An Empirical Examination of Matching Theories: The One Child Policy, Partner Choice and Matching Intensity in Urban China

3 January 2012

Abstract

This paper introduces an index that facilitates the testing of differing matching theories based on the degree of overlap between a theoretically generated matching joint density and its empirical counterpart. The index is asymptotically Normal, consequently permitting inference. To demonstrate its use, the paper examines the effect the One Child Policy had on matching patterns in the marriage market in China. To distinguish between confounding policies of the period, a static general equilibrium model is introduced. It predicts that constraining marital output in the child quantity dimension may raise the marginal benefit of positive assortative matching and investment in child quality, thereby increasing the intensity with which they are pursued and concomitantly reducing the marriage rate. Upon verifying that the policy was binding via a Poisson model, using the matching index, significant support for increases in positive assortative matching and reductions in negative assortative matching were found.

JEL Code: C14, J12, J13

Key Words: Overlap Measure; Matching; Marriage

1 Introduction

Introduced in 1979, the One Child Policy (OCP) of China represented a considerable intervention in the household choice process, through its limiting the production of offspring by families, often through fines and various forms of coercion (Cooney and Li 1994; Li 1995). Its primary rationale was to avert the foreseen strain on the limited economic and agricultural resources as the economy sought economic growth (Greenhalgh 2003). It was implemented most stringently in Urban China and remains one of the most far reaching population control policies in recent history, and has been the subject of much study, ranging from human capital investment (Qian 2009; Rosenzweig and Zhang 2009), intergenerational transmission issues Anderson and Leo (2009), and other demographic issues (Ahn 1994; Ebenstein 2010; Zimmer and Kwong 2003).

Insofar as children are a vital component of marital output which depends upon parental match and input, the OCP provides an exogenous change in the environment which facilitates testing Becker's (1973) predictions regarding parental investments in children (Anderson and Leo 2009) and parental matching, provided all else is constant[1]. Unfortunately not all else is equal and the presence of other intervening policies or trends preclude consideration of the OCP as a pure natural pseudo experiment each of which have to be addressed in the analysis.

Prior to the implementation of the OCP, the government had been encouraging later births and lower birth rates among its populace, so that the policy was not at odds with the background against which it was introduced. Fertility (number of live births per married woman aged 20-44) was already in considerable decline prior to the OCP, having fallen to 2.2 in 1980 from 6.4 in 1965. This phenomenon could also be rationalized as a result of urbanization[2], which diminishes preference for larger families (Therborn 2004), consequently implying that the policy may not be binding for some of the urban populace, for which there is some evidence in terms of completed families with one or fewer children pre-policy.

[1] Li, Han, and Zhao (2012) in examining the empirical effects of a 1998 social policy that allowed Chinese urbanites of both genders to pass on their *hukou* status to their children, regardless of their spouse's status, found greater inter-*hukou* matches, reflecting unintended effects of social policy such as marital matching and male child preference.

[2] In 1949 7.3% of the population was urbanized, however by 1990 20.1% was urbanized (Anderson and Ge 2005)

The Chinese tradition of patrilocal residence of married sons means that considerable old age security benefits accrue to parents of sons as opposed to daughters. This engenders an apparent preference for sons (the usual boy/girl sex ratios at birth are 104/100 in 1995, China's was 117/100 (Peng and Guo 2000)), which within the premise of a binding OCP[3] has to be contended with (Therborn (2004) suggests that expression of these preferences has been facilitated by the development and availability of fetus gender detection and selective abortion techniques[4].). Such preferences however may well be more strongly held in rural as opposed to urban locales. The primary difficulty this presents is that it reduces the individual's ability to achieve their desired matches, and consequently prevent the detection of possible effects on matching due to the OCP.

Finally, the OCP was introduced in tandem with the Economic Reforms of 1979 which precipitated a well documented increase in the incomes of the populace. Should this increase have the propensity to bring about similar changes in family structure, it would not be possible to distinguish the effects of these two policies. Essentially this presents an identification problem which, as will become apparent, is to some degree resolved by the theoretical model developed for the study, since it predicts opposing effects for economic growth and the OCP, with respect to spousal choice decisions.

Notwithstanding the intervening circumstances, this social policy remains not only an excellent test bed for Becker's (1973) prediction[5], but an important example of the unintended effects that social policies could have beyond its good intentions. To our knowledge, the OCP's impact on marital partner choice has not previously been examined. By comparing the overlap measure against both perfect *positive assortative matching* (PAM) and *negative assortative matching* (NAM) over three urban cohorts within six provinces that straddled the implementation of the OCP, the paper provides an ordered matching index through which changes in matching patterns and intensity can be tracked.

Empirical work regarding Becker's (1973) predictions have thus far focused on the estimation of marriage models[6] and the testing of the PAM outcome generated by Becker's

[3]Within the rural context, Zhang (2002) found that the OCP did present a binding constraint.

[4]Doherty et al. (2001) found significant reduction in usage of pre-natal and obstetrics care among women with the implementation of the OCP not explained by price changes, suggesting a less intense expression of preference for male babies.

[5]Siow (2009) provides considerable insight into the theoretical implications summarized in Becker (1993).

[6]See Bergstrom and Lam (1994), Choo and Siow (2006a,b), Dagsvik (2000) and Wong (2003).

(1973) model[7]. The latter remains a developing field, which we augment. Siow (2009) comes closest in spirit to this paper, where he devised a measure related to the principal determinants of the joint density matrix for testing the PAM prediction and its identification as a consequence of supermodularity, as opposed to pure preference by the constituents of the marriage market for potential spouses that are similar in attribute. The approach developed is based on stochastic ordering in statistics (Shaked and Shanthikumar 2007) which relies on investigating the total positivity, or more precisely the determinant of a joint density matrix under consideration.

The measure introduced here provides a stochastic ordering to differing degrees of assortative matching. The primary innovation of the measure is that it is nonparametric and therefore capable of detecting spousal matches that are potentially nonlinear (unlike Pearson's correlation measure which is limited to linear relationships) . Unlike Siow (2009) which is structural since it is developed for testing Becker's (1973) predictions, our measure is of a reduced form nature, capable of discerning between any matching pattern a researcher may propose, since each matching scheme generates a differing joint density. The approach draws on Anderson, Ge, and Leo (2010) and Anderson, Linton, and Whang (2012). Relying on quantifying the proximity between an observed joint density to a hypothesized matching density, the asymptotically normal Overlap Measure unlike previous approaches does not require the joint density matrices to be square . As the empirical density tends towards mimicking a hypothesized density, the overlap measure tends toward one, and zero otherwise, thereby providing an index of how well a proposed matching model fits the observed data, which facilitates inference across different matching schemes.

To address the issue of other confounding interventions or trends, we provide a model that merges the matching and familial choice concerns, which helps to identify the effects of the policies. Specifically we examine how a policy constraining the quantity of children choice would affect the manner in which individual spousal choice is made, when prospects are encountered randomly under the joint concerns of spousal quality, and the quantity-quality of children. Although models of matching and familial human capital investment have been the subject of investigations for decades since Becker (1973), little theoretical work merging both concerns have been developed. Some examples include Laitner (1991)

[7] See Abowd et al. (1999), Fernandez et al. (2005), Galichon and Selanie (2009), Lise et al. (2008), Liu and Lu (2006), de Melo (2008), Mendes et al. (2007), Siow (2009) and Suen and Lui (1999)

which focused on child quality through wealth transfers, and how this affects matching decisions, while Kremer (1997) examined how matching affected intergenerational transmission which was further build upon in Fernandez and Rogerson (2001), Fernandez et al. (2005) and Greenwood et al. (2003). However, most of these were largely computational in nature.

In summary, this paper makes three contributions. Firstly, we introduce an innovative measure for quantifying matching patterns that is easy to use. This is then applied to a study of the manner in which the OCP may have affected marital matching patterns in China, which has not previously been examined. Finally, we provide a simple model merging matching and child quantity-quality concerns to discern between the effects of differing social and economic polices. In lieu of the intervening policies and demographic trends that may confound the analysis, the model predicts that a policy that limits family size would lead to an increase in PAM and decrease in NAM, and consequently a decrease in match rate particularly among low attribute men. As a preamble to the empirical application, we found significant support for the predictions.

In the following, section 2 introduces the Overlap Measure and it's applicability in testing matching hypotheses. Section 3 provides the historical context of the OCP and familial choice trends which motivates the theoretical model therein. Section 4 provides a data summary and establishes the sense in which the number of children in the family have been effectively rationed. Hypotheses about partner choice decisions are examined empirically in section 5, followed finally by a discussion and conclusion.

2 Method of Testing Matching Hypotheses

2.1 The Overlap Measure

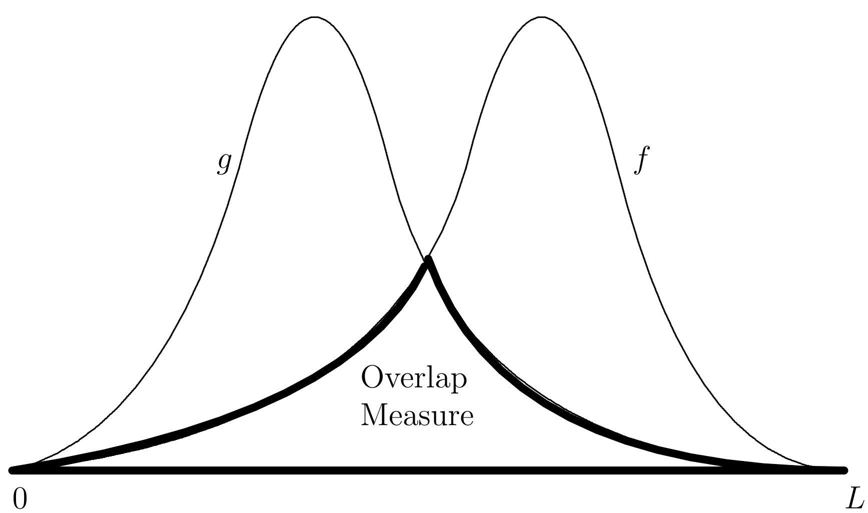
For simplicity, consider a marriage market where individuals are matched on a single attribute t_g^k where $g \in \{h, w\}$ (h denotes a male, and w denotes a female) and $k = \{1, 2, ..., n\}$, such that $t_g^1 < t_g^2 < ... < t_g^n$, in other words the attribute has n ordered realizations. Let $\mathbf{J}_e$ be the empirical/observed joint density matrix of matches in the marriage market with $j_{i,k}^e$ $i, k = \{1, 2, ..., n\}$ being a typical element. Similarly, the theoretically generated joint density matrix is $\mathbf{J}_t$ with $j_{i,k}^t$ being it's typical element. Then the Overlap Measure proposed is of the form,

$$\mathbf{OV}_t = \sum_{i=1}^{n} \sum_{k=1}^{n} \min \left\{ j_{i,k}^{t}, j_{i,k}^{e} \right\} \tag{1}$$

Although this measure has a continuous counterpart (Anderson et al. 2012), because attributes or indices considered in the matching literature are typically discrete in nature, the Overlap Measure as proposed by Anderson et al. (2010) is the natural candidate measure. At one extreme when the two densities are a perfect match $\mathbf{OV}_t = 1$, while at the other when the two densities have no common support then $\mathbf{OV}_t = 0$, consequently the support of $\mathbf{OV}_t$ is $[0, 1]$.

To provide a better grasp of the idea, consider two univariate densities from two differing populations. Intuitively the Overlap Measure quantifies the degree of overlap as depicted in figure 1 in two dimensions. As the degree of overlap of the densities tends to one, the greater is the explanatory power of the matching theory.

Figure 1: Overlap Between Densities f and g, $\mathbf{OV}$



An often overlooked detail is that there is no *a priori* reason why the support of the attributes on both sides of the market should be the same. In other words, the joint density of matches need not generate a square matrix. Whereas most other measures employed thus far require the joint density matrix to be square, the overlap measure is still viable when it is not. The measure is also amenable to use in multivariate domains, although to our knowledge, there has been no attempt in generating an equilibrium matching array.

Finally, the Overlap Measure is asymptotically Normally distributed, which consequently allows for inference (A short explanation is provided in appendix A.1.). For instance, in considering two matching models, the researcher could determine if one yields a significantly better fit. Further, as will be discussed in more detail below, with a suitably indexed hypothetical joint density matrix (corresponding to perfect positive/negative assortative matching for example), the researcher could determine if the marriage market is trending in a particular manner. In other words this presents a reduced form method of determining if there is increased or decreased intensity in a particular type of matching pattern.

2.2 Natural Ordering of the Matching Matrix

Matching intensity can be ordered in a reduced form fashion by relying on what can be achieved under perfectly positive/negative assortative matching. This method of ordering allows the researcher to examine how the intensity of matches within a marriage market has changed across the population and time.

For simplicity, suppose the attribute space of both husbands and wives are partitioned into five mutually exclusive realizations such that $t_g \in \{t_g^1, t_g^2, ..., t_g^5\}$ where $g \in \{h, w\}$ and $t_g^1 < t_g^2 < ... < t_g^5$. If the partitions are matched such that $\Pr(t_h = t_h^k) = \Pr(t_w = t_w^k)$ for all $k \in \{1, 2, ..., 5\}$, letting the row index denote the male type partitions and the columns denote the female type partitions, then the joint density under a null of perfect assortative matching[8] is of the form,

[8]Under the null hypothesis of negative assortative matching, $\mathbf{J}_n$ is a counter-diagonal matrix, with the highest type individuals matching with the lowest type from the other gender. In the perfectly matched marginal density case it follows that,

$$\mathbf{J}_n = \begin{bmatrix} 0 & \dots & 0 & \Pr(t_i = t_i^1) \\ 0 & \dots & \Pr(t_i = t_i^2) & 0 \\ \vdots & \ddots & \vdots & \vdots \\ \Pr(t_i = t_i^5) & \dots & 0 & 0 \end{bmatrix}$$

$$\mathbf{J}_p = \begin{bmatrix} \Pr(t_i = t_i^1) & 0 & \dots & 0 \\ 0 & \Pr(t_i = t_i^2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Pr(t_i = t_i^5) \end{bmatrix} \quad (2)$$

Under this scenario of a perfectly matched marriage market, this matrix would be the "theoretical" joint density matrix against which the empirical density matrix is compared. Then the closer the empirical matrix $\mathbf{J}_e$ is to $\mathbf{J}_p$, the more positively assortatively matched would the sample be.

Sometimes it is not possible to partition the support of types as above, for example when both sides of the marriage market are not matched in terms of the attribute, the perfectly positive assortative matching matrix would have it's elements spill onto the off-diagonal cells. Suppose that the partition is not matched such that $\sum_{k=1}^{m} \Pr(t_h = t_h^k) \leq \sum_{k=1}^{m} \Pr(t_w = t_w^k)$ for all $m \in \{1, 2, .., 5\}$, that is men stochastically dominate women in the attribute realizations. Then the joint density matrix under perfect positive assortative matching, assuming offers are made by men and that higher type men can always outbid lower type men for a potential match, would be of the form:

$$\mathbf{J}_p = \begin{bmatrix} \Pr(t_h = t_h^1) & 0 & \dots & 0 & 0 \\ \Pr(t_h \geq t_h^2) - \Pr(t_w \geq t_w^2) & \Pr(t_w \geq t_w^2) - \Pr(t_h \geq t_h^3) & \dots & 0 & 0 \\ 0 & \Pr(t_h \geq t_h^3) - \Pr(t_w \geq t_w^3) & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & \Pr(t_h = t_h^5) - \Pr(t_w = t_w^5) & \Pr(t_w = t_w^5) \end{bmatrix} \quad (3)$$

Estimates of such a matrix can be constructed from the empirical marginal distributions of men's and women's attribute realizations. Although only results using the above matrix are reported in the following application using a single spousal attribute/quality, there are other methods of arriving at the positive and negative assortative matching matrix which were examined as well in the application that follows, namely when the offers are made by women to men (For a detailed discussion of the difference this generates, see Roth and Sotomayor (1990), particularly theorem 2.13 due to Knuth (1976)) and when the preference for own type is strongest (that is matching clears the diagonal first)[9]. One of the

[9]These results are available from the authors upon request. They did not produce as close a fit to the empirical joint density as equation (3), suggesting that it is closer to the prevailing matching mechanism underlying equation (3).

advantages of the Overlap Measure is it's power, which allows it to discern between joint density matrices generated by differing matching theories, because the measure places equal weight on all cells of the matrix.

3 Model of Spousal Matching Intensity and Familial Choice

The substitutability between family size and the quality of their offspring has been demographically examined in Caldwell (1982) who argued that fertility was high when children are an asset to their parents, and low when they become a liability. Becker and Lewis (1973) and Becker (1993) formalized this idea in their model where both quantity and quality of children feature as part of the household decision process, and can be used to rationalize the effects of urbanization and the preference for sons at birth. An important feature of Becker's analysis is that *quantity* and *quality* choices are simultaneous, with each influencing the other to an extent[10]. He demonstrated that while quantity and quality are likely to be substitutes, they cannot be close substitutes, because of the convexity of the budget constraint between quantity and quality.

On the theoretical exposition of matching within the context of the family, Becker's (1973) theory of positive assortative matching showed that PAM will be the consequence when the marriage output function or the gains to marriage are complementary/supermodular in the qualities of the individuals entering into the marriage. The task then is to synthesize the separate concerns of marital matching and familial choices within a marriage to examine their behavior when they interact with each other.

The following is a static general equilibrium model of marital matching that merges the dual concerns, where choice of a spousal match is dependent on the individual's measurable continuously distributed attribute as well as the consequent choices in child quality and quantity. This permits the examination of how a binding constraint on child quantity

[10] Family formation has most frequently been discussed as an adjunct to the study of female labour supply. The issue being whether fertility should or should not be an argument in the labour supply equation, which in turn hinges on the nature of the planning horizon. One practice in modelling female labour supply is to assume that lifetime fertility decisions are made early in life, "at marriage is the most popular choice"(Browning 1992). The alternative is to assume a simultaneous model, where the agent attempts to have more children while making her labour supply decision.

(Neary and Roberts 1980; Deaton 1981) affects spousal choice endogenously. Intuitively, if individuals on both sides of the marriage market are forward looking, the policy will affect the choice of partner decision by rendering the owner of only childrearing attributes less of a comparative advantage relative to someone with greater income generating attributes all other things equal.

3.1 Model Setup

Consider a model where an individual lives for 2 periods, one as a child, and the other as an adult. At the beginning of the adult period, agents choose to marry or remain single (there is no divorce in this model). The focus is on gains from marriage and how it affects matching and child investment decisions, thus without loss of generality we solve the problem from the perspective of men, apportioning all the rents from marriage to them. The effects of sharing rules or bargaining are not addressed. Clearly bargaining power may be indirectly affected by both the OCP and the reforms, in essence it would further restrict the space of eligible spousal types further, but not change the substance of the results we obtain at the expense of increased complexity.

Adults meet someone of the opposite gender randomly. Marriage is dependent on the attribute/type realization of the man and woman, and utility is assumed to be transferable. Note that for the rest of the discussion, we will refer to the individual's type realization t_g as type or attribute interchangeably. Let the agent's type t_g, $g \in \{h, w\}$, be continuous on a support $\left[\underline{t}, \bar{t}\right]$, $\underline{t}, \bar{t} \in \mathbb{R}$, and distributed with density $f(.)$ and distribution $F(.)$ for both male and female. Upon marriage, they will make their quantity and quality choice in children. The aspect of utility derived from children is described by a function $q(.)$ dependent on the type of the parents, the number of children n, and the value of investment per child k, that is $q \equiv q\left(k, n, t_h, t_w\right)$, such that $q \mapsto \{0\} + \mathbb{R}_+$ is increasing and concave in all its inputs. In other words, q is the composite public good derived from both the quantity and quality of children, while parental type is the technology in the production. This function therefore resembles that adopted in De Tray (1973) with the exception of the dependence on parental attribute realizations. The attribute realization is intended to be a general variable which include ideas such as genetically transmitted talent, and time allocated towards child bonding.

The other aspect of a married individual's utility is derived from personal consumption c_g, $g \in \{h, w\}$. Let the utility function be multiplicatively separable in the utility derived from the children and that from own consumption,

$$u^h \equiv u(q, c_h) = q\,(t_h, t_w, k, n)\, c_h$$

Note that the concavity of q with respect to its inputs will ensure that u^h is likewise concave. In other words, $u^h_{t_g t_g} = q_{t_g t_g}(t_h, t_w, k, n) c_g \leq 0$, for $t_g \in \left[\underline{t}, \overline{t}\right]$, $g \in \{h, w\}$, and the same is true for the quantity and quality of children choices, and the agent's utility function is well behaved. If instead the individual chooses to remain single, utility will only be derived from personal consumption which in turn is dependent on his/her own type,

$$s^g \equiv s(c_g) \quad = \quad c_g$$

for $g \in \{h, w\}$.

The income realization of the family or individual is assumed to be dependent on the type of match and the individual's attribute respectively. Specifically, family income is assumed to be $yx(t_h, t_w)$, and income for a single individual is $yv(t_g)$, $g \in \{h, w\}$, where y is the average income within the economy, $x : (t_h, t_w) \mapsto \{0\} + \mathbb{R}_+$ and $v : t_g \mapsto \{0\} + \mathbb{R}_+$. Let both x_{t_g} and v_{t_g} be nonnegative, so that income realizations are increasing in attribute realization. In addition, let x be concave in both its inputs, $x_{t_g t_g} \leq 0$, while v is assumed to be convex, $v_{t_g t_g} \geq 0$.[11] This setup thus abstracts from redistributive concerns arising from any policy, which is beyond the scope of this paper. The convexity of v accentuates the costs in the form of transfer payments from a man to a prospective spouse of increasingly higher type in order to induce marriage, failing which we may observe low type men matching with high type women leading to NAM. This formulation of income together with the range of q also ensures that for some matches and individual attributes, the choice of remaining single will be made. That is the set of singles by attribute is non-empty.

3.2 The Effect of OCP on Child Quality

If an individual of type t_g, $g \in \{h, w\}$, chooses to remain single, he solves,

$$\max_{c_g} c_g$$

[11] This is stronger than is needed, but is assumed for simplicity of exposition. Our results hold as long as q is more concave than v with respect to t_g (which would result in a greater incidence of NAM).

subject to

$$yv(t_g) \geq c_g$$

The equilibrium utility of this single individual is,

$$\widehat{s}^g \equiv \widehat{s}(t_g) = c_g^* = yv(t_g) \tag{4}$$

so that his income is a proportion of the average income dependent on his attribute.

On the other hand, should the individual find a suitable match, he maximizes his utility subject to his budget and participation constraints:

$$V(t_h, t_w | u^w) = \max_{k,n,c_h,c_w} q(k, n, t_h, t_w)c_h$$

subject to the budget constraint,

$$c_h + c_w + nk \quad \leq \quad yx(t_h, t_w)$$

and the participation constraint,

$$u^w = q(k, n, t_h, t_w)c_w \quad \geq \quad yv(t_w) = \widehat{s}^w$$

where c_h, and c_w are the consumption choices, and t_h and t_w are the attribute realization for the husband and wife respectively. By the usual non-satiation argument, the budget constraint holds with equality, and since the husband can always make himself better off by just meeting the participation constraint, it holds with equality as well. Therefore,

$$c_h \quad = \quad yx(t_h, t_w) - nk - \frac{yv(t_w)}{q(k, n, t_h, t_w)}$$

and he solves,

$$V(t_h, t_w | u^w) = \max_{n,k} q(k, n, t_h, t_w)(yx(t_h, t_w) - nk) - yv(t_w) \tag{5}$$

The first order conditions are thus,

$$q_n(k, n, t_h, t_w)(yx(t_h, t_w) - n^*k^*) \quad = \quad q(k, n, t_h, t_w)k^* \tag{6}$$

$$q_k(k, n, t_h, t_w)(yx(t_h, t_w) - n^*k^*) \quad = \quad q(k, n, t_h, t_w)n^* \tag{7}$$

where k^* and n^* are the optimal value of investment per child, and number of children respectively. In equilibrium, the following condition will hold,

$$\frac{q_n(k, n, t_h, t_w)}{k^*} = \frac{q_k(k, n, t_h, t_w)}{n^*} \tag{8}$$

A sufficient condition for there to be an interior solution is for $q_n \geq q_k$ for $k^* \geq n^*$. In other words, the marginal utility from children should be greater than that from an additional unit of investment in child quality, failing which we would observe large families with little child investment which runs counter to evidence even as it relates to contemporary China.

However, under a situation created by the OCP where n is no longer a choice variable, only (7) would prevail, hence the effect of changes in n on the optimal choice of k can be examined as if n were a parameter. If the OCP were to not bind, the first order conditions stand. Whether OCP is a binding policy or otherwise is an empirical question which is addressed in the following sections. Let $\widetilde{n}$ denote the exogenously constrained number of children, and let the respective optimal choice of investment for each child be k' in that scenario.

The child quantity-quality trade-off (see Schultz (1997) for a discussion) is not a focus of this paper, and we simply assume they are substitutes. Formally,

Assumption 1 *: **Substitutability of Quantity and Quality:** Investment in children, k, and the choice of the number of children, n, are substitutes in the function $q(.,.)$. That is $q_{k,n}(k, n, t_h, t_w) \leq 0$.*

Given this substitutability, a binding policy that impinges on a family's choice in one dimension should yield an increase in the remaining dimension.

Proposition 1 *: Given assumption 1, an exogenously enforced reduction in the number of children raises equilibrium investment in children.*

As is apparent in the proof in appendix A.2, assumption 1 simplifies the proof and creates the tradeoff between the quantity and quality choice. However, notice that in fact the quantity and quality of children can be complementary to some extent without sacrificing the thrust of proposition 1. In other words, the result is more general than suggested by assumption 1.

The same analysis is performed to examine the consequences of the success of the Economic Reform of 1979, which raised the incomes and consequently the quality of lives among the Chinese populace, and should therefore raise familial investments in children, assuming children are *normal goods*. That the reform came at the same time as the OCP would accentuate the increase in investments (holding the nominal cost of investments constant), and consequently child quality.

Proposition 2 *: An exogenous increase in income would increase the number of children born into the family and/or the level of investment per child.*

Propositions 1 and 2 imply that the OCP and Economic Reform of 1979 would have reinforced each other, preventing identification of the cause of changes in investment in children if any[12].

3.3 The Effect of OCP on Assortative Matching Intensity

To analyze assortative matching, Becker (1973) and Siow (2009) are followed in assuming that the marital output function is complementary in spousal attributes.

Assumption 2 *: (**Complementarity of Types**) Assume that* $q_{t_g t_{g'}} \geq 0$ *and* $x_{t_g t_{g'}} \geq 0$*, so that* $u^h_{t_g t_{g'}} \geq 0$*,* $g \neq g'$*,* $g, g' \in \{h, w\}$*. Further, assume that both* $q(k, n, t_h, t_w)$ *and* $x(t_h, t_w)$ *be sufficiently concave in* t_h *and* t_w *so that,*

$$q_{t_g t_g}(k, n, t_h, t_w)(yx(t_h, t_w) - nk) + q(k, n, t_h, t_w)yx_{t_g t_g}(t_h, t_w) \leq -2q_{t_g} yx_{t_g}(t_h, t_w) \qquad (9)$$

such that $t^* = \arg \max_{t_w \in [\underline{t}, \bar{t}]} V(t_h, t_w | u^w) \Leftrightarrow t^* = t_h = t_w$*, for* $t_h, t_w \in \{\underline{t}, \bar{t}\}$.

It is important to note that positive assortative matching is generated by the complimentarity in spousal type in q and x. Assumption 2 implies that agents prefer to be matched with the same or higher attribute realization. The latter half of the assumption is purely to ensure symmetry in our analysis.[13] Just as in Becker (1973) and Siow (2009), this would be sufficient for generating PAM in the model. It should be noted that it is not our intent to explain why different matching patterns occur, rather we examine the degree to which various patterns occur. Thus like Choo and Siow (2006a,b) we find PAM more likely than NAM. Together with the convexity assumption of the utility function of singles, the above ensures that agents would always prefer to match with someone closer to their own type, since the concavity of u^g, $g \in \{h, w\}$, in own attribute and that of the spouse, and the convexity of v in own attribute ensures that the net marital output attains a maxima. Note that the convexity of v is not necessary for positive assortative matching to occur in this model.

[12] Anderson and Leo (2009) examined the effects on intergenerational mobility as a result of the OCP.

[13] An example of such a function is when $q(.)$ and $x(.)$ are quadratic with respect to $(t_h - t_w)$ on $t_h, t_w \in [0, 1]$. A model with search costs that diminish as agent type increases probably produce similar results to those that follow.

To see how the support of an individual's potential spouse is determined, note that since $\widehat{u}^h(t_w) = V(t_h, t_w|u_w) = \max_{n,k} q(k, n, t_h, t_w)(yx(t_h, t_w) - nk) - yv(t_w)$ and $\widehat{s}^h \equiv yv(t_h)$, then a type t_h man's second period utility is,

$$U_h = \max\{\widehat{u}^h(t_w), \widehat{s}^h\} \tag{10}$$

The minimum reservation type of his potential spouse is determined by

$$\begin{aligned} \widehat{u}^h(\underline{t_w^R}) &= \widehat{s}^h \\ \Rightarrow q(\underline{k}, \underline{n}, t_h, \underline{t_w^R})(yx(t_h, \underline{t_w^R}) - \underline{nk}) - yv(\underline{t_w^R}) &= yv(t_h) \end{aligned} \tag{11}$$

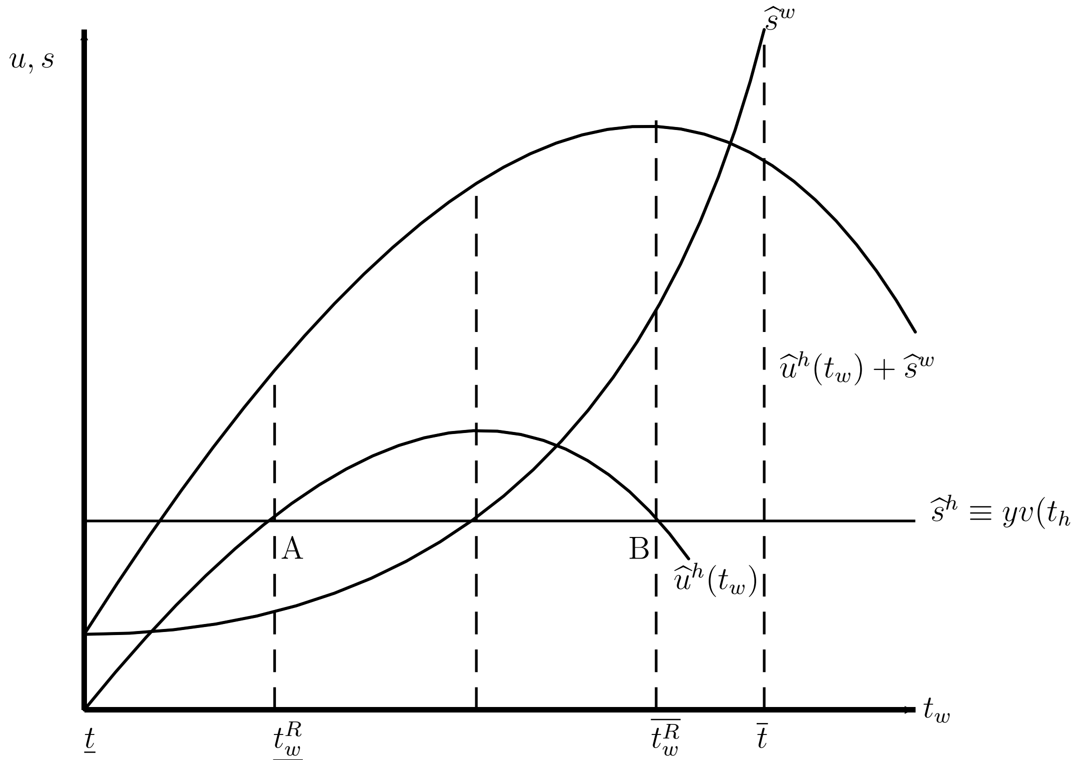
where $\underline{n}$ and $\underline{k}$ are the optimal values for a match between a man of type t_h, and a woman of type $\underline{t_w^R}$. Letting $\underline{t_w^R} \equiv \underline{t_w^R}(t_h)$, from figure 2 it may be observed that (11) determines only the reservation at point A. For spousal types below $\underline{t_w^R}$, although he may be collecting all the rents, he obtains no net benefit from marriage. It is only above $\underline{t_w^R}$ that marital utility would exceed his utility from remaining single.

Men of sufficiently low attribute realization may have an upper bound on the type of his spouse, $\overline{t_w^R}$, beyond which the marital gains from the match may not be sufficient for him to compensate her. She obtains at least $\widehat{s}(\overline{t_w^R}) \equiv \widehat{s}^w$, the utility she would get from remaining single. This upper threshold is therefore determined by

$$\begin{aligned} \widehat{u}^h(\overline{t_w^R}) &= \widehat{s}^h \\ \Rightarrow q(\overline{k}, \overline{n}, t_h, \overline{t_w^R})(yx(t_h, \overline{t_w^R}) - \overline{n}\overline{k}) - yv(\overline{t_w^R}) &= yv(t_h) \end{aligned} \tag{12}$$

The upper bound is point B in figure 2. The type of woman that would present as the optimal spousal type occurs when the marginal gain in gross marital utility from choosing a higher type spouse equates with the marginal increase in cost he would have to transfer to meet her participation constraint. This is where the slope of the gross utility and $\widehat{s}^w \equiv yv(t_w)$ equates, and coincides at the man's own attribute realization by assumption 2. Beyond this optimal realization, his net marital gains start decreasing, and eventually falls below his upper reservation value for marriage. It is important to note that even though by assumption 2, where the marriage output is optimized when his spouse has the same attribute, there will still be matches "off the diagonal" given random matching. The man will choose marriage to a woman as long as her type is within these bounds, since marriage yields a higher utility then his remaining single, and since this is a one period model, he will have no further opportunities of meeting the ideal spouse. The higher curve is meant to show that for sufficiently high type men, their upper bound coincides with that of the support. A similar argument can be made for sufficiently low type men.

Figure 2: Reservation Values given Type



Spousal choice remains a venue through which individuals could adjust to the enactment of the OCP to maintain the gains to marriage. Child outcomes are dependent on both ongoing investment as well as genetically endowed qualities from their parents. Thus the exogenous rationing of child quantity could have also accentuated the importance of good spousal match when PAM is the norm. Note again that the existence of PAM is not disputed, rather the *intensity* of PAM may have been altered. The analysis leads to the following proposition.

Proposition 3 *: When the number of children is fixed below the optimal choice that a married couple would have chosen given their attribute, then:*

1. *for all men, the lower bound on the reservation type of a prospective spouse would rise, while the upper bound would fall, and*

2. *agents who choose to marry would exhibit increased assortative matching.*

To illustrate proposition 3, let there be two broad groups of men, those who benefit from marriage, but who would never be able to attract high type spouses relative to their own type, $t_h = M$, and those who are coveted by all spousal types, $t_h = H$. Figure 3 shows that with a binding family size policy, matches with lower type women yield lower marital output in the post policy regime, consequently shifting the lower bound on the reservation type closer to one's own type. On the other hand, a match with a higher type spouse does not yield sufficient gains to marriage for the man to offer the minimum utility to attract her. This process is depicted as a fall in $\widehat{u}^h\left(t_w|t_h = H\right)$ for a man of type H. For a sufficiently low type man, this may even mean a complete withdrawal from the marriage market, depicted as a fall in $\widehat{u}^h\left(t_w|t_h = M\right)$ for a man of type M in figure 3. The latter observation is reflected in the following corollary.

Corollary 1 *: 1. A binding Family Size Policy which reduces the number of children born into a family reduces the marriage rate for all types of men. 2. Single men remaining in the market are predominantly of low attribute realization.*

The impact of economic growth is analyzed in the model largely to achieve some degree of identification of the matching effects of OCP. Unfortunately income growth could slacken or strengthen the need for a good spousal match, and is dependent upon the relative marginal gains under married and single states. Thus the economic reforms have two possible effects. If the gains to marriage are higher, an increase in the marriage rate would be observed together with a fall in PAM. However, if the gains are lower, a decrease in the marriage rate together with an increase in PAM would be observed. Formally,

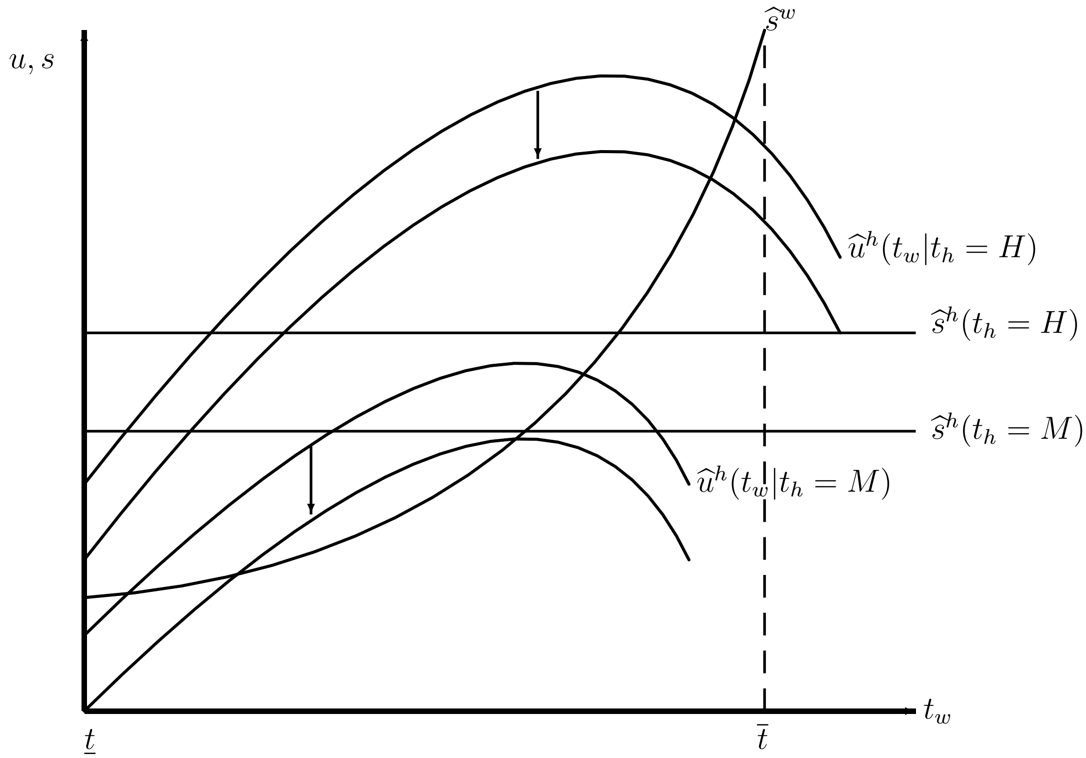
Proposition 4 *: For $\frac{\partial u^h}{\partial y}, \frac{\partial s^h}{\partial y} > 0$, $u^h(y = 0) \leq s^h(y = 0)$, if $\frac{\partial u^h}{\partial y} \geq \frac{\partial s^h}{\partial y}$, then for an increase in y, the average (real) income in the economy, leads to the following:*

1. *for all men, the lower bound on the reservation type of a prospective spouse would fall, while the upper bound would rise and,*

2. *agents who choose to marry would exhibit decreased assortative matching.*

Corollary 2 *An increase in average income increases marriage rates.*

Figure 3: Impact of Binding Family Size Policy on Spousal Type



Proposition 5 *: For $\frac{\partial u^h}{\partial y}, \frac{\partial s^h}{\partial y} > 0$, $u^h(y = 0) \geq s^h(y = 0)$, if $\frac{\partial u^h}{\partial y} \leq \frac{\partial s^h}{\partial y}$, then for an increase in y, the average (real) income in the economy, leads to the following:*

1. *for all men, the lower bound on the reservation type of a prospective spouse would rise, while the upper bound would fall and,*

2. *agents who choose to marry would exhibit increased assortative matching.*

Corollary 3 *1. An increase in average income reduces marriage rates. 2. Single men remaining in the market are predominantly of high attribute realization.*

It is worth noting that in the scenario of proposition 4, the remaining singles would be low attribute individuals whereas under the proposition 5's scenario they would be high attribute. Nonetheless, with some further caveats to be discussed later some degree of identification is afforded.

Proposition 4 says that at the status quo, on the margin of spousal type, for men indifferent to marriage and remaining single, an increase in income available to him cannot make his potential spouse any less attractive. However, if it makes her more attractive, then an expansion of potential spousal types would enhance his chances of finding a match, and consequently raises his welfare. Further, corollary 2 follows and the remaining singles would be low attribute individuals. The assumptions of proposition 5 implies that at the lower bound of attribute realization, the utility from marriage is higher than that from remaining single. Here, men on the margin would rather remain single with an increase in income, consequently giving rise to PAM and a fall in marriage rate, similar to the OCP. However, unlike the singles generated due to the OCP, the singles are individuals with high attribute, thus permitting some degree of identification. This is because under the scenario of proposition 5, the rate of increase to utility under the single state is greater than that under the marriage state.

Finally to show that the marriage market clears, since each individual meets only one potential spouse in their lifetime, and marriage takes place only if the potential spouse is within the reservation attribute bounds, the probability of marriage for a man of type t_h is P such that,

$$P = \Pr(\underline{t_w^R} \leq t_w \leq \overline{t_w^R}) = \int_{\underline{t_w^R}}^{\overline{t_w^R}} f(t_w)dt_w = F(\overline{t_w^R}) - F(\underline{t_w^R}) \tag{13}$$

It is clear that $P \in [0, 1]$. Let there be a unit mass of males and females. Then the marriage rate in the marriage market M is,

$$\begin{aligned} M &= \int_{\underline{t}}^{\bar{t}} \left\{ F(\overline{t_w^R}(t_h)) - F(\underline{t_w^R}(t_h)) \right\} f(t_h)dt_h \\ &< \left\{ F(\overline{t_w^R}(\bar{t})) - F(\underline{t_w^R}(\bar{t})) \right\} \int_{\underline{t}}^{\bar{t}} f(t_h)dt_h \\ &= \left\{ F(\overline{t_w^R}(\bar{t})) - F(\underline{t_w^R}(\bar{t})) \right\} < 1 \end{aligned}$$

from which it is clear that $M \in (0, 1)$ and the market clears.

3.4 Discussion of Other Possible Mechanisms

The model has examined two *direct* venues through which matching in the marriage market could have been affected, through constraining family size by the OCP and the increase in income by the Economic Reforms. Three additional possible venues through which both policies could have affected matching are discussed here.

One possible *indirect* effect on matching via the Economic Reforms is through changes in the returns to education. Essentially as the gains to human capital investment increase, the marginal density of attribute realizations for both sides of the marriage market will be altered. This would necessarily alter the *probability* of an individual meeting her potential spouse over the *entire range of potential spouses* in the marriage market, but *not the desired choice set itself* which is what this paper examines and attempts to measure. The same can be said about the OCP, since among parents who are cognizant of its effects on the gains to marriage and, given that a "good" marriage entered into by their children would raise their own utility, it is in their own interests to ensure that their children's potential gains to marriage do not suffer (See for example Peters and Siow (2002) for a model on premarital investments in children.).

A second potential confounding possibility is the traditional male child preference quite apart from the OCP, which creates a marriage squeeze for men. Grossbard-Shechtman (1993) provided convincing arguments that such a marriage squeeze should create less incentive for female labor force participation, or more generally for investment in human capital, so that the capacity for PAM along the educational attainment and/or income dimension should be lowered, if the incidence of male child preference worsens under the OCP. As will be revealed in the following empirical analysis, we find little evidence that the OCP has exacerbated this phenomenon although the practice remains prevalent. More importantly this bias against the female child and its effect on the sex ratio does not explain how individual spousal choice is made, rather its effect on the equilibrium in the marriage market is through indirectly affecting the demand and supply of spouses.

The agenda underlying the Economic Reforms were reinforced further in 1992−1993, where greater autonomy was granted to employers towards employment and remuneration determination (Naughton 1995; Zhang et al. 2008). This has engendered increasing evidence of a growing male−female wage differential (Gustafsson and Li 2000; Appleton et al. 2005; Zhang et al. 2008) through possibly discriminatory employment practices. These changes will possibly affect the incentives for human capital accumulation in both genders, and as before for the case of male−child preference, its effect is indirect. Further, since these changes occurred more than a decade after the OCP and Economic Reforms, the human capital investments in our sample would have been completed so that the results here will not suffer from confounding through this venue.

4 Description of Data

The observations for the empirical analysis are drawn from an annual urban household survey of six provinces in China from 1989 and 1991 to 2001; Jilin, Shandong, Hubei, Guangdong, Sichuan, and Shaanxi[14]. The matching attribute examined is educational attainment, the classification of which is based on the pre-1986 eight year compulsory educational system, since the youngest set of individuals in our sample, those born in 1969 would have completed their compulsory education prior to the implementation of the new educational laws[15]. Specifically, educational attainment is integer indexed from 1 to 5, with 5 representing college graduates and above, 4 for individuals with technical education, 3 for high school, 2 for middle school, and 1 for primary school and below.

The sample is divided into 3 cohorts, the first are couples with husbands born between 1940 and 1949, the second cohort have husbands born between 1950 and 1959, and the last from 1960 to 1969. This was done because the joint density matrix generated with the assumption that offers of marriage are made by men to women yielded a closer overlap with the observed joint density matrix. The first cohort is construed as the pre-OCP cohort, the last being the post-OCP cohort, with the 1950s cohort straddling the OCP. Finally, we assumed that marriage markets are closed within each province, so the analysis will proceed by province.

[14]This data was obtained from the National Bureau of Statistics as part of the project on Income Inequality during China's Transition organized by Dwayne Benjamin, Loren Brandt, John Giles and Sangui Wang.

[15]China implemented a nine year compulsory educational system, divided into primary (five to six years) and junior secondary (3 to 4 years). Upon completion, the children may then attend senior secondary lasting 3 years. China Education and Research Network.

Table 1: Summary of Parental Characteristics

		Panel A: Summary Statistics by Province & Cohort											
Cohort	Variable/Province	Jilin		Shandong		Hubei		Guangdong		Sichuan		Shannxi	
		Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.
40s Cohort	Number of Children	1.34	0.83	1.47	0.76	1.39	0.73	1.55	0.73	1.06	0.73	1.34	0.83
	Father's Education	3.21	1.38	3.13	1.33	3.19	1.34	3.06	1.41	3.12	1.38	3.25	1.31
	Mother's Education	2.72	1.27	2.57	1.25	2.54	1.23	2.52	1.26	2.53	1.24	2.67	1.15
	Proportion of Families w/o Children	0.16		0.09		0.1		0.03		0.21		0.15	
	Observations	1405		1312		1914		1822		2588		1336	
50s Cohort	Number of Children	1.15	0.45	1.14	0.38	1.09	0.38	1.14	0.41	1.01	0.31	1.16	0.47
	Father's Education	3.17	1.16	3.33	1.24	3.22	1.19	3.22	1.23	2.97	1.28	3.27	1.25
	Mother's Education	3.02	1.07	2.93	1.09	2.93	1.04	2.97	1.06	2.75	1.08	2.91	1.01
	Proportion of Families w/o Children	0.03		0.01		0.02		0.01		0.04		0.03	
	Observations	2687		2976		3282		2690		4433		1804	
60s Cohort	Number of Children	0.99	0.28	1.02	0.27	0.99	0.22	1.01	0.35	0.96	0.28	1.01	0.32
	Father's Education	3.47	1.15	3.86	1.13	3.84	1.12	3.67	1.07	3.71	1.17	3.65	1.14
	Mother's Education	3.33	1.09	3.39	1.09	3.4	1.11	3.37	1.06	3.42	1.13	3.23	1.03
	Proportion of Families w/o Children	0.04		0.03		0.03		0.05		0.06		0.04	
	Observations	1647		1860		1570		1077		2077		973	
		Panel B: Difference-in-Difference of Father's vs. Mother's Educational Attainment across Cohorts											
50s-40s	Statistics/P-Value	231.25	0	103.44	0	342.17	0	232.11	0	446.3	0	138.47	0
60s-40s	Statistics/P-Value	213.83	0	47.37	0	150.62	0	130.06	0	275.96	0	76.45	0
60s-50s	Statistics/P-Value	13.28	0	-77.1	1	-140.98	1	-33.99	1	-78.9	1	-34.43	1

1=Elementary School & Lower, 2=Middle School, 3=High School, 4=Technical Education, 5=College

Table 1 summarizes some of the characteristics of married couples within our sample. Notice the ubiquitous fall in the number of children over the decades, and particularly among the 1960s cohort. Yet, the number of married couples without any children are also lower post OCP than prior. Next, note the increase in educational attainment over the decades for both genders, which may be due to increased returns to education with the economic reforms, or the result of increased investments in children by parents, or the regime shift from post-Cultural Revolution China. On closer examination, observe the faster rate with which women raised their educational attainment relative to men since the 1940s cohort, although educational attainment among men remain higher. This is observed in both panels A and B of table 1, the latter being the difference-in-difference of men's versus women's mean educational attainment. Nonetheless, there is a slow down in these gains among women since the 1950s.

4.1 Is the OCP a Binding Constraint?

To establish the notion that the OCP constituted a binding constraint to the 1960s cohort, it is posited that any child born to a household after the first would be an accident rather than a matter of choice[16]. Consequently an appropriate statistical model for the number of children after the first would be an accidents model such as the Poisson model. If this model is rejected, it would be because children subsequent to the first were not accidents but a matter of systematic choice on the part of parents. The test is reported in panel A of table 2. A secondary concern is the preference for male children which may be revealed in an accidents model that conditions on the gender of the first child. Children subsequent to a male first child would still be accidents, but those subsequent to a female first child may not be. These tests are reported in panels B and C respectively.

From panel A, the Poisson "accidents" model is rejected for all provinces at the 1% level for the 1940s cohort, but the 1950s and 1960s cohorts yielded only three rejections of the model (Shandong and Shaanxi for the 1950s cohort and Shaanxi for the 1960s cohort). When the sample is differentiated by the gender of the first born, similar results prevail among male first born families. However, among female first born families, the model is rejected for all provinces for the 1940s cohort, and not rejected for all provinces for the 1950s and 1960s cohorts at the 1% level. Overall these results are evidence that

[16]The effect that OCP had on the quantity choice can also be gleaned from the marginals of the quantity choice over the 3 cohorts in the appendix, table A.1

post-OCP births after the first child are well described by a Poisson accidents model and confirms the efficacy of the OCP.

Parenthetically, although it was not addressed in the model, the data may shed light on the gender selection issue, which may in itself affect future equilibrium marriage rates. Table 3 presents Standard Normal Tests of the null hypothesis that *the proportion of first born children that are male is at most that of the natural rate*. The hypothesis is rejected for the 1940s cohort at the 5% level in 4 of the 6 provinces, but it is rejected twice for the 1950s cohort (for Shandong and Sichuan) and 1960s cohort (for Shandong and Guangdong), which is not sufficient evidence to suggest that the OCP had exacerbated the gender selection issue for our urban sample[17].

The analysis can be taken further by comparing *the number of children after the first child*, conditioning on the gender of that child. The suggestion here is that for at least the 1950s and 1960s cohorts, if the desire for male offspring was prevalent, but children subsequent to the first were "accidents", a first child being female would increase the chance of such an "accident" occurring. Table 4 presents the Standard Normal Tests for that comparison. At the 5% level of significance, households in two provinces (Guangdong and Shaanxi) among the 1940s cohort had significantly more children if their first born was female, while all six provinces among the 1950s cohort showed this propensity. The suggested upward trend stands in contrast to the 1960s cohort with only four provinces showing the same propensity.

Thus it may be concluded that the OCP or modernization appears to not have exacerbated the traditional male child preference in that the degree to which the $\frac{\text{male}}{\text{female}}$ first birth ratio is skewed has diminished. This is perhaps not surprising since there is some evidence that pre-natal and obstetrics care usage had fallen subsequent to the OCP (Doherty et al. 2001). As far as subsequent children are concerned, it seemed to initially increase the propensity for an "accident" among families whose first child was female among the 1950s cohort, but this had begun to creep closer to pre-OCP levels by the 1960s. Nonetheless, it must be emphasised that families with more than one child for later cohorts are only a small proportion of the total sample, revealing the efficacy of the OCP.

[17] A possible explanation is the observed reduction in obstetrics care usage Doherty et al. (2001) subsequent to the OCP.

Table 2: Pure Poisson Model χ^2 Goodness of Fit Tests and Upper Tail Probabilities by Cohort & Province

Province	Panel A: Sample with Children			Panel B: Sample with First Child Male			Panel C: Sample with First Child Female		
	1940s Cohort	1950s Cohort	1960s Cohort	1940s Cohort	1950s Cohort	1960s Cohort	1940s Cohort	1950s Cohort	1960s Cohort
Jilin	95.06	9.82	1.01	48.70	8.11	0.09	46.67	4.84	1.31
	[0.00]***	[0.02]**	[0.80]	[0.00]***	[0.04]***	[0.99]	[0.00]***	[0.18]	[0.73]
	1.61	1.19	1.04	1.59	1.15	1.01	1.62	1.23	1.06
	1174	2606	1573	610	1327	810	564	1279	763
Shandong	87.76	12.07	0.58	38.53	7.81	0.41	51.40	7.89	1.88
	[0.00]***	[0.01]***	[0.90]	[0.00]***	[0.05]**	[0.94]	[0.00]***	[0.05]**	[0.60]
	1.61	1.15	1.05	1.60	1.11	1.03	1.61	1.19	1.06
	1198	2947	1812	648	1549	966	550	1398	846
Hubei	167.38	1.86	0.32	93.63	0.19	0.15	76.44	1.33	0.17
	[0.00]***	[0.60]	[0.96]	[0.00]***	[0.98]	[0.98]	[0.00]***	[0.72]	[0.98]
	1.55	1.12	1.02	1.54	1.09	1.02	1.55	1.15	1.02
	1726	3210	1521	935	1674	760	791	1536	761
Guangdong	68.62	6.88	0.79	53.77	9.66	1.36	32.28	3.65	3.06
	[0.00]***	[0.08]*	[0.85]	[0.00]***	[0.02]**	[0.71]	[0.00]***	[0.30]	[0.38]
	1.60	1.15	1.07	1.54	1.13	1.04	1.68	1.17	1.11
	1763	2673	1018	972	1362	565	791	1311	453
Sichuan	19.03	3.60	0.49	4.07	2.02	0.07	18.41	2.33	0.53
	[0.00]***	[0.31]	[0.92]	[0.25]	[0.57]	[0.99]	[0.00]***	[0.51]	[0.91]
	1.34	1.06	1.02	1.34	1.04	1.01	1.34	1.07	1.03
	2040	4240	1960	1104	2219	1010	936	2021	950
Shaanxi	37.02	11.79	35.66	13.39	4.55	70.14	29.65	9.13	0.29
	[0.00]***	[0.01]***	[0.00]***	[0.00]***	[0.21]	[0.00]***	[0.00]***	[0.03]**	[0.96]
	1.57	1.20	1.05	1.52	1.16	1.05	1.63	1.25	1.05
	1142	1742	932	606	871	441	536	871	491

→The first two rows report the χ^2 test for the fit of the Pure Poisson regression. The first row reports the statistic, while the second row reports the $\Pr(X \geq x)$ in brackets. Degrees of freedom for all the tests is 3.

→The third and fourth rows report the mean number of children and the number of observations respectively.

→*** represents rejection of the Poisson model at the 1% level of significance, ** at 5% and * at 10%.

Table 3: Standard Normal Test Statistics (H_0: The proportion of first born children that are male is less than or equal to the natural rate of $\frac{104}{100}$)

Province	1940s Cohort	1950s Cohort	1960s Cohort
Jilin	0.67	-0.06	0.41
	[0.25]	[0.52]	[0.34]
Shandong	2.16	1.72	1.99
	[0.02]**	[0.04]**	[0.02]**
Hubei	2.66	1.33	-0.79
	[0.00]***	[0.09]*	[0.79]
Guangdong	3.51	-0.03	2.90
	[0.00]***	[0.51]	[0.00]***
Sichuan	2.84	1.77	0.49
	[0.00]***	[0.04]**	[0.31]
Shaanxi	1.41	-0.82	-2.24
	[0.08]*	[0.79]	[0.99]

→ $\Pr(Z \geq z)$ are in brackets

→*** represents rejection of H_0 at the 1% level of significance, ** at 5% and * at 10%.

Table 4: Standard Normal Test Statistics (H_0: No Difference Between Number of Children Given First Child Male and Female)

Province	Δ for 1940s Cohort	Δ for 1950s Cohort	Δ for 1960s Cohort
Jilin	-0.73	-5.27	-4.51
	[0.77]	[1.00]***	[1.00]***
Shandong	-0.30	-6.34	-2.92
	[0.62]	[1.00]***	[1.00]***
Hubei	-0.30	-5.52	-0.18
	[0.62]	[1.00]***	[0.57]
Guangdong	-4.12	-2.86	-4.56
	[1.00]***	[1.00]***	[1.00]***
Sichuan	0.06	-4.50	-3.14
	[0.47]	[1.00]***	[1.00]***
Shaanxi	-2.54	-4.55	-0.20
	[0.99]***	[1.00]***	[0.58]

→ $\Pr(Z \geq z)$ are in brackets

→*** represents rejection of H_0 at the 1% level of significance, ** at 5% and * at 10% for a one-sided test.

4.2 Capacity for Assortative Matching

The extent to which the OCP influenced partner choice decisions depends upon the degree to which positive or negative assortative pairing prevailed prior to the inception of the OCP, and how it changed thereafter. The comparative statics predict an increase in the incidence of PAM (decrease in NAM) with the onset of the OCP, in the sense that the range of values of a particular attribute an individual is willing to entertain in a partner has narrowed around his own attribute realization. It also predicts a drop in the marriage rate. However these predictions need qualification in terms of the supply and demand conditions within the marriage market, since they are always predicated on the availability of desirable partners. To put it another way, whether there is a natural potential for PAM depends on how well aligned the marginal densities of potential spouses are, as well as in terms of their numbers in the marriage market. The greater the alignment on the matching variable (which within the current analysis is educational attainment) or the greater the overlap of their density functions in terms of that variable, the greater is the capacity for PAM.

To accomodate the imbalance in numbers on both sides of the market when we include both married and single men and women, we introduce a modified version of the Overlap Measure of Anderson et al. (2010) and Anderson et al. (2012). Let $f(.)$ and $g(.)$ be the marginal densities of women and men respectively, where as before t is their continuous attribute realization on the support $t \in [\underline{t}, \overline{t}]$,

$$\mathbf{OV}_f = \frac{n_f}{n_g}\left\{\int_{\underline{t}}^{\overline{t}} \min\left(f(t), \frac{n_g}{n_f} g(t)\right) dt\right\} \tag{14}$$

$$\mathbf{OV}_g = \frac{n_g}{n_f}\left\{\int_{\underline{t}}^{\overline{t}} \min\left(\frac{n_f}{n_g} f(t), g(t)\right) dt\right\} \tag{15}$$

Table 5: Marginals of Married Individuals

		Jilin		Shandong		Hubei		Guangdong		Sichuan		Shaanxi	
		Males	Females	Males	Females	Males	Females	Males	Females	Males	Females	Males	Females
40s Cohort	1	0.085	0.140	0.086	0.195	0.073	0.195	0.139	0.222	0.117	0.215	0.061	0.115
	2	0.327	0.431	0.340	0.398	0.340	0.411	0.298	0.369	0.309	0.384	0.326	0.447
	3	0.170	0.135	0.170	0.145	0.169	0.144	0.177	0.178	0.146	0.135	0.170	0.166
	4	0.129	0.157	0.172	0.162	0.154	0.156	0.138	0.126	0.194	0.183	0.190	0.192
	5	0.290	0.137	0.232	0.101	0.265	0.095	0.248	0.105	0.233	0.083	0.253	0.080
Overlap $\mathbf{OV}_g$, $\mathbf{OV}_f$		0.813	0.811	0.834	0.831	0.807	0.804	0.846	0.844	0.828	0.826	0.825	0.822
# Observations		1408	1405	1316	1312	1920	1914	1826	1822	2593	2588	1340	1336
50s Cohort	1	0.013	0.024	0.028	0.045	0.028	0.037	0.059	0.066	0.076	0.082	0.032	0.047
	2	0.344	0.348	0.308	0.376	0.312	0.347	0.244	0.260	0.406	0.408	0.317	0.303
	3	0.335	0.357	0.241	0.297	0.307	0.383	0.355	0.431	0.196	0.286	0.281	0.452
	4	0.090	0.129	0.154	0.163	0.125	0.115	0.101	0.125	0.116	0.132	0.096	0.086
	5	0.218	0.142	0.269	0.119	0.228	0.118	0.241	0.117	0.205	0.093	0.274	0.112
Overlap $\mathbf{OV}_g$, $\mathbf{OV}_f$		0.939	0.916	0.852	0.849	0.891	0.877	0.890	0.875	0.892	0.885	0.817	0.808
# Observations		2753	2687	2986	2976	3336	3282	2736	2690	4466	4433	1824	1804
60s Cohort	1	0.010	0.010	0.002	0.006	0.006	0.011	0.006	0.014	0.015	0.012	0.010	0.017
	2	0.217	0.239	0.156	0.236	0.138	0.220	0.128	0.190	0.168	0.225	0.179	0.223
	3	0.372	0.382	0.278	0.338	0.323	0.371	0.480	0.430	0.331	0.352	0.367	0.433
	4	0.130	0.151	0.151	0.200	0.123	0.155	0.089	0.148	0.124	0.154	0.120	0.165
	5	0.271	0.219	0.412	0.219	0.410	0.243	0.298	0.218	0.362	0.257	0.324	0.161
Overlap $\mathbf{OV}_g$, $\mathbf{OV}_f$		1.000	0.843	0.854	0.791	0.927	0.787	0.945	0.705	0.950	0.840	0.964	0.807
# Observations		1953	1647	2008	1860	1849	1570	1443	1077	2348	2077	1162	973

1=Elementary School & Lower, 2=Middle School, 3=High School, 4=Technical Education, 5=College

On the other hand, for discrete attribute realization $t_i \in \mathbf{T} = [\underline{t}, \overline{t}]$, the measure would be,

$$\mathbf{OV}_f = \frac{n_f}{n_g}\left\{\sum_{t_i\in\mathbf{T}} \min\left(f(t_i), \frac{n_g}{n_f}g(t_i)\right)\right\} \tag{16}$$

$$\mathbf{OV}_g = \frac{n_g}{n_f}\left\{\sum_{t_i\in\mathbf{T}} \min\left(\frac{n_f}{n_g}f(t_i), g(t_i)\right)\right\} \tag{17}$$

The new Overlap Measure thus rescales the distributions to reflect the relative population sizes in order to compare the degree of matching, and undoing the scaling. In $\mathbf{OV}_f$, $g(t)$ is rescaled to conform to an $f(t)$ base, the term in braces reflecting the extent to which the population described by $f(t)$ can be matched. For $\mathbf{OV}_g$, the idea is reversed. Within an unbalanced market, each measure will produce differing numbers dependent on whichever population is relatively greater. With respect to the marriage market in China where men outnumber women, then $\mathbf{OV}_f < \mathbf{OV}_g$. The rescaling of $\mathbf{OV}_f$, and the consequent measure reflects the lower number of men that can and will be matched in light of their larger numbers. On the other hand, $\mathbf{OV}_g$ reflects the larger numbers of women who will be matched. The Overlap Measure results together with the educational attainment marginals by gender and province are reported in table 5.

Notice the marginals of male attainments stochastically dominate those of females. It is then to be expected that if marriage is indeed beneficial, well educated men in the earlier birth cohorts may adapt through lower incidences of PAM choices. As attainment rose among the general populace, the possibility of increase in PAM would have increased among men with higher attainment. Next, observe that among the 40's cohort, the overlap is very similar across gender and all provinces for both $\mathbf{OV}_f$ and $\mathbf{OV}_g$ reflecting the balance in the marriage market prior to the implementation of the OCP, and this remained largely similar through to the 50's cohort. However, by the 50's cohort there is a significant increase in potential for PAM as reflected by the increase in both the $\mathbf{OV}_f$ and $\mathbf{OV}_g$, with the exception of Shandong and Shaanxi. This potential continued to increase through to the 60's cohort, now including Shaanxi as reflected by the higher $\mathbf{OV}_g$. The lower $\mathbf{OV}_f$ implies however that with the imbalance in numbers on both sides of the market, that there should be more single men. This imbalance in terms of attainment proportions is true for all the provinces with the exception of Jilin and Guangdong, with the mismatch greatest between the middle and high school attainment and college groups. Indeed, the latter two provinces seem very well aligned in terms of their capacity for PAM. Further, reflecting on the theoretical exposition in the previous section, this implies therefore that, should the incidence of PAM increase together with a rise in single men of low educational attainment, it is very possible that it is due to the OCP. However, a fall in tendency towards PAM or an increase in PAM coupled with high proportions of single men with high attainment, would suggest that it is the economic reforms that had the primary effect on **spousal choice**.

5 Testing the Matching Hypotheses

5.1 Empirical Strategy

Besides the predictions of the previously discussed model in section 3, there are additional elements that will serve to identify the effects of the OCP. Firstly, the noted changes in the capacity for PAM can help identify the cause. For example if we observe an increase in PAM in the face of a decrease capacity, it would suggest that the OCP could be the underlying cause. Secondly, the differential impact of both the OCP and Economic Reforms across provinces allows the use of inter-provincial differences in intensity of PAM to identify the cause of the change. Finally, since the OCP and Economic Reforms affected about half of the samples, the differences across cohorts will likewise help in identification.

To elaborate, since *a priori* it is not known if PAM or NAM existed as the status quo, to conclude that PAM rose, it has to be ascertained that there was a significant *increase* in overlap between the empirical joint density and that generated by PAM, coupled with a significant *decrease* in overlap with the density generated by NAM across the cohorts. In other words, the hypotheses that will be tested for each province are for increased PAM via:

$$vs. \begin{array}{rl} H_0: & \Delta\mathbf{OV}_p = \mathbf{OV}_p^t - \mathbf{OV}_p^{t'} \leq 0 \\ H_1: & \Delta\mathbf{OV}_p = \mathbf{OV}_p^t - \mathbf{OV}_p^{t'} > 0 \end{array}$$

and decreased NAM via:

$$vs. \begin{array}{rl} H_0: & \Delta\mathbf{OV}_n = \mathbf{OV}_n^t - \mathbf{OV}_n^{t'} \geq 0 \\ H_1: & \Delta\mathbf{OV}_n = \mathbf{OV}_n^t - \mathbf{OV}_n^{t'} < 0 \end{array}$$

where $t < t'$ and $t, t' \in \{1940s, 1950s, 1960s\}$.

Further for each province, comparisons of the difference in the changes in assortative matching patterns and its intensity, allows us to examine the trends in matching. In the absence of any trends towards PAM (possibly as a result of preference for smaller family sizes due to urbanization), changes in the matching pattern could be due to either the OCP or Economic Reforms depending on the direction of change in assortative matching as discussed previously. However, should there be a "linear" trend towards PAM, the effect that is due to the OCP or the Economic Reforms can be gleaned from examining the difference in the measures from two comparisons, 1940s versus 1950s, and 1950s versus 1960s, which is similar to a *difference-in-difference analysis*. Letting the trend be identified by the change in overlap between the 1940s and 1950s, then a greater relative increase in PAM between the 1950s and 1960s could be identified as being a result of the OCP, and the Economic Reforms otherwise.

5.2 Results

The empirical joint densities of the data are reported in Table 6. Notice the higher diagonal probabilities of the joint densities, which provide some evidence of increased assortative pairing between the cohorts born in the 1940s and 1950s which is not surprising given the capacity for PAM has increased between the two cohorts (The comparison between these two cohorts is akin to examining the marital effects due to the Cultural Revolution which took place between 1966 and 1969.). Interestingly, this was also true among provinces where the mismatch in the marriage market for the 1960s cohort rose. Referencing the marginal densities in table 5 and the previous discussion reveals this as an interesting pattern in lieu of the mismatch on both the demand and supply sides of the marriage markets in all the provinces. Further, the increases in PAM within the 1960s cohorts are among individuals with higher attainment, with the off diagonal matches associated with NAM being much lower compared to previous cohorts. Not reported here are the marginal distributions of single men[18]. They were virtually non-existent for the first two cohorts (this is due to the universality of marriage in China (Yi et al. (1985))), but were substantial for the 1960's cohorts with proportions 0.72 for Jilin, 0.70 for Shandong, 0.55 for Hubei, 0.75 for Guangdong, 0.73 for Sichuan, and 0.72 for Shaanxi for less than technical education categories. As discussed in the theoretical section, this helps identify to some degree whether matching patterns were due to the OCP rather than the Economic reforms, and strongly favours the former.

The corresponding indices and tests for PAM and NAM using the overlap measure

[18]The results are available from the authors upon request.

are reported in Table 7. It must be noted that because the 1950s cohort consists of mainly individuals who made their spousal choice prior to the implementation of the OCP, while the 1960s cohort were those most likely affected, the identification of the impact of the OCP hinges on the increase in PAM within the 1960s cohort over the other two cohorts. From panel A in table 7 note that in all instances, the overlap measures are all statistically significantly different from complete overlap, where complete overlap occurs with the measure at or close to one. Further, the empirical joint density is a closer match to PAM, since the measure is closer to one under PAM than NAM.

The relative change in the overlap measure between two cohorts are reported in panel B. Examining the change in assortative matching between the 1940s and 1950s cohort, the PAM hypothesis remains a better description of the empirical joint densities. Considering the fact that the capacity for PAM rose between the two cohorts with the exception of Shandong and Shaanxi (since both $\mathbf{OV}_f$ and $\mathbf{OV}_g$ rose), the outcomes are not surprising and may be explained as the effects of increased educational attainment in the general populace, and a trend towards increased PAM. Likewise between the 1950s and 1960s cohorts, the PAM hypothesis remained the dominant description of the matching pattern, with the exception of Guangdong. Combined with the high proportion of singles of lower attainment, this is reasonably strong evidence that the OCP may be the cause of the tendency towards PAM.

Finally, the results examining the relative change in the difference between the overlap measures are reported in panel C of table 7. This comparison controls for trends towards increased PAM. The null hypothesis here is,

$$\begin{array}{ll} & H_0: \quad (\mathbf{OV}_p^{60s} - \mathbf{OV}_p^{50s}) - (\mathbf{OV}_p^{50s} - \mathbf{OV}_p^{40s}) \leq 0 \\ vs. & H_1: \quad (\mathbf{OV}_p^{60s} - \mathbf{OV}_p^{50s}) - (\mathbf{OV}_p^{50s} - \mathbf{OV}_p^{40s}) > 0 \end{array}$$

for change in PAM, and for NAM,

$$\begin{array}{ll} & H_0: \quad (\mathbf{OV}_n^{60s} - \mathbf{OV}_n^{50s}) - (\mathbf{OV}_n^{50s} - \mathbf{OV}_n^{40s}) \geq 0 \\ vs. & H_1: \quad (\mathbf{OV}_n^{60s} - \mathbf{OV}_n^{50s}) - (\mathbf{OV}_n^{50s} - \mathbf{OV}_n^{40s}) < 0 \end{array}$$

Since all the provinces had experienced an increase in capacity for PAM between the 1940s to the 1950s (with the exception of Shandong and Shaanxi), we can test whether the increase in PAM between the 1950s and 1960s cohorts is significantly greater than that between the 1940s and 1950s. This is the primary comparison since the increase in capacity between the 1940s and 1950s is the largest, and should *overstate* the trend towards PAM. On the other hand, due to the proximity in time between the 1950s and 1960s cohort, it should *understate* the effect of the OCP. Shandong, Hubei and Sichuan all experienced a significantly higher rate of increase in PAM between the 1950s and 1960s. All three provinces also recorded a slower decline in NAM, revealing the effects from the economic reforms. Nonetheless the punchline remains, that the tendency for PAM as a result of the OCP dominated that of the Economic Reforms, strongly suggesting that these three provinces were significantly affected by the OCP.

Although there was no significant increase towards PAM for Jilin over the three decades, what is suggestive here of the effect of the OCP is the significant fall in the tendency towards NAM, reinforcing the previous conclusion. On the other hand, Guangdong and Shaanxi experienced a significant slow down in PAM, and increases in NAM. One possible reason would be that in these two provinces, the effects of Economic Reforms dominated through economic growth and urbanization, which would have particularly affected the 1950s and 1960s cohorts. In other words, the fall in PAM there would be tempered by the income effect. Taken together with the strong alignment in marginals for Jilin and Guangdong, this is suggestive evidence that the effects of the Economic Reforms may have had spillover effects on the marriage market as well.

As a supplement to the above discussion, if the OCP indeed induced an increase in PAM, it would also reduce the likelihood of a high attainment individual choosing a lower attainment individual as partner. This means that there may be a stochastic dominant shift in the cumulative distribution of spouses across the cohorts[19], which was confirmed with the high attainment group data (lower attainment data analysis is hindered by shrinkage in support). With single men remaining in the market being predominantly from lower attainment groups, evidence suggests the OCP to be the dominant influence.

[19]We thank Aloysius Siow for suggesting this approach, and a complete discussion of which is available from the authors on request.

Table 6: Empirical Joint Density of Matching by Province, and Cohort

			Jilin					Shandong					Hubei				
			Females					Females					Females				
			1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1940s Cohort	Males	1	0.05	0.03	0.00	0.00	0.00	0.05	0.02	0.01	0.01	0.00	0.04	0.03	0.00	0.00	0.00
		2	0.03	0.23	0.03	0.03	0.01	0.07	0.16	0.04	0.05	0.01	0.07	0.18	0.03	0.05	0.01
		3	0.02	0.07	0.06	0.01	0.01	0.02	0.08	0.03	0.02	0.01	0.02	0.07	0.04	0.03	0.01
		4	0.02	0.05	0.02	0.03	0.01	0.03	0.08	0.03	0.03	0.01	0.03	0.06	0.02	0.02	0.02
		5	0.01	0.07	0.03	0.08	0.10	0.02	0.06	0.04	0.05	0.07	0.03	0.08	0.05	0.05	0.06
1950s Cohort	Males	1	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.01	0.02	0.01	0.00	0.00
		2	0.01	0.25	0.05	0.01	0.01	0.01	0.18	0.06	0.03	0.02	0.01	0.19	0.08	0.02	0.01
		3	0.00	0.05	0.22	0.04	0.02	0.01	0.09	0.11	0.03	0.01	0.01	0.07	0.18	0.02	0.02
		4	0.00	0.02	0.03	0.03	0.01	0.00	0.05	0.04	0.04	0.02	0.00	0.04	0.05	0.03	0.01
		5	0.00	0.02	0.06	0.05	0.10	0.01	0.05	0.08	0.06	0.07	0.00	0.04	0.07	0.04	0.08
1960s Cohort	Males	1	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		2	0.00	0.15	0.06	0.01	0.00	0.00	0.07	0.04	0.02	0.01	0.01	0.07	0.03	0.01	0.01
		3	0.00	0.06	0.22	0.03	0.03	0.00	0.08	0.13	0.04	0.02	0.01	0.09	0.17	0.03	0.03
		4	0.00	0.01	0.05	0.06	0.02	0.00	0.03	0.04	0.06	0.02	0.00	0.03	0.04	0.03	0.02
		5	0.00	0.01	0.06	0.06	0.16	0.00	0.04	0.12	0.09	0.18	0.00	0.04	0.12	0.08	0.18
			Guangdong					Sichuan					Shaanxi				
			Females					Females					Females				
			1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
1940s Cohort	Males	1	0.07	0.05	0.02	0.01	0.00	0.06	0.05	0.00	0.01	0.00	0.02	0.03	0.00	0.01	0.00
		2	0.08	0.14	0.03	0.04	0.01	0.08	0.15	0.03	0.04	0.01	0.05	0.19	0.04	0.03	0.01
		3	0.03	0.07	0.05	0.01	0.01	0.02	0.06	0.04	0.02	0.00	0.01	0.06	0.05	0.03	0.01
		4	0.03	0.05	0.03	0.03	0.01	0.03	0.07	0.03	0.05	0.01	0.02	0.07	0.03	0.07	0.00
		5	0.02	0.06	0.05	0.04	0.07	0.03	0.06	0.04	0.06	0.06	0.02	0.09	0.04	0.05	0.05
1950s Cohort	Males	1	0.02	0.03	0.01	0.00	0.00	0.03	0.04	0.01	0.00	0.00	0.01	0.01	0.01	0.00	0.00
		2	0.03	0.11	0.08	0.02	0.00	0.03	0.23	0.10	0.04	0.01	0.02	0.17	0.10	0.01	0.00
		3	0.02	0.08	0.21	0.03	0.02	0.01	0.05	0.10	0.02	0.02	0.01	0.05	0.20	0.01	0.02
		4	0.00	0.02	0.04	0.02	0.01	0.01	0.04	0.03	0.03	0.01	0.00	0.02	0.04	0.03	0.01
		5	0.00	0.03	0.09	0.05	0.08	0.00	0.05	0.05	0.05	0.06	0.00	0.04	0.10	0.04	0.09
1960s Cohort	Males	1	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		2	0.00	0.06	0.04	0.01	0.00	0.00	0.08	0.05	0.01	0.01	0.00	0.10	0.06	0.00	0.00
		3	0.01	0.10	0.27	0.04	0.04	0.00	0.09	0.17	0.03	0.02	0.01	0.08	0.21	0.04	0.02
		4	0.00	0.02	0.03	0.02	0.01	0.00	0.02	0.04	0.04	0.03	0.00	0.02	0.05	0.04	0.01
		5	0.00	0.02	0.09	0.08	0.16	0.00	0.02	0.10	0.07	0.19	0.00	0.03	0.11	0.08	0.13

→ 1=Elementary School & Lower, 2=Middle School, 3=High School, 4=Technical Education, 5=College

Table 7: Matching by Birth Cohort

	Province	Jilin		Shandong		Hubei		Guangdong		Sichuan		Shaanxi	
		Panel A: Overlap Measure by Cohort & Province											
		Positive	Negative	Positive	Negative	Positive	Negative	Positive	Negative	Positive	Negative	Positive	Negative
1940s Cohort	OV Measure	0.66	0.32	0.58	0.50	0.60	0.50	0.61	0.42	0.61	0.44	0.60	0.38
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
	No. of Obs.	1405		1312		1914		1822		2588		1336	
1950s Cohort	OV Measure	0.74	0.39	0.61	0.37	0.66	0.44	0.69	0.41	0.61	0.39	0.72	0.43
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
	No. of Obs.	2687		2976		3282		2690		4433		1804	
1960s Cohort	OV Measure	0.77	0.37	0.72	0.33	0.76	0.44	0.72	0.49	0.70	0.39	0.77	0.45
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)
	No. of Obs.	1647		1860		1570		1077		2077		973	
		Panel B: Difference in Overlap Measure Across Cohorts by Province											
Change in Assortative Mating	1950s-1940s	4.87	4.61	1.99	-8.15	4.17	-4.53	5.67	-0.46	0.05	-4.69	7.40	2.92
		[1.00]***	[1.00]	[0.98]**	[0.00]***	[1.00]***	[0.00]***	[1.00]***	[0.32]	[0.52]	[0.00]***	[1.00]***	[1.00]
	1960s-1950s	2.09	-0.98	7.86	-2.81	7.72	0.01	1.91	4.65	7.44	0.37	2.62	0.78
		[0.98]***	[0.16]	[1.00]***	[0.00]***	[1.00]***	[0.50]	[0.97]**	[1.00]	[1.00]***	[0.65]	[1.00]***	[0.78]
	1960s-1940s	6.25	3.31	8.18	-9.84	10.49	-3.82	6.36	4.00	6.70	-3.61	8.98	3.23
		[1.00]***	[1.00]	[1.00]***	[0.00]***	[1.00]***	[0.00]***	[1.00]***	[1.00]	[1.00]***	[0.00]***	[1.00]***	[1.00]
		Panel C: Difference-in-Difference by Province											
Difference-in-Difference	(60s-50s)-(50s-40s)	-2.26	-3.99	3.54	4.36	2.39	3.11	-2.34	3.87	5.30	3.48	-3.36	-1.37
		[0.01]	[0.00]***	[1.00]***	[1.00]	[0.99]***	[1.00]	[0.01]	[1.00]	[1.00]***	[1.00]	[0.00]	[0.09]*

→ Standard errors are in parenthesis and $\Pr(Z \leq z)$ are in brackets.

→*** represents rejection of H_0 at the 1% level of significance, ** at 5% and * at 10% for a one-sided test.

6 Conclusion

Although there has been much significant work on estimating Becker's (1973) theory of matching, work related to testing the predictions is still preliminary in nature. Generally little has been done in developing indices with known statistical properties that can be used to test matching theories. This paper fills that important void by presenting a matching index, the Overlap Measure developed by Anderson et al. (2010) and Anderson et al. (2012), that is amenable to the examination of differing matching theories with respect to observed matching patterns, and facilitates inference due to its asymptotically Normal properties.

In demonstrating the measure's ease of use, the paper examined the possibility that the One Child Policy (OCP) instituted in China in 1979 affected matching patterns, a question that has not yet received consideration. Due to potential confounding effects arising from the concurrent Economic Reform policy of 1979, this paper developed a static general equilibrium model that examines how the two policies could have competing effects on matching in the marriage market.

It is well understood that marital output has several dimensions and that when one dimension is exogenously constrained below the private optimal choice, agents adjust in other dimensions. What is perhaps less well understood is that the imposition of such a constraint may change the way agents choose their spouses. Here the consequences for partner choice due to the imposition of the OCP have been explored within the context of the urban populace of six provinces. As a guide to the analysis, the model of family formation developed predicts an increase in the marginal benefits and consequently incidence of PAM, but a reduction in the number of matches and the increase in investment in child quality. Importantly for identification reasons, the model also predicts a reduction in the intensity of PAM with economic growth, and/or an increase in single men predominantly of high type.

The matching predictions were empirically examined via annual samples of urban households in six Chinese provinces. The index developed for measuring the intensity of PAM was based upon the degree of overlap between the hypothetical perfectly positive assortative and empirical joint density of matches, on the single dimension of educational attainment. By pooling the samples into three cohorts, those who made family structure decisions prior to the OCP, those whose decisions spanned the introduction of the OCP and those whose decisions were made after the OCP, it was possible to evaluate how matching patterns changed over the introduction of the OCP.

After establishing, via a Poisson "accidents" model, that the OCP did present a binding constraint to families who desired more than one child[20], the intensity of PAM and NAM was examined. The index indicated significant increases (decreases) in the intensity of positive (negative) assortative matching in three of the six provinces, and this was accompanied by a significant increase in the proportion of single men with low educational attainment, all of which accorded with the predictions of the model. Thus the evidence here suggests that the OCP may have precipitated an increase (decrease) in positive (negative) assortative matching.

References

Abowd, J. M., Kramarz, F., and Margolis, D. N. (1999). High Wage Works and High Wage Firms. *Econometrica*, 67(2), 251–334.

Ahn, N. (1994). Effects of the One-Child Family Policy on Second and Third Births in Hebei, Shaanxi and Shanghai. *Journal of Population Economics*, 7(1), 63–78.

Anderson, G. and Ge, Y. (2005). The Size Distribution of Chinese Cities. *Regional Science and Urban Economics*, 35(6), 756–776.

Anderson, G., Ge, Y., and Leo, T. W. (2010). Distributional Overlap: Simple, Multivariate, Parametric and Non-Parametric Tests for Alienation, Convergence and General Distributional Difference Issues. *Econometric Reviews*, 29(3), 247–275.

Anderson, G. and Leo, T. W. (2009). Child Poverty, Investment in Children and Generational Mobility: The Short and Long Term Wellbeing of Children in Urban China after the One Child Policy. *Review of Income and Wealth*, 55(s1), 607–629.

Anderson, G., Linton, O., and Whang, Y.-J. (2012). Nonparametric Estimation and Inference about the Overlap of Two Distributions. *Journal of Econometrics*, 171(1), 1–23.

[20]A bi-product of this analysis was evidence that the OCP suppressed the extent to which gender selection of children occurred, though there was evidence that the probability of having an "accident" after a female first born was greater than the probability of an "accident" after a male first born.

Appleton, S., Song, L., and Xia, Q. (2005). Has China Crossed the River? The Evolution of Wage Structure in Urban China during Reform and Retrenchment. *Journal of Comparative Economics*, 33, 644–663.

Becker, G. S. (1973). A Theory of Marriage: Part I. *Journal of Political Economy*, 81(4), 813–846.

Becker, G. S. (1993). *Treatise on the Family*. Cambridge, Massachusetts: Harvard University Press.

Becker, G. S. and Lewis, H. G. (1973). On the Interaction Between the Quantity and Quality of Children. *Journal of Political Economy*, 81(2), S279–S288.

Bergstrom, T. and Lam, D. (1994). The Effects of Cohort Size on Marriage Markets in Twentieth Century Sweden. In J. Ermisch and N. Ogawa (Eds.), *The Family, the Market, and the State in Ageing Societies*. Oxford: Oxford University Press, (46–63).

Browning, M. (1992). Children and Household Economic Behavior. *Journal of Economic Literature*, 30, 1434–1475.

Caldwell, J. (1982). *Theory of Fertility Decline*. London Academic Press.

Choo, E. and Siow, A. (2006a). Estimating a Marriage Matching Model with Spillover Effects. *Demography*, 43(3), 463–488.

Choo, E. and Siow, A. (2006b). Who Marries Whom and Why. *Journal of Political Economy*, 114(1), 175–201.

Cooney, R. S. and Li, J. (1994). Household Registration Type and Compliance with the "One Child" Policy in China, 1979-1988. *Demography*, 31(1), 21–32.

Dagsvik, J. (2000). Aggregation in Matching Markets. *International Economic Review*, 41(1), 27–57.

de Melo, R. L. (2008). Sorting in the Labor Market: Theory and Measurement. Yale University Manuscript.

De Tray, D. N. (1973). Child Quality and the Demand for Children. *Journal of Political Economy*, 81(2), S70–S95.

Deaton, A. (1981). Theoretical and Empirical Approaches to Consumer Demand Under Rationing. In A. Deaton (Ed.), *Essays in the Theory and Measurement of Consumer Behaviour: In Honour of Sir Richard Stone*. Cambridge: Cambridge University Press, (55–72).

Doherty, J. P., Norton, E. C., and Veney, J. E. (2001). China's One-Child Policy: The Economic Choices and Consequences Faced by Pregnant Women. *Social Science and Medicine*, 52, 745–761.

Ebenstein, A. (2010). The "Missing Girls" of China and the Unintended Consequences of the One Child Policy. *Journal of Human Resources*, 45(1), 87–115.

Fernandez, R., Guner, N., and Knowles, J. (2005). Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality. *The Quarterly Journal of Economics*, 120(1), 273–344.

Fernandez, R. and Rogerson, R. (2001). Sorting and Long Run Inequality. *The Quarterly Journal of Economics*, 116(4), 1305–1341.

Galichon, A. and Selanie, B. (2009). Matching with Trade-offs: Revealed Preferences over Competing Characteristics. Ecole Polytechnique Manuscript.

Greenhalgh, S. (2003). Science, Modernity, and the Making of China's One-Child Policy. *Population and Development Review*, 29(2), 163196.

Greenwood, J., Guner, N., and Knowles, J. A. (2003). More on Marriage, Fertility, and the Distribution of Income. *International Economic Review*, 44(3), 827–862.

Grossbard-Shechtman, S. (1993). *On the Economics of Marriage: A Theory of Marriage, Labor and Divorce*. Boulder, Colorado: Westview Press.

Gustafsson, B. and Li, S. (2000). Economic Transformation and the Gender Earnings Gap in Urban China. *Journal of Population Economics*, 13, 305–329.

Knuth, D. (1976). *Marriages Stables*. Montreal: Les Presses de l'Universite de Montreal.

Kremer, M. (1997). How Much Does Sorting Increase Inequality. *The Quarterly Journal of Economics*, 112(1), 115–139.

Laitner, J. (1991). Modelling Marital Connections among Family Lines. *The Journal of Political Economy*, 99(6), 1123–1141.

Li, J. (1995). China's One-Child Policy: How and How Well Has it Worked? A Case Study of Hebei Province, 1979-88. *Population and Development Review*, 21(3), 563–585.

Li, T., Han, L., and Zhao, Y. (2012). How Status Inheritance Affects Marriage Segregation: Evidence from China. Hong Kong University of Science and Technology Working Paper.

Lise, J., Meghir, C., and Robin, J.-M. (2008). Matching, Sorting and Wages. University College of London Manuscript.

Liu, H. and Lu, J. (2006). Measuring the Degree of Assortative Mating. *Economic Letters*, 92(3), 317–322.

Mendes, R., Berg, G. J. V. D., and Lindeboom, M. (2007). An Empirical Assessment of Assortative Matching in the Labor Market. IZA Discussion Paper No. 3053.

Naughton, B. (1995). *Growing Out of the Plan: Chinese Economic Reform*. New York: Cambridge University Press.

Neary, P. and Roberts, K. (1980). The Theory of Household Behaviour Under Rationing. *European Economic Review*, 13, 25–42.

Peng, X. and Guo, Z. (2000). *The Changing Population of China*. John Wiley and Sons.

Peters, M. and Siow, A. (2002). Competing Pre-Marital Investments. *Journal of Political Economy*, 110, 592–608.

Qian, N. (2009). Quantity-Quality and the One Child Policy:The Only-Child Disadvantage in School Enrollment in Rural China. NBER Working Paper No. 14973.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*. John Wiley and Sons.

Rosenzweig, M. R. and Zhang, J. (2009). Do Population Control Policies Induce More Human Capital Investment? Twins, Birth Weight and China's "One-Child" Policy. *Review of Economic Studies*, 76(3), 1149–1174.

Roth, A. E. and Sotomayor, M. A. O. (1990). *Two-Sided Matching: A Study in Game-Theoretic Modelling and Analysis*. Econometric Society Monographs No. 18. Cambridge: Cambridge University Press.

Schultz, T. P. (1997). Demand for Children in Low Income Countries. In M. R. Rosenzweig and O. Stark (Eds.), *Handbook of Population and Family Economics*, vol. 1A. Amsterdam: Elsevier, (349–430).

Shaked, M. and Shanthikumar, G. J. (2007). *Stochastic Orders*. Springer.

Siow, A. (2009). Testing Becker's Theory of Positive Assortative Matching. University of Toronto Working Paper 356.

Suen, W. and Lui, H.-K. (1999). A Direct Test of the Efficient Marriage Market Hypothesis. *Economic Inquiry*, 37(1), 29–46.

Therborn, G. (2004). *Between Sex and Power: Family in the World 1900-2000*. Routledge.

Wong, L. (2003). Structural Estimation of Marriage Models. *Journal of Labor Economics*, 27(3), 699–728.

Yi, Z., Vaupel, J. W., and Yashin, A. I. (1985). Marriage and Fertility in China: A Graphical Analysis. *Population and Development Review*, 11(4), 721–736.

Zhang, J., Han, J., Liu, P.-W., and Zhao, Y. (2008). Trends in the Gender Earning Differential in Urban China, 1988-2004. *Industrial and Labor Relations Review*, 61(2), 224–243.

Zhang, W. (2002). *Economic Reforms and Fertility Behaviour: A Study of a North China Village*. Routledge Curzon.

Zimmer, Z. and Kwong, J. (2003). Family Size and Support of Older Adults in Urban and Rural China: Current Effects and Future Implications. *Demography*, 40(1), 23–44.

A Appendix

A.1 A Brief Discussion about the Overlap Measure

To see that the Overlap Index is asymptotically normally distributed, define

$$\mathbf{V} = \sqrt{n}\begin{bmatrix} \frac{j_{1,1}-\pi_{1,1}}{\sqrt{\pi_{1,1}}} & \frac{j_{1,2}-\pi_{1,2}}{\sqrt{\pi_{1,2}}} & \cdots & \frac{j_{1,N}-\pi_{1,N}}{\sqrt{\pi_{1,N}}} \\ \frac{j_{2,1}-\pi_{2,1}}{\sqrt{\pi_{2,1}}} & \frac{j_{2,2}-\pi_{2,2}}{\sqrt{\pi_{2,2}}} & \cdots & \frac{j_{2,N}-\pi_{2,N}}{\sqrt{\pi_{2,N}}} \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ \frac{j_{M,1}-\pi_{M,1}}{\sqrt{\pi_{M,1}}} & \frac{j_{M,2}-\pi_{M,2}}{\sqrt{\pi_{M,2}}} & \cdots & \frac{j_{M,N}-\pi_{M,N}}{\sqrt{\pi_{M,N}}} \end{bmatrix} \tag{A-1}$$

where $\pi_{m,n}$, $m \in \{1, 2, ..., M\}$ and $n \in \{1, 2, ..., N\}$, is the true probability of event $\{m, n\}$ occurring, and is the typical element of Π. Then denote $\mathbb{V} = \text{vec}\mathbf{V}$. Next define

$$\mathbb{v}' = \left(\sqrt{\pi_{1,1}}, ..., \sqrt{\pi_{1,N}}, ..., \sqrt{\pi_{M,1}}, ..., \sqrt{\pi_{M,N}}\right) \tag{A-2}$$

and

$$\Omega = \mathbf{I} - \mathbb{v}\mathbb{v}' \tag{A-3}$$

Then by the results in Rao (1973) pages 383 and 391, and Anderson et al. (2010), we have

$$\mathbb{V} \xrightarrow{a} N_{MN}(\mathbf{0}, \Omega) \tag{A-4}$$

Define the matrix of estimated probabilities as $\mathbf{J}$, and let $\mathbf{j} = \text{vec}\mathbf{J}$ and $\pi = \text{vec}\mathbf{\Pi}$ where vec is the vec-operator. Then,

$$\mathbf{j} \xrightarrow{a} N_{MN}\left(\pi, \frac{1}{n}(\text{dg}(\mathbb{v}))\Omega(\text{dg}(\mathbb{v}))'\right) \tag{A-5}$$

$$\Rightarrow \mathbf{i}'\mathbf{j} \xrightarrow{a} N\left(\mathbf{i}'\pi, \frac{1}{n}\mathbf{i}'(\text{dg}(\mathbb{v}))\Omega(\text{dg}(\mathbb{v}))'\mathbf{i}\right) \tag{A-6}$$

where $\mathbf{i}$ is a vector of ones. Let $\mathbf{j}^p$ and $\mathbf{j}^e$ be the vectorized joint density under positive assortative matching and the empirical counterpart respectively. Define $\mathbf{j}^{\min} = \min\{\mathbf{j}^p, \mathbf{j}^e\}$. Likewise, let π^p and π^e be the corresponding vectorized true probabilities (from $\text{vec}\Pi^p$ and $\text{vec}\Pi^e$ respectively), and let $\pi^{\min} = \min\{\pi^p, \pi^e\}$. Then the Overlap Index is $\mathbf{OV}_p = \mathbf{i}'\mathbf{j}^{\min}$. It is clear then asymptotically by equation (A-6),

$$\mathbf{OV}_p := \mathbf{i}'\mathbf{j}^{\min} \xrightarrow{a} N\left(\mathbf{i}'\pi^{\min}, \frac{1}{n^{\min}}\mathbf{i}'(\text{dg}(\mathbb{v}^{\min}))\Omega^{\min}(\text{dg}(\mathbb{v}^{\min}))'\mathbf{i}\right) \tag{A-7}$$

where $\Omega^{\min} = \mathbf{I} - \mathbb{v}^{\min}\mathbb{v}^{\min\,\prime}$ and

$$\mathbb{v}^{\min\,\prime} = \left(\sqrt{\pi_{1,1}^{\min\,\prime}}, ..., \sqrt{\pi_{1,N}^{\min\,\prime}}, \sqrt{\pi_{2,1}^{\min\,\prime}}, ..., \sqrt{\pi_{2,N}^{\min\,\prime}}, \sqrt{\pi_{3,1}^{\min\,\prime}}, ..., \sqrt{\pi_{M,N}^{\min\,\prime}}\right) \tag{A-8}$$

Note that the variance-covariance matrix can be estimated by replacing $\mathbb{v}^{\min}$ with $\mathbf{j}^{\min}$.

A.2 Proof of Propositions

Proof of Proposition 1 *Let k' be the optimal level of investment per child with $\widetilde{n}$ children in the family. Differentiating k' with respect to $\widetilde{n}$ from (7),*

$$\frac{\partial k'}{\partial \widetilde{n}} = \frac{q_n \widetilde{n} + q + q_k k' - q_{kn}\left(yx - \widetilde{n}k'\right)}{q_{kk}\left(yx - \widetilde{n}k'\right) - 2q_k \widetilde{n}} \leq 0 \tag{A-9}$$

Given assumption 1, a binding constraint on the number of children, i.e. one that is lower than what the parents would have chosen, would increase investments in children.

Proof of Proposition 2 *Differentiating (6) and (7) with respect to y respectively gives,*

$$\frac{\partial n^*}{\partial y} = -\frac{q_n x}{\left(q_{nn}\left(yx - n^* k^*\right) - 2q_n k^*\right)} \geq 0 \tag{A-10}$$

$$\frac{\partial k^*}{\partial y} = -\frac{q_k x}{\left(q_{kk}\left(yx - n^* k^*\right) - 2q_k k^*\right)} \geq 0 \tag{A-11}$$

Therefore, an increase in income would increase the number of children in the family, and/or the level of investment per child.

Proof of Proposition 3 *For the proof of point 1, differentiating $\underline{t_w^R}$ in (11) with respect to the number of children $\widetilde{n}$,*

$$\frac{\partial \underline{t_w^R}}{\partial \widetilde{n}} = \frac{qk' - q_{\widetilde{n}}(yx - \widetilde{n}k')}{q_{\underline{t_w^R}}\left(yx - \widetilde{n}k'\right) + qyx_{\underline{t_w^R}} - yv_{\underline{t_w^R}}} \leq 0 \tag{A-12}$$

where k' is the optimal choice of k given $t_w = \underline{t_w^R}$, t_h and $\widetilde{n}$. Since $\widetilde{n}$ is binding from below, by revealed preference the marginal benefit would be greater than the marginal cost, and the numerator is non-positive. By assumption 2, and $\underline{t_w^R} \leq t_h$, the greater the type of an individual, the greater the gains to marriage, so the denominator is positive.

For the upper bound on the reservation value, we differentiate $\overline{t_w^R}$ in (12) with respect to $\widetilde{n}$ as above.

$$\frac{\partial \overline{t_w^R}}{\partial \widetilde{n}} = \frac{qk'' - q_{\widetilde{n}}\left(yx - \widetilde{n}k''\right)}{q_{\overline{t_w^R}}\left(yx - \widetilde{n}k''\right) + qyx_{\overline{t_w^R}} - yv_{\overline{t_w^R}}} \geq 0 \tag{A-13}$$

Where k'' is the optimal choice of k given $t_w = \overline{t_w^R}$, t_h and $\widetilde{n}$. The numerator as before is non-positive. By assumption 2, and $\overline{t_w^R} \geq t_h$, the denominator is negative, and point 1 follows.

Since there is a narrowing in the range of potential matches around the agents type, incidences of assortative matches rise. Formally, let a man of type t_h be matched with and married to a woman of type t_w^*. Then

$$\Pr(\underline{t_w^R} \leq t_w^* \leq \overline{t_w^R}) = 1 \tag{A-14}$$

It follows that,

$$\int_{\underline{t_w^R}}^{\overline{t_w^R}} f(t_w^*|t_h)dt_w^* = \frac{1}{f(t_h)}\int_{\underline{t_w^R}}^{\overline{t_w^R}} g(t_w^*, t_h)dt_w^* = \frac{1}{f(t_h)}\left[G(\overline{t_w^R}, t_h) - G(\underline{t_w^R}, t_h)\right] = 1 \tag{A-15}$$

where $g(.)$ *and* $G(.)$ *are respectively the joint density and joint distribution functions. The total differential of (A-15) with respect to* $\widetilde{n}$ *may be written as,*

$$\frac{1}{f(t_h)}\left[\frac{\partial G(\overline{t_w^R}, t_h)}{\partial \overline{t_w^R}}\frac{\partial \overline{t_w^R}}{\partial \widetilde{n}} - \frac{\partial G(\underline{t_w^R}, t_h)}{\partial \underline{t_w^R}}\frac{\partial \underline{t_w^R}}{\partial \widetilde{n}}\right] d\widetilde{n} + \frac{1}{f(t_h)}\frac{\partial\left[G(\overline{t_w^R}, t_h) - G(\underline{t_w^R}, t_h)\right]}{\partial \widetilde{n}} d\widetilde{n} = 0 \tag{A-16}$$

Since

$$\frac{1}{f(t_h)} > 0, \frac{\partial G(t_w, t_h)}{\partial t_w} > 0, \frac{\partial \overline{t_w^R}}{\partial \widetilde{n}} \geq 0, \frac{\partial \underline{t_w^R}}{\partial \widetilde{n}} \leq 0 \tag{A-17}$$

It may be observed that

$$\frac{\partial\left[G(\overline{t_w^R}, t_h) - G(\underline{t_w^R}, t_h)\right]}{\partial \widetilde{n}} \leq 0 \tag{A-18}$$

Proof of Proposition 4 *As in the proof of proposition 3, differentiate* $\underline{t_w^R}$ *and* $\overline{t_w^R}$ *in (11) and (12) with respect to* y *respectively.*

$$\frac{\partial \underline{t_w^R}}{\partial y} = \frac{-qx + v(\underline{t_w^R}) + v(t_h)}{q_{\underline{t_w^R}}(yx - \underline{n}\underline{k}) + qyx_{\underline{t_w^R}} - yv_{\underline{t_w^R}}} \leq 0 \tag{A-19}$$

First note that by the complimentarity assumption of assumption 2, and $\underline{t_w^R} \leq t_h$*, the greater the type of an individual, the greater the gains to marriage, so the denominator is positive. Secondly, if* $u^h(y = 0) \leq s^h(y = 0)$ *and* $\frac{\partial u^h}{\partial y} \geq \frac{\partial s^h}{\partial y}$*, the numerator is negative, and the inequality follows.*

$$\frac{\partial \overline{t_w^R}}{\partial y} = \frac{-qx + v(\overline{t_w^R}) + v(t_h)}{q_{\overline{t_w^R}}(yx - \overline{n}\overline{k}) + qyx_{\overline{t_w^R}} - yv_{\overline{t_w^R}}} \geq 0 \tag{A-20}$$

By assumption 2, and $\overline{t_w^R} \geq t_h$*, the denominator is negative. Here, similarly if* $u^h(y = 0) \leq s^h(y = 0)$ *and* $\frac{\partial u^h}{\partial y} \geq \frac{\partial s^h}{\partial y}$*, the numerator is negative, and the inequality follows. The rest of the arguments are similar to proposition 3.*

Proof of Proposition 5 *The proof mirrors that of proposition 4 with the exception that with* $u^h(y = 0) \geq s^h(y = 0)$ *and* $\frac{\partial u^h}{\partial y} \leq \frac{\partial s^h}{\partial y}$ *the signs of the numerators in equations (A-19) and (A-20) are now both positive, so that the signs of (A-19) and (A-20) are switched.*

Table A.1: Density of Number of Children Among Married Couples

Province	# of Children	40s Cohort	50s Cohort	60s Cohort	Province	# of Children	40s Cohort	50s Cohort	60s Cohort
	0	0.16	0.03	0.04		0	0.03	0.01	0.05
	1	0.39	0.80	0.92		1	0.48	0.86	0.88
	2	0.38	0.16	0.03		2	0.40	0.12	0.06
Jilin	3	0.06	0.01	0.00	Guangdong	3	0.07	0.01	0.00
	4	0.00	0.00	0.00		4	0.01	0.00	0.00
	5	0.00	0.00	0.00		5	0.00	0.00	0.00
	# of Obs.	1405	2687	1647		# of Obs.	1822	2690	1077
	0	0.09	0.01	0.03		0	0.21	0.04	0.06
	1	0.43	0.85	0.93		1	0.55	0.90	0.92
	2	0.41	0.14	0.05		2	0.22	0.05	0.02
Shandong	3	0.07	0.00	0.00	Sichuan	3	0.02	0.00	0.00
	4	0.00	0.00	0.00		4	0.00	0.00	0.00
	5	0.00	0.00	0.00		5	0.00	0.00	0.00
	# of Obs.	1312	2976	1860		# of Obs.	2588	4433	2077
	0	0.10	0.02	0.03		0	0.15	0.03	0.04
	1	0.46	0.87	0.95		1	0.44	0.78	0.92
	2	0.40	0.10	0.02		2	0.34	0.18	0.03
Hubei	3	0.04	0.01	0.00	Shaanxi	3	0.06	0.01	0.01
	4	0.00	0.00	0.00		4	0.01	0.00	0.00
	5	0.00	0.00	0.00		5	0.00	0.00	0.00
	# of Obs.	1914	3282	1570		# of Obs.	1336	1804	973