac{dP^H}{dA^M} \\ (b - V^M)\frac{dP^M}{dA^H} & (\delta)^2 C''(\delta\alpha^M) + \frac{dP^M}{dA^M}(b - V^M) \end{bmatrix} \begin{bmatrix} \frac{d\alpha^H}{d\lambda} \\ \frac{d\alpha^M}{d\lambda} \end{bmatrix}$$
$$= \begin{bmatrix} V^H \frac{dP^H}{d\lambda} \\ -(b - V^M)\frac{dP^M}{d\lambda} \end{bmatrix}.$$

The determinant of the matrix on the left-hand side is given by

$$\nabla = \det \begin{bmatrix} (\delta)^2 C''(\delta\alpha^H) - V^H \frac{dP^H}{dA^H} & -V^H \frac{dP^H}{dA^M} \\ (b - V^M)\frac{dP^M}{dA^H} & (\delta)^2 C''(\delta\alpha^M) + \frac{dP^M}{dA^M}(b - V^M) \end{bmatrix}$$
$$= (\delta)^2 C''(\delta\alpha^H)(\delta)^2 C''(\delta\alpha^M) + (\delta)^2 C''(\delta\alpha^H)\frac{dP^M}{dA^M} - V^H \frac{dP^H}{dA^H}(\delta)^2 C''(\delta\alpha^M).$$

Clearly, $\nabla < 0$ for $(\delta)^2 C''(\delta\alpha^H) < V^H \frac{dP^H}{dA^H}$ and $(\delta)^2 C''(\delta\alpha^M) < -(b - V^M)\frac{dP^M}{dA^M}$ as stated in Assumption 1.

Similarly, we have:

$$\det \begin{bmatrix} V^H \frac{dP^H}{d\lambda} & -V^H \frac{dP^H}{dA^M} \\ -(b - V^M)\frac{dP^M}{d\lambda} & (\delta)^2 C''(\delta\alpha^M) + \frac{dP^M}{dA^M}(b - V^M) \end{bmatrix}$$
$$= V^H \frac{dP^H}{d\lambda}(\delta)^2 C''(\delta\alpha^M) < 0,$$

and

$$\det \begin{bmatrix} (\delta)^2 C''(\delta\alpha^H) - V^H \frac{dP^H}{dA^H} & V^H \frac{dP^H}{d\lambda} \\ (b - V^M)\frac{dP^M}{dA^H} & -(b - V^M)\frac{dP^M}{d\lambda} \end{bmatrix}$$
$$= -(\delta)^2 C''(\delta\alpha^H)(b - V^{M,H})\frac{dP^M}{d\lambda} > 0.$$

Therefore, we get:

$$\frac{d\alpha^H}{d\lambda} = \frac{V^H \frac{dP^H}{d\lambda}(\delta)^2 C''(\delta\alpha^M)}{\nabla} > 0,$$
$$\frac{d\alpha^M}{d\lambda} = \frac{-(\delta)^2 C''(\delta\alpha^H)(b - V^M)\frac{dP^M}{d\lambda}}{\nabla} < 0.$$

Since there is a one-to-one mapping from α to π, we get:

$$\frac{d\pi^H}{d\lambda} > 0 \text{ and } \frac{d\pi^M}{d\lambda} < 0.$$

■

Proof of results M1 and M2 Next, we want to verify that:

$$\text{(M1):} \qquad \frac{d\pi^H}{dS^H} > 0 \text{ and } \frac{d\pi^M}{dS^H} < 0; \frac{d^2\pi^H}{dS^H db} > 0 \text{ but } \frac{d^2\pi^M}{dS^H db} > 0.$$

$$\text{(M2):} \qquad \frac{d\pi^M}{dS^M} > 0 \text{ but } \frac{d^2\pi^M}{dS^M db} < 0.$$

Proof. Similar to the case above, we know $\frac{dP^H}{dA^H}$,$\frac{dP^H}{dA^M}$,$\frac{dP^M}{dA^M}$ and $\frac{dP^M}{dA^H}$. We also know that:

$$\begin{aligned}
\frac{dP^H}{dS^H} &= \frac{(1-A^H)\lambda(1-A^M)(1-\lambda)S^M}{[(1-A^H)\lambda S^H + (1-A^M)(1-\lambda)S^M]^2}, \\
\frac{dP^M}{dS^H} &= -\frac{(1-A^M)\lambda S^M(1-A^H)(1-\lambda)}{[(1-A^M)(1-\lambda)S^M + (1-A^H)\lambda S^H]^2}.
\end{aligned}$$

Therefore, we can solve for $\frac{d\alpha^H}{dS^H}$ and $\frac{d\alpha^M}{dS^H}$ from:

$$\begin{aligned}
(\delta)^2 C''(\delta\alpha^H)\frac{d\alpha^H}{dS^H} &= V^H\{\frac{dP^H}{dS^H} + \frac{dP^H}{dA^H}\frac{d\alpha^H}{dS^H} + \frac{dP^H}{dA^M}\frac{d\alpha^M}{dS^H}\} \\
(\delta)^2 C''(\delta\alpha^M)\frac{d\alpha^M}{dS^H} &= -(b-V^M)\{\frac{dP^M}{dS^H} + \frac{dP^M}{dA^H}\frac{d\alpha^H}{dS^H} + \frac{dP^M}{dA^M}\frac{d\alpha^M}{dS^H}\},
\end{aligned}$$

which can be written on matrix form:

$$\begin{bmatrix} (\delta)^2 C''(\delta\alpha^H) - V^H\frac{dP^H}{dA^H} & -V^H\frac{dP^H}{dA^M} \\ (b-V^M)\frac{dP^M}{dA^H} & (\delta)^2 C''(\delta\alpha^M) + (b-V^M)\frac{dP^M}{dA^M} \end{bmatrix} \begin{bmatrix} \frac{d\alpha^H}{dS^H} \\ \frac{d\alpha^M}{dS^H} \end{bmatrix}$$
$$= \begin{bmatrix} V^H\frac{dP^H}{dS^H} \\ -(b-V^M)\frac{dP^M}{dS^H} \end{bmatrix}.$$

The determinant

$$\nabla = \det \begin{bmatrix} (\delta)^2 C''(\delta\alpha^H) - V^H\frac{dP^H}{dA^H} & -V^H\frac{dP^H}{dA^M} \\ (b-V^M)\frac{dP^M}{dA^H} & (\delta)^2 C''(\delta\alpha^M) + (b-V^M)\frac{dP^M}{dA^M} \end{bmatrix}$$

is again negative under Assumption 1.

Thus,

$$\frac{d\alpha^H}{dS^H} = \frac{V^H\frac{dP^H}{dS^H}(\delta)^2 C''(\delta\alpha^M)}{\nabla} > 0,$$

and

$$\frac{d\alpha^M}{dS^H} = \frac{-(\delta)^2 C''(\delta\alpha^H)(b - V^M)\frac{dP^M}{dS^H}}{\nabla} < 0.$$

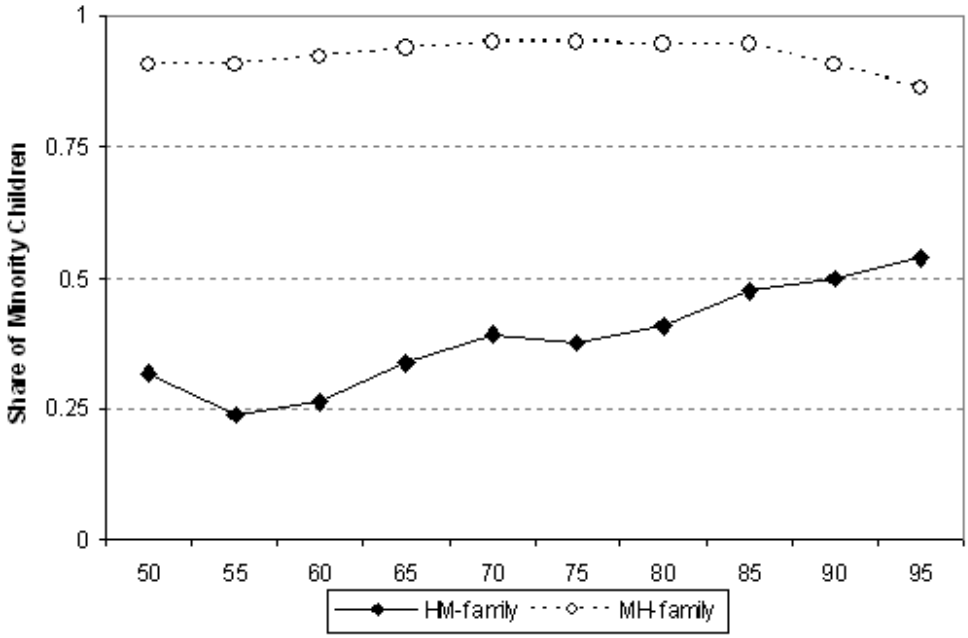
Then we have: $\frac{d\pi^H}{dS^H} > 0$ and $\frac{d\pi^M}{dS^H} < 0$.

Moreover, $\frac{d^2\pi^H}{dS^H db} > 0$ because V^H is increasing in b and $\frac{d^2\pi^M}{dS^H db} > 0$ because $-(b - V^M)$ is decreasing in b.

$$\frac{d\alpha^M}{dS^M} = \frac{-(b-V^M)\frac{dP^M}{dS^M}(\delta)^2 C''(\delta\alpha^H)}{\nabla} > 0.$$

Therefore, (M1) and (M2) follow. ■

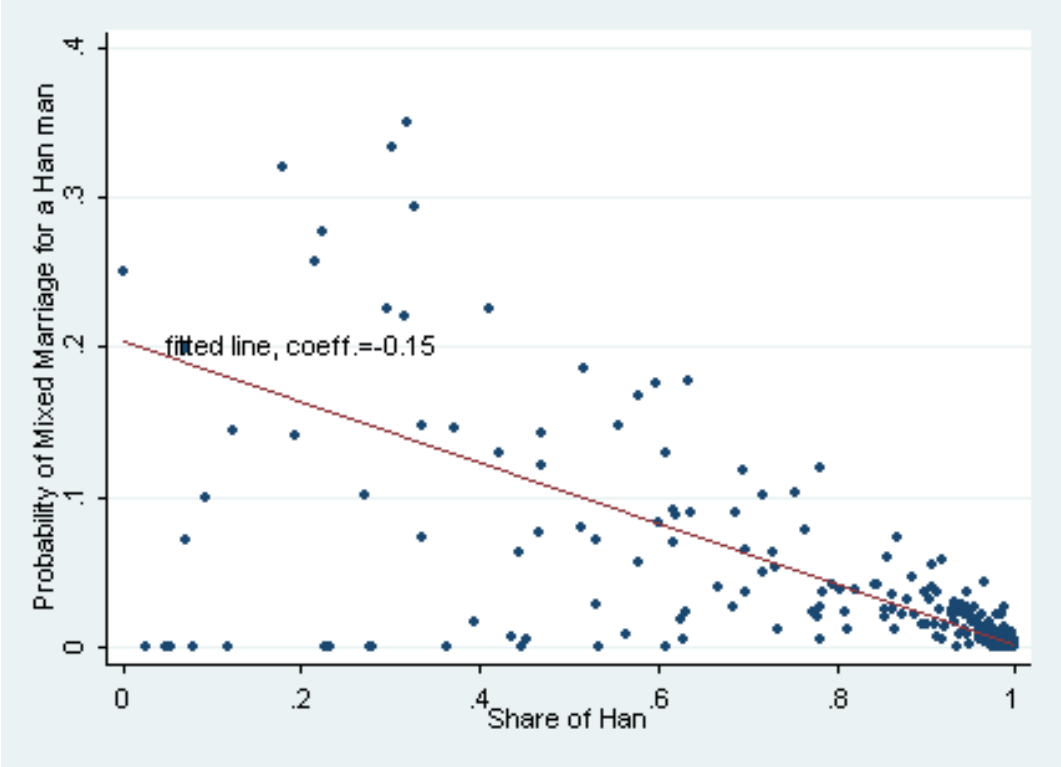
Figure 1: The Share of Minority Children By Birth Cohorts



Notes: This figure displays the share of minority children in mixed marriages by cohorts. It shows that (1) the children are more likely to be a minority in mixed marriages with a male-minority and (2) there is a increasing trend of minority chidren in mixed marriages with a male-Han.

Figure 2: Han Population Share and Mixed Marriages

(a) Han Men



(b) Minority Men

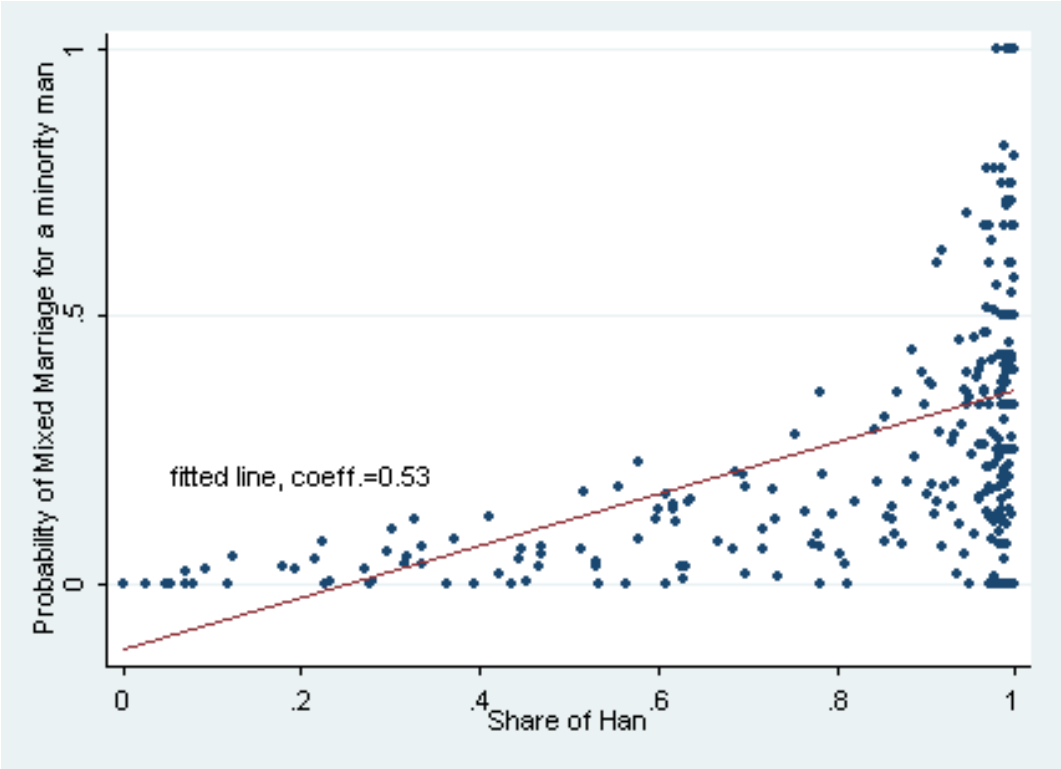
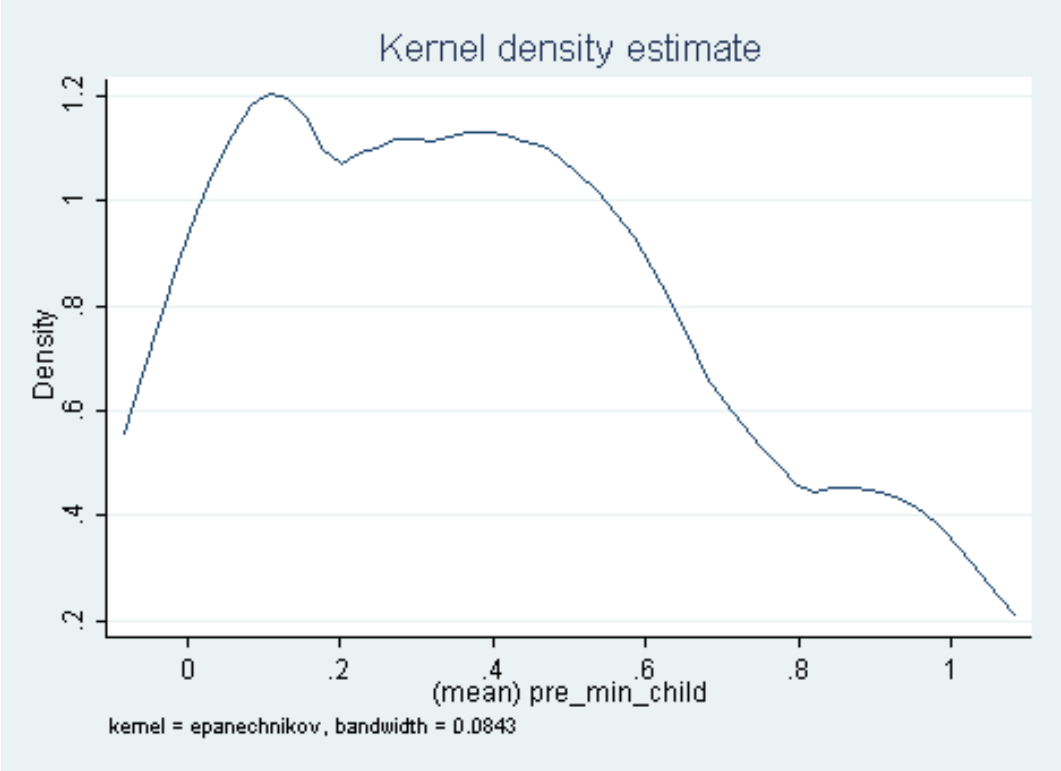


Figure 3: Distribution of Social Norms

(a) HM-Families



(b) MH-Families

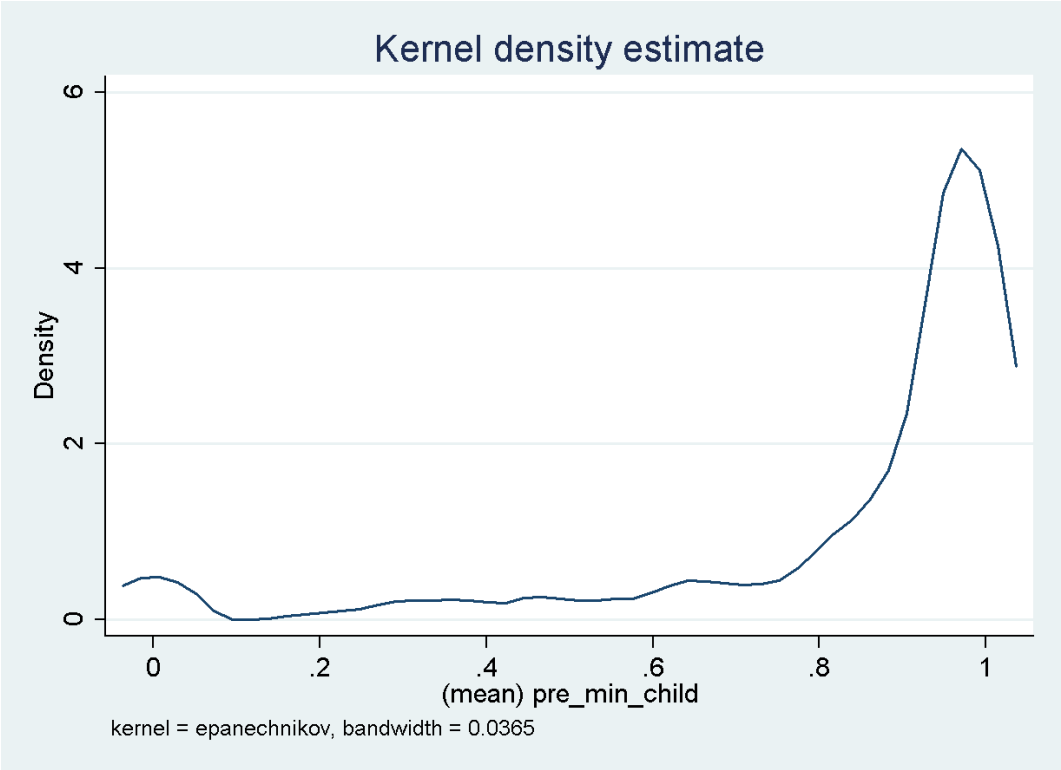
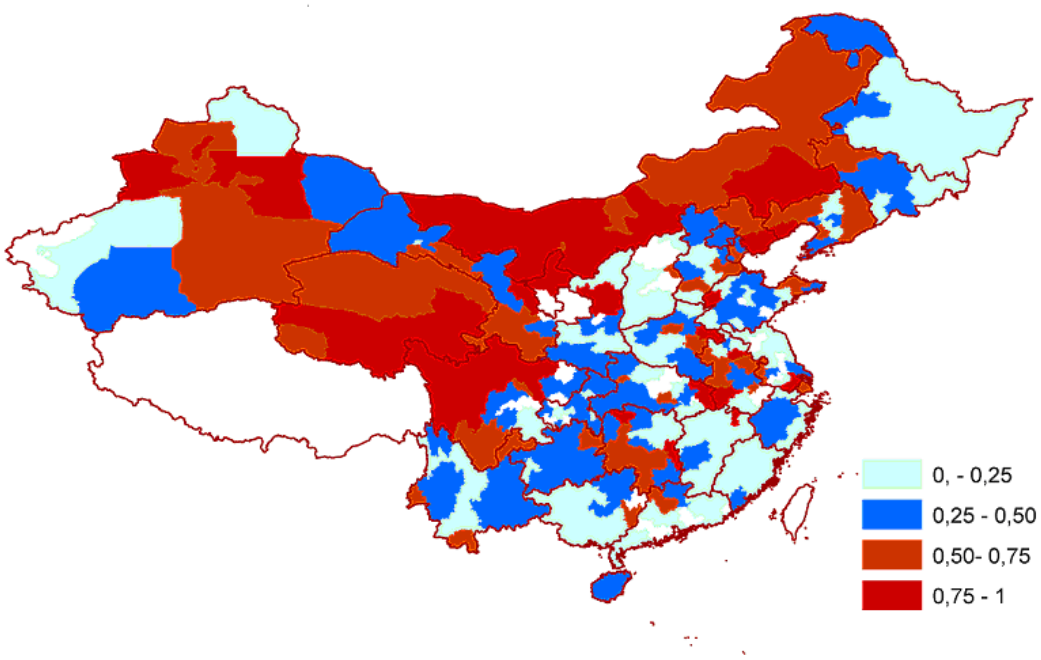


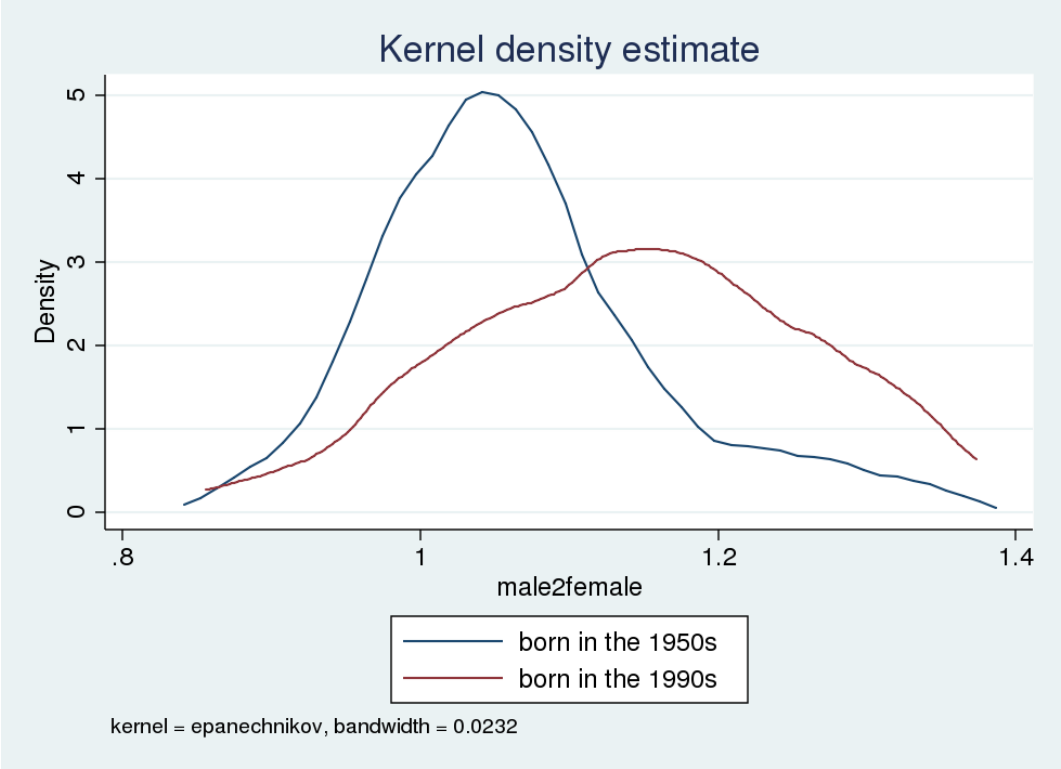
Figure 4: Spatial Distribution of Social Norms (for the 1970s cohort)



Notes: This figure maps the average share of minority children in mixed marriages with a male-Han. The share is calculated based on the 1970s cohorts.

Figure 5: Distribution of Sex Ratios

(a) Han Sex Ratios (across prefectures)



(b) Minority Sex Ratios (across province-ethnicities)

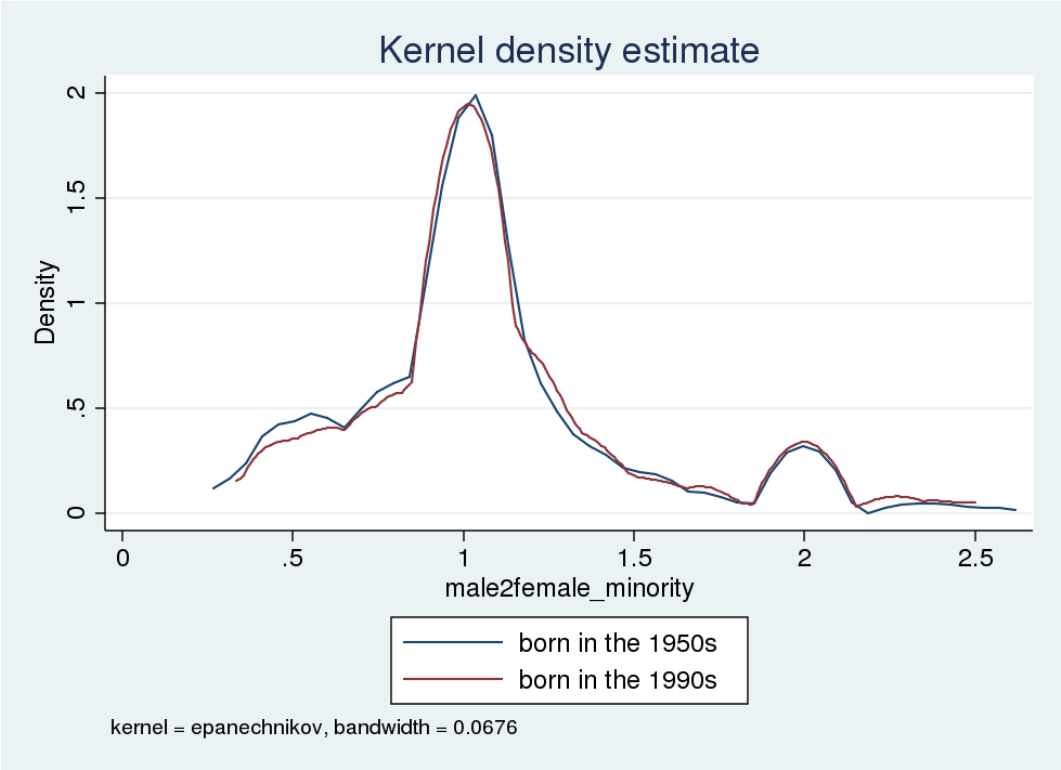
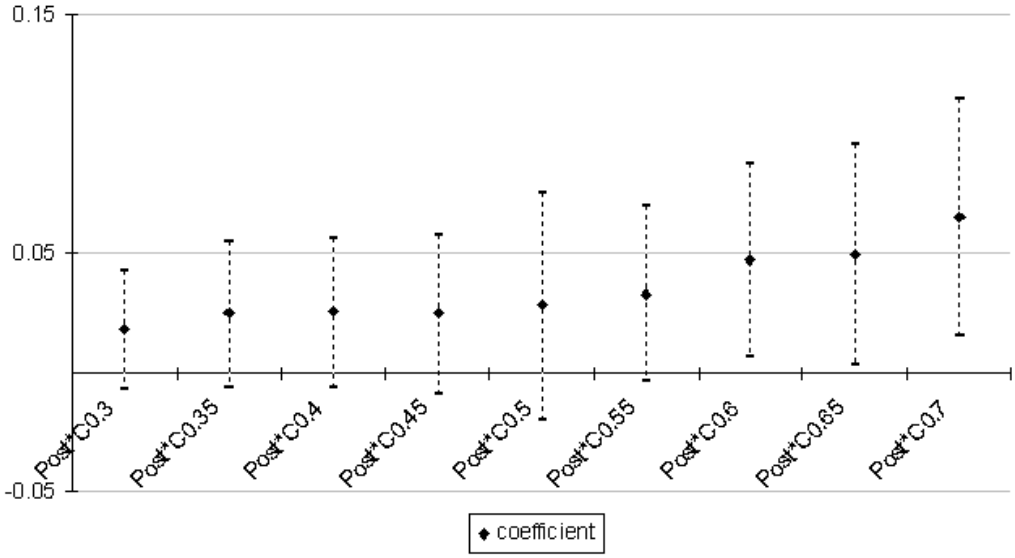


Figure 6: The effect of material benefits * social norms



Notes: This figure plots the results for Prediction F1 using different cutoff values. The dimonds indicate the coefficients and the dashed lines indicate the 95% confidence intervals.

Table 1: Summary Statistics

	(1)	(2)
Panel A: Children in Mixed Families (Censuses 1982-2005)		
	HM-family	MH-family
Minority Child	0.40	0.94
	(0.49)	(0.24)
Born after 1980	0.43	0.38
	(0.50)	(0.49)
Minority Child in 1970s	0.39	0.95
	(0.26)	(0.10)
Zhuang Wife	0.17	
	(0.38)	
Religious Wife	0.17	
	(0.37)	
Observations	97399	94420
Panel B: Mixed Marriages (Censuses 2000-2005)		
	Han Man	Minority Man
Mixed Marriage	0.014	0.118
	(0.118)	(0.332)
College Education	0.08	0.06
	(0.27)	(0.24)
Observations	735875	73478
Urban Identity	0.32	0.20
	(0.47)	(0.40)
Observations	735447	73424

Table 2A: Material Benefits and Social Norms I: Norms are defined by prefecture-the 1970s cohort

	(1) MinorChild	(2) MinorChild	(3) MinorChild	(4) MinorChild	(5) MinorChild
I(=0.45)*Born after 1980	0.025				
	(0.017)				
I(=0.50)*Born after 1980		0.029*			
		(0.017)			
I(=0.55)*Born after 1980			0.033*		
			(0.019)		
I(=0.60)*Born after 1980				0.047**	
				(0.021)	
I(=0.65)*Born after 1980					0.050**
					(0.023)
Prefecture FE	Y	Y	Y	Y	Y
Birth Cohort FE	Y	Y	Y	Y	Y
Province Trends	Y	Y	Y	Y	Y
# clusters	346	346	346	346	346
# observations	97399	97399	97399	97399	97399

Notes: The standard errors are clustered at the prefecture level. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 2B: Material Benefits and Social Norms II: Norms are defined by prefecture-previous cohort

	(1) MinorChild	(2) MinorChild	(3) MinorChild	(4) MinorChild	(5) MinorChild
I(=0.45)*Born after 1980	0.022				
	(0.015)				
I(=0.50)*Born after 1980		0.023			
		(0.016)			
I(=0.55)*Born after 1980			0.015		
			(0.015)		
I(=0.60)*Born after 1980				0.038**	
				(0.018)	
I(=0.65)*Born after 1980					0.040**
					(0.020)
Prefecture FE	Y	Y	Y	Y	Y
Birth Cohort FE	Y	Y	Y	Y	Y
Province Trends	Y	Y	Y	Y	Y
# clusters	346	346	346	346	346
# observations	97399	97399	97399	97399	97399

Notes: The standard errors are clustered at the prefecture level. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 2C: Material Benefits and Social Norms for Rural children: norms are defined by prefecture-cohort-residency

	(1) MinorChild	(2) MinorChild	(3) MinorChild	(4) MinorChild	(5) MinorChild
I(=0.45)*Born after 1980	0.045**				
	(0.021)				
I(=0.50)*Born after 1980		0.040*			
		(0.021)			
I(=0.55)*Born after 1980			0.053***		
			(0.016)		
I(=0.60)*Born after 1980				0.038**	
				(0.019)	
I(=0.65)*Born after 1980					0.048**
					(0.024)
Prefecture FE	Y	Y	Y	Y	Y
Birth Cohort FE	Y	Y	Y	Y	Y
Province Trends	Y	Y	Y	Y	Y
# clusters	299	299	299	299	299
# observations	8786	8786	8786	8786	8786

Notes: The standard errors are clustered at the prefecture level. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 2D: Material Benefits and Social Norms for Urban children: norms are defined by prefecture-cohort-residency

	(1) MinorChild	(2) MinorChild	(3) MinorChild	(4) MinorChild	(5) MinorChild
I(=0.45)*Born after 1980	0.160***				
	(0.031)				
I(=0.50)*Born after 1980		0.154***			
		(0.028)			
I(=0.55)*Born after 1980			0.154***		
			(0.028)		
I(=0.60)*Born after 1980				0.173***	
				(0.030)	
I(=0.65)*Born after 1980					0.166***
					(0.030)
Prefecture FE	Y	Y	Y	Y	Y
Birth Cohort FE	Y	Y	Y	Y	Y
Province Trends	Y	Y	Y	Y	Y
# clusters	272	272	272	272	272
# observations	3164	3164	3164	3164	3164

Notes: The standard errors are clustered at the prefecture level. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 3: Heterogenous material benefits

	(1) MinorChild	(2) MinorChild	(3) MinorChild	(4) MinorChild	(5) MinorChild
Zhuang Wife*Post1980	-0.060***	-0.054***	-0.026**	-0.022**	
	(0.014)	(0.014)	(0.011)	(0.011)	
Zhuang Wife*Post Policy					-0.044***
					(0.011)
Zhuang Wife	-0.133***	-0.134***	-0.144***	-0.157***	-0.138***
	(0.039)	(0.039)	(0.034)	(0.036)	(0.035)
Born after 1980	0.092***				
	(0.012)				
Post Policy					0.022***
					(0.008)
Prefecture FE	Y	Y	Y	Y	Y
Birth Cohort FE		Y	Y	Y	Y
Province Trends			Y	Y	Y
# clusters	346	346	346	338	339
# observations	97399	97399	97399	73835	95753

Notes: The table shows that having a Zhuang Wife (and hence enjoying fewer material benefits with a minority child) cuts the effect of material benefits (measured by the post 1980 dummy). Column (5) uses the one-child policy timing in Edlund et al. (2013) to measure material benefits.
The standard errors are clustered at the prefecture level. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 4: Material Benefits and Intrinsic Costs

	(1) MinorChild	(2) MinorChild	(3) MinorChild	(4) MinorChild	(5) MinorChild	(6) MinorChild	(7) MinorChild	(8) MinorChild
Son * Born after 1980	-0.012**	-0.016***	-0.007					
	(0.006)	(0.006)	(0.006)					
Son * Post Policy				-0.007				
				(0.006)				
Religious Wife * Born after 1980					-0.035**	-0.037**	-0.009	
					(0.016)	(0.016)	(0.012)	
Religious Wife * Post Policy								-0.018
								(0.015)
Son	-0.004	0.000	-0.009**	-0.010***				
	(0.004)	(0.004)	(0.004)	(0.004)				
Religious Wife					0.110***	0.111***	0.093***	0.088***
					(0.017)	(0.017)	(0.017)	(0.017)
Born after 1980	0.087***				0.087***			
	(0.011)				(0.012)			
Post Policy				0.017**				0.017**
				(0.008)				(0.008)
Prefecture FE	Y	Y	Y	Y	Y	Y	Y	Y
Birth Cohort FE		Y	Y	Y		Y	Y	Y
Province Trends			Y	Y			Y	Y
# clusters	346	346	346	339	346	346	346	339
# observations	97399	97399	97399	95753	95578	95578	95578	93932

Notes: Columns (4) and (8) use the one-child policy timing in Edlund et al. (2013) to measure material benefits.
The standard errors are clustered at the prefecture level. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 5: The effect of Han Sex ratios on a Han man's marriage

	(1) HM*100	(2) HM*100	(3) HM*100	(4) HM*100	(5) HM*100	(6) HM*100	(7) HM*100
Male2Female * Post1980				2.140**	2.154**	2.145**	
				(1.029)	(1.024)	(1.031)	
Male2Female * Post Policy							3.111***
							(1.036)
Male2Female Ratio	1.707***	1.730***	1.806***	0.504	0.519	0.604	0.691
	(0.502)	(0.500)	(0.506)	(0.643)	(0.640)	(0.644)	(0.608)
Post Policy							-3.040***
							(1.062)
Urban * Post1980						-0.026	0.046
						(0.083)	(0.089)
College * Post1980						0.041	0.008
						(0.133)	(0.163)
Urban		0.052	0.050		0.054	0.069	-0.026
		(0.058)	(0.058)		(0.059)	(0.074)	(0.065)
College		0.569***	0.590***		0.568***	0.560***	0.607***
		(0.073)	(0.074)		(0.073)	(0.115)	(0.137)
Prefecture FE	Y	Y	Y	Y	Y	Y	Y
Marirage Cohort FE	Y	Y	Y	Y	Y	Y	Y
Province Trends			Y			Y	Y
# clusters	336	336	336	336	336	336	327
# observations	714079	713674	713674	714079	713674	713674	638588

Notes: Column (7) uses the one-child policy timing in Edlund et al. (2013) to measure material benefits.
The standard errors are clustered at the prefecture level. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 6: The effect of Han Sex ratios on a minority man's marriage

	(1) MH*100	(2) MH*100	(3) MH*100	(4) MH*100	(5) MH*100	(6) MH*100	(7) MH*100
Male2Female * Post1980				5.324	5.513	7.275	
				(5.740)	(5.702)	(5.711)	
Male2Female * Post Policy							-2.147
							(4.254)
Male2Female Ratio	-1.973	-1.245	-1.042	-5.210	-4.596	-5.718*	-0.923
	(2.328)	(2.422)	(2.430)	(3.540)	(3.543)	(3.446)	(3.203)
Urban * Post1980						4.210***	4.399***
						(0.957)	(0.916)
College * Post1980						1.135	0.152
						(2.212)	(2.075)
Urban		10.076***	9.950***		10.079***	7.127***	7.235***
		(0.899)	(0.877)		(0.900)	(1.091)	(1.108)
College		3.690***	3.922***		3.684***	2.511	3.007*
		(1.104)	(1.094)		(1.103)	(2.051)	(1.762)
Prefecture FE	Y	Y	Y	Y	Y	Y	Y
Marriage cohort FE	Y	Y	Y	Y	Y	Y	Y
Province Trends			Y			Y	Y
# clusters	326	326	326	326	326	326	318
# observations	56777	56728	56728	56777	56728	56728	54205

Notes: Column (7) uses the one-child policy timing in Edlund et al. (2013) to measure material benefits.
The standard errors are clustered at the prefecture level. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table 7: The effect of minority Sex ratios on a minority man's marriage

	(1) MH*100	(2) MH*100	(3) MH*100	(4) MH*100	(5) MH*100	(6) MH*100	(7) MH*100
Male2Female * Post 1980				-24.705**	-25.061**	-23.913**	
				(10.240)	(10.146)	(10.215)	
Male2Female * Post Policy							-22.935**
							(8.959)
Male2Female Ratio	11.840**	12.046**	11.830**	25.981***	26.385***	25.607***	22.294***
	(5.668)	(5.485)	(5.477)	(7.831)	(7.676)	(7.673)	(6.781)
Urban * Post 1980						4.731***	5.435***
						(0.941)	(0.813)
College * Post 1980						0.697	-0.192
						(1.754)	(1.733)
Urban		8.842***	8.699***		8.831***	5.305***	5.121***
		(0.853)	(0.837)		(0.854)	(0.951)	(0.942)
College		3.352***	3.557***		3.376***	2.464	2.577*
		(0.942)	(0.933)		(0.945)	(1.561)	(1.457)
Prefecture FE	Y	Y	Y	Y	Y	Y	Y
Birth cohort FE	Y	Y	Y	Y	Y	Y	Y
Province Trends			Y			Y	Y
# clusters	321	321	321	321	321	321	314
# observations	66046	65999	65999	66046	65999	65999	64670

Notes: Column (7) uses the one-child policy timing in Edlund et al. (2013) to measure material benefits.
The standard errors are clustered at the prefecture level. *** significant at 1%, ** significant at 5%, * significant at 10%.

A Web Appendix

Table W1: Ethnicity of Children in HM families versus in MH families

	(1) MinorChild	(2) MinorChild	(3) MinorChild
MH Marriage	0.538***	0.519***	0.520***
	(0.031)	(0.030)	(0.030)
Prefecture FE		Y	Y
Birth Cohort FE			Y
# clusters	348	348	348
# observations	191819	191819	191819

Notes: The table shows that the first fact displayed in Figure 1 is also true at the individual level.
The standard errors are clustered at the prefecture level. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table W2: Ethnicity of Children in HM families versus in MH families

	(1) MinorChild	(2) MinorChild	(3) MinorChild	(4) MinorChild
Born in the 1950s	-0.029		0.019	
	(0.057)		(0.038)	
Born in the 1960s	0.023		0.043	
	(0.053)		(0.037)	
Born in the 1970s	0.090*		0.055	
	(0.053)		(0.036)	
Born in the 1980s	0.137**		0.048	
	(0.054)		(0.036)	
Born in the 1990s	0.207***		0.007	
	(0.056)		(0.036)	
Born after 1980		0.081***		-0.008
		(0.010)		(0.006)
Prefecture FE	Y	Y	Y	Y
# clusters	346	346	336	336
# observations	97399	97399	94420	94420

Notes: The table shows that the second fact displayed in Figure 1 is also true at the individual level.
The standard errors are clustered at the prefecture level. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table W3: Mixed Marriage for a Han man and a Minority Man

	(1) Mixed Marriage	(2) Mixed Marriage	(3) Mixed Marriage	(4) Mixed Marriage
Minority Man	0.104***	0.092***	0.091***	0.092***
	(0.001)	(0.014)	(0.014)	(0.014)
Urban				0.009***
				(0.001)
College				0.009***
				(0.001)
Prefecture FE		Y	Y	Y
Birth cohort FE			Y	Y
# clusters		345	345	345
# observations	809353	809353	809353	808871

Notes: The table shows that the third fact is also true at the individual level.
The standard errors are clustered at the prefecture level. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table W4: Han Share and Mixed Marriages

	(1) HM	(2) HM	(3) HM	(4) MH	(5) MH	(6) MH
Han Share	-0.224***	-0.223***	-0.223***	0.436***	0.334***	0.307***
	(0.002)	(0.023)	(0.023)	(0.011)	(0.051)	(0.051)
Urban			0.000			0.109***
			(0.001)			(0.012)
College			0.005***			0.056***
			(0.001)			(0.012)
Province FE		Y	Y		Y	Y
Birth cohort FE		Y	Y		Y	Y
# clusters		312	312		304	304
# observations	599617	599617	599245	37528	37528	37493

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The table shows that the fourth fact displayed in Figure 2 is also true at the individual level. As population shares are stable within a prefecture, we control for province fixed effects rather than prefecture fixed effects.
The standard errors are clustered at the prefecture level. *** significant at 1%, ** significant at 5%, * significant at 10%.

Table W5: Use One-Child Policy Timing to Measure Material Benefits

	(1) MinorChild	(2) MinorChild	(3) MinorChild	(4) MinorChild	(5) MinorChild
I(=0.45)*Post	0.027**				
	(0.013)				
I(=0.50)*Post		0.031**			
		(0.013)			
I(=0.55)*Post			0.032***		
			(0.012)		
I(=0.60)*Post				0.037***	
				(0.011)	
I(=0.65)*Post					0.036***
					(0.011)
Prefecture FE	Y	Y	Y	Y	Y
Birth Cohort FE	Y	Y	Y	Y	Y
Province Trends	Y	Y	Y	Y	Y
# clusters	339	339	339	339	339
# observations	95753	95753	95753	95753	95753

Notes: The table shows that results using one-child policy timing in Edlund et al. (2013) to measure material benefits. These results show that the estimates in Table 2A in the main text are robust.
The standard errors are clustered at the prefecture level. *** significant at 1%, ** significant at 5%, * significant at 10%.