Schooling Expansion and the Female Marriage Age: Evidence from Indonesia *

Danyan Zha†

June 28, 2019

Click here for the latest version.

Abstract

This paper analyzes how education distribution affects the marriage market (in particular, female marriage age) by exploiting a massive primary school construction program in Indonesia in the late 1970s as a quasi-natural experiment. Using the variation across regions in the number of schools constructed and the variation across birth cohorts, I show that in densely populated areas, primary school construction did not affect primary school attainment rate. Moreover, the program decreased secondary school attainment rate for both men and women due to a crowding out of teacher resources. Using this change in the education distribution as a source of variation and taking advantage of the large average spousal age gap (five years), I show a woman marries earlier when average education of other women decreases holding their potential husbands education distribution unchanged. I then develop a novel two-to-one dimensional matching model with transferable utilities in an OLG framework and show that the empirical finding suggests that in Indonesia, male education is complementary to both characteristics of women: education and youth.

Keywords: Marriage market; Female marriage age; Spousal age gap; Schooling expansion

JEL classifications: D10, I24, I25, I28, J12, O12, O15

*I am very grateful to Pierre-André Chiappori, Bernard Salanié, and Cristian Kiki Pop-Eleches for their continuing guidance and support. I am indebted to Esther Duflo for sharing the INPRES school construction intensity data with me. I also thank So Yoon Ahn, Michael Best, Suresh Naidu, Rodrigo Soares, Jack Willis and participants at the Applied Microeconomics Theory Colloquium, the Development Colloquium at Columbia University, the 2018 North American Summer Meeting of the Econometric Society at UC Davis, the 2018 European Summer Meeting of the Econometric Society at University of Cologne, and the SEA 88th Annual Meetings at Washington, DC for helpful discussions and comments. All errors are my own.

†Department of Economics, Columbia University. Contact: dz2266@columbia.edu

1 Introduction

Recent decades have witnessed the expansion of schooling world-wide, especially in low- and middle-income countries (World Bank, 2018). There is a large literature studying the impact of these policies on individuals' education, wage, income, wealth and health outcomes (Malamud et al., 2018; Jürges et al., 2011; Akresh et al., 2018). However, few researchers have investigated the impact on marital outcomes, an important dimension of individual life. How should the marriage market respond to changes in the education distribution? Matching theories suggest that individuals' marital outcomes depend on various marriage market conditions including the characteristic distributions of men and women. Hence, a schooling expansion policy may affect an individual's marriage outcome via changing his/her education *and* that of others.

In this paper, I exploit the setting of primary school construction in Indonesia in the late 1970s as a quasi-natural experiment to answer the following question: how does a market-level shock to the education distribution affect marital outcomes, in particular, female marriage age and spousal age gap? I find a woman marry earlier and the spousal age gap increases when fewer women in her birth cohort graduate from secondary school and the education distribution of their potential husbands does not change. To explain the empirical finding, I then develop a two-to-one dimensional matching model with transferable utilities (TU) in an overlapping generations (OLG) framework in which men differ in education and women differ in both education and age. I then show that the empirical finding is consistent with an assumption of complementarity between men's education and women's education, and a complementarity between men's education and women's young age in the marital surplus a couple can generate.

The massive size of the primary school construction program in Indonesia makes it a good setting to answer the question by providing a large exogenous shock to the education distribution. When oil price increased in 1972, the Indonesian government has experienced a tremendous revenue increase. This facilitated one of the largest education expansion programs in the world: INPRES Sekolah Dasar. Approximately one primary school was built per 500 primary school-aged children between 1972/73 and 1978/79. This creates potential variation in the education distribution across regions and birth cohorts.

Using both variations, my identification strategy is difference-in-differences similar to other papers studying this program. (Duflo, 2001; Ashraf et al., 2016) One difference comes from the difference in construction intensity across regions, defined as the average number of schools constructed per 1000 children between 1972/73 and 1978/79; the other difference comes from individual birth cohorts. INPRES SD is a program that targets equality, and hence, more schools were built in regions where there was initially a larger number of un-enrolled school-aged children. Children attend primary school between ages 7 and 12. Therefore children aged 13 or older in 1974 would not have been impacted by the primary school construction program. Children who were younger than 13 would have been impacted by this program.

The first part of my paper focuses on the impact of the school construction program on individuals' education, which is my source of variation in the analysis of marriage market outcomes. This part builds upon earlier studies using the same schooling expansion program (Duflo, 2001; Ashraf et al., 2016). I replicate some of their findings but also find some surprising results not mentioned in the previous literature. Consistent with previous findings, there is a positive effect on primary school attainment rate[1] for men but not for women. However, I find a negative effect on secondary school attainment rate for women in the full-sample analysis. As suggested in Duflo (2001), the program may have different effect in sparsely populated and densely populated regions. Exploring the heterogeneity of the effects depending on population densities of the regions considered, I find that in sparsely populated regions, the school construction program had a positive effect on both primary school and secondary school attainment rates for men but not for women; in densely populated regions, the school construction program did not affect primary school attainment rate but had a negative effect on secondary school attainment rate for both men and women.

I then investigate two potential mechanisms leading to the negative secondary school attainment result: (1) a decrease in secondary school quality due to resources being crowded out; (2) a decrease in primary school quality due to the massive scale of construction. The analysis supports the first mechanism. Building a primary school increases the total demand for teachers in a region, which may affect the availability of teachers in secondary education. Moreover, demand for teachers

[1]Primary school attainment rate is defined as the percentage of people who complete primary school or above. Similarly, secondary school attainment rate is defined as the percentage of people who complete secondary school or above.

can be more competitive in densely populated regions than sparsely populated regions since it's earlier to relocate for teachers in the former area. I show that total number of teachers and average number of teachers in secondary school increase less in the regions where more primary schools were constructed after the launch of the school construction program. The negative effect on teacher availability in secondary education in future years only exists in densely populated regions, not in sparsely populated regions. Moreover, rapidly constructing primary schools could also decrease the quality of primary education. Using the education level of primary school teachers in the censuses as a proxy for school quality, I show that teacher education increases less in regions where more primary schools were constructed. However, I do not find a heterogeneity effect between sparsely and densely populated regions. In summary, the negative result on secondary school attainment rate is due primarily to the crowding out of teacher resources in secondary education because of the primary school construction.

The second part of my paper employs a theoretical framework to understand how female marriage age reacts to the change in the education distributions of men and women across cohorts. To incorporate marriage age as a choice for women, I build a two-period OLG model in which women can choose to seek partners either in the first period or the second, but men all marry in the second period. In any given year, the marriage market unfolds as in Choo and Siow (2006), where the marital surplus generated by a couple depends on their types and some idiosyncratic draws modeled by random vectors. Women differ in two dimensions (education and age) while men only differ in one dimension (education). In a stationary equilibrium, a woman's expected return from the marriage market should be equalized between choosing to marry in the first period or the second.

How the percentage of women choosing to marry in the first period changes with respect to the education distributions of men and women will depend on the assumptions on how male education interacts with female age and female education in generating martial surplus. If there is no interaction between male education and female age in generating marital surplus, female marriage age choice does not depend on the education distribution of men or women. Intuitively, an individual's gain from marriage comes from his/her marginal contribution to the martial surplus. Hence, in this case, women will fully capture the contribution of their age to the martial surplus in their own utilities. Therefore, the marriage age decision would be fully determined by how young age and old age contribute differently to marital surplus.

Suppose that women marrying at a young age is "good" for marital surplus; then, in the case in which there is complementarity between men's education and women's young age, the model predicts that an increase in the proportion of educated men would decrease the female marriage age (i.e., increase the percentage of women marrying in first period). If there is also complementarity between men's education and women's education, an increase in the proportion of educated women would have the opposite effect: an increase in the female marriage age (i.e. a decrease in the percentage of females marrying in first period). Intuitively, an increase in the share of educated women would create a relative shortage of educated men when there is complementarity between men's education and women's education. Hence it would have the opposite effect of an increase in educated men.

In the last part of this paper, I examine the impact of the school construction program on female age at first marriage and the spousal age gap. The average spousal age gap in Indonesia is 5 years; hence, for the first few cohorts of women who were impacted by the school construction program, the education level of their potential husbands was minimally impacted. Therefore, by comparing these female cohorts with the older cohorts who were not impacted by the program, I am able to observe how the female marriage age reacts to the change in the female education distribution while holding the male education distribution unchanged. In sparsely populated regions where there is no effect on female education, as expected, I do not observe any effect on female age at first marriage or the spousal age gap. In densely populated regions where there is a negative effect on secondary school attainment rate for women, I find a decrease in female age at first marriage and an increase in the spousal age gap. This is consistent with the model prediction when women having a young age is "good", and there exists complementarity both between men's education and women's education and between men's education and women's young age in generating marital surplus.

I then proceed to quantify the effect by first estimating the impact of a change in female education distribution on female age at first marriage using the school construction program as an instrument variable for the percentage of female secondary school graduates and find that a 10-

percentage-point decrease in secondary school attainment rate leads to a decrease in the average female marriage age of 1.1 years and an increase in the average spousal age gap of 0.35 years. After removing the impact of individual education, the average female marriage age decreases by 0.7 years and the spousal age gap increases by 0.2 years.

This paper is related to several distinct literatures. The modeling approach in this paper is built on previous research studying marriage age using OLG models (Bhaskar, 2015; Iyigun and Lafortune, 2016; Zhang, 2018) and a model of matching with TU with separable idiosyncratic preferences in marital surplus.(Choo and Siow, 2006; Chiappori et al., 2017; Galichon and Salanié, 2015). Of these OLG papers,, some only focus on age (Bhaskar, 2015), while others simultaneously study individuals' educational and marriage age decisions. Making education endogenous is an attractive feature by itself, but it is less attractive in answering the current research question of how marriage age responds to a change in the education distribution in the marriage market.

My paper contributes to a growing literature studying the impact of education reform on marriage market. Hener and Wilson (2018) studies a compulsory reform in UK and finds that women decrease the marital age gap to avoid marrying less-qualified men. André and Dupraz (2018) studies school construction in Cameroon and finds that education increases the likelihood of being in a polygamous union for both men and women. In contrast to both of these papers, the present paper analyzes the effect via a general equilibrium framework.

My paper complements a large literature on the impact of marriage market conditions on individuals' outcomes. Most of the existing literature focuses on the sex ratio in the marriage market. (e.g. Abramitzky et al., 2011; Angrist, 2002; Charles and Luoh, 2010) I focus on a distinct but equally important dimension of marriage market conditions: the education distributions of men and women.

Finally, the paper also contributes to the large literature studying Indonesia's INPRES program. (Duflo, 2001; Breierova and Duflo, 2004; Ashraf et al., 2016; Dominguez, 2014). While previous papers mention only the negative effect on secondary education, this paper complements those works by providing evidence for plausible mechanisms.

The remainder of this paper proceeds as follows. The next section discusses the model and its predictions. Section 3 describes the school construction program and background information on Indonesia. Section 4 describes the data and my identification strategy. The main results on educational and marital outcomes are presented in Sections 5 and 6. Section 7 concludes the paper.

2 Model

In this section, I develop a two-period OLG matching model with Transferable Utility (TU) to study how a change in the education distribution across birth cohorts may affect marriage market outcomes, in particular, female marriage age. There are several important features:

- Individuals get utility from participating in the marriage market.
- Individuals' education affect the marital surplus, for both men and women.
- Individuals' age play an asymmetric role for men and women. Women's age matters but not men's. in the surplus function. Much research has documented that female youth is more important than male youth in the marriage market, this could be due either to the fundamental difference of female age and male age in the household production function related to fertility, or due to a stronger male preference for youth related beauty. (Low, 2017; Siow, 1998; Edlund, 2006; Dessy and Djebbari, 2010; Zhang, 2018; Arunachalam and Naidu, 2006)
- Women are allowed to choose to participate in the marriage market either early or late. However, a woman who participated in period 1 cannot enter into the marriage market in period 2, whether she remains married or single. This can be rationalized as the existence of a stigma associated with women who have tried to seek partners in an early period.
- Each marriage market is modeled as a matching model with TU with idiosyncratic random preference draws. The existence of random preference draws allows the existence of couples of all types with respect to male education, female education and female age, which suits the reality more compared to the static model. In each marriage market, women differ in both education and age, while men only differ in education.

Two-period OLG

There is an infinite number of periods, r=1,2.... At the beginning of each period, a unit mass of men and a unit mass of women enter the economy. Assume people can only make marriage decisions in the first two periods, therefore the problem is simplified to a two-period OLG problem. Furthermore, to focus on female marriage age decision, I assume that women choose whether they want to seek partners in period 1 (when they are young) or delay this process to period 2 (when they are old). Men always seek partners in period 2. Individuals differ in their education type, L or H. In the model, let's focus on the utilities individuals obtain from the marriage market.

Marriage market at one period

I will first discuss how marriage market unfolds given women's marriage timing choices in any given period.

Individual types

Women can choose to participate in one of the two periods, hence in any period, there are at most four types of women: Low education and Young (L_1), Low education and Old (L_2), High education and Young (H_1), and High education and Old (H_2). Men only participate in period 2, hence there are two types of men in any period: Low education (L) and High education (H).

Utilities and matching surplus

Denote x as the type of women and X as the type set, i.e. $x \in X = \{L_1, L_2, H_1, H_2\}$. Similarly, denote y as the type of men and Y as the type set, i.e. $y \in Y = \{L, H\}$. To include the possibility of being single, denote $X_0 = X \cup \emptyset$, $Y_0 = Y \cup \emptyset$. Suppose that a woman i with type x and a man j with type y form a couple. I assume their lifetime utilities are as following:

$$\text{woman i's utility: } u_{ij} = \alpha_{xy} + \tau_{ij} + \varepsilon_{iy}$$

$$\text{man j's utility: } v_{ij} = \gamma_{xy} - \tau_{ij} + \eta_{xj}$$

α_{xy}, γ_{xy} indicate the systematic part of the utility each individual gets from the marriage depending on their types. τ_{ij} represents the transfer between i and j, which is going to be determined in equilibrium. [2] $\varepsilon_{iy}, \eta_{xj}$ represent the individuals' idiosyncratic tastes in partner types. Notice they only depend on the partners' types.

For individual singles, their utilities will be:

$$u_{i\emptyset} = \alpha_{x\emptyset} + \varepsilon_{i\emptyset}$$

$$v_{\emptyset j} = \gamma_{\emptyset y} + \eta_{\emptyset j}$$

Without loss of generality, we can normalize $\alpha_{x\emptyset} = 0$ and $\gamma_{\emptyset j} = 0$. [3] Then α_{xy} and γ_{xy} can be interpreted as the net systematic gain from marriage.

There are three important assumptions underlying my specification of individuals'utility: [4]

- There exists a transfer technology among a couple to transfer their utilities one to one without loss, which is the basic feature of a matching model with TU.
- Both transfer and the random taste terms are additive to the systematic part.
- The random terms are individual specific but only depend on the partner's type.

This utility specification may seem restrictive, but it allows for "matching on unobservables" and allows model tractability. What it rules out is the "chemistry" term between two individuals conditional on their types, i.e., some unobserved preferences of one individual towards some unobserved characteristic of one partner.

Stable Matching

Given the population and type distribution,G_x, G_y in a marriage market, a matching is defined as a measure μ on set $X \times Y$ and a set of payoffs $\{u_i, v_j, i \in I, j \in J\}$ such that $u_i + v_j = \alpha_{xy} + \gamma_{xy} + \varepsilon_{iy} + \eta_{xj}$ for any matched couple (i, j). In other words, a matching specifies who marries with whom

[2] τ can be either positive or negative.

[3] Because we can always define $\tilde{\alpha}_{xy} = \alpha_{xy} - \alpha_{x\emptyset}; \tilde{\gamma}_{xy} = \gamma_{xy} - \gamma_{\emptyset y}$, as the systematic utility surplus an individual obtain from marriage compared to being single.

[4] This is the "Separability" assumption in Galichon and Salanié (2015). As noted in that paper, what matters in the model is the surplus a couple can jointly achieve, i.e. $\alpha_{xy} + \gamma_{xy} + \varepsilon_{iy} + \eta_{xj}$ in our case here. How we attribute this surplus to male preference or female preference doesn't matter. For example, it can be the case that women don't have any random taste for men and their utilities without any transfer is α_{xy}. Men's utilities are $\gamma_{xy} + \varepsilon_{iy} + \eta_{xj}$, indicating that man j not only has a random draw η_{xj} depending on women's type, but also has own-type specific random taste for a particular woman i, represented by ε_{iy}. The solution to the model is the same no matter how we interpret the joint surplus into people's preference. The same assumption is also imposed in Choo and Siow (2006) and Chiappori et al. (2017).

and how each mathched couple divides the surplus. Notice that the female type distribution G_x is endogenously determined by female marriage timing choices and the exogenous type distribution, denoted as $E_f = (n_L, n_H)$. And the male type distribution G_y is the same as the exogenous type distribution, denoted as $E_m = (m_L, m_H)$.

In a stable matching, there are two requirements:

- (Individual rationality) Any matched individual is weakly better off than being single.

$$u_i \geq \varepsilon_{i0}, v_j \geq \eta_{0j}, \forall i \in I, j \in J$$

- (No blocking pair) There doesn't exist any two individuals, woman i and man j, who are currently not matched to each other but would both rather match to each other compared with their current condition.

$$u_i + v_j \geq \alpha_{xy} + \gamma_{xy} + \varepsilon_{iy} + \eta_{xj}, \forall i \in I, j \in J$$

Therefore, in any stable matching and given equilibrium transfers τ_{ij}, the following conditions hold true:

$$\text{Woman i chooses } j^*(i) : j^*(i) = \max_{j \in J_0} u_{ij}$$

$$\text{Man j chooses } i^*(j) : i^*(j) = \max_{i \in I_0} v_{ij}$$

where J_0 represent all men and the possibility of being single, I_0 represent all women and the possibility of being single.

Lemma 1. *For any stable matching, there exists two vectors U^{xy} and V^{xy} such that:*

(i) Woman i of type x achieves utility:

$$\tilde{u_i} = \max_{y \in Y_0} (U^{xy} + \varepsilon_{iy})$$

and she matches some man whose type y achieves the maximum;

(ii) Man j of type y achieves utility:

$$\tilde{v_j} = \max_{x \in X_0}(V^{xy} + \eta_{xj})$$

and he matches some woman whose type x achieves the maximum.

(iii) If there exist women of type x matched with men of type y at equilibrium, then

$$U^{xy} + V^{xy} = \alpha_{xy} + \gamma_{xy}$$

This lemma has been proved in Chiappori et al. (2017); Galichon and Salanié (2015). I'll write a short version of the proof in the appendix. With TU, the additive structure and type-specific heterogeneity, this two-sided matching problem is simplified to a one-sided discrete choice problem.

Solutions with Gumbel distribution

If we further assume Gumbel distribution for ε, η, a closed form solution of the stable matching and the expected utilities of each type can be derived. From now on, let's assume the random terms $\varepsilon_{iy}, \eta_{xj}$ follow independent Gumbel distributions $G(-k, 1)$, with $k \simeq 0.5772$ being the Euler constant. With the properties of the Gumbel distribution and Lemma 1, for a given woman i of type x,

$$\begin{aligned}\mu_{y|x} &:= \text{Pr (Woman i (of type x) matched with a man of type y)} \\ &= \frac{exp(U^{xy})}{1 + \sum_{y \in Y} exp(U^{xy})}\end{aligned}$$

$$\begin{aligned}\mu_{\emptyset|x} &:= \text{Pr (Woman i (of type x) is single)} \\ &= \frac{1}{1 + \sum_{y \in Y} exp(U^{xy})}\end{aligned}$$

Therefore,

$$\frac{\mu_{y|x}}{\mu_{\emptyset|x}} = exp(U^{xy}), \forall x \in X$$

Similar logic applies to the other side: men:

$$\frac{\mu_{x|y}}{\mu_{\emptyset|y}} = exp(V^{xy}), \forall y \in Y$$

Denote n_x, m_y as the population of each type. Note that n_x depends on women's participation choices. Denote μ_{xy} as the mass of matched couples between woman of type x and man type y, note that $\mu_{xy} = \mu_{yx}$ by construction since it's a one-to-one match; denote μ_{x0} as the mass of single women of type x, μ_{0y} as the mass of single men of type y; then we have:

$$\frac{\mu_{xy}^2}{\mu_{x0}\mu_{0y}} = exp(U^{xy} + V^{xy}) = exp(\alpha_{xy} + \gamma_{xy})$$

Denote $\Phi_{xy} = \alpha_{xy} + \gamma_{xy}$. Then given Φ_{xy}, the previous equation provides a matching function between the mass of any couple type and the probabilities of singlehood. With the following feasibility constraints, we can construct a system of equations with $|X|+|Y|$ unknowns (probabilities of singlehood for each type) and $|X| + |Y|$ equations. Decker et al. (2013) shows the existence and uniqueness of the solution to this system.

$$\mu_{x0} + \mu_{xL} + \mu_{xH} = n_x, \forall x \in \{L_1, L_2, H_1, H_2\}$$

$$\mu_{0y} + \mu_{L_1y} + \mu_{L2y} + \mu_{H_1y} + \mu_{H_2y} = m_y, \forall y \in \{L, H\}$$

Moreover, we can recover the expected utilities each type gets from participating in this marriage market. With the properties of Gumbel distributions,

$$u_x := E[\tilde{u_i}] = E[\max_{y \in Y_0}(U^{xy} + \varepsilon_{iy})] = ln(1 + \sum_{y \in Y} exp(U^{xy})) = -ln(\mu_{\emptyset|x}) = -ln(\frac{\mu_{x0}}{n_x})$$

$$v_y := E[\tilde{v_j}] = E[\max_{x \in X_0}(V^{xy} + \eta_{xj})] ln(1 + \sum_{y \in Y} exp(U^{xy})) = -ln(\mu_{\emptyset|y}) = -ln(\frac{\mu_{0y}}{m_y})$$

In this case, the expected utility has one-to-one correspondence with the single rate in this case. The smaller the single rate is, the larger the expected utility is. [5]

[5] This is a specific property of the Gumbel distribution.

Stationary equilibrium with OLG

Before participating in any marriage market, the strategic choice for each woman in the model is to choose when to enter into the marriage market, given the predetermined education distribution of women and men, denoted by (E_f, E_m). For a woman with education e, if she chooses to enter in period 2 instead of period 1, this increases the expected marital return of all women in period 1 marriage market and decreases the expected return of all women in period 2 marriage market. [6] In a stationary equilibrium, the percentage of women who choose to wait until period 2 equates women's expected returns in the two marriage markets. Denote the percentage of women with education e who choose to seek partners in period 1 (or period 2) as q_e^1 (or q_e^2), assume $e \in \{L, H\}$. Of course, $q_e^1 + q_e^2 = 1, \forall e$.

We say the marriage market with distribution of female types and male types as (G_x, G_y) is the induced marriage market of a strategy vector $\mathbf{q}$ if the distribution of female types (four) and male types (two) in the marriage market is (G_x, G_y) when women adopt strategy $\mathbf{q}$. Note that for male distribution, $G_y = G_m, \forall \mathbf{q}$.

Definition 1. *Strategy vector $\mathbf{q} = \left\{q_H^1, q_H^2, q_L^1, q_L^2\right\}$ forms a stationary equilibrium if $u_{H_1} = u_{H_2}$ and $u_{L_1} = u_{L_2}$ in the induced marriage market, where u_{e_1} (u_{e_2}) is the expected marriage payoff of women with education e who choose to enter the marriage market in period 1 (period 2).*

Denote $\Phi_{xy} = \alpha_{xy} + \gamma_{xy}$. We have woman's type $x \in \{L_1, L_2, H_1, H_2\}$, man's type $y \in \{L, H\}$.

Proposition 1. *There exists a unique stationary equilibrium, and the equilibrium strategy $\mathbf{q}$ satisfy:*

$$\min(\Phi_{L_1L} - \Phi_{L_2L}, \Phi_{L_1H} - \Phi_{L_2H}) \leq \ln(\frac{q_L^1}{q_L^2}) \leq \max(\Phi_{L_1L} - \Phi_{L_2L}, \Phi_{L_1H} - \Phi_{L_2H})$$

$$\min(\Phi_{H_1L} - \Phi_{H_2L}, \Phi_{H_1H} - \Phi_{H_2H}) \leq \ln(\frac{q_H^1}{q_H^2}) \leq \max(\Phi_{H_1L} - \Phi_{H_2L}, \Phi_{H_1H} - \Phi_{H_2H})$$

Proof. See the appendix □

Intuitively, the equilibrium percentage of women who decide to participate in period 1 depends on the marital surplus difference between marrying in period 1 and period 2 given any partner

[6] As proved in Galichon and Salanié (2017), an addition of one woman hurts all women and benefits all men; an addition of one man hurts all men and benefits all women.

type. The larger the difference, the higher the percentage of women seeking partners in period 1.

One corollary of Proposition 1 is that the equilibrium strategy $\mathbf{q}$ satisfies the following conditions: $0 < q_e^1 < 1, 0 < q_e^2 < 1, \forall e \in \{L, H\}$. In equilibrium, it will never happen that all women of the same education type choose to participate in period 1 or period 2, as long as the surplus Φ terms are bounded. Intuitively, if all women of one education type choose to participate in period 1, a woman could benefit by choosing to participate in period 2, which makes her the only older woman with that education. The scarcity of this type would earn large marital returns for the woman. Since the support of Gumbel distribution is $\mathbb{R}$, the potential return could be large enough such that being the only one of older type in period 2 is more rewarded than participating in period 1 no matter how large the surplus difference $\Phi_{e_1y} - \Phi_{e_2y}$ is as long as it is finite. [7]

Proposition 2. *If given education type $e \in \{L, H\}$, $\Phi_{e_1H} - \Phi_{e_2H} = \Phi_{e_1L} - \Phi_{e_2L}$, then q_e^1, q_e^2 are uniquely pinned down by:*

$$q_e^1 = \frac{\exp(\Phi_{e_1L})}{\exp(\Phi_{e_1L}) + \exp(\Phi_{e_2L})}, q_e^2 = \frac{\exp(\Phi_{e_2L})}{\exp(\Phi_{e_1L}) + \exp(\Phi_{e_2L})}$$

Proof. See the appendix □

$\Phi_{e_1H} - \Phi_{e_2H} = \Phi_{e_1L} - \Phi_{e_2L}$ indicates that the gain of female youth in surplus is independent of men's education. [8] This means that male education and female youth don't interact in the marital surplus, hence the marginal contribution of female youth in the surplus doesn't depend on their partner's education type either. In a matching model, individuals' marital gain come from their marginal contributions to the surplus. In this case, women get all the benefit (or cost) of female youth if they choose to participate in period 1. Their choice of marriage market is fully pinned down by this difference in marital surplus independent of the education distribution of both sides.

[7] One can also understand this in terms of the probability of singlehood. In the model, single probability has one-to-one correspondence with the expected utility: the lower the single probability, the higher the expected marital return. For a woman who is the only one of older type in period 2, she would almost for sure get married since the men who have very large draws for this particular older type would compete fiercely among themselves and want to marry her.

[8] It can depend on female education, e. For example, the return of female youth is larger for less educated women than more educated women, or the other way around. The empirical observation that less educated women marry earlier supports the case that the gain is larger for less educated women.

Comparative statics

School construction would lead to a dynamic change in the population education. However, unlike in Bhaskar (2015), the current model doesn't focus on the transitory period, which is of less interest in this paper. I will concentrate instead on how the stationary equilibrium changes in response to the change in population education. For simplicity, let's assume male population and female population are equal. Without loss of generality, I can also normalize the population of each side to 1 since the model has constant returns to scale. Let us analyze how female marriage age decision would change when the education distribution of men or women changes, respectively.

Proposition 3. *Denote female education distribution as* $G_f = (n_L, 1 - n_L)$ *and male education distribution as* $G_m = (m_L, 1 - m_L)$.

Keeping n *constant,* $\forall y \in \{L, H\}$*, a decrease in* m_L *would*

- *increase* q_e^1*, if* $\Phi_{e_1 H} - \Phi_{e_2 H} > \Phi_{e_1 L} - \Phi_{e_2 L}$*;*
- *decrease* q_e^1*, if* $\Phi_{e_1 H} - \Phi_{e_2 H} < \Phi_{e_1 L} - \Phi_{e_2 L}$*.*

Proof. See the appendix. □

If the percentage of more-educated men increases, the equilibrium percentage of women marrying in period 1 increases if male education and female youth are complementary [9] in the marital surplus; the equilibrium percentage of women marrying in period 1 decreases if instead male education and female maturity are complementary in the marital surplus. Notice that whether the marital surplus is super-modular in male education and female education does not matter.

A stable matching maximizes the total social surplus in a TU framework. (Shapley and Shubik, 1971) When male education and female youth are complementary, the social surplus is larger if we pair more educated men with younger women. Hence when there is a decrease in m_L, the existence of more educated men would induce more women to marry in period 1 to take advantage of the higher social surplus. Vice versa.

[9] There are at least four ways to interpret the complementarity between male education and female youth. For example, (1) all men prefer female youth and more educated men value female youth more than less educated men. (2) All women prefer more educated men and younger women value male education more than older women. (3) All men dislike female youth and more educated men dislike female youth less than less educated men. (4) All women dislike more educated men and younger women dislike more educated men less than older women. Of course, the first and second seem to be more plausible than the last two.

Proposition 4. *Denote female education distribution as* $N_f = (n_L, 1 - n_L)$ *and male education distribution as* $N_m = (m_L, 1 - m_L)$.

Further assume super-modularity in men's education and women's education: holding m *constant,* $\forall e \in \{L, H\}$, *a decrease in* n_L *would*

- *decrease* q_e^1, *if* $\Phi_{e_1H} - \Phi_{e_2H} > \Phi_{e_1L} - \Phi_{e_2L}$
- *increase* q_e^1, *if* $\Phi_{e_1H} - \Phi_{e_2H} < \Phi_{e_1L} - \Phi_{e_2L}$

Proof. See the appendix □

A change in female education distribution affects the equilibrium female choice by affecting the potential gain of female youth via affecting the potential distribution of men a woman can marry to. If n_L decreases, for a given woman, other women are more educated. They are more likely to marry with more educated men due to the complementarity in education. Therefore, on the market, more educated men are more scarce, which will discourage all women from participating in period 1 as predicted in Proposition 3 if male education and female youth are complementary.

3 Background

3.1 INPRES Primary School Construction Program in Indonesia

The Indonesian government has consistently sought to broaden educational opportunity since the country's independence in 1945. However, due to financial difficulties and political conflict, in the country's early years, Indonesia remained backward relative to neighboring countries and to countries with similar levels of income. As late as the 1971 population census, only 62% of primary school-aged children (ages 7-12 inclusive) were enrolled in any kind of school, while only 54% appeared on the rolls of public and private schools reporting to the Ministry of Education. (see Snodgrass, 1984). Due to increased oil production and the first OPEC-engineered price rise in 1972-1973, which unexpectedly raised government revenue, a primary school construction aid program (Program Bantuan Pembangunan Sekolah Dasar), known as INPRES Sekolah Dasar and more informally as INPRES SD, was inaugurated in 1973. In addition to school construction, the government also provided textbooks and teacher training to ensure that the buildings were for education purposes. By 1983, nearly all Indonesian children had at least begun to enroll in primary school, while the percentage of 7-12 year olds enrolled exceeded 90%. INPRES SD has been a successful case of education policies in developing countries.

Between 1973/74 and 1978/79, 62,000 primary schools were scheduled to be built. Each school consists of three classrooms, and each classroom has one teacher and can accommodate 40 pupils. The allocation rule every year is as follows: (a) ensure that each district(*kecamatan* in Indonesian, one level below the regency and two levels below the province level) was allocated at least one school and each province at least 50, (b) the remainder were distributed according to the estimated population of unenrolled 7-12 year old children. This creates variation in the construction intensity that I exploit in my empirical analysis.

3.2 Education System in Indonesia

In Indonesia, the education system consists of six years of primary school (sekolah dasar, SD), three years of middle school (sekolah menengah pertama, SMP) and three years of high school (sekolah menengah atas, SMA), followed by various kinds of higher education. Children generally begin primary school at age 7. Two ministries are responsible for managing the education system, with 84 percent of schools being under the Ministry of National Education and the remaining 16 percent being under the Ministry of Religious Affairs. In the 2000 census, although 86.1 percent of the Indonesian population was registered as Muslim, only 15 percent of school-aged individuals attended religious schools. (Library of Congress).

INPRES 1973 initiated Indonesia's program of compulsory education, but six-year compulsory education for primary school-aged children (7-12 age group) was not fully implemented until 1984. In May 1994, nine-year compulsory education for the 7 to 15 age group was introduced. Of all pupils, 92% were enrolled in public schools for primary education, and 50% were enrolled in public schools for secondary education. The Indonesian government focused more on primary education than on the secondary level. In 1985, of public spending on education, 62% went to primary education, while 27% went to secondary education. (see Tan and Mingat, 1992, table 3.1, table 6.5)

In the 1980s, although all children began primary school, only approximately 62% of pupils entering primary school actually graduated from grade 6. Transition between primary school and junior secondary school was also low, at approximately 60%. (see Jones and Hagul, 2001, table 1, figure 2). Transition between junior secondary and senior secondary was also low: 53%. However, the survival rates of junior secondary school and senior secondary school are fairly high in Indonesia, at more than 90%. (see Tan and Mingat, 1992, table4.5, table 4.6, Table A.1)

The primary school enrollment rate increased from 62% in 1972 to 90% in 1983. However, regency-specific plan data are available only for 1973/74-1978/79; hence I limited my sample to individuals born at or before 1972 who were older than age 7 in 1979. For those born after1972, I am unable to identify the primary school construction intensity they were exposed to at age 7.

Link Regency Code between Censuses. Indonesia has experienced a substantial increase in the number of regions (*Pemekaran Daerah*) since the enactment of Law No.22 of 1999 concerning regional autonomy. The number of regencies increased from 271 in 1971 to 304 in 1995 to 437 in 2005 and to 494 in 2010. Hence, I use the GIS shapefiles provided by IPUMS across census years to link regencies of birth in 2005 and 2010 back to the regency of birth variable in 1995 to assign the proper program intensity to each individual. Since most of this expansion is in the form of dividing existing regencies into several small regencies, I can link most of the regencies.

School intensity data are available for 290 unique regencies in Duflo(2001), which were coded using 1995 labels. There were 304 regencies in 1995, and the 14 lost regencies were in East Timor, which became part of Indonesia as the 27th province in 1976.

3.3 Identification Strategy

Education. To analyze how the education distribution was impacted across regencies and birth cohorts, my empirical strategy is difference-in-differences, as used in citeDuflo2001. One difference comes from the school construction intensity, defined as the average number of primary schools built between 1973 and 1978 in one regency per 1000 children aged $5 \sim 14$ in 1971. The other difference comes from birth cohorts. In Indonesia, children begin attending primary school at ages $7 \sim 12$. Those aged 13 or above in 1974 would not have been impacted by the program because they were already out of primary school. For those aged less than or equal to 12 in 1974, the younger they were, the more exposed they were to this school construction program.

The quantitative effect of the school construction program on individuals born in birth cohort

k and regency j can be estimated with the following specification:

$$y_{jk} = \alpha_j + \beta_k + \sum_{l=2}^{12}(P_j d_{kl})\gamma_l + \sum_{l=14}^{21}(P_j d_{kl})\gamma_l + \sum_{l=2}^{21}(\boldsymbol{C_j} d_{kl})\delta_l + \varepsilon_{jk}$$

where y_{jk} is the percentage of individuals completing primary school (secondary school) born in regency j, and in birth cohort k, d_{kl} is a dummy that indicates whether birth cohort k individuals are age l in 1974 (year-of-birth dummy). α_j denotes the regency fixed effect, and β_k denotes the birth cohort fixed effect. P_j is the school construction intensity in regency j. ε_{jk} is the error term. C_j represents other region-specific variables.

The coefficients γ_l are the coefficients of interest. They represent the effect of one additional primary school constructed on the dependent variable for individuals of age l in 1974. There is a testable restriction on coefficients γ_l. A valid identification strategy would require that $\gamma_l = 0$ if $l > 13$, i.e., the variation in the outcome variable is not correlated with the primary school available starting in 1974 for the children who were already out of primary school in 1974. I should expect that for $l \leq 12$, $\gamma_l > 0$, and that γ_l decreases with l, in other words, a higher impact on the younger generation.

Marriage Market One difficulty in this empirical analysis of the marriage market is that both men and women are potentially simultaneously affected by school construction; hence, I may not be able to identify which side drives my results. However, the large positive spousal age gap in my sample provides a novel setting in which only the female education distribution in the marriage market changes, while that of men does not. Because women marry older husbands, for the first few cohorts of women whose education is impacted, their potential husbands are older than they are and would not have been impacted by the program. The larger the average spousal age gap norm is, the more birth cohorts of women I can attribute to the experiment in which only the women's, not the men's, education distribution changes in the marriage market.

The reduced-form regression specification for marriage market outcomes for women is the same as the previous specification for education.

4 Empirical Results on Education

In this section, I present my empirical results on education, which is my source of variation for marriage market outcomes. I first present the results for the full sample, then show the results for two subsamples depending on population density. Finally, I provide further evidence for the mechanisms behind the different results observed in the subsamples.

4.1 First-Stage Effect on Education for the Full Sample

First, I examine the impact on education for the full sample so that I can introduce shifts in the distribution of education for my later analysis of marriage market outcomes.

In Figure 3, I plot γ_l when the dependent variable is the percentage of individuals who complete at least primary school for men (or women), i.e., the effect of one additional primary school constructed per 1000 children on primary school attainment rate for men (or women) with age l in 1974. To simplify the graph, For simplicity, I combine three birth cohorts together on the graph.

Two important results stand out from Figure 3. First, γ_l is not significantly different from 0 for l larger than 13 for both men and women. This lends confidence in the identification assumption: the birth cohort trend in the primary school attainment rate does not differ across regions with different school construction intensities. Secondly, γ_l is positive for men with age $l \leq 12$ in 1974, indicating a positive effect on the primary school attainment rate for men; γ_l is zero for women except the youngest cohorts, aged $l \leq 3$, indicating a lagged effect on female primary school attainment rate. Both results are consistent with previous findings in Duflo (2001) and Ashraf et al. (2016).

Difference-in-differences estimates are provided in columns (1)-(3) in Table 2. Following Duflo (2001), the sample includes individuals born between 1950 and 1961 who are older than 12 in 1974, and individuals born between 1968 and 1972, who are younger than 7 in 1974. "Post" indicates individuals born between 1968 and 1972. Column (3) suggests that one additional school increases male primary school attainment rate by 0.6 percentage points. This is smaller than the estimate in Figure 2 in Duflo (2001) where it's shown that approximately 1.5% more individuals had at least 6 years of schooling between high program regions (where on average 2.44 schools were built) and low program regencies (where on average 1.54 schools were built). My estimate is smaller, and one potential reason for this divergence is the inclusion of more controls in my analysis compared to Equation (4) in Duflo (2001).

In Figure 4, I plot the coefficients of the interactions of age in 1974 and program intensity for completing secondary school or above. Surprisingly, I find a negative impact on secondary school attainment rate, especially for women. This is surprising because, if anything, I should expect positive spillover effects from primary school completion to secondary school completion. This finding is also mentioned for men in Duflo (2001) but not discussed in detail there. The difference-in-differences estimates in columns (4)-(6) in Table 2 suggest one additional school being built decreases women's secondary school attainment rate by 0.53 percentage points.

4.2 Heterogeneity Results on Education

Further insight into the effect of the program can be obtained by examining its impact on different types of regions. In this section, I repeat the previous exercise on two subsamples divided by population density: sparsely populated regions with densities below the medium density and densely populated regions with densities above the medium. Population density is calculated as the population in the 1971 census divided by the area of each region in 1971. The median density (the density for the region of birth for the median person in the weighted sample) is 470 inhabitants per square kilometer. There are 183 regions in the sparsely populated subsample, and the average number of schools constructed per 1000 children is 2.1. There are 91 regions in the densely populated subsample, and the average number of schools constructed per 1000 children is 1.67, which is somewhat lower than that in the sparsely populated subsample.

In Figure 5 and Figure 6, I plot the coefficients on education γ_l for both sparsely populated and densely populated subsamples. The difference-in-differences estimates are shown in Table 3.

As Figure 5 shows, in sparsely populated areas, the program increased the primary school attainment rate (top) and secondary attainment rate (bottom) for men but did not affect women's education. Difference-in-differences estimates are provided in Panel A of Table 3. For men, one additional school constructed per 1000 children increased the percentage completing primary school or above by 1 percentage point and the percentage completing secondary school or above by 0.69 percentage points.

As Figure 6 shows, in densely populated areas, the program did not affect the primary school attainment rate (top) but decreased the secondary school attainment rate (bottom) for both men and women. Difference-in-differences estimates in Panel B of Table 3 suggests that one additional school being built per 1000 children decreased the secondary school attainment rate by 2.3 percentage points for both men and women.

These heterogeneous effects are consistent with the finding in Duflo (2001) that the program increased years of schooling in sparsely populated areas but not in densely populated areas for men. Duflo (2001) interprets this as evidence that the program increased men's education mainly through decreasing average school distance. This could explain the difference in the results on the primary school attainment rate across the two subsamples, but has no explanatory power for the negative result on the secondary school attainment rate in densely populated regions.

4.3 Mechanism

In this section, I investigate further the surprising finding of a negative effect on secondary school attainment rate in densely populated regions. There are at least two possibilities: (1) building primary schools crowds out resources available to secondary schools and deteriorates secondary school quality and (2) a sudden increase in primary school availability may decrease primary education quality and hence the quality of primary school graduates. I explore the heterogeneity in the results for sparsely and densely populated regions and show that the first conjecture is more plausible.

Deterioration in Secondary Education Quality? Teacher scarcity is always a challenge in Indonesia's education system. Building primary schools increases the aggregate demand for teachers. This could affect the availability of secondary school teachers. To test this conjecture, I use the total number and average number of teachers per school in secondary education across regions in the years after the INPRES-SD program and to check whether there is a differential change in regions where more primary schools were constructed. Specifically, I estimate the following specification:

$$y_{jt} = \alpha_j + \beta_t + \sum_{l=2}^{6}(P_j d_{tl})\gamma_l + \sum_{l=2}^{6}(\boldsymbol{C_j} d_{tl})\delta_l + \varepsilon_{jt}$$

where j denotes region, and t denotes the survey year: 1 indicates year 1973/74, 2 indicates year 1978/79, 3 indicates year 1983/84, 4 indicates year 1988/89, 5 indicates year 1993/94, and 6 indicates 1995/96. y_{jt} indicates the total or average number of secondary school teachers in year t in region j. d_{tl} is a year dummy indicating whether $t = l$. α_j denotes the regency fixed effect, β_j denotes the year fixed effect. P_j is the school construction intensity in regency j. ε_{jk} is the error term. C_j represents other region-specific variables. The baseline year is 1973/74 ($t = 1$).

The results are presented in Table 4. The omitted baseline year is 1973/74. The negative coefficients in column 1 and column 2 suggest that in regions where more schools were constructed, a smaller increase is observed in the total number and the average number of teachers per school in secondary education in later years. Reassuringly, column (3) shows a positive effect of the program on the total number of teachers in primary school education, which is consistent with the teacher crowding out story.

Moreover, since the negative effect on secondary school attainment is only observed in densely populated regions, this negative effect on the number of teachers in secondary school should also only exist in densely populated regions if this is the mechanism responsible. Figure 7 separately plots the coefficients before the interaction term of the year dummy and school construction intensity from the previous specification for sparsely and densely populated regions. A negative effect on average number of teachers in secondary education is found for densely populated regions but not for sparsely populated regions. This confirms my conjecture that primary school construction increases the demand for teachers, which crowds out teacher resources available for secondary school education and leads to a negative effect on the secondary school attainment rate. Moreover, this phenomenon exists only in densely populated regions.

Deterioration in Primary Education Quality? A second conjecture is that the deterioration in primary school quality leads to a decrease in student quality among primary school graduates, and this in turn induces a lower secondary school attainment rate. To meet the surge in demand for teachers created by the school expansion, primary school teacher quality may have been sacrificed. Jalal et al. (2009); Bharati et al. (2018) To proxy for teacher quality in primary school, I adapt the method used in Behrman and Birdsall (1983); Bharati et al. (2018): calculating the percentage of primary teachers who completed secondary school or some college across regions in the Indonesian censuses of 1971, 1980, and 1990 and the inter-censuses of 1976 and 1985.

The empirical specification is similar to those above. The outcome variable is the percentage of primary school teachers who completed secondary school (or some college) in one regency in that census year. The baseline year is 1971, before the school expansion program started.

Figure 8 shows the coefficients of the interaction term between the year fixed effect and school construction intensity, separately for sparsely and densely populated regions, for the two proxies: the percentage of teachers completing secondary school (top) and completing some college (bottom). Consistent with the results in Bharati et al. (2018), I observe a negative impact of the program on teacher quality in 1976, but not for later years. However, I do not find different patterns between sparsely and densely populated regions. Therefore, this suggests that deterioration in primary education quality is not the main reason for the negative impact on the secondary school attainment rate.

4.4 Summary.

Here is a summary of the results on Education.

Result 1: The program has a positive effect on primary school attainment rate for men and a surprising negative effect on secondary school attainment rate for women.

Result 2: In sparsely populated regions, there is a positive effect on primary school attainment and secondary school attainment rate for men but zero effect for women.

Result 3: In densely populated regions, for both men and women, there is no effect on primary school attainment rate, but negative effect on secondary school attainment rate.

In light of the different effects on education in sparsely and densely populated regions, I should expect different results on marriage market outcomes in sparsely and densely populated regions. Moreover, I should expect zero effect on female marriage age or spousal age gap in sparsely populated regions since female education is not impacted.

5 Results on Marriage Market Outcomes

In this section, I present my empirical results on female marriage age and the spousal age gap. I first show reduced-form event study results on the impact of school construction on female marriage age and the spousal age gap for the treated female cohorts, separately for sparsely and densely populated regions. I then provide the 2SLS estimate of how female marriage age and the spousal age gap change with respect to the female education distribution using the school construction program as an instrument variable for female education distribution.

5.1 Reduced-Form Results

The empirical specification for the reduced-form results is the same as the previous specification for the education results.

Figure 9 presents the coefficients of the interaction between the birth cohort dummy and school construction intensity on female age at first marriage (top) and the spousal age gap (bottom) by female age group in 1974 in sparsely populated regions. All coefficients of the interaction between the birth cohort dummy and school construction intensity are not significantly different from zero. This is expected since female education was not substantially affected by the school construction program in sparsely populated regions. The results for densely populated regions are presented in Figure 10. The top panel shows a negative effect on female age at first marriage for one additional primary school being built in the region. Correspondingly, the bottom panel shows a positive effect on the spousal age gap.

Difference-in-differences estimates are presented in Table 5. The sample includes women born between 1953 and 1961 who were older than 12 in 1974, and women born between 1965 and 1970. "Post" indicates women born between 1965 and 1970. Columns (1) and (2) show the estimates for sparsely populated regions. Neither female age at first marriage nor the spousal age gap was impacted. Columns (3) and (4) present the estimates for densely populated regions and suggest that one additional school being constructed decreased the average female age at first marriage by 0.25 years and increased the spousal age gap by 0.075 years.

5.2 2SLS estimate

Since I lack first stage results for female education in sparsely populated regions, in this subsection, l focus on densely populated regions. Consider the following equation that characterizes how own education and the education distribution may affect an individual's choice of marriage age and the spousal age gap:

$$y_{ijk} = \alpha_j + \beta_k + D_{ijk}c + E_{jk}b + \nu_{ijk}$$

where α_j is a region fixed effect, β_k is a birth cohort fixed effect. y_{ijk} denotes the marriage age or spousal age gap of a woman i born in year k in region j, D_{ijk} is a dummy variable denoting whether woman i completes secondary school, and E_{jk} denotes the female secondary school attainment rate for birth cohort k in region j.

The coefficient of interest is b, indicating the impact of an increase in the proportion of educated women on female marriage age and the spousal age gap. However, ordinary least-squares (OLS) estimates of this equation may lead to biased estimates if there is correlation between E_{jk} and ν_{ijk} or there is correlation between D_{ijk} and ν_{ijk}. Unobserved individual characteristics such as ability or family attitudes could affect both her education attainment and marriage decisions, leading to a correlation between D_{ijk} and ν_{ijk}. Unobserved region cohort specific characteristics such as a construction of entertainment facilities or a promotion of family planning policies could affect the education attainment and marriage decisions of a few cohorts in the region, leading to a correlation between E_{jk} with ν_{ijk}.

To address this issue, let us take the average across individuals i given birth cohort k and region j:

$$\bar{y}_{jk} = \alpha_j + \beta_k + E_{jk}(b + c) + \bar{\nu}_{jk}$$

The school construction program provides a good instrument variable for E_{jk}, and hence I can obtain a valid estimate of $(b + c)$. OLS and 2SLS estimates of this specification are shown in Panel A of Table 6 for female age at first marriage and the spousal age gap. The IV estimate for female age at first marriage, although imprecisely estimated, indicates that increasing the share of female secondary graduates by 10 percentage points would increase the average female marriage age by 1.09 years. The IV estimate for the spousal age gap indicates that increasing the share of female secondary graduates by 10 percentage points would decrease the average spousal age gap by 0.35 years.

Separating the Effects of Own Education and the Education Distribution. From the previous specification, we know that:

$$E(y_{ijk}|D_{ijk} = 0) = \alpha_i + \beta_k + E_{jk}b$$

$$E(y_{ijk}|D_{ijk} = 1) = \alpha_i + \beta_k + c + E_{jk}b$$

Hence, $c = E(y_{ijk}|D = 1) - E(y_{ijk}|D = 0)$, which can be empirically estimated as the difference in the outcome variables conditional on education level. From the summary statistics, we know that the difference in age at first marriage age between female secondary school graduates and female primary school graduates is 4 years, while the difference in the spousal age gap between secondary school graduates and primary school graduates is (-1.5) years. Comparing this with previous estimates indicates that when controlling for a woman's education, increasing the percentage of female secondary graduates by 10 percentage points in her birth cohort would increase her first marriage age by 0.69 years and decrease the spousal age gap by 0.2 years.

5.3 Interpretation

My findings on the marriage market are consistent with the model when there is complementarity between higher education of husbands and younger age of wives in the marital surplus. A decrease in the percentage of female secondary school graduates creates a relative abundance of secondary school graduate men, which would encourage more women to marry earlier. In the Indonesian setting, the regions with more school constructed experienced a smaller increase in female secondary graduates, which created a relative abundance of male secondary graduates in the marriage market, and this encouraged even more women to marry earlier.

6 Conclusion

In this paper, I show that women adjust their marriage age to the education of other women in the local marriage market. Exploiting a massive school construction program in the late 1970s in Indonesia, I first document the empirical finding that in densely populated regions, the secondary school attainment rate declines for both men and women due to a crowding out of teacher resources in secondary education due to primary school construction. I then analyze the age at first marriage of the first few cohorts of women who were exposed to the school construction program. Since the spousal age gap is on average 5 years, these women's potential husbands' educations were minimally impacted. I find that women decrease their marriage age when there is a decrease in the average secondary school attainment rate of other women in the same cohort. To explain this, I construct a two-to-one dimensional matching model embedding female choice of marriage age into a two-period OLG framework and show that if with respect to marital surplus, (1) there exists complementarity between men's education and women's education, (2) there exists complementarity between men's education and women's young age, then women will decrease their marriage age in response to a decrease in other women's education. Intuitively, when the education of other women decreases, they tend to marry less-educated men due to the complementarity with education, and this creates an abundance of more-educated men. Due to the complementarity between men's education and women's young age, the abundance of more-educated men induce women to marry earlier.

This study is a step toward further understanding the effect of market conditions on individuals' marriage decisions and outcomes. Education expansion policies have been observed around the world. The empirical finding that female marriage age responds to other women's education has direct policy implications. When evaluating education policies with potential market-level impacts, we as researchers should consider both the direct effect on individuals and the indirect effect via changing market conditions.

References

Abramitzky, R., A. Delavande, and L. Vasconcelos (2011). Marrying up: the role of sex ratio in assortative matching. *American Economic Journal: Applied Economics 3*(3), 124–57.

Akresh, R., D. Halim, and M. Kleemans (2018). Long-term and intergenerational effects of education: Evidence from school construction in indonesia. Technical report, National Bureau of Economic Research.

André, P. and Y. Dupraz (2018). Education and polygamy: Evidence from cameroon. Technical report.

Angrist, J. (2002). How do sex ratios affect marriage and labor markets? evidence from america's second generation. *The Quarterly Journal of Economics 117*(3), 997–1038.

Arunachalam, R. and S. Naidu (2006). The price of fertility: marriage markets and family planning in bangladesh. *University of California, Berkeley*.

Ashraf, N., N. Bau, N. Nunn, and A. Voena (2016). Bride price and female education. Technical report, National Bureau of Economic Research.

Behrman, J. R. and N. Birdsall (1983). The quality of schooling: quantity alone is misleading. *The American Economic Review 73*(5), 928–946.

Bharati, T., S. Chin, and D. Jung (2018). Recovery from an early life shock through improved access to schools: Evidence from indonesia.

Bhaskar, V. (2015). The demographic transition and the position of women: A marriage market perspective.

Breierova, L. and E. Duflo (2004). The impact of education on fertility and child mortality: Do fathers really matter less than mothers? Technical report, National bureau of economic research.

Charles, K. K. and M. C. Luoh (2010). Male incarceration, the marriage market, and female outcomes. *The Review of Economics and Statistics 92*(3), 614–627.

Chiappori, P.-A., B. Salanié, and Y. Weiss (2017). Partner choice, investment in children, and the marital college premium. *American Economic Review 107*(8), 2109–67.

Choo, E. and A. Siow (2006). Who marries whom and why. *Journal of political Economy 114*(1), 175–201.

Decker, C., E. H. Lieb, R. J. McCann, and B. K. Stephens (2013). Unique equilibria and substitution effects in a stochastic model of the marriage market. *Journal of Economic Theory 148*(2), 778–792.

Dessy, S. and H. Djebbari (2010). High-powered careers and marriage: can women have it all? *The BE Journal of Economic Analysis & Policy 10*(1).

Dominguez, C. (2014). Aggregate effects on the marriage market of a big increase in educational attainment. Dissertation, Yale University.

Duflo, E. (2001). Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *American Economic Review 91*(4), 795–813.

Edlund, L. (2006). Marriage: past, present, future? *CESifo Economic Studies 52*(4), 621–639.

Galichon, A. and B. Salanié (2015). Cupid's invisible hand: Social surplus and identification in matching models.

Galichon, A. and B. Salanié (2017). The econometrics and some properties of separable matching models. *American Economic Review 107*(5), 251–55.

Hener, T. and T. Wilson (2018). Marital age gaps and educational homogamy-evidence from a compulsory schooling reform in the uk. Technical report, Ifo Working Paper.

Iyigun, M. and J. Lafortune (2016). Why wait? a century of education, marriage timing and gender roles.

Jalal, F., M. Samani, M. C. Chang, R. Stevenson, A. B. Ragatz, and S. D. Negara (2009). *Teacher certification in Indonesia: A strategy for teacher quality improvement.* Departemen Pendidikan Nasional, Republik Indonesia.

Jones, G. W. and P. Hagul (2001). Schooling in indonesia: crisis-related and longer-term issues. *Bulletin of Indonesian Economic Studies 37*(2), 207–231.

Jürges, H., S. Reinhold, and M. Salm (2011). Does schooling affect health behavior? evidence from the educational expansion in western germany. *Economics of Education Review 30*(5), 862–872.

Low, C. (2017). A reproductive capital model of marriage market matching. *Manuscript, Wharton School of Business*.

Malamud, O., A. Mitrut, and C. Pop-Eleches (2018). The effect of education on mortality and health: Evidence from a schooling expansion in romania. Technical report, National Bureau of Economic Research.

Shapley, L. S. and M. Shubik (1971). The assignment game i: The core. *International Journal of game theory 1*(1), 111–130.

Siow, A. (1998). Differntial fecundity, markets, and gender roles. *Journal of Political Economy 106*(2), 334–354.

Snodgrass, D. (1984, October). *Development Program Implementation Studies No.5: Inpres Sekolah Dasar*. Cambridge, MA.: Harvard Institute for International Development.

Tan, J. P. and A. Mingat (1992). *Education in Asia : a comparative study of cost and financing (English).* World Bank regional and sectoral studies. Washington, DC : The WorldBank.

World Bank (2018). *World Development Report, LEARNING to Realize Educations Promise.*

Zhang, H. (2018, October). Human Capital Investments, Differential Fecundity, and the Marriage Market. Working Papers 2018-7, Michigan State University, Department of Economics.

Table 1: Summary Statistics

2010 Census:	Old Cohorts Born between 1950 and 1961		Young Cohorts Born between 1962 and 1974	
	Males	Females	Males	Females
Education Attainment				
Some School	0.22	0.33	0.10	0.16
Primary School	0.56	0.54	0.53	0.57
Secondary School	0.17	0.11	0.30	0.23
University or above	0.05	0.02	0.07	0.05
Marriage outcomes				
Never Married (= 1)	0.01	0.02	0.04	0.03
Separated (= 1)	0.01	0.04	0.02	0.04
Husband Age minus Wife Age	5.83	4.84	4.45	4.60
Married Couples				
Husband More Educated (=1)	0.26	0.27	0.27	0.25
Same Education (=1)	0.62	0.63	0.57	0.59
Wife More Educated (=1)	0.13	0.11	0.16	0.16
Years of Schooling Gap (Husband's minus Wife's)	0.57	0.68	0.44	0.37
Observations	1,275,648	1,231,961	2,148,572	2,128,266

‘

Source: Indonesian Census 2010.

Figure 1: Marriage Frequencies (left) and Marriage Proportions (right) by Education for Females

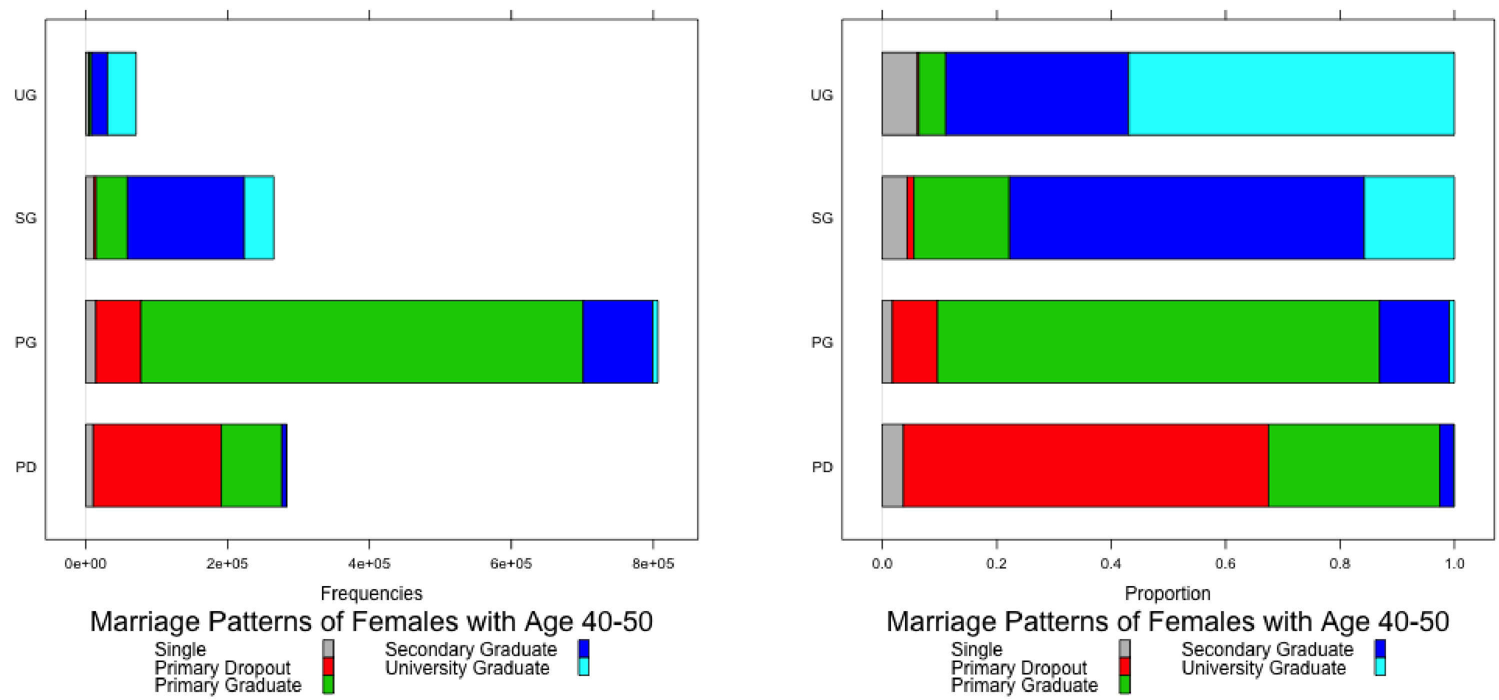


Figure 2: Primary School Construction Intensity Map

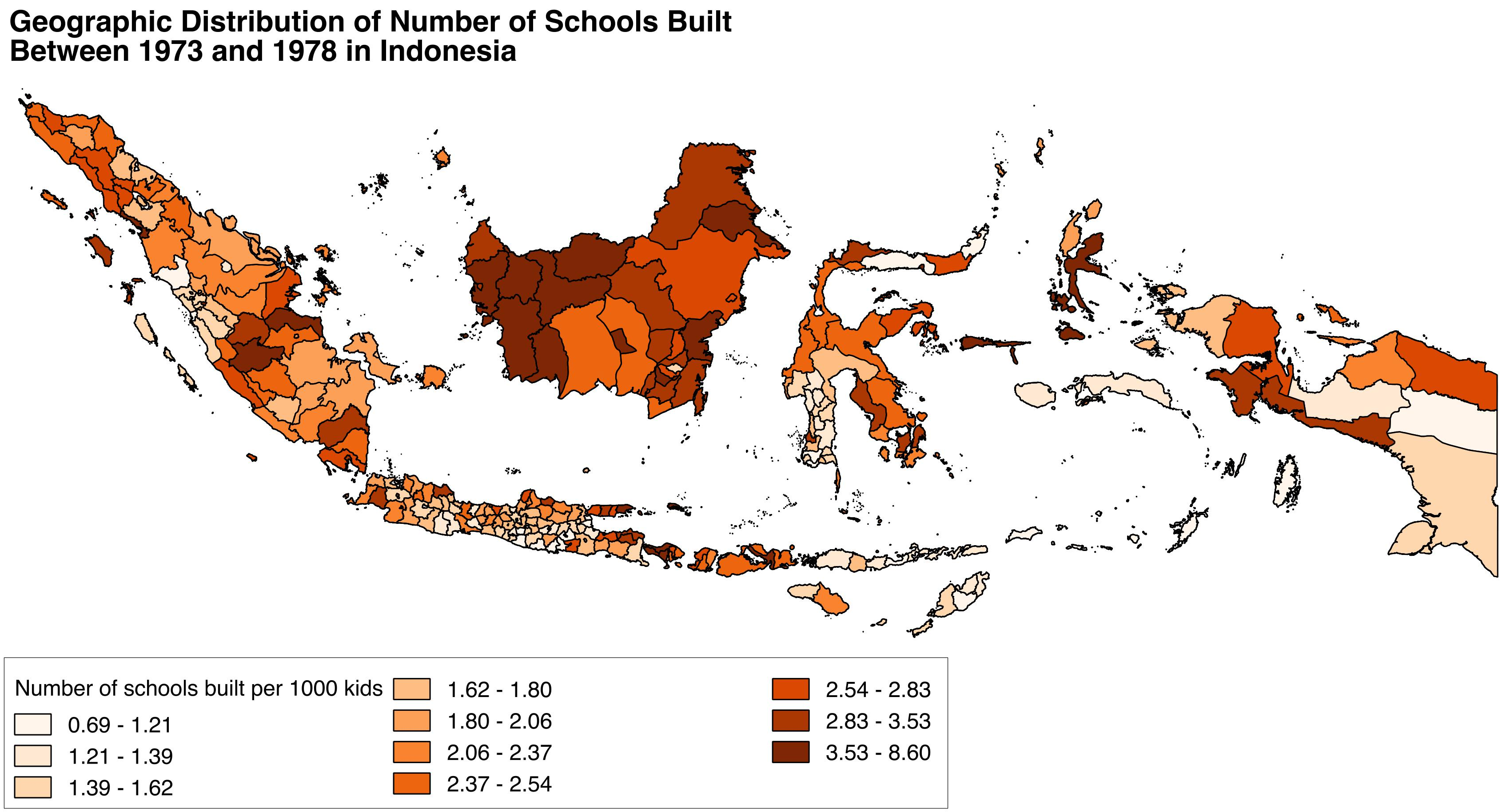
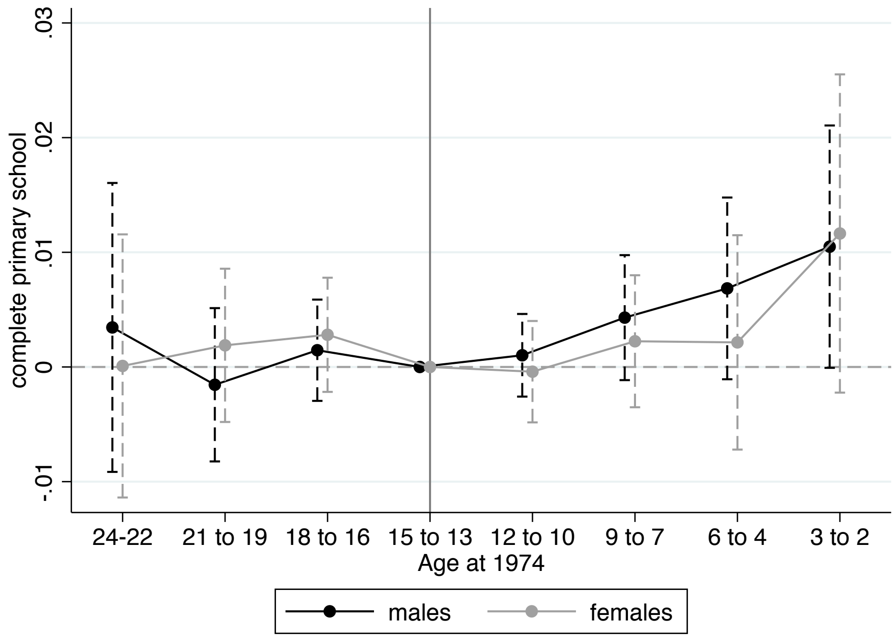
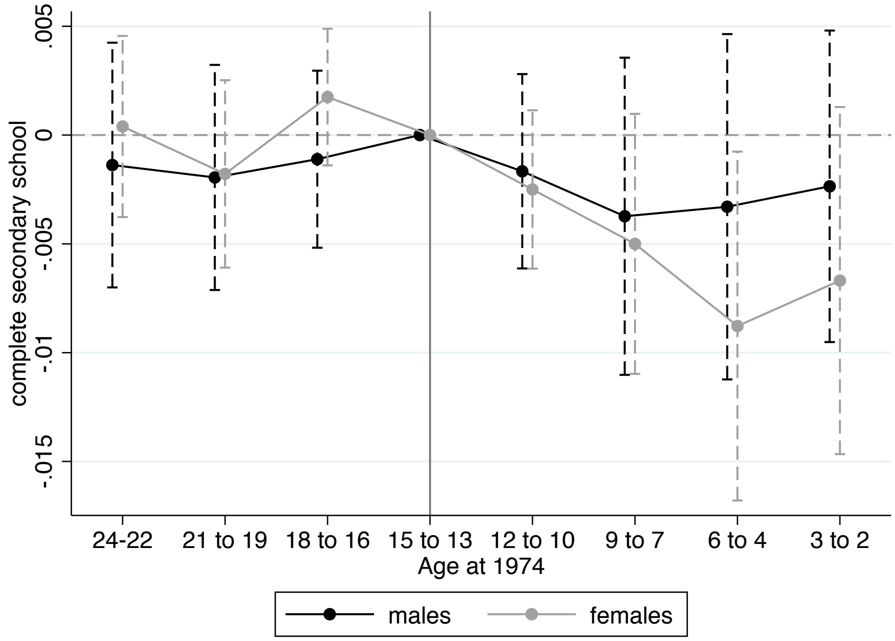


Figure 3: Coefficients of the Interactions: Age in 1974 * Program Intensity in the Region of Birth in the Education Equation for Completing Primary School



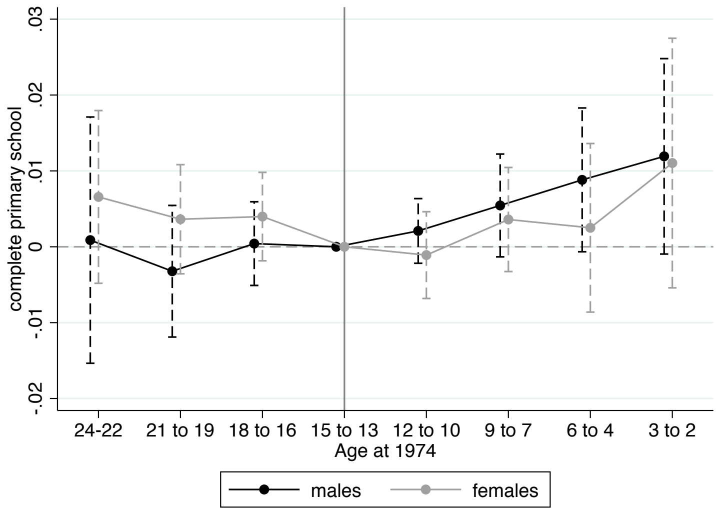
Note: This figure reports estimates of the effect of school construction on primary school completion for 3-year cohorts separately for males and females born in one regency. Dependent variable is the percentage of individuals completing primary school when observed in 2010. The x-axis reports the age range (in 1974) for each cohort and the y-axis reports the estimated coefficient, which can be interpreted as the effect of one additional primary school built per 1000 kids on primary school attainment rate in that regency. The sample consists of individuals born between 1950 and 1972 observed in 2010 Indonesian census. The vertical line indicates the youngest cohort that did not receive any treatment from school construction, since they were out of primary school at 1974, when the first round of constructed primary schools became available. Confidence intervals of 95% were plotted. The figure shows zero effect for individuals older than 13 at 1974, but an increasing positive effect for males younger than 13. For females, the effect is smaller.

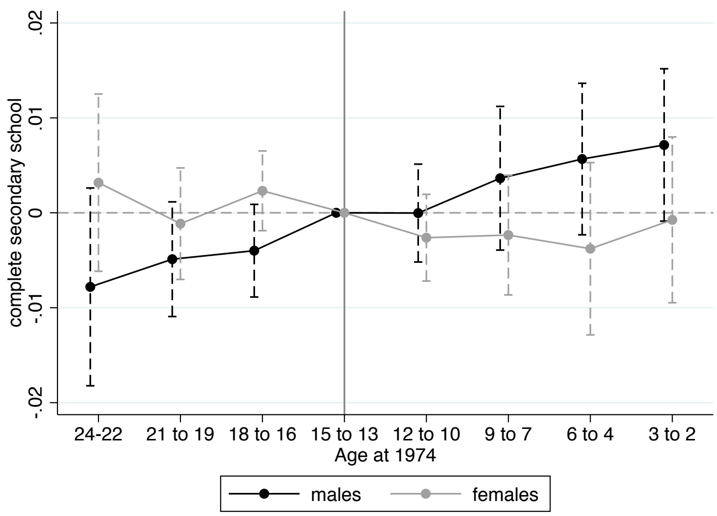
Figure 4: Coefficients of the Interactions: Age in 1974 * Program Intensity in the Region of Birth in the Education Equation for Completing Secondary School



Note: This figure is built like Figure 3 but considers secondary school attainment rate. It reports estimates of the effect of school construction on secondary school completion for 3-year cohorts separately for males and females born in one regency. Dependent variable is the percentage of individuals completing secondary school when observed in 2010. The x-axis reports the age range (in 1974) for each cohort and the y-axis reports the estimated coefficient, which can be interpreted as the effect of one additional primary school built per 1000 kids on primary school attainment rate in that regency. The sample consists of individuals born between 1950 and 1972 observed in 2010 Indonesian census. The vertical line indicates the youngest cohort that did not receive any treatment from school construction, since they were out of primary school at 1974, when the first round of constructed primary schools became available. Confidence intervals of 95% were plotted.

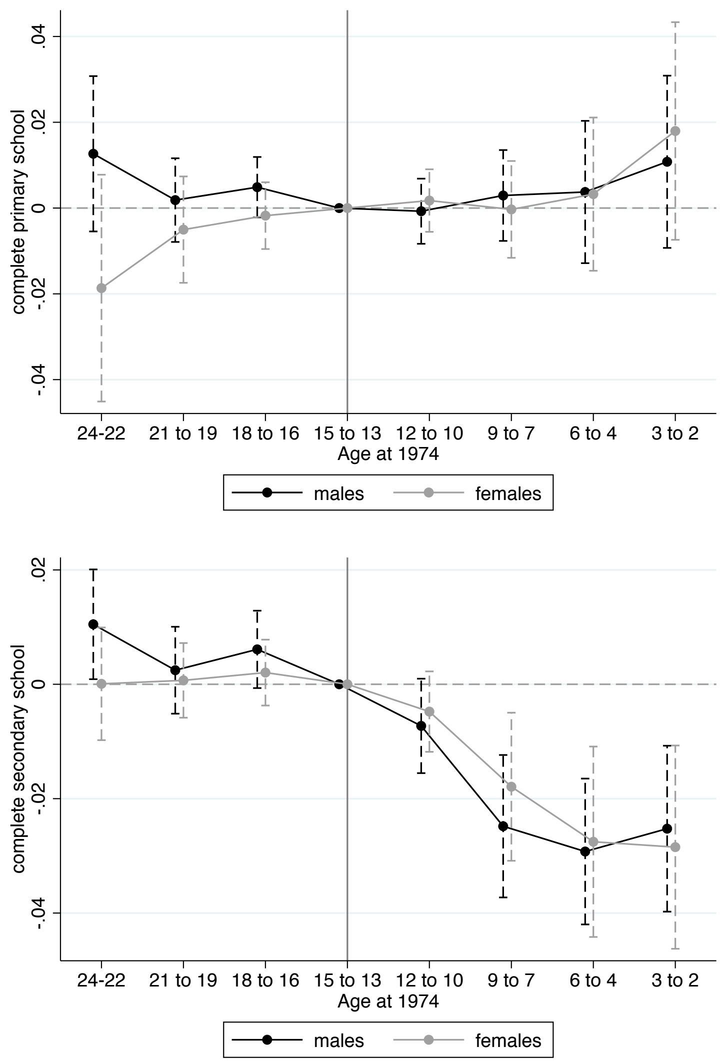
Figure 5: Effect of Program on Completing Primary School (top) and Completing Secondary School (bottom) in Sparsely Populated Areas





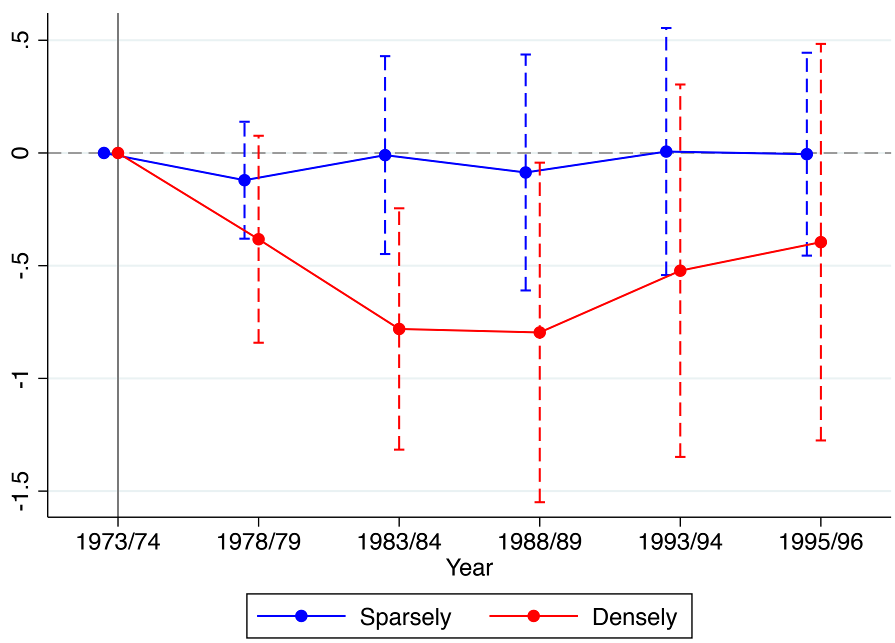
Note: This figure is similar to Figure 3 and Figure 4 but focuses on a subgroup: sparsely populated regions. It reports estimates of the effect of school construction on primary school completion (top) and secondary school completion (bottom) for 3-year cohorts separately for males and females in this subgroup. Sparsely populated regions are defined as those regions with population density smaller than the weighted medium density in 1971. This figure shows that:

Figure 6: Effect of Program on Completing Primary School (top) and Completing Secondary School (bottom) in Densely Populated Areas



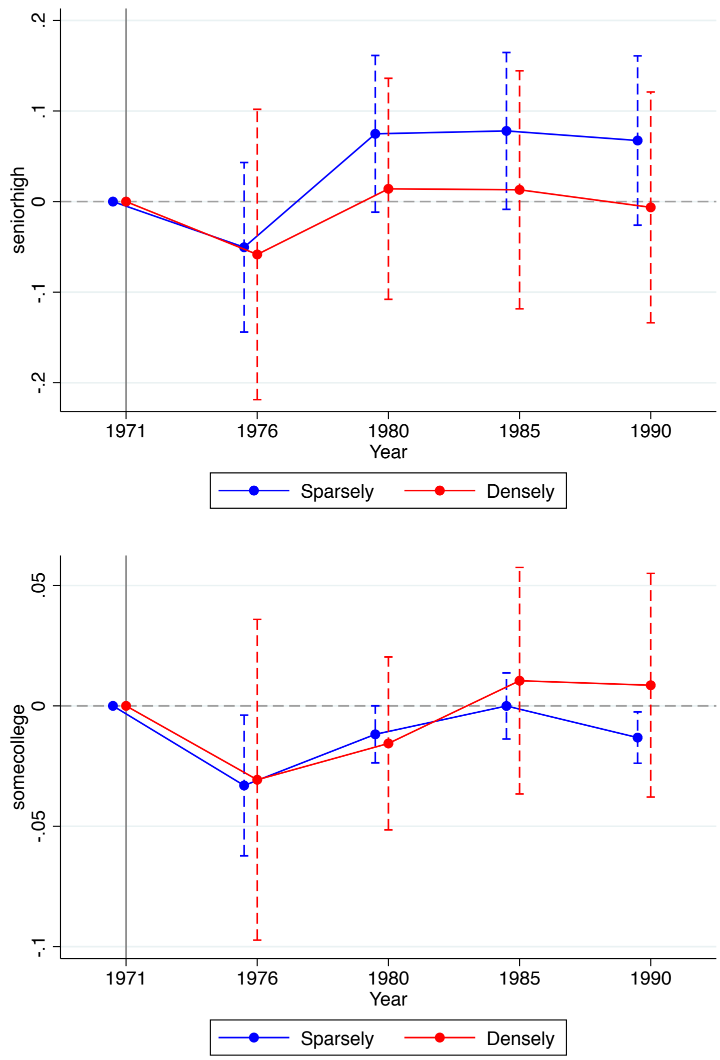
Note: This figure is built like Figure 5 but focuses on the other subgroup: densely populated regions. It reports estimates of the effect of school construction on primary school completion (top) and secondary school completion (bottom) for 3-year cohorts separately for males and females in this subgroup. Densely populated regions are defined as those regions with population density larger than the weighted medium density in 1971. This figure shows that:

Figure 7: Coefficients of the Interactions: Census Year * Program Intensity in the Regency in Average Number of Secondary School Teachers Equation



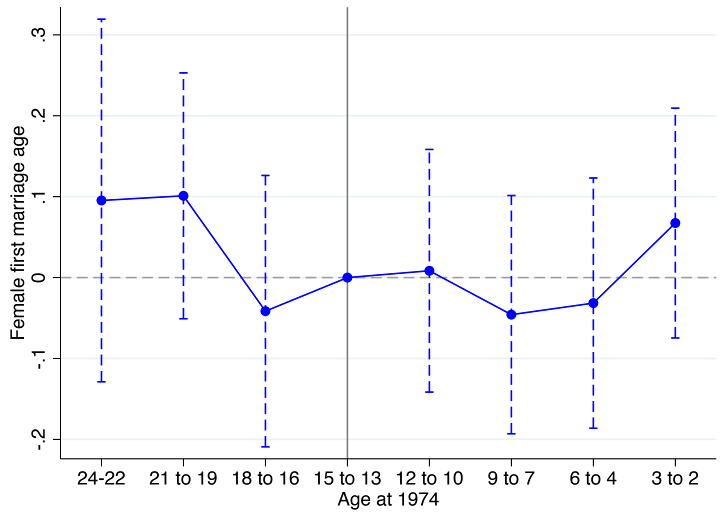
Note: This figure reports estimates of the effect of school construction on average number of teachers in secondary school across different years in sparsely populated areas and densely populated areas. The baseline year is 1973/74. The data was provided by Indonesian Education Ministry and was collected in Duflo (2001). The dependent variable was the average number of teachers in secondary school across different regencies. This figure supports the argument that the negative effect on secondary school attainment is due to teacher resource crowding out in densely populated regions because of primary school construction.

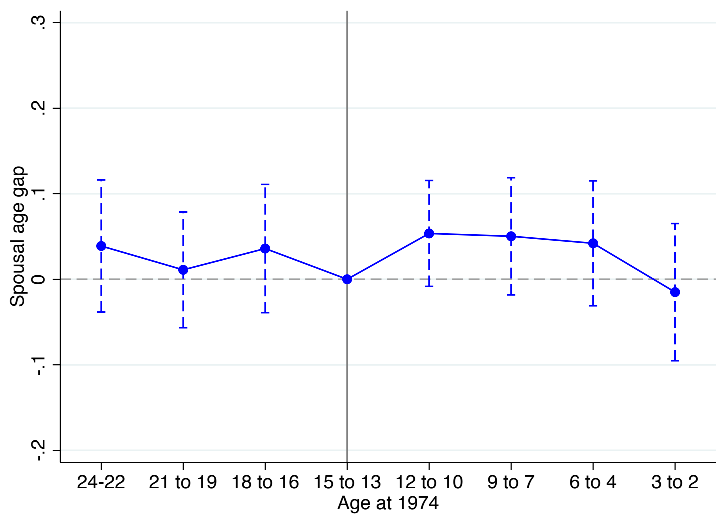
Figure 8: Effect of Program on Primary Teacher Education



Note: This figure reports estimates of the effect of school construction on the education level of primary school teacher in the two subsamples: sparsely and densely populated regions. Dependent variable for the top panel is a dummy indicating the teacher completes secondary school, for the bottom panel is a dummy indicating the teacher has some post-secondary education. The baseline year is 1971. Primary school teacher information for each region is obtained from identifying those individuals who claim their occupation is primary school teacher in the census year.

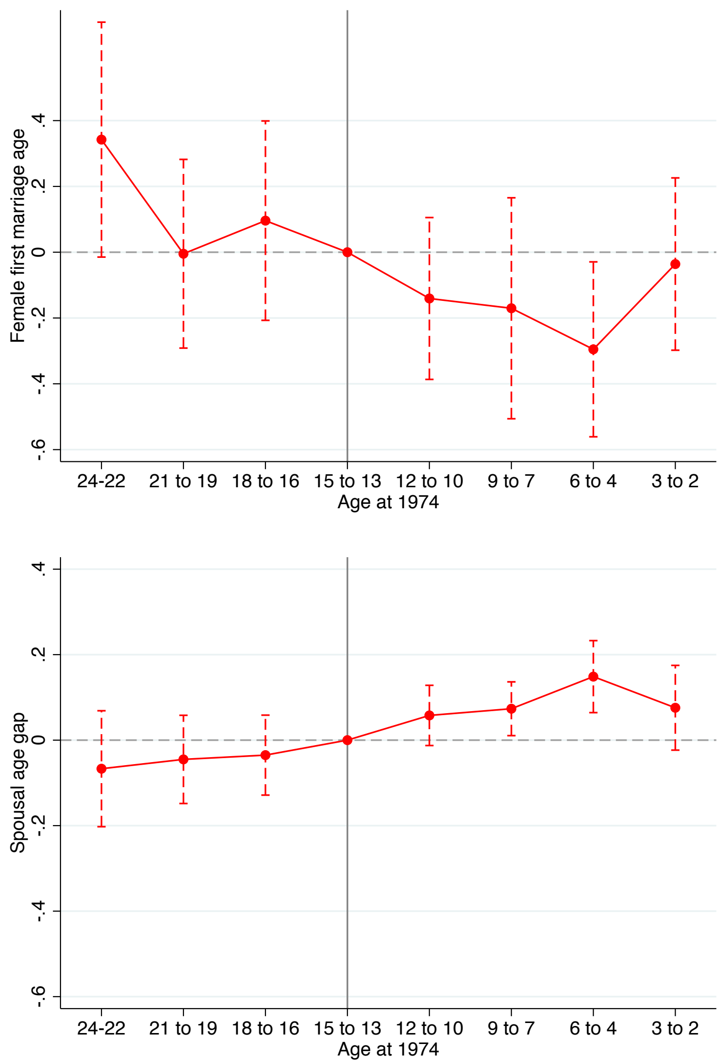
Figure 9: Effect of Program on Female Age at First Marriage (top) and the Spousal Age Gap (bottom) in Sparsely Populated Areas





Note: This figure reports estimates of the effect of school construction on female first marriage age (top) and spousal age gap (bottom) for 3-year cohorts females in sparsely populated areas. The x-axis reports the age range (in 1974) for each cohort and the y-axis reports the estimated coefficient, which can be interpreted as the effect of one additional primary school built per 1000 kids on primary school attainment rate in that regency.

Figure 10: Effect of Program on Female Age at First Marriage (top) and the Spousal Age Gap (bottom) in Densely Populated Areas



Note: This figure reports estimates of the effect of school construction on female first marriage age (top) and spousal age gap (bottom) for 3-year cohorts females in densely populated areas. The x-axis reports the age range (in 1974) for each cohort and the y-axis reports the estimated coefficient, which can be interpreted as the effect of one additional primary school built per 1000 kids on primary school attainment rate in that regency.

Table 2: Effect of School Construction on Education

All sample:	Indicator for Completing at least:					
	Primary School			Secondary School		
Males:	(1)	(2)	(3)	(4)	(5)	(6)
Post × Intensity	0.022***	0.0079*	0.0062**	-0.0077**	-0.0012	-0.0014
	(0.0060)	(0.0042)	(0.0025)	(0.0033)	(0.0030)	(0.0018)
Dep. var. mean			0.847			0.306
Observations	6509	6302	6302	6509	6302	6302
Clusters	283	274	274	283	274	274
Adjusted R-squared	0.917	0.951	0.974	0.961	0.959	0.974
Duflo Controls:	No	Yes	Yes	No	Yes	Yes
Log-linear Trend:	No	No	Yes	No	No	Yes
Females:						
Post × Intensity	0.020***	0.0027	0.0041	-0.021***	-0.0064*	-0.0053**
	(0.0065)	(0.0052)	(0.0034)	(0.0049)	(0.0033)	(0.0023)
Dep. var. mean			0.766			0.211
Observations	6509	6302	6302	6509	6302	6302
Clusters	283	274	274	283	274	274
Adjusted R-squared	0.934	0.956	0.979	0.949	0.960	0.976
Duflo Controls:	No	Yes	Yes	No	Yes	Yes
Log-linear trend:	No	No	Yes	No	No	Yes

Notes: This table displays results on the effect of school building on education attainment (completing primary school and completing secondary school) for males and females. Following the strategy of Duflo (2001), the sample consists of individuals born between either 1968 and 1972 or 1950 and 1961. Post refers to the treated cohort, born between 1968 and 1972, while the untreated cohort was born between 1950 and 1961. Educational attainment data are taken from the Indonesian 2010 Census. Intensity is the number of schools built in a region per 1,000 kids in the school-aged population. All columns include district fixed effect, school year fixed effect, school year interacted with number of children at 1971. Duflo Controls consist of school year interacted with enrollment rate at 1971 and school year interacted with water sanitization program. Standard errors are clustered at the birthplace district level. Significance levels: * 10%, ** 5%, *** 1%.
Source: Indonesian Census 2010

Table 3: Heterogeneous Effect of School Construction on Education

Panel A: Density < Medium:	Indicator for Completing at least:			
	Primary School		Secondary School	
	(1) Male	(2) Female	(3) Male	(4) Female
Post × Intensity	0.010**	0.0066	0.0069**	-0.00066
	(0.0051)	(0.0062)	(0.0032)	(0.0037)
Dep. var. mean	0.820	0.736	0.270	0.183
Observations	4209	4209	4209	4209
Clusters	183	183	183	183
Adjusted R-squared	0.949	0.952	0.937	0.941
Panel B: Density > Medium:				
Post × Intensity	0.0054	-0.0060	-0.023***	-0.023***
	(0.0066)	(0.0073)	(0.0055)	(0.0073)
Dep. var. mean	0.873	0.795	0.341	0.239
Observations	2093	2093	2093	2093
Clusters	91	91	91	91
Adjusted R-squared	0.957	0.967	0.975	0.975
Duflo Controls:	Yes	Yes	Yes	Yes
Log-linear trend:	Yes	Yes	Yes	Yes

Notes: This table is similar to Table 2 and displays the heterogeneity effect of school building on education attainment in sparsely and densely populated regions. All columns include district fixed effect, school year fixed effect, school year interacted with number of children at 1971. Duflo Controls consist of school year interacted with enrollment rate at 1971 and school year interacted with water sanitization program. Standard errors are clustered at the birthplace district level. Significance levels: * 10%, ** 5%, *** 1%. Source: Indonesian Census 2010

Table 4: Effect of School Construction on Number of Teachers in Secondary and Primary Education

	Secondary School		Primary School	
	(1) Total number	(2) Average number	(3) Total number	(4) Average number
INPRES Intensity ×				
year=1978/79	-13.9	-0.19*	28.0***	-0.16**
	(8.48)	(0.11)	(9.91)	(0.063)
year=1983/84	-43.5	-0.14	61.1**	-0.069
	(31.8)	(0.18)	(29.2)	(0.11)
year=1988/89	-65.1	-0.20	89.4*	-0.031
	(49.1)	(0.21)	(45.5)	(0.066)
year=1993/94	-59.3	-0.086	95.1*	-0.062
	(52.0)	(0.22)	(57.0)	(0.075)
year=1995/96	-51.7	-0.074	177.4**	0.29*
	(60.0)	(0.19)	(68.6)	(0.17)
Dep. var. mean in 1973/74	555.996	14.723	1529.996	6.762
Dep. var. mean in 1995/96	2583.821	22.989	4207.180	8.345
Observations	1,656	1,656	1,664	1,664
R-squared	0.928	0.929	0.942	0.829
Duflo Controls:	Yes	Yes	Yes	Yes

Notes: This table displays the effect of school construction on the number of teachers in secondary and primary education in the future years. Baseline year is 1973/74. All columns include district fixed effect, school year fixed effect, school year interacted with number of children at 1971. Duflo Controls consist of school year interacted with enrollment rate at 1971 and school year interacted with water sanitization program. Standard errors are clustered at the birthplace district level. Significance levels: * 10%, ** 5%, *** 1%.

Source: Indonesian Education Ministry

Table 5: Reduced-form Effect of School Construction on Female Marriage Outcomes

	Density < Medium		Density > Medium	
Panel A: **Female First Marriage Age:**				
	(1)	(2)	(3)	(4)
Post × Intensity	-0.054	0.024	-0.27**	-0.25
	(0.076)	(0.078)	(0.12)	(0.16)
Dep. var. mean		19.231		19.153
Observations	2664	2664	1365	1365
Clusters	183	183	91	91
Adjusted R-squared	0.673	0.702	0.710	0.717
Panel B: **Spousal age gap**				
Post × Intensity	0.030	0.057	0.18***	0.075*
	(0.025)	(0.040)	(0.044)	(0.040)
Dep. var. mean		4.838		4.776
Observations	2745	2745	1365	1365
Clusters	183	183	91	91
Adjusted R-squared	0.858	0.879	0.895	0.917
Duflo Controls:	Yes	Yes	Yes	Yes
Log-linear trend:	No	Yes	No	Yes

Notes: This table displays the reduced-form effect of school construction on female first marriage age (top) and spousal age gap (bottom) in sparsely populated regions (left) and densely populated regions (right). Post refers to the first few treated cohorts that were affected by the school construction program, i.e., those born between 1965 and 1970, while the untreated cohort was born between 1953 and 1961. Female first marriage data is taken from Indonesian SUPAS 2005. Spousal age gap data is taken from the Indonesian 2010 Census. All columns include district fixed effect, school year fixed effect, school year interacted with number of children at 1971. Duflo Controls consist of school year interacted with enrollment rate at 1971 and school year interacted with water sanitization program. Standard errors are clustered at the birthplace district level. Significance levels: * 10%, ** 5%, *** 1%.
Source: Indonesian SUPAS 2005, Indonesian Census 2010

Table 6: Results of Female Education Distribution on Female Marriage Outcomes

Panel A: OLS and IV	Female first marriage age		Spousal age gap	
	(1) OLS	(2) IV	(3) OLS	(4) IV
Percentage of females with secondary degree	1.92	10.9*	-2.71***	-3.49*
	(1.37)	(6.53)	(0.38)	(2.03)
First Stage F statistics		12.929		12.929
Dep. var. mean	19.647		4.550	
Observations	1365	1365	1365	1365
Clusters	91	91	91	91
Adjusted R-squared	0.763	0.754	0.920	0.920
Duflo Controls:	Yes	Yes	Yes	Yes
Log-linear trend:	Yes	Yes	Yes	Yes

Panel B: First stage and Reduced form	Complete Secondary	Female age at first marriage	Spousal age gap
Post × Intensity	-0.021***	-0.23*	0.075*
	(0.0059)	(0.14)	(0.040)
Dep. var. mean	0.261	19.647	4.550
Observations	1365	1365	1365
Clusters	91	91	91
Adjusted R-squared	0.988	0.763	0.917
Duflo Controls:	Yes	Yes	Yes
Log-linear trend:	Yes	Yes	Yes

Notes: This table displays the OLS and IV estimates of the effect of female education distribution on marriage market outcomes. All columns include district fixed effect, school year fixed effect, school year interacted with number of children at 1971. Post refers to the first few treated cohorts that were affected by the school construction program, i.e., those born between 1965 and 1970, while the untreated cohort was born between 1953 and 1961. Duflo Controls consist of school year interacted with enrollment rate at 1971 and school year interacted with water sanitization program. Standard errors are clustered at the birthplace district level. Significance levels: * 10%, ** 5%, *** 1%.
Source: Indonesian SUPAS 2005, Indonesian Census 2010

7 Appendix

A. Proof for Lemma 1

Proof. Denote $\tilde{u_i}, \tilde{v_j}$ the equilibrium utility individuals get. We know that if woman i and man j match in equilibrium, then $\tilde{u_i} + \tilde{v_j} = \alpha_{xy} + \gamma_{xy} + \varepsilon_{iy} + \eta_{xj}$.

For woman i of type x,

$$\begin{aligned}\tilde{u_i} &= \max_{j \in J} \{\alpha_{xy} + \gamma_{xy} + \varepsilon_{iy} + \eta_{xj} - \tilde{v_j}, \varepsilon_{i0}\} \\ &= \max_{y \in Y} \left\{ \max_{j \text{ where } j_i = y} (\alpha_{xy} + \gamma_{xy} + \eta_{xj} - \tilde{v_j}) + \varepsilon_{iy}, \varepsilon_{i0} \right\}\end{aligned}$$

Define $U^{xy} = \max_{j \text{ where } j_i = y} (\alpha_{xy} + \gamma_{xy} + \eta_{xj} - \tilde{v_j}), U^{x0} = 0$, then we get:

$$\tilde{u_i} = \max_{y \in Y_0} (U^{xy} + \varepsilon_{iy})$$

Moreover,

$$\tilde{u_i} \geq U^{xy} + \varepsilon_{iy}, \forall y \in Y_0$$

and it achieves equality when the set of women of type x matched with men of type y is nonempty.

With similar notations, define $V^{xy} = \max_{i \text{ where } x_i = x} (\alpha_{xy} + \gamma_{xy} + \varepsilon_{iy} - \tilde{u_i}), V^{0y} = 0$, then:

$$\tilde{v_j} = \max_{x \in X_0} (V^{xy} + \eta_{xj})$$

$$\tilde{v_j} \geq V^{xy} + \eta_{xj}, \forall x \in X_0$$

and it achieves equality when the set of men of type y matched with women of type x is nonempty.

If there exist women of type x matched with men of type y,

$$\tilde{u_i} = U^{xy} + \varepsilon_{iy}$$

$$\tilde{v_j} = V^{xy} + \eta_{xj}$$

Hence $U^{xy} + V^{xy} = \alpha_{xy} + \gamma_{xy}$ $\square$

B. An important lemma

To prove the propositions, I'll first establish an important lemma related to how probabilities of singlehood change related to the shift of marginals in types.

Lemma 2. *Assume the idiosyncratic tastes follow Gumbel distributions. Assume there are two types for each side, denote the female marginal as $n = (x, 1-x)$ and male marginal as $m = (y, 1-y)$, the surplus matrix as:*

$$\Phi = \begin{bmatrix} \Phi_{LL} & \Phi_{LH} \\ \Phi_{HL} & \Phi_{HH} \end{bmatrix}$$

denote the mass of singles of females (males) in equilibrium as: $\mu_{L0}, \mu_{H0}(\mu_{0L}, \mu_{0H})$ then:

(a)

$$\frac{\partial \mu_{L0}}{\partial x} > 0, \frac{\partial \mu_{H0}}{\partial x} < 0$$

(b) If the marital surplus function is super-modular, i.e., $\Phi_{LL} + \Phi_{HH} > \Phi_{LH} + \Phi_{HL}$, then

(b1)

$$\frac{\partial \mu_{0L}}{\partial x} > 0 \Rightarrow \frac{\partial \mu_{0H}}{\partial x} > 0$$

$$\frac{\partial \mu_{0H}}{\partial x} < 0 \Rightarrow \frac{\partial \mu_{0L}}{\partial x} < 0$$

(b2) There exists some $\underline{\delta_x}, \bar{\delta_x}, \underline{\delta_y}, \bar{\delta_y}$, such that if $\underline{\delta_x} < x < \bar{\delta_x}, \underline{\delta_y} < y < \bar{\delta_y}$, then:

$$\frac{\partial \frac{\mu_{0L}}{\mu_{0H}}}{\partial x} < 0$$

Proof. Denote $a = \exp(\frac{\Phi_{LL}}{2}), b = \exp(\frac{\Phi_{LH}}{2}), c = \exp(\frac{\Phi_{HL}}{2}), d = \exp(\frac{\Phi_{HH}}{2})$;

denote $s_{L0} = \sqrt{\mu_{L0}}, s_{H0} = \sqrt{\mu_{H0}}, s_{0L} = \sqrt{\mu_{0L}}, s_{0H} = \sqrt{\mu_{0H}}$;

denote $D_{L0} = \frac{\partial s_{L0}}{\partial x}, D_{H0} = \frac{\partial s_{H0}}{\partial x}, D_{0L} = \frac{\partial s_{0L}}{\partial x}, D_{0H} = \frac{\partial s_{0H}}{\partial x}$.

Then we can rewrite the feasibility constraints with the matching function as:

$$s_{L0}^2 + s_{L0}s_{0L}a + s_{L0}s_{0H}b = x$$

$$s_{H0}^2 + s_{H0}s_{0L}c + s_{H0}s_{0H}d = 1 - x$$

$$s_{0L}^2 + s_{L0}s_{0L}a + s_{H0}s_{0L}c = y$$

$$s_{0H}^2 + s_{L0}s_{0H}b + s_{H0}s_{0H}d = 1 - y$$

In the four equations above, taking the derivative with respect to x, we get:

$$(2s_{L0} + as_{0L} + bs_{0H})D_{L0} + s_{L0}(aD_{0L} + bD_{0H}) = 1 \tag{1}$$

$$(2s_{H0} + cs_{0L} + ds_{0H})D_{H0} + s_{H0}(cD_{0L} + dD_{0H}) = -1 \tag{2}$$

$$(2s_{0L} + as_{L0} + cs_{H0})D_{0L} + s_{0L}(aD_{L0} + cD_{H0}) = 0 \tag{3}$$

$$(2s_{0H} + bs_{L0} + ds_{H0})D_{0H} + s_{0H}(bD_{L0} + dD_{H0}) = 0 \tag{4}$$

Hence we can express D_{0L}, D_{0H} using D_{L0}, D_{H0} from Equation 3 and Equation 4:

$$D_{0L} = -\frac{s_{0L}(aD_{L0} + cD_{H0})}{2s_{0L} + as_{L0} + cs_{H0}} \tag{5}$$

$$D_{0H} = -\frac{s_{0H}(bD_{L0} + dD_{H0})}{2s_{0H} + bs_{L0} + ds_{H0}} \tag{6}$$

Plugging in Equation 1 and Equation 2, we get:

$$(2s_{L0}+\frac{as_{0L}(2s_{0L} + cs_{H0})}{2s_{0L} + as_{L0} + cs_{H0}}+\frac{bs_{0H}(2s_{0H} + ds_{H0})}{2s_{0H} + bs_{L0} + ds_{H0}})D_{L0}-(\frac{acs_{L0}s_{0L}}{2s_{0L} + as_{L0} + cs_{H0}}+\frac{bds_{L0}s_{0H}}{2s_{0H} + bs_{L0} + ds_{H0}})D_{H0} = 1 \tag{7}$$

$$(2s_{H0}+\frac{cs_{0L}(2s_{0L} + as_{L0})}{2s_{0L} + as_{L0} + cs_{H0}}+\frac{ds_{0H}(2s_{0H} + bs_{L0})}{2s_{0H} + bs_{L0} + ds_{H0}})D_{H0}-(\frac{acs_{H0}s_{0L}}{2s_{0L} + as_{L0} + cs_{H0}}+\frac{bds_{H0}s_{0H}}{2s_{0H} + bs_{L0} + ds_{H0}})D_{L0} = -1 \tag{8}$$

Add Equation 7 and Equation 8, we get:

$$(2s_{L0}+\frac{2as_{0L}^2}{2s_{0L} + as_{L0} + cs_{H0}}+\frac{2bs_{0H}^2}{2s_{0H} + bs_{L0} + ds_{H0}})D_{L0}+(2s_{H0}+\frac{2cs_{0L}^2}{2s_{0L} + as_{L0} + cs_{H0}}+\frac{2ds_{0H}^2}{2s_{0H} + bs_{L0} + ds_{H0}})D_{H0} = 0 \tag{9}$$

Hence D_{L0} and D_{H0} have opposite signs. With Equation 7, we know:

$$D_{L0} > 0, D_{H0} < 0$$

This completes the proof for (a).

For part (b1) of the lemma, with super-modularity, we know:

$$a * d > b * c$$

Since $D_{L0} > 0$:

$$\frac{a}{c} D_{L0} > \frac{b}{d} D_{L0}$$

$$\Rightarrow: \frac{a}{c} D_{L0} + D_{H0} > \frac{b}{d} D_{L0} + D_{H0}$$

Hence:

$$aD_{L0} + cD_{H0} < 0 \quad \Rightarrow \quad bD_{L0} + dD_{H0} < 0$$

$$bD_{L0} + dD_{H0} > 0 \quad \Rightarrow \quad aD_{L0} + cD_{H0} > 0$$

Recall Equation 5 and Equation 6, we have:

$$\frac{\partial \mu_{0L}}{\partial x} > 0 \Rightarrow \frac{\partial \mu_{0H}}{\partial x} > 0$$

$$\frac{\partial \mu_{0H}}{\partial x} < 0 \Rightarrow \frac{\partial \mu_{0L}}{\partial x} < 0$$

Proof for (b1) is complete.

Now let's prove part (b2):

$$\frac{\partial \frac{s_{0L}}{s_{0H}}}{\partial x} = \frac{D_{0L} s_{0H} - D_{0H} s_{0L}}{s_{0H}^2}$$

Using Equation 5 and Equation 6,

$$\begin{aligned} D_{0L} s_{0H} - D_{0H} s_{0L} &= -\frac{s_{0L} s_{0H} (aD_{L0} + cD_{H0})}{2s_{0L} + as_{L0} + cs_{H0}} + \frac{s_{0L} s_{0H} (bD_{L0} + dD_{H0})}{2s_{0H} + bs_{L0} + ds_{H0}} \\ &= \frac{s_{0L} s_{0H} ([b(2s_{0L} + cs_{H0}) - a(2s_{0H} + ds_{H0})] D_{L0} + [d(2s_{0L} + as_{L0}) - c(2s_{0H} + bs_{L0})] D_{H0})}{(2s_{0L} + as_{L0} + cs_{H0})(2s_{0H} + bs_{L0} + ds_{H0})} \end{aligned}$$

It has the same sign as:

$$[2bs_{0L} - 2as_{0H} + (bc - ad)s_{H0}] D_{L0} + [2ds_{0L} - 2cs_{0H} + (ad - bc)s_{L0}] D_{H0}$$

$$= 2s_{0L}(bD_{L0} + dD_{H0}) - 2s_{0H}(aD_{L0} + cD_{H0}) - (ad - bc)(D_{L0} s_{H0} - D_{H0} s_{L0})$$

We know that $(ad-bc)(D_{L0}s_{H0}-D_{H0}s_{L0})>0$, since $ad-bc>0, D_{L0}>0, D_{H0}<0$.

According to (b1), there are only three cases:

(Case 1): $aD_{L0}+cD_{H0}>0, bD_{L0}+dD_{H0}<0$; it's straightforward to show:

$$\frac{\partial \frac{s_{0L}}{s_{0H}}}{\partial x}<0$$

(Case 2): $aD_{L0}+cD_{H0}>0, bD_{L0}+dD_{H0}>0$

in this case, from Equation 9, we know $s_{L0}D_{L0}+s_{H0}D_{H0}<0$, hence:

$$\frac{a}{c}>\frac{b}{d}>\frac{s_{L0}}{s_{H0}}$$

Since we know $\frac{s_{L0}}{s_{H0}}$ increases with x, to satisfy previous inequality, we know that x is also relatively small in this case.

There exists some $\underline{\delta_x}, \bar{\delta_y}$ such that for $x>\underline{\delta_x}, y<\bar{\delta_y}$,

$$\frac{\partial \frac{s_{0L}}{s_{0H}}}{\partial x}<0$$

(Intuition: we need x to be away from 0 and y to be away from 1 to avoid large value of s_{0L} and small value of s_{0H}.)

(Case 3): $aD_{L0}+cD_{H0}<0, bD_{L0}+dD_{H0}<0$

in this case, from equation (9), we know $s_{L0}D_{L0}+s_{H0}D_{H0}>0$, hence:

$$\frac{s_{L0}}{s_{H0}}>\frac{a}{c}>\frac{b}{d}$$

x is relatively large in this case. There exists some $\bar{\delta_x}, \underline{\delta_y}$ such that for $x<\bar{\delta_x}, y>\underline{\delta_y}$,

$$\frac{\partial \frac{s_{0L}}{s_{0H}}}{\partial x}<0$$

(Intuition: we need x to be away from 1 and y to be away from 0 to avoid small value of s_{0L} and large value of s_{0L}.)

Proof for part (b2) is complete. $\square$

Lemma 3. *An extension of Lemma 2:*

Suppose there are two types on one side, and there are $K > 2$ types on the other side, denote the marginals as $n = (x_1, x_2, ..., x_K), m = (y, 1-y)$, where $\sum_k x_k = r$, where r is a constant. The surplus matrix is:

$$\Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ ... & ... \\ \Phi_{K1} & \Phi_{K2} \end{bmatrix}$$

denote the mass of singles in equilibrium as: $\mu_{k0}, \mu_{01}, \mu_{02}$ then:

(a)

$$\frac{\partial \mu_{01}}{\partial y} > 0, \frac{\partial \mu_{02}}{\partial y} < 0$$

(b) For any two types k_1, k_2, if we increase k_1 by decreasing k_2, then $\mu_{x_{k1}}$ increases and $\mu_{x_{k2}}$ decreases.

(c) For any two types k_1, k_2, if $\Phi_{k_1} + \Phi_{k_2 2} > \Phi_{k_2 1} + \Phi_{k_1 2}$, then there exist values $\underline{\delta_{x_1}}, \bar{\delta_{x_1}}, \underline{\delta_{x_2}}, \bar{\delta_{x_2}}, \underline{\delta_y}, \bar{\delta_y}$:

$$x_{k_1} \in (\underline{\delta_{x_1}}, \bar{\delta_{x_1}})$$

$$x_{k_2} \in (\underline{\delta_{x_2}}, \bar{\delta_{x_2}})$$

$$y \in (\underline{\delta_y}, \bar{\delta_y})$$

such that: $\frac{\mu_{01}}{\mu_{02}}$ decreases if we shift some mass from type k_2 to type k_1, i.e.:

$$\frac{\mu_{01}}{\mu_{02}}|_{(n=(...,x_{k_1}+\Delta, x_{k_2}-\Delta,...),m)} < \frac{\mu_{02}}{\mu_{01}}|_{(n=(...,x_{k_1},x_{k_2},...),m)}, \forall \Delta > 0$$

Proof. The proof is very similar to the proof of Lemma 2. WLOG, assume we shift the mass from type 2 to type 1 and denote $x_1 = x, x_2 = \gamma - x$, then $n = (x, \gamma - x, x_3, ..., x_k)$, and $m = (y, 1-y)$.

Denote $s_{i0} = \sqrt{\mu_{i0}}, s_{0j} = \sqrt{\mu_{0j}}$. First, write down the feasibility conditions:

$$s_{10}^2 + s_{10}s_{01}\phi_1 + s_{10}s_{02}\widetilde{\phi}_1 = x$$
$$s_{20}^2 + s_{20}s_{01}\phi_2 + s_{20}s_{02}\widetilde{\phi}_2 = \gamma - x$$
$$\vdots$$
$$s_{K0}^2 + s_{K0}s_{01}\phi_K + s_{K0}s_{02}\widetilde{\phi}_K = x_K$$
$$s_{01}^2 + s_{10}s_{01}\phi_1 + s_{20}s_{01}\phi_2 + ... + s_{K0}s_{01}\phi_K = y$$
$$s_{02}^2 + s_{10}s_{02}\widetilde{\phi_1} + s_{20}s_{02}\widetilde{\phi_2} + ... + s_{K0}s_{02}\widetilde{\phi}_K = 1 - y$$

To prove part (a), let's take the derivative with respect to y for all $K+2$ equations and denote $D_{i0} = \frac{\partial s_{i0}}{\partial y}, D_{0j} = \frac{\partial s_{0j}}{\partial y}$.

$$D_{10}(2s_{10} + \phi_1 s_{01} + \widetilde{\phi}_1 s_{02}) + s_{10}(\phi_1 D_{01} + \widetilde{\phi}_1 D_{02}) = 0 \tag{10}$$
$$D_{20}(2s_{20} + \phi_2 s_{01} + \widetilde{\phi}_2 s_{02}) + s_{20}(\phi_2 D_{01} + \widetilde{\phi}_2 D_{02}) = 0 \tag{11}$$
$$\vdots$$
$$D_{K0}(2s_{K0} + \phi_K s_{01} + \widetilde{\phi}_K s_{02}) + s_{K0}(\phi_K D_{01} + \widetilde{\phi}_K D_{02}) = 0 \tag{12}$$
$$D_{01}(2s_{01} + \phi_1 s_{10} + \phi_2 s_{20} + \cdots + \phi_K s_{K0}) + s_{01}(\phi_1 D_{10} + \phi_2 D_{20} + \cdots + \phi_K D_{K0}) = 1 \tag{13}$$
$$D_{02}(2s_{02} + \widetilde{\phi}_1 s_{10} + \widetilde{\phi}_2 s_{20} + \cdots + \widetilde{\phi}_K s_{K0}) + s_{02}(\widetilde{\phi}_1 D_{10} + \widetilde{\phi}_2 D_{20} + \cdots + \widetilde{\phi}_K D_{K0}) = -1 \tag{14}$$

We can rearrange Equation 10 - Equation 12 to express D_{k0} as a function of D_{01}, D_{02}

$$D_{k0} = -\frac{s_{k0}(\phi_k D_{01} + \widetilde{\phi}_k D_{02})}{2s_{k0} + \phi_k s_{01} + \widetilde{\phi}_k s_{02}}, \forall k = 1, 2, ..., K \tag{15}$$

We can substitute Equation 15 to Equation 13 and Equation 14:

$$D_{01}(2s_{01} + \sum_{k=1}^{K} \frac{\phi_k s_{k0}(2s_{k0} + \widetilde{\phi}_k s_{02})}{2s_{k0} + \phi_k s_{01} + \widetilde{\phi}_k s_{02}}) - D_{02} \sum_{k=1}^{K} \frac{s_{01}\phi_k s_{k0}\widetilde{\phi}_k}{2s_{k0} + \phi_k s_{01} + \widetilde{\phi}_k s_{02}} = 1 \tag{16}$$

$$D_{02}(2s_{02}+\sum_{k=1}^{K}\frac{\widetilde{\phi}_k s_{k0}(2s_{k0}+\phi_k s_{01})}{2s_{k0}+\phi_k s_{01}+\widetilde{\phi}_k s_{02}})-D_{01}\sum_{k=1}^{K}\frac{s_{02}\widetilde{\phi}_k s_{k0}\phi_k}{2s_{k0}+\phi_k s_{01}+\widetilde{\phi}_k s_{02}}=-1 \tag{17}$$

Add Equation 16 and Equation 17,

$$D_{01}(2s_{01}+\sum_{k=1}^{K}\frac{\phi_k 2s_{k0}^2}{2s_{k0}+\phi_k s_{01}+\widetilde{\phi}_k s_{02}}))+D_{02}(2s_{02}+\sum_{k=1}^{K}\frac{\widetilde{\phi}_k 2s_{k0}^2}{2s_{k0}+\phi_k s_{01}+\widetilde{\phi}_k s_{02}})=0 \tag{18}$$

Therefore D_{01} and D_{02} should have negative signs. Moreover, with Equation 16, we know:

$$D_{01}>0, D_{02}<0$$

Part (a) is proved.

Now let's prove part (b) Let me abuse the use of the notation D_{i0} and D_{0j}. For the proof of part (b), denote $D_{i0}=\frac{\partial s_{i0}}{\partial x}, D_{0j}=\frac{\partial s_{0j}}{\partial x}$. Let's take the derivative with respect to x for all $K+2$ feasibility equations:

$$D_{10}(2s_{10}+\phi_1 s_{01}+\widetilde{\phi}_1 s_{02})+s_{10}(\phi_1 D_{01}+\widetilde{\phi}_1 D_{02})=1 \tag{19}$$

$$D_{20}(2s_{20}+\phi_2 s_{01}+\widetilde{\phi}_2 s_{02})+s_{20}(\phi_2 D_{01}+\widetilde{\phi}_2 D_{02})=-1 \tag{20}$$

$$D_{30}(2s_{30}+\phi_3 s_{01}+\widetilde{\phi}_3 s_{02})+s_{30}(\phi_3 D_{01}+\widetilde{\phi}_3 D_{02})=0 \tag{21}$$

$$\vdots$$

$$D_{K0}(2s_{K0}+\phi_K s_{01}+\widetilde{\phi}_K s_{02})+s_{K0}(\phi_K D_{01}+\widetilde{\phi}_K D_{02})=0 \tag{22}$$

$$D_{01}(2s_{01}+\phi_1 s_{10}+\phi_2 s_{20}+\cdots+\phi_K s_{K0})+s_{01}(\phi_1 D_{10}+\phi_2 D_{20}+\cdots+\phi_K D_{K0})=0 \tag{23}$$

$$D_{02}(2s_{02}+\widetilde{\phi}_1 s_{10}+\widetilde{\phi}_2 s_{20}+\cdots+\widetilde{\phi}_K s_{K0})+s_{02}(\widetilde{\phi}_1 D_{10}+\widetilde{\phi}_2 D_{20}+\cdots+\widetilde{\phi}_K D_{K0})=0 \tag{24}$$

Rearrange Equation 21 - Equation 22 to express D_{k0} as a function of D_{01}, D_{02} for $k>2$:

$$D_{k0}=-\frac{s_{k0}(\phi_k D_{01}+\widetilde{\phi}_k D_{02})}{2s_{k0}+\phi_k s_{01}+\widetilde{\phi}_k s_{02}}, \forall k=3,...,K \tag{25}$$

Substitute Equation 25 to Equation 23 and Equation 24:

$$
\begin{aligned}
D_{01}(2s_{01}+\phi_1 s_{10}+\phi_2 s_{20}+\sum_{k=3}^{K}\frac{\phi_k s_{k0}(2s_{k0}+\widetilde{\phi}_k s_{02})}{2s_{k0}+\phi_k s_{01}+\widetilde{\phi}_k s_{02}}) \\
-D_{02}\sum_{k=3}^{K}\frac{s_{01}\phi_k s_{k0}\widetilde{\phi}_k}{2s_{k0}+\phi_k s_{01}+\widetilde{\phi}_k s_{02}} \\
+s_{01}(\phi_1 D_{10}+\phi_2 D_{20})=0
\end{aligned} \tag{26}
$$

$$
\begin{aligned}
D_{02}(2s_{02}+\widetilde{\phi}_1 s_{10}+\widetilde{\phi}_2 s_{20}+\sum_{k=3}^{K}\frac{\widetilde{\phi}_k s_{k0}(2s_{k0}+\phi_k s_{01})}{2s_{k0}+\phi_k s_{01}+\widetilde{\phi}_k s_{02}}) \\
-D_{01}\sum_{k=3}^{K}\frac{s_{02}\widetilde{\phi}_k s_{k0}\phi_k}{2s_{k0}+\phi_k s_{01}+\widetilde{\phi}_k s_{02}} \\
+s_{02}(\widetilde{\phi}_1 D_{10}+\widetilde{\phi}_2 D_{20})=0
\end{aligned} \tag{27}
$$

Then (Equation 19 + Equation 20)- (Equation 26 + Equation 27) gives us:

$$
D_{10}2s_{10}+D_{20}2s_{20}-D_{01}(2s_{01}+\sum_{k=3}^{K}\frac{2\phi_k s_{k0}^2}{2s_{k0}+\phi_k s_{01}+\widetilde{\phi}_k s_{02}})-D_{02}(2s_{02}+\sum_{k=3}^{K}\frac{2\widetilde{\phi}_k s_{k0}^2}{2s_{k0}+\phi_k s_{01}+\widetilde{\phi}_k s_{02}})=0 \tag{28}
$$

Moreover, from Equation 26 and Equation 27, we can express D_{01} and D_{02} as a linear combination of D_{10} and D_{20}. Denote We can also show that the coefficents are all negative. Combing Equation 28, D_{10} and D_{20} should have negative signs. Therefore $D_{10} > 0, D_{20} < 0$. Part (b) is proved.

Now let's prove part (c). Let's follow the notation of the proof for part (b): $D_{i0} = \frac{\partial s_{i0}}{\partial x}, D_{0j} = \frac{\partial s_{0j}}{\partial x}$.

Rearrange Equation 19 - Equation 22 to express D_{k0} as a function of D_{01}, D_{02}:

$$
D_{10}=\frac{1-s_{10}(\phi_1 D_{01}+\widetilde{\phi}_1 D_{02})}{2s_{10}+\phi_1 s_{01}+\widetilde{\phi}_1 s_{02}} \tag{29}
$$

$$
D_{20}=\frac{-1-s_{20}(\phi_2 D_{01}+\widetilde{\phi}_2 D_{02})}{2s_{20}+\phi_2 s_{01}+\widetilde{\phi}_2 s_{02}} \tag{30}
$$

$$
D_{k0}=-\frac{s_{k0}(\phi_k D_{01}+\widetilde{\phi}_k D_{02})}{2s_{k0}+\phi_k s_{01}+\widetilde{\phi}_k s_{02}}, \forall k=3,...,K \tag{31}
$$

Substitute Equation 29 - Equation 31 to Equation 23 and Equation 24:

$$\begin{aligned} D_{01}(2s_{01} + \sum_{k=1}^{K} \frac{\phi_k s_{k0}(2s_{k0} + \widetilde{\phi}_k s_{02})}{2s_{k0} + \phi_k s_{01} + \widetilde{\phi}_k s_{02}}) - D_{02} \sum_{k=1}^{K} \frac{s_{01}\phi_k s_{k0}\widetilde{\phi}_k}{2s_{k0} + \phi_k s_{01} + \widetilde{\phi}_k s_{02}} \\ = s_{01}(\frac{\phi_2}{2s_{20} + \phi_2 s_{01} + \widetilde{\phi}_2 s_{02}} - \frac{\phi_1}{2s_{10} + \phi_1 s_{01} + \widetilde{\phi}_1 s_{02}}) \end{aligned} \tag{32}$$

$$\begin{aligned} D_{02}(2s_{02} + \sum_{k=1}^{K} \frac{\widetilde{\phi}_k s_{k0}(2s_{k0} + \phi_k s_{01})}{2s_{k0} + \phi_k s_{01} + \widetilde{\phi}_k s_{02}}) - D_{01} \sum_{k=1}^{K} \frac{s_{02}\widetilde{\phi}_k s_{k0}\phi_k}{2s_{k0} + \phi_k s_{01} + \widetilde{\phi}_k s_{02}} \\ = s_{02}(\frac{\widetilde{\phi}_2}{2s_{20} + \phi_2 s_{01} + \widetilde{\phi}_2 s_{02}} - \frac{\widetilde{\phi}_1}{2s_{10} + \phi_1 s_{01} + \widetilde{\phi}_1 s_{02}}) \end{aligned} \tag{33}$$

Denote:

$$A = 2\frac{s_{01}}{s_{02}} + \sum_{k=1}^{K} \frac{\phi_k s_{k0} 2\frac{s_{k0}}{s_{02}}}{2s_{k0} + \phi_k s_{01} + \widetilde{\phi}_k s_{02}}$$

$$B = \sum_{k=1}^{K} \frac{\widetilde{\phi}_k s_{k0}\phi_k}{2s_{k0} + \phi_k s_{01} + \widetilde{\phi}_k s_{02}}$$

$$C = 2\frac{s_{02}}{s_{01}} + \sum_{k=1}^{K} \frac{\widetilde{\phi}_k s_{k0} 2\frac{s_{k0}}{s_{01}}}{2s_{k0} + \phi_k s_{01} + \widetilde{\phi}_k s_{02}}$$

$$F = s_{01}(\frac{\phi_2}{2s_{20} + \phi_2 s_{01} + \widetilde{\phi}_2 s_{02}} - \frac{\phi_1}{2s_{10} + \phi_1 s_{01} + \widetilde{\phi}_1 s_{02}}) \tag{34}$$

$$G = s_{02}(\frac{\widetilde{\phi}_2}{2s_{20} + \phi_2 s_{01} + \widetilde{\phi}_2 s_{02}} - \frac{\widetilde{\phi}_1}{2s_{10} + \phi_1 s_{01} + \widetilde{\phi}_1 s_{02}}) \tag{35}$$

we know $A > 0, B > 0, C > 0$, moreover:

$$D_{01}s_{02}(A + B) - D_{02}s_{01}B = F \tag{36}$$

$$D_{02}s_{01}(C + B) - D_{01}s_{02}B = G \tag{37}$$

Therefore:

$$D_{01}s_{02} - D_{02}s_{01} < 0 \iff C * F - A * G < 0$$

One sufficient condition for $CF - AG < 0$ is that $F < 0, G > 0$. One sufficient condition for

$F < 0, G > 0$ when $\frac{\widetilde{\phi}_2}{\widetilde{\phi}_1} > \frac{\phi_2}{\phi_1}$ is that:

$$\frac{\phi_2}{\phi_1} < \frac{s_{20}}{s_{10}} < \frac{\widetilde{\phi}_2}{\widetilde{\phi}_1}$$

since we can arrange Equation 34 and Equation 35:

$$F = s_{01}\left(\frac{1}{\frac{2s_{20}}{\phi_2} + s_{01} + \frac{\widetilde{\phi}_2}{\phi_2}s_{02}} - \frac{1}{2\frac{s_{10}}{\phi_1} + s_{01} + \frac{\widetilde{\phi}_1}{\phi_1}s_{02}}\right) \tag{38}$$

$$G = s_{02}\left(\frac{\widetilde{\phi}_2}{2s_{20} + \phi_2 s_{01} + \widetilde{\phi}_2 s_{02}} - \frac{\widetilde{\phi}_1}{2s_{10} + \phi_1 s_{01} + \widetilde{\phi}_1 s_{02}}\right) \tag{39}$$

Hence there exists $\underline{\delta_{x_1}}, \bar{\delta_{x_1}}, \underline{\delta_{x_2}}, \bar{\delta_{x_2}}$, when

$$x \in (\underline{\delta_{x_1}}, \bar{\delta_{x_1}})$$

$$(\gamma - x) \in (\underline{\delta_{x_2}}, \bar{\delta_{x_2}})$$

we have:

$$\frac{\partial \frac{s_{01}}{s_{02}}}{\partial x} < 0$$

$\square$

C. Proof for the Propositions

Proof for Proposition 1

Proof. To prove the existence of a stationary equilibrium, we need to show that there is a solution to the following equilibrium conditions given G_f, G_m, Φ, denote $G_f = (n_L, n_H), G_m = (m_L, m_H)$:

$$\mu_{0y}+\sqrt{\mu_{L_1 0}\mu_{0y}}\exp(\frac{\Phi_{L_1 y}}{2})+\sqrt{\mu_{L_2 0}\mu_{0y}}\exp(\frac{\Phi_{L_2 y}}{2})+\sqrt{\mu_{H_1 0}\mu_{0y}}\exp(\frac{\Phi_{H_1 y}}{2})+\sqrt{\mu_{H_2 0}\mu_{0y}}\exp(\frac{\Phi_{H_2 y}}{2}) = m_y, \forall y \in \{L, H\} \tag{40}$$

$$\mu_{e_1 0} + \sqrt{\mu_{e_1 0}\mu_{0L}}\exp(\frac{\Phi_{e_1 L}}{2}) + \sqrt{\mu_{e_1 0}\mu_{0H}}\exp(\frac{\Phi_{e_1 H}}{2}) = q_e^1 * n_e, \forall e \in \{L, H\} \tag{41}$$

$$\mu_{e_2 0} + \sqrt{\mu_{e_2 0}\mu_{0L}}\exp(\frac{\Phi_{e_2 L}}{2}) + \sqrt{\mu_{e_2 0}\mu_{0H}}\exp(\frac{\Phi_{e_2 H}}{2}) = q_e^2 * n_e, \forall e \in \{L, H\} \tag{42}$$

$$q_e^1 + q_e^2 = 1, \forall e \in \{L, H\} \tag{43}$$

$$\exp(-u_{e_1}) = \frac{\mu_{e_1 0}}{q_e^1 * n_e}, \forall e \in \{L, H\} \tag{44}$$

$$\exp(-u_{e_2}) = \frac{\mu_{e_2 0}}{q_e^2 * n_e}, \forall e \in \{L, H\} \tag{45}$$

$$u_{e_1} = u_{e_2}, \forall e \in \{L, H\} \tag{46}$$

Equation 40-Equation 42 characterize the equilibrium conditions of marriage market stability for given $\mathbf{q}$ strategy under the assumption of Gumbel distribution. Equation 44-Equation 45 characterize the expected marital utilities of females. Equation 43 comes from the property of stationarity. Equation 46 guarantees that women are indifferent between choosing to marry at period 1 or period 2.

Re-arrange Equation 41 and Equation 42 , we can get:

$$\frac{\mu_{e_1 0}}{q_e^1} + \sqrt{\mu_{0L}}\sqrt{\frac{\mu_{e_1 0}}{q_e^1}}\frac{1}{\sqrt{q_e^1}}\exp(\frac{\Phi_{e_1 L}}{2}) + \sqrt{\mu_{0H}}\sqrt{\frac{\mu_{e_1 0}}{q_e^1}}\frac{1}{\sqrt{q_e^1}}\exp(\frac{\Phi_{e_1 H}}{2}) = n_e$$

$$\frac{\mu_{e_2 0}}{q_e^2} + \sqrt{\mu_{0L}}\sqrt{\frac{\mu_{e_2 0}}{q_e^2}}\frac{1}{\sqrt{q_e^2}}\exp(\frac{\Phi_{e_2 L}}{2}) + \sqrt{\mu_{0H}}\sqrt{\frac{\mu_{e_2 0}}{q_e^2}}\frac{1}{\sqrt{q_e^2}}\exp(\frac{\Phi_{e_2 H}}{2}) = n_e$$

Combining with Equation 44-Equation 46, we can get:

$$\begin{aligned}\sqrt{\frac{q_e^1}{q_e^2}} &= \frac{\sqrt{\mu_{0L}}\exp(\frac{\Phi_{e_1L}}{2}) + \sqrt{\mu_{0H}}\exp(\frac{\Phi_{e_1H}}{2})}{\sqrt{\mu_{0L}}\exp(\frac{\Phi_{e_2L}}{2}) + \sqrt{\mu_{0H}}\exp(\frac{\Phi_{e_2H}}{2})} \\ &= \exp(\frac{\Phi_{e_1L} - \Phi_{e_2L}}{2})\frac{\sqrt{\mu_{0L}} + \sqrt{\mu_{0H}}\exp(\frac{\Phi_{e_1H}-\Phi_{e_1L}}{2})}{\sqrt{\mu_{0L}} + \sqrt{\mu_{0H}}\exp(\frac{\Phi_{e_2H}-\Phi_{e_2L}}{2})} \\ &= \exp(\frac{\Phi_{e_1L} - \Phi_{e_2L}}{2})\frac{1 + \sqrt{\frac{\mu_{0H}}{\mu_{0L}}}\exp(\frac{\Phi_{e_1H}-\Phi_{e_1L}}{2})}{1 + \sqrt{\frac{\mu_{0H}}{\mu_{0L}}}\exp(\frac{\Phi_{e_2H}-\Phi_{e_2L}}{2})}\end{aligned} \tag{47}$$

There are three cases:

1. $\Phi_{e_1H} - \Phi_{e_1L} = \Phi_{e_2H} - \Phi_{e_2L}$
2. $\Phi_{e_1H} - \Phi_{e_1L} > \Phi_{e_2H} - \Phi_{e_2L}$
3. $\Phi_{e_1H} - \Phi_{e_1L} < \Phi_{e_2H} - \Phi_{e_2L}$

Case one: In the first case, we have:

$$\sqrt{\frac{q_e^1}{q_e^2}} = \exp(\frac{\Phi_{e_1L} - \Phi_{e_2L}}{2}) \tag{48}$$

Hence equilibrium strategy $\mathbf{q}$ is pinned down by Equation 48 and Equation 43. Moreover, we know that given $\mathbf{q}$, Equation 40-Equation 42 has a unique equilibrium solution according to Decker et al. (2013). Hence stationary equilibrium exists in this case and is unique.

Case two: In the second case, $\frac{q_e^1}{q_e^2}$ is an increasing function of $\frac{\mu_{0H}}{\mu_{0L}}$ in Equation 47. Moreover, according to Lemma 3, we know that when $\Phi_{e_1H} - \Phi_{e_1L} > \Phi_{e_2H} - \Phi_{e_2L}$ indicating there is a complementarity between male High type and female marrying at period 1, an increase in $\frac{q_e^1}{q_e^2}$ would lead to a decrease in $\frac{\mu_{0H}}{\mu_{0L}}$ from Equation 40-Equation 43.

Moreover, from Equation 47, we know that:

$$\sqrt{\frac{q_e^1}{q_e^2}} \to \exp(\frac{\Phi_{e_1L} - \Phi_{e_2L}}{2}), \text{ as } \frac{\mu_{0H}}{\mu_{0L}} \to 0$$

$$\sqrt{\frac{q_e^1}{q_e^2}} \to \exp(\frac{\Phi_{e_1H} - \Phi_{e_2H}}{2}), \text{ as } \frac{\mu_{0H}}{\mu_{0L}} \to +\infty$$

While from Equation 40 - Equation 43, we know $\frac{\mu_{0H}}{\mu_{0L}}$ is bounded by finite positive number when $\exp(\frac{\Phi_{e_1L}-\Phi_{e_2L}}{2}) \leq \sqrt{\frac{q_e^1}{q_e^2}} \leq \exp(\frac{\Phi_{e_1H}-\Phi_{e_2H}}{2})$.

Hence equilibrium exists and is unique.

Case three: In the third case, $\frac{q_e^1}{q_e^2}$ is a decreasing function of $\frac{\mu_{0H}}{\mu_{0L}}$ in Equation 47. Moreover, according to Lemma 3, we know that when $\Phi_{e_1H} - \Phi_{e_1L} < \Phi_{e_2H} - \Phi_{e_2L}$ indicating there is a complementarity between male L type and female marrying at period 1, an increase in $\frac{q_e^1}{q_e^2}$ would lead to an increase in $\frac{\mu_{0H}}{\mu_{0L}}$ from Equation 40 - Equation 43. Applying the same logic as in case two, equilibrium exists and is unique.

Moreover, we know that equilibrium strategy satisfies:

$$\min(\Phi_{L_1L} - \Phi_{L_2L}, \Phi_{L_1H} - \Phi_{L_2}) \leq \ln(\frac{q_L^1}{q_L^2}) \leq \max(\Phi_{L_1L} - \Phi_{L_2L}, \Phi_{L_1H} - \Phi_{L_2H})$$

$$\min(\Phi_{H_1L} - \Phi_{H_2L}, \Phi_{H_1H} - \Phi_{H_2}) \leq \ln(\frac{q_H^1}{q_H^2}) \leq \max(\Phi_{H_1L} - \Phi_{H_2L}, \Phi_{H_1H} - \Phi_{H_2H})$$

□

Proof for Proposition 2

Proof. This is our first case in the previous proof of Proposition 1. Hence from Equation 47 equation (17), we know:

$$\frac{q_e^2}{q_e^1} = \exp(\Phi_{e_2L} - \Phi_{e_1L})$$

with $q_e^1 + q_e^2 = 1$, we have:

$$q_e^2 = \frac{\exp(\Phi_{e_2L})}{\exp(\Phi_{e_2L}) + \exp(\Phi_{e_1L})}, \quad q_e^1 = \frac{\exp(\Phi_{e_1L})}{\exp(\Phi_{e_2L}) + \exp(\Phi_{e_1L})}$$

□

Proof for Proposition 3 and Proposition 4

Proof. From the proof of proposition 1, we know that equilibrium strategy is pinned down by both Equation 47 and Equation 40-Equation 43. Hence how equilibrium strategies change depend on whether $\Phi_{e_1H} - \Phi_{e_2H} > \Phi_{e_1L} - \Phi_{e_2L}$ or $\Phi_{e_1H} - \Phi_{e_2H} < \Phi_{e_1L} - \Phi_{e_2L}$, and how $\frac{\mu_{0H}}{\mu_{0L}}$ changes in equilibrium.

Let's first prove Proposition 3, according to Lemma 3 result (a), an increase in m_H would increase μ_{0H} and decrease μ_{0L}, which increases $\frac{\mu_{0H}}{\mu_{0L}}$ given any strategy q_y, hence an increase in m_H would

- increase q_e^1, if $\Phi_{e_1H} - \Phi_{e_2H} > \Phi_{e_1L} - \Phi_{e_2L}$
- decrease q_e^1, if $\Phi_{e_1H} - \Phi_{e_2H} < \Phi_{e_1L} - \Phi_{e_2L}$

Then let's prove Proposition 4, according to Lemma 3(b), an increase in n_H would decrease $\frac{\mu_{0H}}{\mu_{0L}}$ given any strategy q_y if the following condition holds:

$$\exp(\frac{\Phi_{e_1H} - \Phi_{e_2H}}{2}) \leq \sqrt{\frac{\mu_{e_10}}{\mu_{e_20}}} \leq \exp(\frac{\Phi_{e_1L} - \Phi_{e_2L}}{2})$$

Moreover, we know that:

$$\frac{\mu_{e_10}}{\mu_{e_20}} = \frac{q_e^1}{q_e^2}$$

and

$$\exp(\frac{\Phi_{e_1H} - \Phi_{e_2H}}{2}) \leq \sqrt{\frac{q_e^1}{q_e^2}} \leq \exp(\frac{\Phi_{e_1L} - \Phi_{e_2L}}{2})$$

from Equation 47. Therefore the condition always holds in the neighborhood of the equilibrium. Hence an increase in n_H would

- decrease q_e^1, if $\Phi_{e_1H} - \Phi_{e_2H} > \Phi_{e_1L} - \Phi_{e_2L}$
- increase q_e^1, if $\Phi_{e_1H} - \Phi_{e_2H} < \Phi_{e_1L} - \Phi_{e_2L}$

$\square$