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

Ruixue Jia† and Torsten Persson‡

February 11, 2014

Abstract

This paper provides a framework to link the ethnic choice for children with interethnic marriage. Our model is constructed 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 Roland Benabou, Gerard Roland, Jean Tirole, and particpants in the UCSD development lunch and the LSE/UCL development seminar for helpful comments, and 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 do institutions and government policy interventions shape ethnic identification? The answer to this question has important implications. For instance, conflicts could be exacerbated by political institutions that induce individuals to identify with specific ethnic groups (Horowitz, 2000). Ethnic identification is the broad topic of the seminal paper by Bisin and Verdier (2000). Motivated by a large sociological literature, these authors set the task for themselves to theoretically understand why cultural convergence is so slow, even in the US. They model the persistent propensity of ethnic and religious minorities to marry within their own kin and socialize their children in the same mold.

In much of the literature, the desire to identify with a certain ethnicity has immaterial motives with social and psychological roots, such as a desire for social recognition or self esteem. Yet, history is ripe with examples of groups that gradually or suddenly change their identity to reap material benefits. For instance, Bates (1974) discusses how economic and political change drove emerging ethnic groups to compete for the spoils of patronage in post-colonial Africa. The case studies in Vail (1989) describe how people in different parts of southern Africa, when trying to cope with the process of change, came to identify with vaguely defined ethnic groups in colonial times, and how these became major interest groups in post-colonial times. Botticini and Eckstein (2007) demonstrate how material incentives played an important historical role in individual transitions between Judaism and Christianity. Cassan (2012) shows how higher-caste 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.

Such cultural switchovers may still reflect a tradeoff between extrinsic material benefits and intrinsic costs shaped by existing self-images or social norms. Which way intrinsic motivations tilt that tradeoff is far from clear, however. Indeed, recent theoretical work by Benabou and Tirole (2011) shows that extrinsic incentives to make a certain choice may be either crowded out or crowded in by intrinsic incentives.

Most of the existing literature on ethnic policies focuses on choices by a single generation. However, it is easy to imagine that such choices also entail intergenerational aspects: e.g., the ethnicity mixed couples are expected to transmit to their children may affect decisions in the marriage market. An analogous intergenerational link is indeed present in the model which is formulated and structurally estimated by Bisin, Topa and Verdier (2001) of how people choose marriage partners across religious groups and how the resulting families socialize their children in the religious domain.

China is an interesting testing ground when it comes to government ethnic policies and family choices. A multiethnic society with 55 officially recognized ethnicities beyond the dominant Han, China is still relatively homogenous, despite some ethnic tensions with occasional riots in Tibet and Xinjiang. Meanwhile, the population share of ethnicities displays a great deal of regional dispersion: the minority share ranges from 0.3% in Jiangxi province to 94% in Tibet. Also, the national and provincial governments have made policy interventions that remind of "affirmative action" for US minorities. Moreover, mixed ethnic couples are free to choose whichever of their two ethnicities for their children and we can observe these choices directly in the data.

A few facts on ethnicity of children and mixed marriages stand out from the Chinese data (the censuses 1982, 1990, 2000 and a mini-census 2005 – see Section 3 for more detail on sources). One is:

F1 *The propensity to choose minority identity for 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.*

The probability of having minority children for minority-man and Han-man mixed marriages are 94 percent vs. 41 percent on average. Figure 1 plots this probability over time, by five-year birth cohorts, for the two types of mixed marriages. The figure illustrates a second fact:

F2 *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 minority identity among the children of such couples is 36 percent in cohorts born before 1980 but 45 percent in cohorts born after 1980. Differently, we observe little change in mixed couples with a minority man – 94 percent have minority children in cohorts born before 1980 against 93 percent after 1980.

[Figure 1 about here]

When it comes to mixed marriages, we observe:

F3 *The frequency of marrying across ethnic lines is much smaller for Han men than for minority men.*

These frequencies are 1.4 percent versus 11.8 percent, where the latter is an average across minority groups. Moreover:

F4 *In a cross-sectional comparison across China's prefectures, the wedge in the frequency of mixed marriage is clearly increasing in the Han population share.*

Panels A and Panels B in Figure 2 plot the share of Han population in a prefecture against the probability of mixed marriage for Han men and minority men for cohorts married in the 1980s (plots look similar for other marriage cohorts). Clearly, the Han population share is negatively associated with mixed marriages for Han men (the slope of the fitted line is around -0.15), but positively associated with mixed marriages for minority men (the slope of the fitted line is around 0.52). Any convincing theoretical explanation of ethnic choices in China should be able to reproduce facts F1 through F4.

[**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 pattern similar to F1, relying on the 0.095-percent sample of the 2000 census. These authors find that the average probability of having a minority child is more than one half, and argue that this raises the 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 distinction between marrying a Han and marrying a minority (regardless of which 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 may be important to allow for differences in population shares and religiosity.

To the best of our knowledge, no existing research has systematically analyzed ethnic decisions 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 by parents of their children's ethnicity and the decision to marry across ethnic lines. Our paper tries to fill these two gaps.

We do this in two steps. First, we set up 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. Our model is constructed to be consistent with facts F1-F4 on the choices of ethnicity for children and mixed marriages. But the model also delivers several auxiliary predictions.

In a second step, we take these auxiliary predictions to Chinese micro-data. For example, we empirically evaluate the interplay between social norms and incentives on ethnic choices. We are not aware of any existing empirical work on how social norms alternatively crowd in or crowd out material incentives, as in the theoretical work by Benabou and Tirole (2011). A similar methodology may also apply to other contexts. We also examine how sex ratios, together with material benefits, affect inter-ethnic marriage. This contributes to an existing literature that evaluates the consequences of imbalanced sex ratios in China, in which Wei and Zhang (2011) show that higher male-female ratios might explain a large part of increased saving rates in China, whereas Edlund et al. (2013) document that higher sex ratios lead to more crimes. Marriage search is an important underlying mechanism of both studies.

In what follows, we next formulate our model, where agents choose how to search in marriage markets and what ethnicity to pick for their children. We show that the model implies facts F1-F4, and spell out a number of additional model predictions. In Section 3, we discuss which data can be used to test these predictions. In Section 4, we confront the model's auxiliary predictions with the data and present our econometric results. Section 5 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 model the determinants of mixed marriages, and the ethnicity choices for children in such marriages. The model has two connected stages: a marriage stage and a child stage. Given their information, agents at the former stage have rational expectations about outcomes at the latter. Hence, we consider the stages in reverse order. For the child 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). But we extend their setup to encompass two different groups. For the marriage stage, we use a framework with costly directed search similar to the one in Bisin and Verdier (2000, 2001) to model directed 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 within ethnic groups. As further discussed in Section 3, we can measure most of these determinants at the regional (province or prefecture) level and some at the individual level. We should thus think about the model as capturing these individual or regional conditions. While the model certainly is highly stylized, it is not only consistent with facts F1-F4, but it also yields a number of additional predictions which we take to the data in Section 4.

2.1 The Child 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

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

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* individual cost of having a minority child (different from the Han man's own ethnicity). Its first component is the average stigma perceived individually by the household when the ethnicity of their child does not coincide with that of the Han man. By definition, this component is common and deterministic to everyone in the same region, but it could differ across regions; it could also differ across ethnicities depending on "cultural or linguistic distance". The second component ε governs the variation in intrinsic cost and constitutes the main source of heterogeneity in the model. We assume that ε 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. We think about ε as a value specific to each match that is only revealed to the household once the man and woman have entered into marriage.

The final term in (1) 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 as 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 relative to the household's individual payoff. 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 household in the same region and birth cohort, but there could also be separate within-region-cohort peer groups, say households with or without higher education, or households in urban vs. rural areas.

Conformity and the cutoff rule For the analysis to follow, it is useful to define the variable

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

Following the terminology in Benabou and Tirole (2011), the first term on the RHS of (2) can be interpreted as the social "honor" within the household's peer group of having a child of the man's own identity. This will be the choice of households with sufficiently high ε. The second term, deducted from this honor, is the social "stigma" for the Han-man household of having a child with identity different than Han. This will be the choice of households with a sufficiently low value of ε. For short, we will use the label of *conformity* when referring to the difference Δ below. Conformity captures the gain in social reputation (or self-image) that the household faces when giving its child the same ethnicity as the Han man rather than minority ethnicity.

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

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

The second equality implicitly defines a cutoff value of ε, below which agents have minority children, as a function of b, e and μ. Based on this cutoff rule we can write equilibrium conformity as

$$\Delta(\varepsilon^*) = E(\varepsilon \mid \varepsilon > \varepsilon^*) - E(\varepsilon \mid \varepsilon < \varepsilon^*) \ . \tag{4}$$

By the definition of truncated means, $\Delta(\varepsilon^*)$ is always positive. Note also that for the whole peer group, social reputation is like a zero-sum game: under a veil of ignorance about ε, the ex ante expected value of $\mu E(\varepsilon \mid m)$ is zero.[1]

The properties of the equilibrium and its comparative statics will crucially reflect the sign of $\frac{d\Delta}{d\varepsilon}$, i.e., the derivative of conformity. Suppose ε^* goes up such 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 conformity Δ has a unique interior minimum, so $\frac{d\Delta}{d\varepsilon} < 0$ for low values of ε^*, when few Han have minority kids, and $\frac{d\Delta}{d\varepsilon} > 0$ for high values of ε^*, when many Han have minority kids. It follows that ethnicity choices for children are strategic complements when $\frac{d\Delta}{d\varepsilon} < 0$, while they are strategic substitutes when $\frac{d\Delta}{d\varepsilon} > 0$. For some of the results below, we also assume that the second derivative is positive $\frac{d^2\Delta}{d\varepsilon^2} > 0$.

[1] Specifically, intergrating the reputational terms across all individuals in equilibrium, we obtain:

$$\int^{\varepsilon^*} E(v \mid \varepsilon < \varepsilon^*) d\varepsilon + \int_{\varepsilon^*} E(\varepsilon \mid \varepsilon > \varepsilon^*) d\varepsilon = 0$$

Minority-Han mixed couples In a M, H mixed couple, where the man is minority rather than Han, the preference function analogous to (1) can be written:

$$v + mb - (1 - m)(e(M) + \varepsilon) + \mu E(\varepsilon \mid m) \ , \tag{5}$$

where $e(M)$ and ε now represent the average and heterogeneous parts of the intrinsic cost of having a Han child, different from the minority man's own ethnicity in analogy with the decision for Han-man mixed household. We specifically assume that both the distribution function G for ε and the weight on social reputation μ are exactly the same in the two types of families in the same locality. This is a strong assumption, although one can think of arguments why μ, say, could be either higher or lower among minorities than majorities – the former may be more eager to fit in or more eager to preserve their identities. We do not pursue this issue further, however. The main argument for this is measurement: since proxies for μ and the distributions of ε would be very hard to find in available data, any prospective theoretical prediction would risk to be empirically empty.

The mixed couple will have a Han child when $b + \mu E(\varepsilon \mid m = 1) > -(e(M) + \varepsilon) + \mu E(\varepsilon \mid m = 0)$. Defining conformity is an analogous way as before – i.e., Δ is the difference between the honor of having a minority child, $\mu E(\varepsilon \mid m = 1)$, and the stigma of having a Han child, $\mu E(\varepsilon \mid m = 0)$ – we can write the condition for having a minority child as

$$\varepsilon > -b - e(M) - \mu\Delta(\varepsilon^*) = \varepsilon^*_M \ .$$

Given the symmetric distribution of ε, this condition is equivalent to:

$$\varepsilon < b + e(M) + \mu\Delta(\varepsilon^*) = \varepsilon^*_M(b, e(M), \mu) \ . \tag{6}$$

The fraction of mixed households with a minority man that will have a minority child is thus $G(\varepsilon^*_M(b, e(M), \mu))$.

Comparison across mixed marriages (ε^*_M versus ε^*_H) Having formulated the child stage of the model, we show that its predictions on the ethnicity choices for children are consistent with facts F1 and F2 noted in the introduction.

It follows from (6) and (3) that $\varepsilon^*_M(b, e(M), \mu) > \varepsilon^*_H(b, e(H), \mu)$. Since the two c.d.f.s are the same, this means that $G(\varepsilon^*_M) > G(\varepsilon^*_M)$ – i.e., minority children are more frequent in mixed marriages where the man is minority rather than Han. The intuition is straightforward: on average, minority men experience both material benefits b and immaterial benefits $e(M)$ of a minority child, so – compared to Han men – more of them will choose minority identity for the children. Clearly, this prediction is consistent with fact F1 about the average status of children in different types of mixed marriages.

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

$$\frac{\partial m^H(b, e, \mu)}{\partial b} = g(\varepsilon^*_H(b, e, \mu)) \frac{1}{1 + \mu \frac{d\Delta(\varepsilon^*_H(b,e,\mu))}{d\varepsilon}} > 0 \ . \tag{7}$$

Similarly, defining $m^M(b, e, \mu) = G(\varepsilon^*_M(b, e, \mu))$, the effect of extrinsic incentives (b) for a minority-Han family is:

$$\frac{\partial m^M(b, e, \mu)}{\partial b} = g(\varepsilon^*_M(b, e, \mu)) \frac{1}{1 + \mu \frac{d\Delta(\varepsilon^*_M(b,e,\mu))}{d\varepsilon}} > 0 \ . \tag{8}$$

Thus, higher material benefits raises the probability of having a minority child in both types of families. But we can say more. Comparing the two expressions, we note that $g(\varepsilon^*_M(b, e, \mu))$ is smaller than $g(\varepsilon^*_H(b, e, \mu))$, i.e., minority-man couples having Han children is more of a tail event than Han-man couples having minority children. Moreover, for the derivatives of conformity, we have $\frac{d\Delta(\varepsilon^*_M(b,e,\mu))}{d\varepsilon} > \frac{d\Delta(\varepsilon^*_H(b,e,\mu))}{d\varepsilon}$, i.e., the marginal Han-man mixed couple is in a region where fewer families have minority children than the marginal minority-man couple. For the marginal Han man having a minority child is thus a strategic complement rather than a strategic substitute; alternatively; or, if it is a strategic substitute, the substitutability is smaller than for the marginal minority man. This means that the immaterial incentives are more likely to crowd in rather than crowd out material incentives for mixed couples with Han men; or crowd them out less than for mixed couples with minority men. Thus, the effect of material incentives is higher for Han-minority families.

This prediction, *together with* the fact that benefits for minority children have gone up over time (see Section 3 for a discussion of the benefits), makes the model consistent with fact F2: a more pronounced trend over time to have minority kids in mixed marriages with Han men than in those with minority men (recall Figure 1).

Having established the link between our model and facts F1 and F2, we turn our interest to three auxiliary predictions from the model. These are the ones we will test empirically.

Material benefits (b) and social norms (Δ_ε) Our first prediction concerns the strength of the interaction effect between material benefits and social norms. Here we will focus on the effects on mixed households with Han men. From (7), we see that net material benefits are crowded in by social reputation – i.e., the social multiplier $\frac{1}{1+\mu\frac{d\Delta(\varepsilon^*_M(b,e,\mu))}{d\varepsilon}}$ is larger than 1 – when few people have minority kids and their ethnicity choices are strategic complements (i.e., when $\frac{d\Delta(\varepsilon^*_H(b,e,\mu))}{d\varepsilon} < 0$). Instead, benefits are crowded out when many people have minority kids ($\frac{d\Delta(\varepsilon^*_H(b,e,\mu))}{d\varepsilon} > 0$).

But the effect of a change in benefits also includes the density $g(\varepsilon^*_M)$ at the cutpoint. Suppose we compare two localities with cutpoints equidistant from zero, i.e., $\varepsilon^*_n > 0$ and $-\varepsilon^*_n$. Because the density is symmetric around zero, we have $g(-\varepsilon^*_n) = g(\varepsilon^*_n)$. In this case, the relative effects of material benefits is only governed by the different social multipliers. This implies a specific prediction across regions (or more generally across peer groups):

C1 *Suppose all prefectures (peer groups) in a province (prefecture) experience the same increase in benefits, due to a provincial policy. Then, comparing prefectures (peer groups) where the share of minority children in male-Han mixed marriages is small with those where it is large, we should see a larger positive effect in the former on the probability of having minority kids.*[2]

[2] Note that we also get different comparative statics for minority-Han families. As $\frac{d\Delta}{d\varepsilon}$ 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 (8), this also requires that

In the data, we will evaluate prediction C1 by comparing prefectures above and below some intermediate share of minority children. Another way to get at the interaction of material benefits and social norms is to consider the comparative statics for cutpoints at different points in the distribution, say different quartiles. To see this, let ε_n^*, $n = 1, 2, 3, 4$ be cutpoints at the midpoint of each quartile in the ε distribution. Since density g is symmetric around its mid-point at 0, we have $\varepsilon_1^* = -\varepsilon_4^*$ and $\varepsilon_2^* = -\varepsilon_3^*$. It immediately follows that $g(\varepsilon_1^*) = g(\varepsilon_4^*) < g(\varepsilon_2^*) = g(\varepsilon_3^*)$. Moreover, under the assumption that $\frac{d^2\Delta}{d\varepsilon^2} > 0$, the first derivatives in the social multiplier are monotonically ordered as: $\frac{d\Delta(\varepsilon_1^*)}{d\varepsilon^*} < \frac{d\Delta(\varepsilon_2^*)}{d\varepsilon^*} < 0 < \frac{d\Delta(\varepsilon_3^*)}{d\varepsilon^*} < \frac{d\Delta(\varepsilon_4^*)}{d\varepsilon^*}$. Using these facts in (7), we obtain another testable prediction over observables:

C1' *Suppose all prefectures* (*peer groups*) *in a province* (*prefecture*) *experience the same increase in benefits, due to a provincial policy. Then, comparing the effect on the share of minority children,* (i) *it is larger in the first, second and third quartile than in the fourth quartile,* (*ii*) *it is larger in the second quartile than in the third quartile,* (*iii*) *the relative effects in the first and second quartiles are ambiguous.*

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

C2 *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*) We have analyzed the effect of higher benefits of minority children on the probability of having minority children among mixed couples. Do the (average) intrinsic costs $e(H)$ of having minority children systematically alter this effect among Han-man mixed couples? In terms of the model notation, this is a question about the interaction effect of b and e on $m^H(b, e, \mu)$. This interaction effect can be written:

---

the density g is relatively flat across the two equilibrium points.) To test this prediction empirically, however, we need enough variation in ε_M^*. This is difficult since $G(\varepsilon_M^*)$ is close to 1 in most cases.

$$\frac{\partial^2 m^H(b,e,\mu)}{\partial b \partial e} = \frac{\partial^2 m^H(b,e,\mu)}{\partial b \partial \varepsilon_H^*} \frac{\partial \varepsilon_H^*}{\partial e} \ .$$

It follows from (7) that the first term on the right-hand side $\frac{\partial^2 m^H(b,e,\mu)}{\partial b \partial \varepsilon_H^*}$ itself includes two effects which depend on the cutoff value ε_H^*. The sign of the first effect depends on the change in the density $\frac{dg(\varepsilon_H^*(b,e,\mu))}{d\varepsilon^*}$, which is positive before the single peak of g and negative thereafter. The sign of the second effect is negative as it depends on the second derivative of conformity $\frac{d^2\Delta}{d\varepsilon^{*2}}$, which we have earlier assumed is positive; thus the social multiplier goes down as the cutoff increases. As for the second term on the right-hand side, $\frac{\partial \varepsilon_H^*}{\partial e}$, we know that it is negative – i.e., with higher intrinsic costs, fewer couples have minority kids. Putting these results together, we have:

C3 *The interaction effect of b and e on the share of male-Han mixed couples that have minority children is more likely to be negative if the share of minority children in the peer group of these couples is small, and the negative effect is weaker if this share is large.*

2.2 The Marriage Stage

To model the marriage market, we use a model of directed search similar to that in Bisin and Verdier (2000, 2001). 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 suppose that they are the active agents in the marriage market and thus we only consider the search behavior 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 child 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. 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 $\alpha^J = A^J$.

The slope parameter δ captures individual level search difficulties. For example, directed search towards your own ethnicity may be cheaper to conduct in an ethnically homogenous rural community than in a mixed city environment, which could be represented by different values of δ. But we will not pursue this line of argument here.

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 about the search technology imply that the probability of a Han man to marry a Han woman is:

$$\pi^H = \alpha^H + (1 - \alpha^H)P^H \ , \tag{9}$$

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. The corresponding probabilities π^M and P^M for minority men are defined accordingly.

An important assumption of the model is that men expend their search effort before any matches have been made. Therefore, they do not observe the match-specific value of ε they will draw together with the partners they will eventually marry.

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 (9), and where

$$V^H = G(\varepsilon_H^*)[b - e(H) - \mu\Delta(\varepsilon_H^*)] - \mu E(\widetilde{\varepsilon} \mid m = 0) = G(\varepsilon_H^*)(b - e(H)) \tag{10}$$

is the continuation value of such marriage which is obtained by taking expectations of the expression in (1). The second equality in (10) follows from fact that p.d.f. g is symmetric around zero. Because of this, the weighted sum of the two truncated means that make up the honor and stigma terms in the nonconformity expression sum to zero, which implies that $G(\varepsilon_H^*)[E(\widetilde{\varepsilon} \mid m = 0) - E(\widetilde{\varepsilon} \mid m = 1)] = -G(\varepsilon_H^*)\Delta(\varepsilon_H^*) = E(\widetilde{\varepsilon} \mid m = 0)$.

Thus, the objective function incorporates the expected outcome from the child stage of the model, given the man's information. 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. Not knowing the household-specific shock ε, the ex ante probability of having a minority child in such a marriage is given by the unconditional probability $G(\varepsilon_H^*)$ derived in the previous subsection. In this event, the man will reap additional extrinsic benefits b and suffer intrinsic cost $e(H)$. Thus, the social norms regarding the ethnicity choice for children – to the degree they affect the cutoff value ε_H^* – 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 made 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:

$$\delta C'(\delta\alpha^H) \geq -(1 - P^H)V^H \quad \text{c.s.} \quad \alpha^H \geqslant 0 \ . \tag{11}$$

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

$$V^M = G(\varepsilon_M^*)b - (1 - G(\varepsilon_M^*))e(M) \ . \tag{12}$$

This continuation payoff, obtained from (5), is different from that of a Han man. The minority man's probability of getting a minority child, and hence benefits b, is given by $\pi^M + (1 - \pi^M)G(\varepsilon_M^*)$, the probability of meeting a minority woman plus the probability of meeting a Han woman times the probability of having a minority child. With probability $(1-\pi^M)(1-G(\varepsilon_M^*))$ he enters a mixed marriage and gets a Han child and suffers an expected cost $e(M)$. (As for the Han man, the terms in social reputation cancel out in expectation.)

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 expression to rewrite π^M in terms of α^M and P^M, we can write the first-order condition to this problem as:

$$\delta C'(\delta\alpha^M) \geq (1-P^M)b - (1-P^M)V^M \quad \text{c.s.} \quad \alpha^M \geqslant 0 \ . \tag{13}$$

We can rewrite the RHS of the inequality as $(1-P^M)(b-V^M) = (1-P^M)(1-G(\varepsilon^*_M))[b+e(M)] > 0$.

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 $e(M)$ and foregoes extrinsic benefits of b.

To get unambiguous signs in the comparative statics for the marriage stage, we postulate the following for the rest of 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. If, as in most regions of China, the Han share of the population is large, we get the result that the frequency of marriages across ethnic lines is higher among minority men than among Han men, i.e., $1-\pi^H < 1-\pi^M$. This follows mechanically from the definitions of P^H and P^M.

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) is to raise the probability to meet a partner of Han ethnicity in the unrestricted (common) pool (P^H) for a Han man. This tends to decrease the probability of mixed marriages $1-\pi^H$. The effect of a higher Han population ratio is the opposite for a minority man.

These results, which are more or less machanical, make the model consistent with facts F3 and F4 in the introduction, about the average mixed-marriage propensity and its pattern across prefectures. Given that our model is consistent with these facts, we examine two 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 is magnified by higher 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 in M1 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 direct effect on a minority man is the opposite. Like the results on population shares, these results are straightforward implications of the relative sizes of the groups.

Less trivially, the interaction effects ($\frac{d^2\pi^H}{dS^H db}$ and $\frac{d^2\pi^M}{dS^H db}$) in M1 reflect the interaction between the child stage and the marriage stage. They are determined by the main effect we have just discussed and the continuation values in the child stage. For example, the continuation value for a Han man to marry a minority is increasing in material benefits b, whereas a higher b diminishes the gap between the continuation value of mixed marriages and within-ethnic marriages for a minority man.

The intuition for the effects of minority sex ratios on the choices among minorities in prediction M2 is similar, and these effects naturally have the opposite sign as the effects of Han sex ratios on the choices among the Han.

3 Data and Measurement

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

Linking of datasets We draw on two sources of data. The first involves three of China's censuses: the 1-percent samples of the 1982 and 1990 censuses, and the 0.095-percent sample of the 2000 census. Our second source is the 2005 population survey that covers about 1 percent of the population, also known as the mini-census. These data provide demographic information and some information on socioeconomic status for altogether about 25 million people. One drawback of the data is that they give the location of the household at the time of the respective census (or mini-census), rather than at the time of marriage or childbirth. Therefore, our results could be biased by migration. We deal with this prospective problem in a couple of ways below.

As in the model, we are interested in the husband-wife-children structure of households. The husband or wife data draws on the information about the gender of the head of household. 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 units we focus on are the areas defined by four-digit census codes: prefectures or cities. Considering that some areas change names and codes over time, we unify the boundaries based on the year 2000 information to end up with 348 prefectures and cities. Since over 330 of these are prefectures, we refer to all of them by this label.

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

In all these 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 1, 41 percent of the children in Han-minority families are minorities whereas 94 percent of the children in minority-Han families are minority. This is fact F1 in the introduction.

[Table 1 about here]

To study the inter-ethnic marriage decisions, we follow the model's marriage stage and ask whether a Han man marries 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 which are 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 2000 and 2005 census the probability of marrying a minority woman for a Han man is 1.4 percent. One reason for this small number is that the average population share of Han is above 90 percent. The probability of marrying a Han woman for a minority man is 11.8 percent. This difference is fact F3 highlighted in the introduction.

Tables W1 and W3 in the Web Appendix show that facts F1 and F3 are true not only at the aggregate level, but also at the individual level (which is the domain of the model), even when we control for prefecture fixed effects and cohort fixed effects.

Measuring material benefits (b in the model) We measure material benefits of minority children in alternative ways. The People's Republic of China (1949-) has employed different policies to the benefit of ethnic minorities. These policies cover three basic aspects: (i) Family planning. When the family-planning policy started in the 1960s, minorities were more favorably treated than the Han majority. Over time, there is also regional variation in the treatment of different minorities. (ii) 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 province. (iii) Employment. The national ethnic policy states that minorities should have favorable treatment in employment. However, explicit quotas for minority employment rarely exist. As minorities are often discriminated in employment, it is unclear that this policy would make people tend to choose minority identity for children.

It is not straightforward to quantify regional variation over time in these policies. Nevertheless, we try to do it in three ways:

(1) Timing. Since the 1980s, family planning was switched more strictly to one-child policy, whereas minority-minority couples are usually allowed to have two or more children.[3] Hence, the benefits of being minorities became larger after 1980. On top of this, the additional benefits of better opportunities in higher education are largely contemporaneous. Our first and basic measure of minority benefits is thus a dummy indicating post-1980 cohorts.

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 fact F2 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.

(2) The one-child policy. We can also measure minority benefits by exploiting the gradual rollout of one-child policy across provinces. The precise timing is based on the year when a province set up a family-planning organization (data is available for 27 provinces, which is used in the working-paper version of Edlund et al., 2013).[4] An advantage of this measure is that it is staggered across provinces as the organizations are established between the 1970s and the 1980s. A 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 related way to measure the minority advantage of the one-child policy

[3]Figure A1 in the web appendix illustrate the differences in fertility by age groups and marriage types. The one-child policy also allows other exceptions. For example, rural families can have a second child if the first child is a girl or is disabled.

[4]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).

is to use revealed fertility. It is only reasonable to define completed fertility rates for women aged over 40 and hence we focus on those born before 1960 as measured in the 2000 census. In particular, those born between 1955 and 1960 are more likely to be affected by the one-child policy, compared with those before before 1955. Therefore, we use the province-specific ratio of completed fertility of minority-minority couples to minority-Han couples for those born between 1955 and 1960 (or those born before 1955) as an alternative proxy for the minority advantage for those married after 1980 (or those married before 1980). The fertility differences are most relevant for minority men. As Figure 3 shows, the difference between the fertility of minority women and Han women that marry minority men has a distinct peak for the 1955-60 cohort (see the yellow and gray curves). Therefore, we only use them when examining the marriage choices of minority men. But for a Han man, the quantity of children does not differ much whether he marries a Han or a minority (see the blue and orange curves).

[**Figure 3 about here**]

(3) Heterogeneous benefits. The third measure we explore exploits 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 above 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 percent.

Measuring the effect of social norms ($\frac{d\Delta}{d\varepsilon}$ in the model) Following the discussion about crowding out or crowding in the model (the sign of $\frac{d\Delta}{d\varepsilon}$), we measure social norms primarily by previous shares of minority children in mixed marriages, separately for male-Han and male-Minority mixed marriages. Obviously, we want the social norms for a particular cohort to be predetermined. To make sure that our results are reasonably robust, we define the peer group relevant for the social norms in a three different ways.

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

(2) **Previous cohort in the same prefecture and the same type of mixed marriage.** Given 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 for the prevailing social norms in a cohort is to use the birth cohort from the previous decade in the same prefecture. For example, the 1980s cohort of mixed Han-men marriages in the prefecture becomes the peer group for the 1990s cohort, and so on.

(3) **Same residency and previous cohort in the same prefecture and the same type of mixed marriage**. Measures (1) and (2) use only ethnicity of the man, 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-ethnicity-cohort-residency level. A limitation of this method is that it implies smaller groups, due to the disaggregation itself and the fact that 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 4 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. Therefore, we focus on the effect of social norms for Han-minority families.

[Figure 4 about here]

Figure 5 further maps the spatial distribution across China of the ethnicity choices (based on the 1970s cohort) by male-Han mixed families. It indicates that social norms 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 $\frac{d\Delta}{d\varepsilon} < 0$ for low values of the cutoff ε^* (when few people have minority kids) and a strategic substitutability $\frac{d\Delta}{d\varepsilon} > 0$ for high values of ε^* (when many have minority kids). We do not observe the distribution of ε and thus cannot measure the critical cutoff value when the sign flips. Instead, we check how the estimates behave, as we vary the assumption about the critical cutoff value.

[Figure 5 about here]

Measuring intrinsic costs (e in the model) A first measure of intrinsic cost e that we use is whether the child is a son or a daughter. Consistent with the Confucian values, the intrinsic costs of having a son with different ethnicity are higher than for a daughter. 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. Of course, the men that marry religious women are a selected sample, but our question concerns how a religious wife shapes the effect of material benefits on ethnic choice for children, rather than the effect of a religious wife itself.

Table 1 shows that the share of male-Han mixed families with a religious wife is about 18 percent. 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 population share of the Han population by prefecture and 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 concern, we trim both the right and the left 5-percent 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 shares among those born in the 1980s.

This information is used to generate fact F4 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 S for the Han population, we calculate sex ratios by prefecture and 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 6 plots the distribution of sex ratios after trimming the upper and lower 5-percent 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-percent tails in the baseline estimates. Panel B of Figure 6 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 6 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 Han men 32 percent have urban identities and 8 percent have a college education. Among the minority men, 20 percent have urban identities and 6 percent have a 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 child stage, we use all censuses as our baseline. In that case, we present the results using only the 2000 and 2005 censuses with individual controls for urban identity and college education as a robustness check.

Migration The variations across prefectures and provinces discussed in this section, are calculated based on the residency of individuals at the census times. However, this residency may be different than birth place, due to migration. Only the 2000 census includes information whether an individual's birth place lies in the same county as his or her current residency (the 1982 and 1990 censuses spells out whether one lived in the same county five years ago, and the 2005 mini-census only has information on whether one lived in the same province one year ago). Based on the 2000 census, over 85 percent of individuals were born in the same county as their current residency, while 94 percent were born in the same province. Given that prefecture is the administrative level above county, these facts suggest that migration is unlikely to make a major difference for our main results. Moreover, Frijters, Gregory and Meng (2013) document that rural-urban migration did not take off until 1997.

We nevertheless conduct robustness checks by limiting the sample to the censuses until 2000, and by excluding individuals whose birth county and residency county are different. This should minimize the potential impact of migration.

4 Empirical Results

This section presents our empirical specifications and estimation results, beginning with the ethnic choices of children followed by the mixed marriages.

4.1 Ethnic Choice for Children

Material incentives and social norms – Predictions C1 and C1' 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 C1 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) rather than crowded out by prevailing social norms. To test this, we ask whether β_b is positive in the specification:

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

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

We use a dummy of for post-1980 cohorts, $Post1980_t$ to measure material benefits. $Cutoff_p$ is a dummy variable which indicates whether the peer group – defined in the three ways discussed in Section 3 – has a share of minority children smaller than some cutoff value. To be flexible, we use a wide range of cutoff values from 0.3 to 0.7. Thus, the parameter of interest β_b measures the difference in the effect of material benefits between prefectures below the assumed cutoff and prefectures above this cutoff.

To allow for an effect of prefecture characteristics that are time-invariant or change slowly over time, we control for prefecture fixed effects ($pref_p$). To hold constant factors that affect the ethnicity choices of different cohorts – including the direct effects of post1980 benefits – we also include birth-cohort fixed effects ($birth_t$, for every ten years). Finally, we include province-specific (linear) trends ($prov \times t$) to control for different evolutions across provinces, such as different growth rates or different provincial policies in other areas than ethnicity.

The results using different ways to measure the peer group relevant for social norms are similar. Here, we present two sets of results using the previous cohort as the relevant peer group. The results when the peer groups are instead defined by (i) the 1970s cohort and (ii) the previous cohort plus rural and urban residence are presented in the web appendix.

To save space, Table 2A presents the results for the cutoff range between 0.5 and 0.7 while Figure 7 visualizes all the results. As shown in column (1), the average effect of $Post1980_t$ is around 0.08. Moreover, the estimated effect of material incentives is indeed generally larger when the share is smaller than the cutoff value. This is consistent with the theoretical prediction that benefit have a larger effect in peer groups where few mixed households have minority children, because they are crowded in by a strategic complementarity (or crowded out less by a weaker strategic substitutability). The estimates suggest that the differences at the two sides of the cutoff can be half the average effect of material benefits – represented by the coefficient on $Post1980_t$ in column (1).

[Table 2A about here]
[Figure 7 about here]

Similarly, to test Prediction C1', we replace Cutoff_p with quartiles. Table 2B presents the quartile results, where the fourth quartile is left as the reference group. Column (1) shows the interaction effect with the the direct effects of post1980 benefits. Column (2) controls for birth-cohort effects while column (3) further controls for province-specific trends. Consistent with prediction C1', the effect is generally larger for the first, second and third quartile. In addition, the effect for the second quartile is larger than that for the third quartile (0.050 vs. 0.036). These effects are large: the difference in effects of higher material benefits in the three first quartiles vs. the fourth quartile is on the order of the average effect estimated in column 1 of Table 2A.

[**Table 2B about here**]

In the web appendix, we present three sets of robustness checks. First, we use alternative ways of measuring social norms. Table W5 reports results using social norms among mixed households with Han men within a prefecture by the share of minority children among such households in the 1970s cohort in the same prefecture. The results are a bit larger than those in Table 2A. Tables W6 and W7 present results for rural-residency and urban-residency members of the same ethnicity-prefecture-cohort, respectively, using the previous cohort to define the peer group. 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. This finding is consistent with the idea that social norms may have a sharper effect the more narrowly (and accurately) the peer group is defined.

Second, to deal with migration concerns, we show results after dropping all data after the 2000 census as well as individuals whose birth county and residency county are different in the 2000 census. The results are presented in Table W8. Compared with the results in Table 2A, the coefficients are smaller and less precisely estimated.

Third, we measure the benefits of minority kids by the provincial timing of the one-child policy rather than by $Post1980_t$. As shown in Table W9, the coefficients are a bit larger than those in Table 2A.

In summary, the data are clearly consistent with Predictions C1 and C1'. The results reported here constitute solid evidence for an important effect of peer-group dependent social norms on ethnicity choices.

Heterogeneity in material benefits – Prediction C2 Auxiliary prediction C2 of our model says that the effect of higher benefits should be smaller for mixed households where the man is Han and the wife is Zhuang rather than some other minority, simply because the Zhuang experienced a smaller increase in minority benefits. To test this, we check whether $\beta_z < 0$ in the specification:

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

The estimates are presented in Table 3. Column (1) shows the result controlling for $Post1980_t$ rather than cohort fixed effects ($birth_t$), whereas column (2) (and subsequent columns) includes these fixed effects. 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. Similar to the robustness checks in the previous test, Column (4) drops all data after the 2000 census as well as individuals whose birth county and residency county is different in the 2000 census. The magnitude of the estimate is similar to the one 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**]

Altogether, the econometric estimates seem consistent with Prediction C2 as well.

Material benefits and intrinsic costs – Prediction C3 Our final prediction about the child stage C3, concerns the interaction effect of material benefits and cultural distance on the choice of a minority child. Given that the average share of minority children for Han-Minority households is small (around 0.4), our model predicts that this interaction effect is negative. To measure intrinsic costs, we use a dummy for whether the child is a son and another dummy for whether the minority wife is religious (although we recognize that the selection of such a wife may not be random). Thus, we estimate:

$$\begin{aligned} MinChild_{i,p,t} &= \beta_s Post1980_t \times Son_i + \delta Son_i + \\ &\quad + pref_p + birth_t + prov \times t + \varepsilon_{i,p,t} \ . \ , \end{aligned}$$

and

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

expecting to find negative values of β_s and β_r.

The estimates are found in Table 4. In Columns (1) and (5), we represent the overall effects of the benefits by the $Post1980_t$ dummy rather than the cohort fixed effects $birth_t$. Columns (1)-(4) present the effect of having a son and columns (5)-(8) the effect of a wife that belongs to a religious minority. Columns (1)-(2) and columns (5)-(6) show that having a son as well as having a religious wife cuts 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. The results limiting the sample until the 2000 census are also similar.

[Table 4 about here]

On balance, we find that the estimates square with Prediction C3.

4.2 Inter-Ethnic Marriage

Han sex ratios and mixed marriages – Predictions M1 To study the links between sex ratios and mixed marriages, we first examine prediction M1 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 $t-20$:

$$
HanMixed_{i,p,t} = \beta_h \left(\frac{H^{men}}{H^{women}}\right)_{p,t-20} + pref_p + marry_t + X_i + prov \times t + \varepsilon_{i,p,t} \ .
$$

where $HanMixed_{i,p,t}$ is a dummy indicating marrying a minority or not for Han man i in prefecture p and cohort t. Since the mean is very low (1.4 percent), we multiply the dummy with 100 such that the results can be interpreted in terms of percentage points. As in the specification for children, we control for prefecture fixed effects and marriage cohort fixed effects

($marry_t$). Finally, 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:

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

Estimates are displayed 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 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 $X_i \times Post1980_t$ 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.

To minimize potential impacts of migration, Table W10 in the Web Appendix uses the 2000 census and excludes individuals whose birth county and residency county is different. The results are similar to those in Table 5.

Thus, when it comes to Han men the results on the effects of sex rations are consistent with Prediction M1.

[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 $MinMixed_{i,p,t}$. 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. Table W11 in the Web Appendix reports the results using the 2000 census excluding individuals whose birth county and residency counties are different. They are similar to those in Table 6.

Unsurprisingly, the results in Tables 5 and 6 say that 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 insignificant for Han men.

[**Table 6 about here**]

In sum, our estimates of how sex ratios among the Han affect mixed marriages have the sign predicted in M1. They are statistically significant for couples with Han men but not for couples with minority men.

Sex ratios across minorities – Predictions M2 Our auxiliary prediction M2 across minority groups says that a higher sex ratio in a certain minority has a positive effect on the probability that a man of that minority enters a mixed marriage, i.e., $\beta_m > 0$ in:

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

Further, this prediction holds that the policy benefits for minorities should have a dampening influence on the effect of sex ratios, i.e., θ_m should be negative if we run:

$$\begin{aligned} MinMixed_{i,p,t} \quad = \quad & \theta_m(\frac{M^{men}}{M^{women}})_{g,t-20} \times Post1980_t + \beta_m(\frac{M^{men}}{M^{women}})_{g,t-20} \\ & + pref_p + marry_t + X_i + X_i \times Post1980_t + prov \times t + \varepsilon_{i,p,t} \ . \end{aligned} \tag{14}$$

The results are presented in Table 7. Consistent with our theory, sex ratios across minority groups have a strong positive effect on the probability that a minority man marries a Han woman. A one standard-deviation increase in the minority sex ratio (0.7) increases the probability of entering a mixed marriage for a man from that minority by about 8 percentage points, which is about 80 percent of the mean probability for a minority man. Also as predicted, the interaction between material benefits and sex ratios has a negative significant effect on the mixed marriage choice for a minority man. As in Tables 5 and 6, both urban identity and college education increase a minority man's chance of marrying a Han woman. Once again, Table W12 in the Web Appendix uses the 2000 census only and excludes individuals whose birth county and residency counties are different and shows a similar pattern to that in Table 7.

[**Table 7 about here**]

Another way to explore the same hypotheses is to use the revealed fertility rates for women born between 1955 and 1960 in minority-minority and minority-Han couples to measure the one-child policy benefits of minority marriage. Specifically, we substitute these province specific fertility ratios for $Post1980_t$ in expression (13) above.[5] Column (1) in Table 8 shows that the minority-minority advantage decreases the likelihood of interethnic marriage for a minority man. Column (2) shows that this advantage indeed mitigates the effect of sex ratios on mixed marriages. Columns (3)-(4) also include invididual characteristics to show that the results are robust.

[**Table 8 about here**]

Together, these estimates are consistent with the predictions in M2.

5 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 we carry out on Chinese microdata generally find support for these predictions. More generally, our results speak to two issues

[5] Specifically, we use the ratio calculated for those born between 1955 and 1960 for those married in 1980 or later and use the ratio based on those born before 1955 for those married before 1980.

that have rarely been studied in the data. One issue is specific to China, namely the interplay between sex ratios and interethnic marriage patterns. The other issue is more general, namely the interplay between material incentives and social norms. Our methodology for empirically investigating this phenomenon could plausibly be applied in other contexts – e.g., in tax evasion – where the interplay between extrinsic and intrinsic motivation is also important.

References

[1] Bates, Robert (1974), "Ethnic Competition and Modernization in Contemporary Africa", *Comparative Political Studies* 6, 457-484.

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

[3] 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.

[4] Bisin, Alberto, Giorgo Topa, and Thierry Verdier (2001), "Religious Intermarriage and Socialization in the United States", *Journal of Political Economy* 112, 615-664.

[5] 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.

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

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

[8] Frijters, Paul, Robert Gregory, and Xin Meng (2013), "The Role of Rural Migrants in the Chinese Urban Economy", in Dustmann, Christian (ed.), *Migration–Economic Change, Social Challenge*, Oxford University Press.

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

[10] 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.

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

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

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

[14] Vail, Leroy (1989), *The Creation of Tribalism in Southern Africa*, University of California Press.

[15] 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 that model is consistent with fact F4 We wish to establish that the model implies fact F4 in the introduction, i.e.,

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

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:

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

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

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

■

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

$$\text{(M1):} \quad \frac{d\pi^H}{dS^H} > 0 \text{ and } \frac{d\pi^M}{dS^H} < 0; \frac{d^2\pi^H}{dS^H db} > 0, \text{ but } \frac{d^2\pi^M}{dS^H db} > 0.$$
$$\text{(M2):} \quad \frac{d\pi^M}{dS^M} > 0, \text{ but } \frac{d^2\pi^M}{dS^M db} < 0 .$$

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:

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

which can be written on matrix form:

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

The determinant

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

is again negative under Assumption 1.

Thus,

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

and

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

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

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

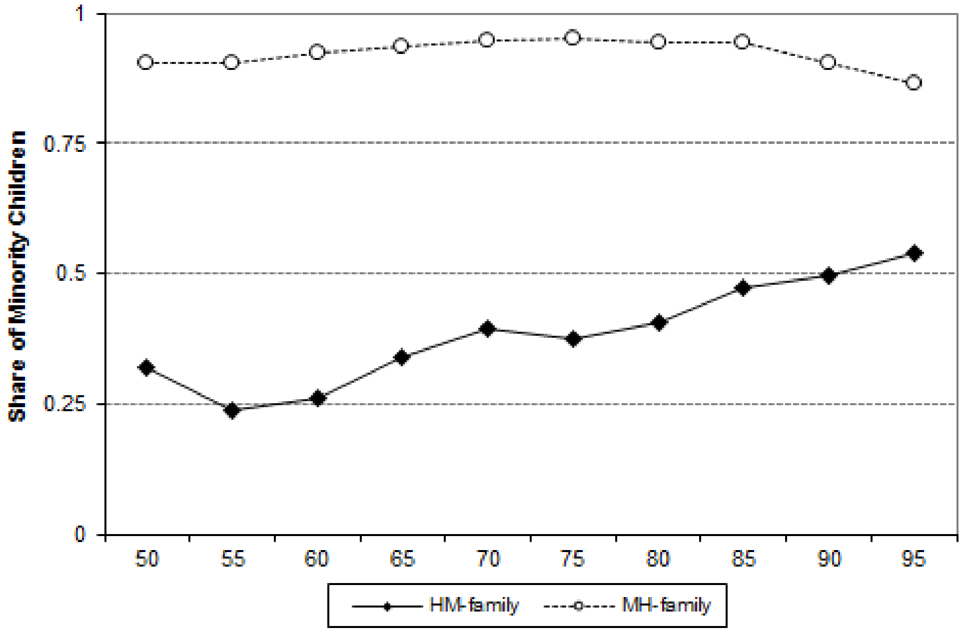
Finally, we have:

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

Also, $\frac{d^2\pi^M}{dS^M db} < 0$ as because $-(b - V^M)$ is decreasing in b.

Therefore, (M1) and (M2) follow. ■

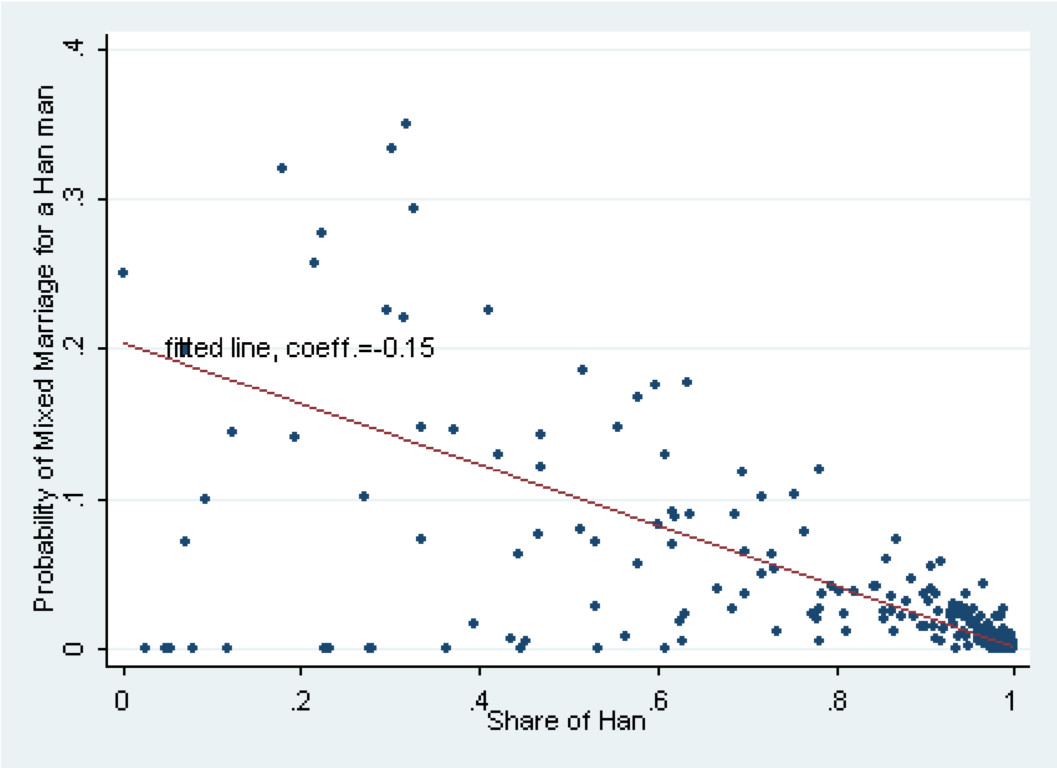
Figure 1: The Share of Minority Children By Birth Cohorts



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

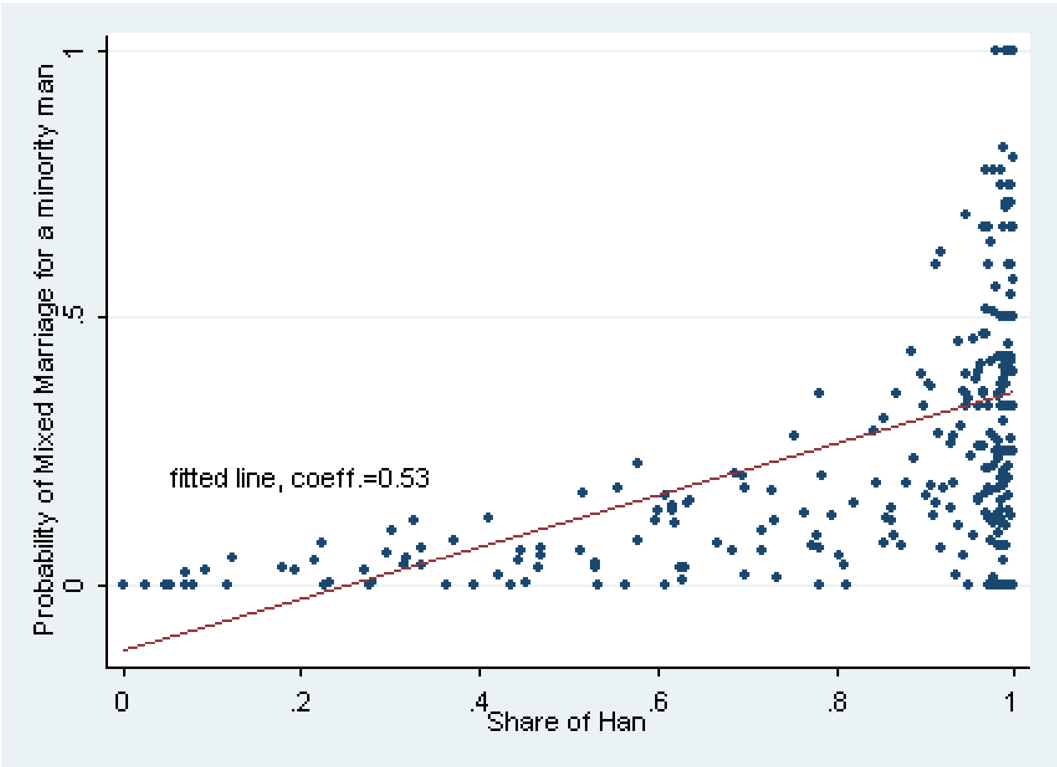
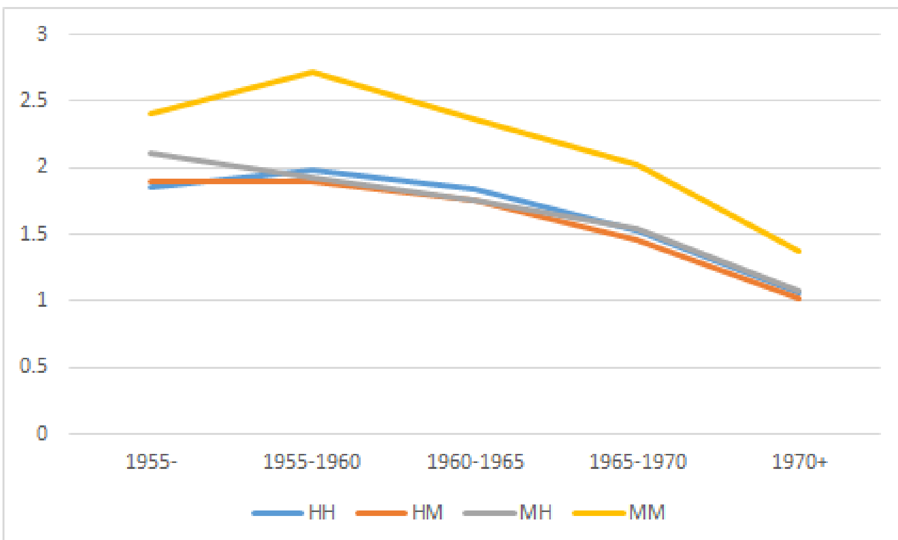


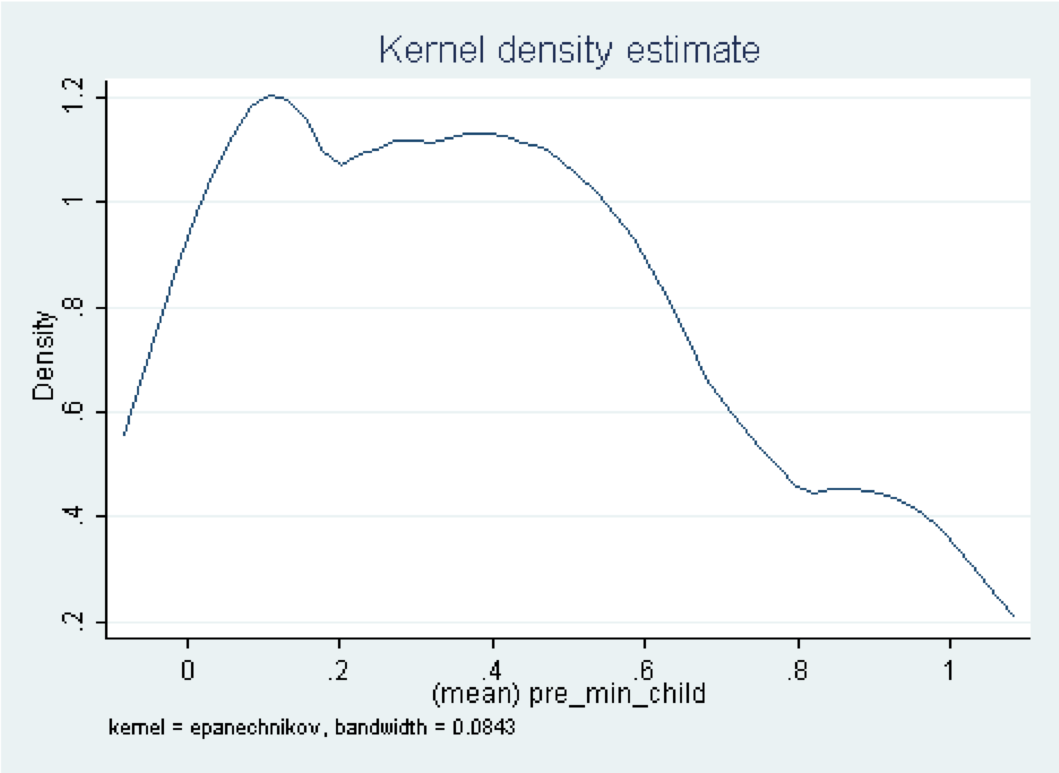
Figure 3: Fertility Across Marriage Types



Notes: The calculation is based on censuses 2000.

Figure 4: Distribution of Social Norms

(a) HM-Families



(b) MH-Families

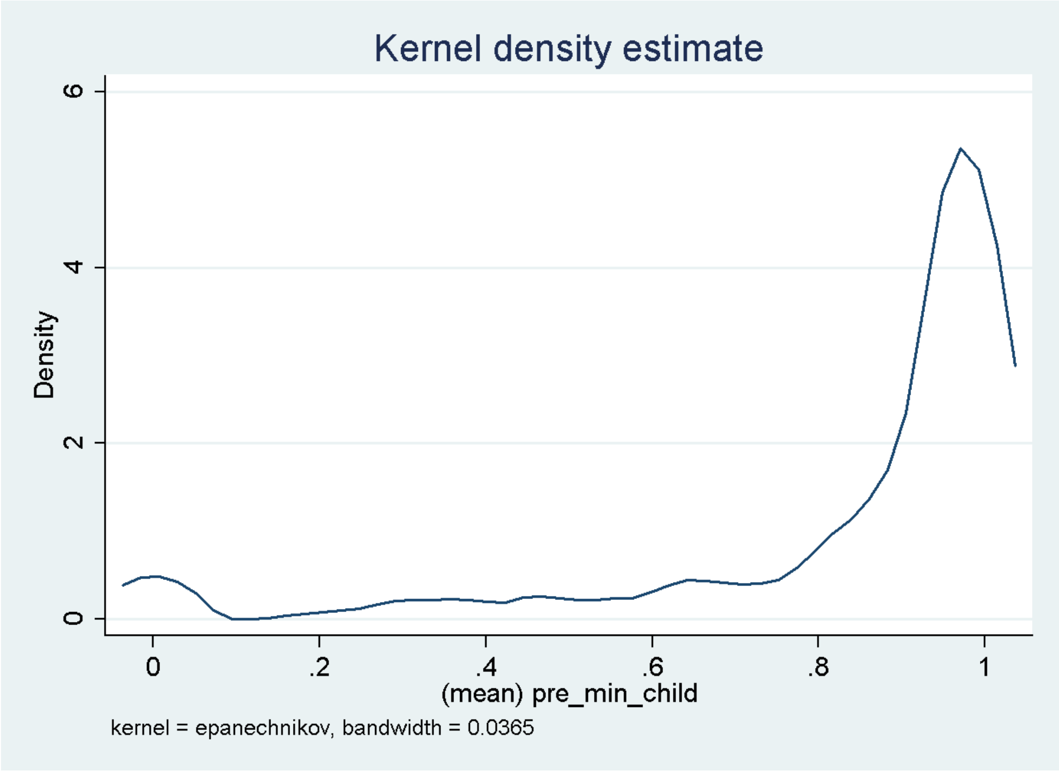
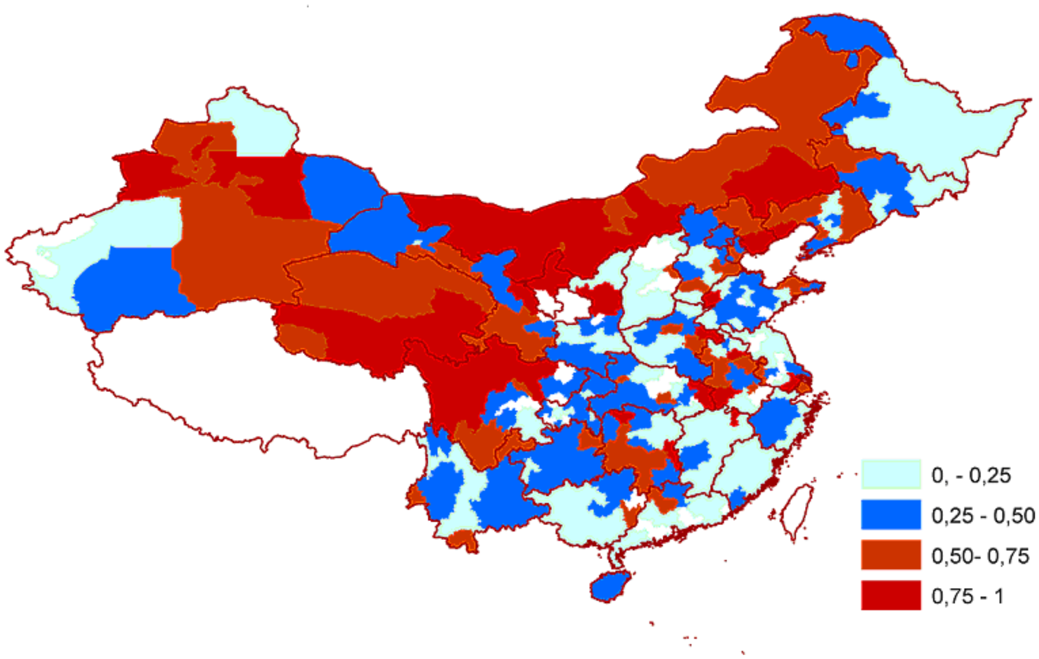


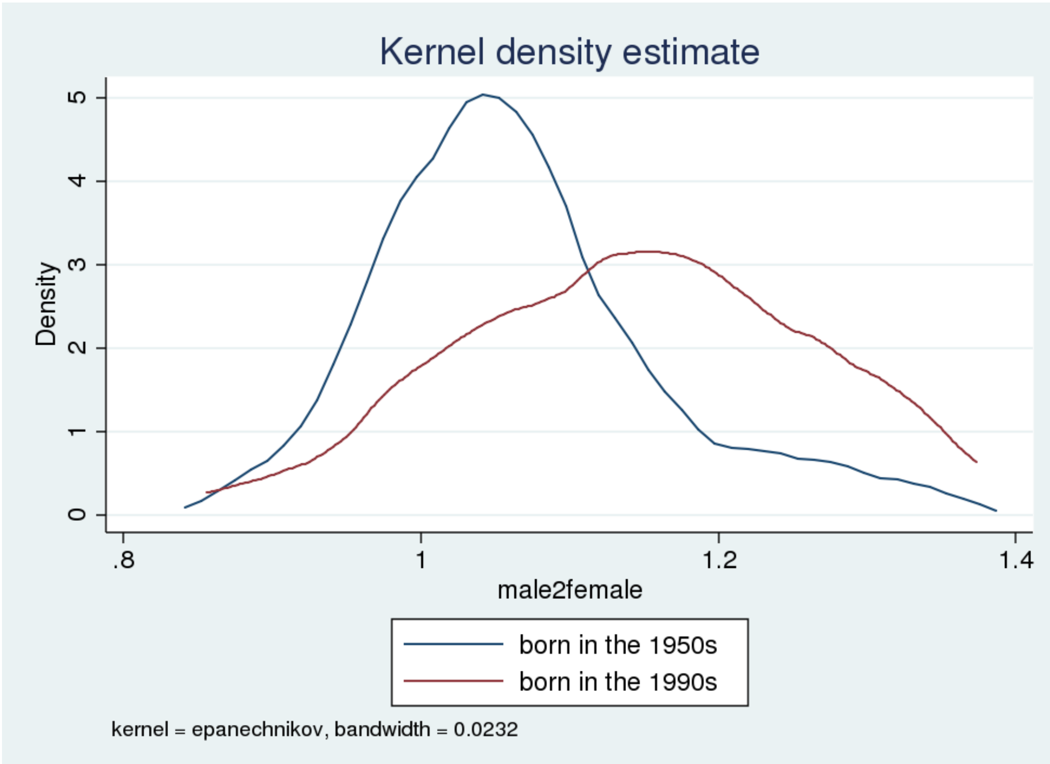
Figure 5: Spatial Distribution of Social Norms (for the 1970s cohort)



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 6: Distribution of Sex Ratios

(a) Han Sex Ratios (across prefectures)



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

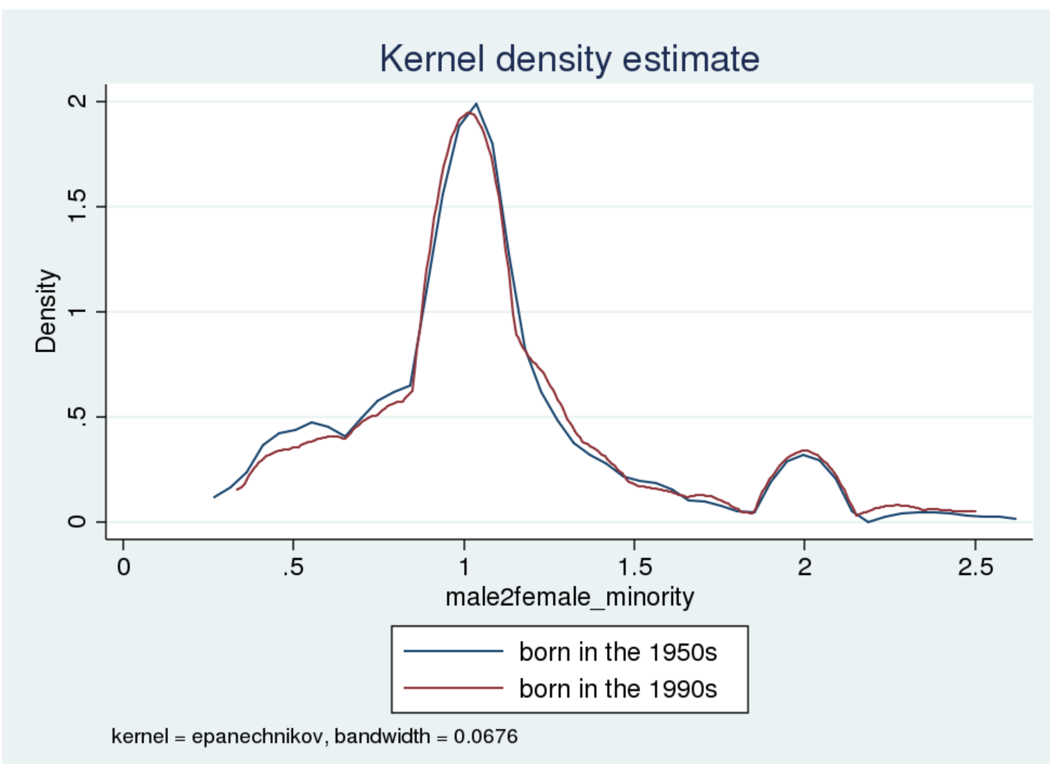
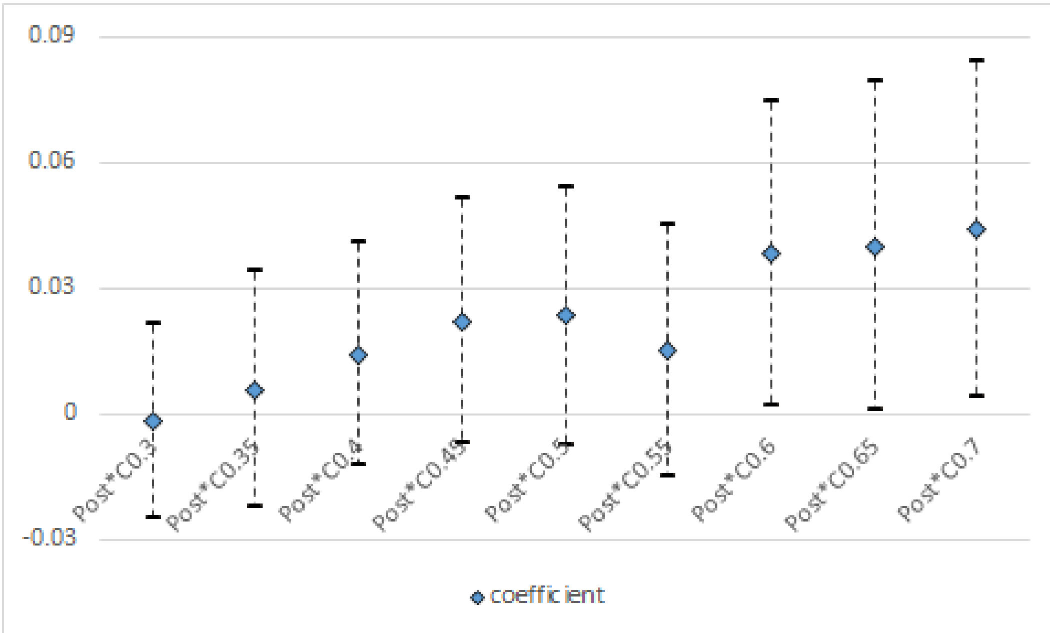


Figure 7: The effect of material benefits * social norms



Notes: This figure plots the results for Prediction C1 using different cutoff values. The diamonds indicate the coefficients and the dashed lines indicate the 95% confidence intervals.

Table 1: Summary Statistics

Panel A: Children in Mixed Families (Censuses 1982-2005)	(1) HM-family	(2) 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
	(0.25)	(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: Norms are defined by prefecture-previous cohort

	(1) MinorChild	(2) MinorChild	(3) MinorChild	(4) MinorChild	(5) MinorChild	(6) MinorChild
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)	
I(=0.70)*Born after 1980						0.044**
						(0.020)
Born after 1980	0.081***					
	(0.010)					
Prefecture FE	Y	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	346
# observations	97399	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: Quartile Results

	(1) MinorChild	(2) MinorChild	(3) MinorChild
I(0-0.25)*Post 1980	0.052**	0.061***	0.046*
	(0.020)	(0.020)	(0.026)
I(0.25-0.5)*Post 1980	0.079**	0.094***	0.050*
	(0.031)	(0.031)	(0.029)
I(0.5-0.75)*Post 1980	0.068**	0.084***	0.036
	(0.034)	(0.030)	(0.035)
Born after 1980	0.011		
	(0.018)		
Prefecture FE	Y	Y	Y
Birth Cohort FE		Y	Y
Province Trends			Y
# clusters	331	331	331
# observations	95528	95528	95528

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**	-0.023*	
	(0.014)	(0.014)	(0.011)	(0.012)	
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)	(0.035)	(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	339	339
# observations	97399	97399	97399	90502	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 (4) reports results using the censuses 1982, 1990 and 2000 while excluding individuals whose birth county and residency county are different. 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				-26.522***	-27.642***	-26.482***	
				(10.118)	(10.113)	(10.092)	
Male2Female Ratio	15.662***	16.366***	16.032***	28.401***	29.638***	29.094***	26.398***
	(5.122)	(5.028)	(4.970)	(6.948)	(6.974)	(6.839)	(6.417)
Male2Female * Post Policy							-25.570***
							(9.384)
Urban * Post 1980						4.432***	5.028***
						(0.892)	(0.822)
College * Post 1980						0.599	0.044
						(1.645)	(1.639)
Urban		8.759***	8.627***		8.756***	5.681***	5.598***
		(0.864)	(0.849)		(0.865)	(0.974)	(0.982)
College		3.427***	3.628***		3.449***	2.583*	2.370*
		(0.942)	(0.933)		(0.945)	(1.498)	(1.375)
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	323	323	323	323	323	323	316
# observations	70231	70181	70181	70231	70181	70181	68827

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

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 8: Exploring Revealed Fertility

	(1) MH	(2) MH	(3) MH	(4) MH
Male2Female Ratio * MM Advantage		-0.562**	-0.589**	-0.585**
		(0.238)	(0.232)	(0.230)
Male2Female Ratio		0.776***	0.811***	0.806***
		(0.270)	(0.266)	(0.263)
MM Advantage	-0.051***	0.541**	0.568**	0.569**
	(0.015)	(0.245)	(0.239)	(0.235)
Urban * MM Advantage				0.008
				(0.032)
College * MM Advantage				-0.092**
				(0.037)
Urban			0.089***	0.079**
			(0.009)	(0.038)
College			0.037***	0.144***
			(0.010)	(0.047)
Prefecture FE	Y	Y	Y	Y
Birth cohort FE	Y	Y	Y	Y
Province Trends	Y	Y	Y	Y
# clusters	318	318	318	318
# observations	66561	66561	66515	66515

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

Notes: MM advantage is defined as the ratio of the number of children in Minority-Minority families to the number of children in Minority-Han families.

A Web Appendix

A.1 Estimation Results for F1-F4

Table W1-W4 present the estimation results for facts F1-F4.

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 fact F1 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 fact F2 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 fact F3 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

Notes: The table shows that fact F4 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%.

A.2 Robustness Checks for C1

Tables W5-W9 present robustness checks for Prediction C1.

Table W5: Material Benefits and Social Norms: Norms are defined by prefecture-the 1970s cohort

	(1) MinorChild	(2) MinorChild	(3) MinorChild	(4) MinorChild	(5) MinorChild	(6) 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)
Born after 1980	0.081***					
	(0.010)					
Prefecture FE	Y	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	346
# observations	97399	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 W6: 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 W7: 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 W8: Use subsample to miminize the migration impact

	(1) MinorChild	(2) MinorChild	(3) MinorChild	(4) MinorChild	(5) MinorChild	(6) MinorChild
I(=0.50)*Born after 1980		0.015				
		(0.016)				
I(=0.55)*Born after 1980			0.013			
			(0.016)			
I(=0.60)*Born after 1980				0.028		
				(0.019)		
I(=0.65)*Born after 1980					0.031	
					(0.021)	
I(=0.70)*Born after 1980						0.040*
						(0.022)
Born after 1980	0.070***					
	(0.010)					
Prefecture FE	Y	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	339
# observations	90502	90502	90502	90502	90502	90502

Notes: The table shows that results using the censuses 1982, 1990 and 2000 while excluding individuals whose birth county and residency county are different. 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%.

Table W9: Use One-Child Policy Timing to Measure Material Benefits

	(1) MinorChild	(2) MinorChild	(3) MinorChild	(4) MinorChild	(5) MinorChild	(6) MinorChild
I(=0.50)*Post Policy		0.028**				
		(0.012)				
I(=0.55)*Post Policy			0.022**			
			(0.010)			
I(=0.60)*Post Policy				0.034***		
				(0.011)		
I(=0.65)*Post Policy					0.033***	
					(0.011)	
I(=0.70)*Post Policy						0.032***
						(0.009)
Post Policy	0.101***					
	(0.013)					
Prefecture FE	Y	Y	Y	Y	Y	Y
Birth Cohort FE		Y	Y	Y	Y	Y
Province Trends		Y	Y	Y	Y	Y
# clusters	346	339	339	339	339	339
# observations	96543	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%.

A.3 Robustness Checks for M1

Tables W10-W11 present robustness checks for Prediction M1.

Table W10: The effect of Han Sex ratios on a Han man's marriage: using subsample

	(1) HM*100	(2) HM*100	(3) HM*100	(4) HM*100	(5) HM*100	(6) HM*100	(7) HM*100
Male2Female * Post1980				2.997**	2.984*	2.999**	
				(1.519)	(1.520)	(1.519)	
Male2Female * Post Policy							2.946*
							(1.540)
Male2Female Ratio	2.284***	2.261***	2.261***	0.696	0.679	0.712	0.974
	(0.841)	(0.839)	(0.839)	(0.956)	(0.954)	(0.955)	(0.876)
Post Policy							-2.833*
							(1.566)
Urban * Post1980						-0.179	-0.165
						(0.139)	(0.150)
College * Post1980						-0.309	-0.395
						(0.335)	(0.368)
Urban		-0.138	-0.138		-0.137	-0.026	-0.055
		(0.110)	(0.110)		(0.110)	(0.102)	(0.105)
College		0.547***	0.547***		0.549***	0.808**	0.893**
		(0.191)	(0.191)		(0.191)	(0.322)	(0.357)
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	329	329	329	329	329	329	320
# observations	187881	187733	187733	187881	187733	187733	177692

Notes: The table shows results using the 2000 census while excluding individuals whose birth county and residency county are different. 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 W11: The effect of Han Sex ratios on a minority man's marriage: using subsample

	(1) MH*100	(2) MH*100	(3) MH*100	(4) MH*100	(5) MH*100	(6) MH*100	(7) MH*100
Male2Female * Post1980				3.464	3.149	3.767	
				(11.099)	(11.081)	(11.249)	
Male2Female * Post Policy							-4.184
							(7.892)
Male2Female Ratio	-8.415*	-8.285*	-8.285*	-10.466	-10.150	-10.599	-4.559
	(4.771)	(4.906)	(4.906)	(6.999)	(6.938)	(6.979)	(5.948)
Urban * Post1980						1.326	1.790
						(2.013)	(1.979)
College * Post1980						2.216	3.021
						(4.766)	(4.896)
Urban		9.117***	9.117***		9.116***	8.269***	8.036***
		(1.256)	(1.256)		(1.256)	(1.853)	(1.833)
College		3.448*	3.448*		3.449*	1.485	0.438
		(2.062)	(2.062)		(2.062)	(4.049)	(4.110)
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	262	262	262	262	262	262	255
# observations	13378	13363	13363	13378	13363	13363	12939

Notes: The table shows results using the 2000 census while excluding individuals whose birth county and residency county are different. 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.4 Robustness Checks for M2

Tables W12 present robustness checks for Prediction M2.

Table W12: The effect of minority Sex ratios on a minority man's marriage: using sub-sample

	(1) MH*100	(2) MH*100	(3) MH*100	(4) MH*100	(5) MH*100	(6) MH*100	(7) MH*100
Male2Female * Post 1980				-27.192	-29.034	-28.524	
				(18.414)	(18.415)	(18.397)	
Male2Female * Post Policy							-25.967*
							(15.581)
Male2Female Ratio	21.133**	21.665**	21.665**	33.606***	34.986***	34.958***	33.494***
	(8.404)	(8.395)	(8.395)	(10.577)	(10.745)	(10.713)	(9.769)
Urban * Post 1980						2.481	3.420*
						(1.836)	(1.801)
College * Post 1980						-1.880	-1.055
						(4.494)	(4.552)
Urban		8.599***	8.599***		8.596***	6.905***	6.389***
		(1.079)	(1.079)		(1.079)	(1.651)	(1.647)
College		4.295**	4.295**		4.339**	5.610	4.742
		(1.963)	(1.963)		(1.962)	(3.827)	(3.833)
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	258	258	258	258	258	258	251
# observations	15448	15431	15431	15448	15431	15431	15189

Notes: The table shows results using the 2000 census while excluding individuals whose birth county and residency county are different. 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%.