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

Gordon Anderson*
University of Toronto

Teng Wah Leo †‡
St. Francis Xavier University

First Version: 1 January, 2007
This Version: 11 August, 2009

*Address: Department of Economics University of Toronto, 150 St. George Street, Toronto, Ontario, M5S 3G7, Canada. Email: anderson@chass.utoronto.ca

†**Corresponding Author**. Address: Department of Economics, St. Francis Xavier University, 1 West Street, Antigonish, Nova Scotia, B2G 2W5, Canada. Telephone: 1 (902) 867 5449. Fax: 1 (902) 867 3610. Email: tleo@stfx.ca

‡We would like to thank Aloysius Siow, Loren Brandt, Lars Osberg, and Kuan Xu for their very helpful comments and suggestions, and seminar participants at CEPA at The University of Toronto, Dalhousie University, the meetings of the Canadian Economics Association, and the Far Eastern & South Asian meeting of the Econometric Society.

Abstract

In 1979 the Chinese government implemented the One Child Policy in an attempt to ameliorate the potential negative economic implications of the nation's perceived population explosion. This paper examines the consequences of this policy on marital matching both theoretically and empirically. A simple general equilibrium model is developed which predicts that constraining marital output in the child quantity dimension raises the marginal benefit of positive assortative matching and investment in child quality, increasing the intensity with which they are pursued, concomitantly it reduces the marriage rate. A Poisson model is employed to check the extent to which the policy was binding and a general index of matching intensity with convenient statistical properties is developed to examine the matching predictions on three cohorts of urban households, one pre-policy, one post-policy and one which spans the period during which the policy was introduced. Significant increases in positive assortative matching and reductions in negative assortative matching were observed, as well as a reduced rate of pairing.

JEL Code: J12, J13

Key Words: Marriage; Matching

1 Introduction

One of the most controversial and far reaching population control policies in recent history is China's One Child Policy (OCP). Introduced in 1979 the OCP represented a considerable intervention in the household choice process, implemented with fines and various other forms of coercion, it encouraged families to limit production of offspring. Such an intervention could have changed fundamentally the nature of both existing and anticipated marriage arrangements and can be expected to have influenced family formation decisions in many dimensions, for instance in the choice of partner, the family size and investments in children. As such it appears to provide a natural pseudo experiment within which, *ceteris paribus*, the nature of family formation choices can be examined before and after the policy. Unfortunately not all else is equal and some context for the policy is appropriate.

Firstly, there is a sense in which the desired outcome of 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 be rationalized as a result of urbanization[1] 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.

Secondly while previous work found that the OCP enforced a binding constraint on family size (Zhang 2002), it should be noted that the policy was introduced in tandem with the Economic Reforms of 1979 which precipitated a well documented increase in the incomes of the population. 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 model discussed in the following section, since it predicts opposing effects for economic growth and the OCP with respect to partner choice decisions.

Thirdly, within the premise of a binding policy, changes in the apparent preference for

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

sons in China have to be contended with (The usual boy/girl sex ratios at birth are around 104/100, while China's in 1995 for example was 117/100 (Peng and Guo 2000)). These preferences are apparently founded upon a tradition of patrilocal residence of married sons who, as opposed to daughters, thus provide considerable old age security benefits for their parents (See Therborn (2004) who suggests that expression of these preferences has been facilitated by the development and availability of fetus gender detection and selective abortion techniques.). Such preferences may well be more strongly held in rural as opposed to urban situations.

With regard to the decline in family size prior to the OCP, the demographer J.C. Caldwell developed a theory which has a distinctly economic flavour (Caldwell 1982). His view was that fertility was high when children are an asset to their parents and low when they become a liability, although empirical verification of the idea encountered difficulties since "...the marginal value of each extra child is impossible to determine..." (Caldwell et al. 1982). Becker (1993) formalized this in developing models where both number of and quality of children and the quality of partners feature as part of the household decision process. Becker's model can be used to rationalize the effect of urbanization and the preference for sons at birth. An important feature of Becker's analysis is that "quantity" and "quality" choices are to some degree simultaneous, with each influencing the other to an extent[2]. He demonstrates that while quantity and quality are likely to be substitutes, they cannot be close substitutes (because the budget constraint between quantity and quality is convex, equilibrium would not exist if the indifference curve between quantity and quality were in some sense "less" convex).

The concern here however is with the change in the *degree of assortative matching* induced by exogenous family size constraints imposed on the family. A simple static general equilibrium model of marital matching is presented, where choice of a spousal match is dependent on the individual's measurable continuously distributed attribute or quality as well as the consequent choices in child quality and quantity, should the marriage

[2]Family formation has most frequently been discussed in the economics literature 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. To some extent this hinges upon the nature of the planning horizon. One culture in developing female labour supply models is to assume that lifetime fertility decisions are made early in life, "at marriage is the most popular choice" observes Browning (1992). An alternative culture is to assume a simultaneous model where the agent attempts to have more children while making her labour supply decision.

take place. This approach permits the examination of how a binding constraint on child quantity(Neary and Roberts (1980); Deaton (1981)) or family size decision 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 childrearing attributes less of a comparative advantage relative to someone with income generating attributes all other things equal. The effect of the OCP on child quality is examined in Anderson and Leo (2009).

The empirical approach in this paper differs from recent work in the empirical matching literature (See Choo and Siow (2006) and Dagsvik (2000)) but builds on the empirical literature in mobility measures (See Dardanoni (1993), Maasoumi (1996), Quah (1996), Shorrocks (1976), and Shorrocks (1978)) and stochastic dominance measures (See Anderson (1996), Atkinson (1970), Bourguignon and Fields (1997), Davidson and Duclos (2000)). A simple and easily applied statistic is provided with which to measure the proximity between an empirical joint density in the matched individual's attribute and that generated by a hypothesized matching scheme, such as positive or negative assortative matching. Further, the statistic is mean invariant and asymptotically normally distributed, which facilitates inferences.

The observations are drawn from a bi-annual urban household survey of six provinces in China from 1989 to 2001; Shaanxi, Jilin, Hubei, Sichuan, Guangdong and Shandong[3]. The spousal choice is considered in terms of the cohort of males and females by year of birth. Specifically, the sample is divided into 3 cohorts, the first with couples where the husband was born between 1940 and 1949, the second cohort where the husband was born between 1950 and 1959, and the last from 1960 to 1969. This was done principally because the empirical 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 implementation of the policy.

To establish the notion that the OCP constituted a binding constraint to the 1960s cohort, a Poisson model of the number of children born after the first child was employed.

[3]These data were 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.

The rationale for this was that if the constraint was binding, yet such children were born they would have been "accidents". While the model was strongly rejected for the 1940s cohort, it was not for the 1960s cohort lending support to the view that the OCP constituted a binding constraint. Partner matching was then studied in terms of an integer index of individual educational attainment. There is significant evidence of increased positive assortative matching amongst the post-OCP cohorts which cannot be explained by the Economic Reforms of China over the past decades, nor the historical regime shift since the Cultural Revolution.

In the following, Section 2 formulates a simple model and develops some comparative statics for the various family formation decisions. Section 3 outlines the empirical strategy, 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 4 and some discussion and conclusions are reported in section 5.

2 A Simple Model

Consider a model where an individual lives for 2 periods, one as a child, and one as an adult. At the beginning of the adult period, agents choose to marry or remain single (there is no divorce in this model)[4]. The rate at which an adult meets someone of the opposite gender is random. Marriage is dependent on the type (quality and type will be used interchangeably) of the man and that of their potential spouse, and utility is assumed transferable. Let the subscript h denote a male, and w denote a female. Let the agent's type t_i, $i \in \{h, w\}$, be continuous on a support $\left[\underline{t}, \overline{t}\right]$, $\underline{t}, \overline{t} \in \mathbb{R}$, and distributed with density $f(.)$ and distribution $F(.)$ for both male and female. If they find a match, they will then choose the number of children to have and the amount of investment in each child. 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 amount of investment per child k, that is $q \equiv q\left(t_h, t_w, k, n\right)$, such that $q \mapsto \{0\} + \mathbb{R}_+$ is increasing and concave in

[4]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 imposition of other sharing rules will not affect the essence of the results presented below.

all it's inputs. The other aspect of a married individual's utility is derived from personal consumption c_i, $i \in \{h, w\}$. Finally, the utility function is multiplicatively separable in the utility derived from the children and that from own consumption, $u_h = q\,(t_w, n, k|t_h)\, c_h$. 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, $s_i = \max_{c_i} c_i$, $i \in \{h, w\}$.

The income realization of the family or individual is assumed to be dependent on the type of match and the individual's type respectively. Specifically, family income is assumed to be $yx(t_h, t_w)$, and income for a single individual to be $yv(t_i)$, $i \in \{h, w\}$, where y is the average income within the economy, $x : (t_h, t_w) \mapsto \{0\} + \mathbb{R}_+$ and $v : t_i \mapsto \{0\} + \mathbb{R}_+$. This setup thus abstracts from redistributive concerns arising from any policy. Further, this formulation of income together with the range of q ensures that for some matches and individual types, the choice of remaining single will be made. That is the set of single individuals by type is non-empty.

The following functional assumptions are also made,

Assumption 1 *: Investment in children, k, and the choice of the number of children, n, are substitutes in the function $q(.)$, which parents derive from having children in their marriage. That is $q_{k,n}(t_w, n, k|t_h) \leq 0$.*

Assumption 2 *: $u_{t_i} \geq 0$, $u_{t_i,t_i} \leq 0$, for $t_i \in \left[\underline{t}, \bar{t}\right]$, $i \in \{h, w\}$.*

Assumption 3 *: (**Complementarity of Types**) $u_{t_i,t_j} \geq 0$, $i \neq j$, $i, j \in \{h, w\}$. Further, let $t^* = \arg \max_{t_w \in \left[\underline{t}, \bar{t}\right]} u\,(t_w, n, k|t_h) \Leftrightarrow t^* = t_h = t_w$, for $t_h, t_w \in \{\underline{t}, \bar{t}\}$.*

Assumption 4 *: (**Convex in Types When Single**): $v_{t_i} \geq 0$, $v_{t_i,t_i} \geq 0$, $i \in \{h, w\}$.*

Assumption 5 *: (**Single Crossing with respect to Average Income**): $\frac{\partial u_h}{\partial y}, \frac{\partial s_h}{\partial y} > 0$, $u_h(y = 0) \leq s_h(y = 0)$, and that $\frac{\partial u_h}{\partial y} \geq \frac{\partial s_h}{\partial y}$.*

Assumption 1 creates the tradeoff between the choice of investment per child, and the number of children in a family. Assumption 2 ensures that $u_i(.)$ $i \in \{h, w\}$ is well behaved on the support of the agent's type. Assumption 3 says that given an agent's type, they would prefer to be matched with someone of the same type or better. Together with assumption 4, this ensures that agents would always prefer to match with someone closer to their own type, since the concavity of $u_i(.)$ $i \in \{h, w\}$ in own type and the type of spouse, and the convexity of $v(.)$ in own type ensures that gross marital output attains a maxima for agents of a sufficiently low type on the support. An example of a function that would meet these assumptions is when $q(.)$ and $x(.)$ are quadratic functions with respect to $(t_h - t_w)$ on $t_h, t_w \in [0, 1]$[5].

Assumption 5 pertains to the effect of income on preferences in the marriage and single state. It ensures that utility gained from marriage increases at a faster rate with respect to income than in the single state, so that once an individual finds marriage desirable at his current income, increases in it will not reduce his desire to be married. In effect it ensures that the utility, with respect to income, from marriage and being single can intersect at most once.

Abstracting from intra-household bargaining and focusing on the total value of marital output, without loss of generality the solution to the individual's problem will be solved from the perspective of the man choosing a prospective wife.

2.1 Single Agent

If an individual of type t_i $i \in \{h, w\}$, chooses to remain single, he solves the following problem,

$$\max_{c_i} c_i$$

subject to

$$yv(t_i) \geq c_i$$

where y is the average income of all individuals within the economy, and $v : t_i \mapsto 0 + \mathbb{R}_+, i \in \{h, w\}$. Then an individual i's income is described by the product of $v(t_i)$, $i \in \{h, w\}$ and y, which means that his income is a proportion of the average income, dependent ultimately on his type. The optimal consumption choice is that the individual spends all his income on himself $c_i = yv(t_i)$. Let the optimized utility of this single individual be,

$$\widehat{s}_i = yv(t_i) \tag{1}$$

[5]We suspect a model with search costs that fall as agent types rises may generate similar results we present below.

2.2 Married Man

If the individual finds a suitable match and chooses marriage, he solves the following problem subject to his budget constraint and the participation constraint in order for his prospective spouse to enter into matrimony with him.

$$\max_{n,c_h,c_w,k} q(t_w, n, k|t_h)c_h$$

subject to

$$\begin{aligned} c_h + c_w + nk &\leq yx(t_h, t_w) \\ q(t_w, n, k|t_h)c_w &\geq yv(t_w) \end{aligned}$$

where c_h, and c_w, are the consumption choices, and t_h and t_w are the type 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, the participation constraint holds with equality as well. Thus

$$\begin{aligned} c_w &= \frac{yv(t_w)}{q(t_w, n, k|t_h)} \\ \Rightarrow c_h &= yx(t_h, t_w) - nk - \frac{yv(t_w)}{q(t_w, n, k|t_h)} \end{aligned}$$

and he solves,

$$\max_{n,k} q(t_w, n, k|t_h)(yx(t_h, t_w) - nk) - yv(t_w) \tag{2}$$

The first order conditions are,

$$q_n(t_w, n^*, k^*|t_h)(yx(t_h, t_w) - n^*k^*) = q(t_w, n^*, k^*|t_h)k^* \tag{3}$$

$$q_k(t_w, n^*, k^*|t_h)(yx(t_h, t_w) - n^*k^*) = q(t_w, n^*, k^*|t_h)n^* \tag{4}$$

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

$$\frac{q_n(t_w, n^*, k^*|t_h)}{k^*} = \frac{q_k(t_w, n^*, k^*|t_h)}{n^*} \tag{5}$$

However, under a situation where n is no longer a choice variable only (4) would prevail, hence the effect of changes in n on the optimal choice of k can be examined as if n is a parameter. For the rest of the paper, let $n = \widetilde{n}$ be for cases where the number of children is exogenously determined and let the respective optimal choice of investment for each child be k' there.

2.3 Comparative Statics

The OCP in China coincided with the Chinese Economic Reforms in 1979 which precipitated considerable economic growth. Should the impact of economic growth on familial choices yield similar outcomes to the OCP, it would not be possible to identify the separate policy effects. This section examines the impact derived from both policies on quantity and quality of children, followed by spousal matching decisions. The following four propositions relate to how the OCP and economic growth might have affected spousal and family size choices (proofs are supplied in appendix A.1).

First, let $\widehat{u}_h = \max_{n,k} q(t_w, n, k|t_h)(yx(t_h, t_w) - nk) - yv(t_w)$ and $\widehat{s}_h = yv(t_h)$, then a type t_h man's second period utility is,

$$U_h = \max\{\widehat{u}_h, \widehat{s}_h\} \tag{6}$$

The reservation type of his potential spouse is determined by

$$\begin{aligned} \widehat{u}_h &= \widehat{s}_h \\ \Rightarrow q(\underline{t_w^R}, \underline{n}, \underline{k}|t_h)(yx(t_h, \underline{t_w^R}) - \underline{nk}) - yv(\underline{t_w^R}) &= yv(t_h) \end{aligned} \tag{7}$$

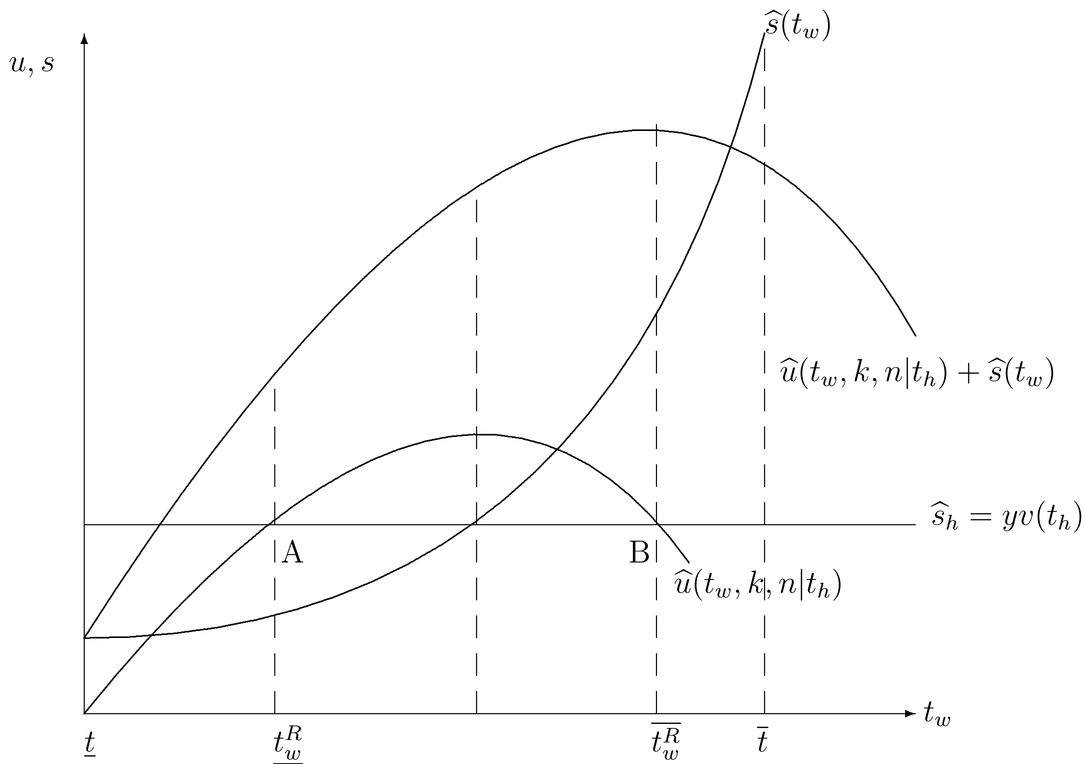
where $\underline{n}$ and $\underline{k}$ are the optimal values for a match between a man of type t_h and woman of type $\underline{t_w^R}$. Letting $\underline{t_w^R} \equiv \underline{t_w^R}(t_h)$, from figure 1 it may be observed that (7) determines only the lower bound of the reservation at point 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 a sufficiently low type 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 $\widehat{s}(\overline{t_w^R})$, in other words the utility she would otherwise get from remaining single. This upper threshold is determined by

$$q(\overline{t_w^R}, \overline{n}, \overline{k}|t_h)(yx(t_h, \overline{t_w^R}) - \overline{n}\overline{k}) - \widehat{s}(\overline{t_w^R}) = \widehat{s}_h = yv(t_h) \tag{8}$$

The upper bound is point B in figure 1. 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 pay to meet her participation constraint. This is where the slope of the gross utility and $yv(t_w)$ equates, and coincides at the man's own type. Beyond this optimal type, his own marital gains start decreasing, and fall below his value of remaining single eventually. Note that by construction, $\underline{t_w^R} \leq t_h \leq \overline{t_w^R}$.

Figure 1: Reservation Values given Type



Intuitively, given that quantity and quality of children are substitutable, a binding policy that impinges on a family's choice in one dimension should yield an increase in the remaining dimension which is stated in the following proposition.

Proposition 1 *: An exogenously enforced reduction in the number of children raises equilibrium investment in children.*

Yet the success of the Economic Reform of 1979, which raised the income, and consequently the quality of lives among the Chinese populace, should similarly 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 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 true cause of changes in investment in children if any, should the impact of the OCP be considered solely from the perspective of child outcome. However the manner in which either policy could have effected spousal choices can also be examined. Intuitively, 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 imposition or rationing of child quantity via the OCP could have also accentuated the importance of good spousal match, assuming positive assortative matching is the norm. Note the existence of positive assortative matching is not disputed, rather the *degree or intensity* of positive assortative matching may have been altered.

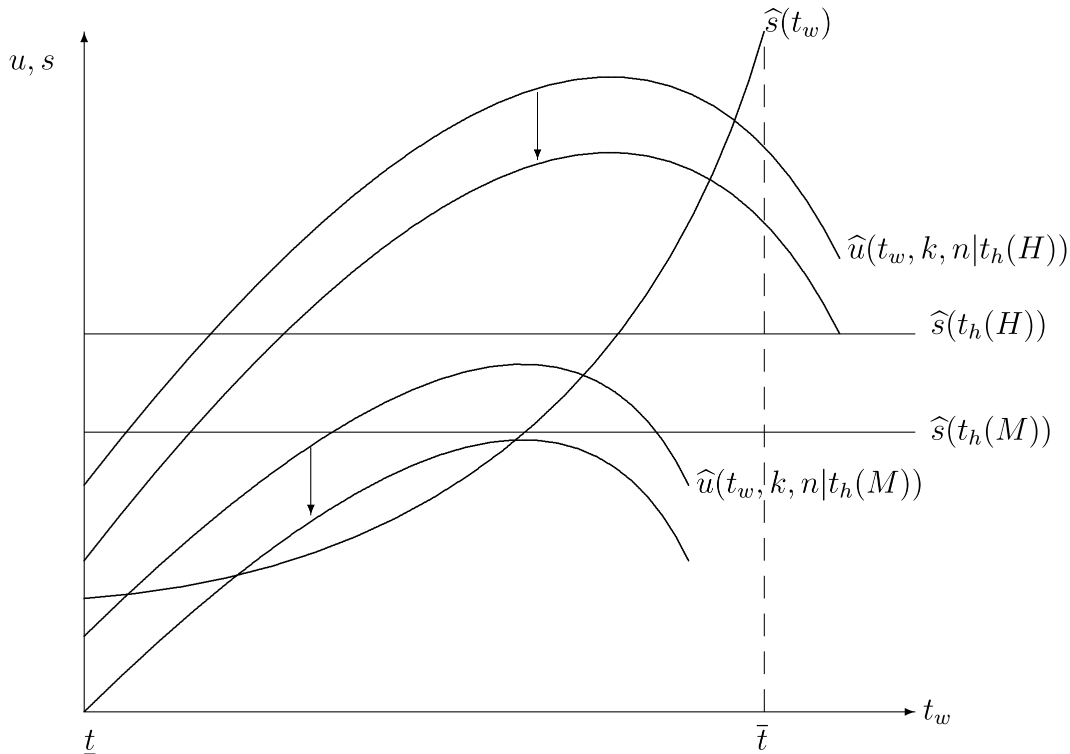
Proposition 3 *: When the number of children is fixed below the optimal choice that a married couple would have chosen given their types, then:*

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 (M), and those who are coveted by all spousal types (H). Figure 2 shows how a binding family size policy might affect choice of spousal type. 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 the potential spouse. This process is depicted as a fall in $\widehat{u}\,(t_w, k, n|t_h(H))$ for a man of type $t_h(H)$, noting further that for a substantial fall in the utility of the man from marriage, he might not be able to attract spouses of higher types. On the other hand, for a sufficiently low type agent, this may even mean a complete withdrawal from the marriage market as shown in figure 2, in the fall of $u\,(t_w, k, n|t_h(M))$ for a man of type $t_h(M)$. The latter observation is reflected in the following corollary.

Corollary 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.*

Figure 2: Impact of Binding Family Size Policy on Spousal Type



On the other hand economic growth, by raising disposable income, could potentially slacken the need for a good spousal match. However, at the same time, economic growth may have also raised the gains to remaining single, thereby reducing the merits of marriage. These possibilities are examined in the following propositions.

Proposition 4 *: 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.*

The intuition to proposition 4 is as follows; if at the status quo on the margin of spousal type, the man is indifferent to marrying or remaining single, an increase in income available to him cannot make his potential spouse any less attractive. However, if it makes her more attractive, by increasing his utility, the marginal prospective spousal type at the lower bound must fall, while the upper bound must increase. (This outcome is facilitated by the assumption that an increased average income within the economy has no redistributive effects, so growth results in a shift of the entire distribution to a new mean income level, while maintaining its shape). The corollary below follows:

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

Assume that each individual meets one and only one potential spouse in their lifetime, so that if a man meets a women within the bounds of a potential spouse, he will marry her. Therefore the probability of marriage for a man of type t_h is P such that,

$$P = \Pr(\underline{t_w^R} \leq t_w \leq \overline{t_w^R}) = \int\limits_{\underline{t_w^R}}^{\overline{t_w^R}} f(t_w)dt_w = F(\overline{t_w^R}) - F(\underline{t_w^R}) \tag{9}$$

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

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

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

The model has explicitly argued that the two venues through which matching in the marriage market could have been *directly* affected were through constraining family size due to the OCP and the increase in income as a result of the Economic Reforms of 1979. However other possible venues through which both policies could have affected matching should be acknowledged, at least conceptually. One possible *indirect* effect that may affect matching via the Economic Reforms is through changes in the returns to education. Essentially as the gains to human capital investment increase, the marriage market would see changes in the composition at various educational attainment levels, that is the marginal distribution of educational attainment 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 choice set itself* which is what this paper examines and attempts to measure. Similarly, with respect to the OCP, if parents 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 Peters and Siow (2002) for a model on premarital investments in children). However analysis of these effects are beyond the purview of this paper.

3 The Overlap Measure and Empirical Strategy

3.1 Elements of a Matching Matrix, the Matching Index and its Asymptotic Distribution

To assess the change in matching behavior via the relationship between the respective attributes of a pairing, the joint density of the attributes of spouses are compared with what could have emerged under a perfect positive (negative) assortative scenario. Anderson et al. (2009) provide a matching index which makes such a comparison.

To illustrate the use of the measure, suppose the type space of both husbands and wives can be partitioned into five mutually exclusive types (as in the data used herein) such that $t_i \in \{t_i^1, t_i^2, ..., t_i^5\}$ where $i \in \{h, w\}$ and $t_i^1 < t_h^2 < ... < t_h^5$. If the type 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 is of the form,

$$\mathbf{J}_p = \begin{bmatrix} \Pr(t_i = t_i^1) & 0 & ... & 0 \\ 0 & \Pr(t_i = t_i^2) & ... & 0 \\ : & : & ::: & : \\ 0 & 0 & ... & \Pr(t_i = t_i^5) \end{bmatrix} \tag{10}$$

On the other hand, it may not always be possible to partition the support of types such that the above joint density matrix is derived (particularly when the realizations of types are discrete). 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 $m \in \{1, 2, .., 5\}$, that is men stochastically dominate women in the type measure. Then the realized 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:

$$\mathbf{J}_p = \begin{bmatrix} \Pr(t_h = t_h^1) & 0 & ... & 0 & 0 \\ \Pr(t_h \geq t_h^2) - \Pr(t_w \geq t_w^2) & \Pr(t_w \geq t_w^2) - \Pr(t_h \geq t_h^3) & ... & 0 & 0 \\ 0 & \Pr(t_h \geq t_h^3) - \Pr(t_w \geq t_w^3) & ... & 0 & 0 \\ . & . & ... & . & . \\ 0 & 0 & ... & \Pr(t_h = t_h^5) - \Pr(t_w = t_w^5) & \Pr(t_w = t_w^5) \end{bmatrix} \tag{11}$$

Estimates of such a matrix can be constructed from the empirical marginal distributions of men's and women's types. Although only results using the above matrix are reported, there are other methods of arriving at the positive and negative assortative matching matrix which were examined as well, 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)[6].

The idea underlying the matching index is that the more intensively is the above pattern of matches pursued, the closer will the empirical joint density of matches be to (11). Thus proximity to complete positive assortative matching can be assessed by comparing the degree of concurrence of the joint density matrices of $\mathbf{J}_p$ constructed from

[6]These results are available from the authors upon request, generally they did not produce as close a fit to the empirical joint density as equation (11) suggesting that the prevailing matching mechanism is the one underlying equation (11).

the marginal densities under the null against the empirical joint density, just as in the case of independence or contingency table tests. Specifically, let the elements of the joint density matrix generated by the null hypothesis be $j^p_{i,k}$, and that for the empirical joint density by $j^e_{i,k}$, where $i, k \in \{1, 2, ..., n\}$, n being the number of mutually exclusive type realizations for both *married* men and women. The measure of the overlap between theoretical and empirical joint density then provides an index of the *degree or intensity* of positive assortative matching[7]. Specifically,

$$OV_p = \sum_{i=1}^{n} \sum_{k=1}^{n} \min \left\{ j^p_{i,k}, j^e_{i,k} \right\} \tag{12}$$

This overlap measure, which lies between 0 and 1, is asymptotically normally distributed (Anderson et al. 2009) and changes in the measure provide evidence of changes in the degree of positive assortative matching (A short explanation is provided in appendix A.2). The attractive feature of these indices is that they can be readily applied when $\mathbf{J}_p$ is not square and they can be implemented in multivariate domains. Further, since they are asymptotically normally distributed, they facilitate inferences about trends toward or away from different matching patterns.

3.2 Empirical Strategy, Data Summary and the Effectiveness of the Constraint

The bi-annual samples are pooled and divided into three cohorts of individuals based on the birth year of men, so that a couple is classified as belonging to the 1940s cohort if the husband is born between 1940 to 1949, likewise for the 1950s and 1960s cohort. This classification follows from the assumption that offers are from men and permits the

[7] Complete negative assortative matching can be examined in a similar fashion, where the joint density matrix $\mathbf{J}_n$ under the null hypothesis of negative assortative matching is a counter-diagonal matrix, where the highest type individuals match with the lowest type from the other gender. In the perfectly matched marginal density case it follows that,

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

examination of family formation decisions prior to the introduction of the OCP (represented by the 1940s cohort), during the introduction of the OCP (the 1950s cohort) and after the introduction of the OCP (the 1960s cohort). Partners are assumed to match on the basis of their educational attainments, which are integer indexed from 1 to 5, with 5 being college graduates and above, 4 being individuals who obtained technical education, 3 being high school, 2 being middle school, and 1 being primary school and lower. It is also assumed that marriage markets are closed within the provinces, so the analysis will proceed by province.

Table 1: Summary of Parental Characteristics

Province	Variable	1940s Cohort		1950s Cohort		1960s Cohort	
		Mean	S.D.	Mean	S.D.	Mean	S.D.
Jilin	Number of Children	1.3547	0.8333	1.1550	0.4618	0.9904	0.2753
	Father's Education	3.1818	1.3794	3.1741	1.1636	3.4738	1.1559
	Mother's Education	2.7004	1.2614	3.0165	1.0724	3.3311	1.0927
	Observations	1342		2723		1661	
Shandong	Number of Children	1.4867	0.7726	1.1478	0.3971	1.0182	0.2585
	Father's Education	3.1128	1.3216	3.3286	1.2541	3.8288	1.1336
	Mother's Education	2.5531	1.2518	2.9259	1.1048	3.3512	1.0745
	Observations	1206		2970		1922	
Hubei	Number of Children	1.4012	0.7608	1.1157	0.4023	0.9927	0.2229
	Father's Education	3.1812	1.3377	3.2202	1.2133	3.7847	1.1258
	Mother's Education	2.5248	1.2449	2.8931	1.0500	3.3681	1.0865
	Observations	1649		3397		1649	
Guangdong	Number of Children	1.5875	0.7460	1.1696	0.4427	1.0152	0.3327
	Father's Education	3.0413	1.4145	3.2011	1.2509	3.6340	1.0717
	Mother's Education	2.4732	1.2586	2.9261	1.0687	3.3612	1.0666
	Observations	1549		2760		1254	
Sichuan	Number of Children	1.0647	0.7603	1.0133	0.3501	0.9744	0.2563
	Father's Education	3.1247	1.3725	2.9652	1.3065	3.6205	1.1736
	Mother's Education	2.5109	1.2525	2.7036	1.0942	3.3821	1.1050
	Observations	2165		4514		2308	
Shaanxi	Number of Children	1.3491	0.8425	1.1680	0.4827	1.0118	0.3199
	Father's Education	3.2498	1.2991	3.2704	1.2642	3.6171	1.1439
	Mother's Education	2.6821	1.1480	2.8971	1.0208	3.1988	1.0151
	Observations	1249		1827		1016	

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

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

Province	Sample with Children			Sample with First Child Male			Sample with First Child Female		
	1940-1949	1950-1959	1960-1969	1940-1949	1950-1959	1960-1969	1940-1949	1950-1959	1960-1969
Jilin	93.6160	10.4290	0.6662	50.9660	10.2570	0.1746	42.9730	3.7461	0.5444
	(0.0000)	(0.0153)	(0.8811)	(0.0000)	(0.0165)	(0.9816)	(0.0000)	(0.2902)	(0.9090)
	1.5949	1.1658	1.0384	1.5815	1.1311	1.0280	1.6092	1.2018	1.0486
	1170	2437	860	607	1243	428	563	1194	432
Shandong	83.8070	6.8248	0.3350	40.1020	6.8786	0.8844	44.5130	4.1572	1.7326
	(0.0000)	(0.0777)	(0.9533)	(0.0000)	(0.0759)	(0.8292)	(0.0000)	(0.2450)	(0.6297)
	1.5951	1.1326	1.0560	1.5891	1.0910	1.0376	1.6022	1.1785	1.0781
	1193	2805	1124	645	1472	612	548	1333	512
Hubei	160.4400	3.9586	0.3219	85.7050	0.4048	0.2199	77.6480	3.1315	0.1114
	(0.0000)	(0.2660)	(0.9559)	(0.0000)	(0.9393)	(0.9743)	(0.0000)	(0.3718)	(0.9904)
	1.5377	1.1057	1.0259	1.5321	1.0820	1.0302	1.5444	1.1314	1.0216
	1722	3046	926	934	1585	464	788	1461	462
Guangdong	62.1030	14.7550	0.7909	50.1570	17.0430	0.1743	28.6870	5.7603	0.7851
	(0.0000)	(0.0020)	(0.8516)	(0.0000)	(0.0007)	(0.9816)	(0.0000)	(0.1239)	(0.8530)
	1.6001	1.1311	1.0585	1.5422	1.1203	1.0369	1.6714	1.1420	1.0874
	1738	2380	427	959	1197	244	779	1183	183
Sichuan	19.5340	1.9527	0.1060	4.2455	1.4128	0.0147	18.8190	1.2121	0.1169
	(0.0002)	(0.5823)	(0.9911)	(0.2361)	(0.7025)	(0.9995)	(0.0003)	(0.7501)	(0.9897)
	1.3414	1.0506	1.0139	1.3443	1.0359	1.0073	1.3380	1.0667	1.0207
	2030	3992	1081	1098	2088	549	932	1904	532
Shaanxi	35.4280	3.5457	15.4910	14.0700	2.4533	46.5000	27.0940	2.0168	0.2537
	(0.0000)	(0.3149)	(0.0014)	(0.0028)	(0.4838)	(0.0000)	(0.0000)	(0.5689)	(0.9685)
	1.5675	1.1769	1.0550	1.5223	1.1439	1.0382	1.6187	1.2101	1.0717
	1140	1622	527	605	813	262	535	809	265

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 parenthesis. 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.

Table 1 summarizes some of the characteristics of married couples within our sample. First note the ubiquitous fall in the number of children over the decades, and particularly among the 1960s cohort. In addition, note the increase in educational attainment over the decades which may be due to increased returns to education with the economic reforms, or it could be from increased investments in children by parents as discussed earlier, or simply due to the regime shift from pre- to post-cultural revolution China.

Table 2 reports the results for a simple Poisson model of the number of children a couple has subsequent to their first child. The first order effect of the OCP is to curtail the demand for children, consequently it may be conjectured that the presence of any additional children after the first child is likely to be an accident, which underlies the Poisson model. The first panel relates to the three cohorts without conditioning on the gender of the first child, the second reports the results when the first child is male, while the third panel reports those for when the first child is female.

Note from the first panel that for the pre-OCP 1940s cohort, the Poisson "accidents" model is rejected for all provinces at the 1% level whereas the 1950s and 1960s cohorts yield only two rejection of the model (Guangdong in the 1950s cohort and Shaanxi in the 1960s cohort). When the sample is split into the gender of the first born child, the same results prevail when the first born is male. However, when the first born is a female, the Poisson model is rejected for all provinces for the 1940s cohort, and not rejected for all provinces for the 1950s and 1960s cohorts. Overall these results must be viewed as strong evidence that post-OCP births after the first child are well described by a Poisson accidents model, confirming the efficacy of the OCP in Urban China.

Parenthetically, although it was not addressed in the model, the data may shed some light on the gender selection issue. Table 3 presents Standard Normal Tests of the null hypothesis that the proportion of first born children that are male is less than or equal to the natural rate. As may be seen, the hypothesis is rejected for the pre-OCP (1940s) cohort at the 5% level in 4 of the 6 provinces, whereas it is rejected only once for the 1950s cohort (for Sichuan) and twice in the 1960s cohort (for Shandong and Guangdong), which is not sufficient evidence to suggest that the OCP has exacerbated the gender selection issue for our urban sample.

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

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	0.6158	0.0246	-0.7116
	[0.2690]	[0.4902]	[0.7616]
Shandong	2.1315	1.5863	2.3258
	[0.0165]	[0.0563]	[0.0100]
Hubei	2.7052	1.1648	-0.5310
	[0.0034]	[0.1220]	[0.7023]
Guangdong	3.5009	-0.6697	2.5473
	[0.0002]	[0.7485]	[0.0054]
Sichuan	2.8014	1.6737	-0.1276
	[0.0025]	[0.0471]	[0.5508]
Shaanxi	1.4115	-0.6905	-0.5809
	[0.0791]	[0.7551]	[0.7194]

$\Pr(Z \geq z)$ are in brackets

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, for the Pre-OCP (1940s) cohort, households in two provinces had significantly more children *if their firstborn was female*, while for the 1950s cohort, households in 5 provinces had significantly more. The suggested trend stands in contrast to the post-OCP (1960s) cohort, where households in 3 provinces had significantly more children.

Thus it may be concluded that the OCP or modernization appears to have suppressed the impact of the traditional preference for males in that the degree to which the $\frac{\text{male}}{\text{female}}$ first birth ratio is skewed has diminished. As far as subsequent children are concerned, it seemed to initially increase the propensity for an "accident" amongst families whose first child was female amongst the 1950s cohort that straddled the OCP, but by the 1960s cohort this had returned to pre-OCP levels.

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

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	-0.7510	-4.4887	-1.5708
	[0.7736]	[1.0000]	[0.9417]
Shandong	-0.3531	-6.6330	-2.9030
	[0.6380]	[1.0000]	[0.9981]
Hubei	-0.4227	-4.1360	0.8160
	[0.6637]	[1.0000]	[0.2074]
Guangdong	-3.9103	-1.3907	-2.2090
	[1.0000]	[0.9178]	[0.9862]
Sichuan	0.2578	-4.3960	-1.8830
	[0.3983]	[1.0000]	[0.9700]
Shaanxi	-2.4423	-3.2883	-1.4851
	[0.9926]	[0.9995]	[0.9310]

$\Pr(Z \geq z)$ are in brackets

incidence of positive assortative matching (decrease in negative assortative matching) with the onset of the OCP, in the sense that the range of values of a particular characteristic one is willing to entertain in a partner has narrowed around his own characteristic. It also predicts a drop in the marriage rate. However these predictions need qualification in terms of the supply and demand conditions the matchers confront in the sense that they are always predicated on the availability of partners with whom the agents wish to match.

Table 5 summarizes the distribution of types by gender and province for married individuals in the three birth cohorts. Together their spousal choices would straddle the implementation of the OCP, permitting an examination of changes in spousal choice as suggested by the model in section 2. Specifically, the 1940s cohort would be the strictly pre-policy cohort, while the 1960s cohort would be the strictly post-policy cohort, with the 1950s cohort straddling the policy period, since the age of individuals born in the 1950s would be between the ages of 20 to 29 when the OCP was implemented.

Table 5 reveals that the marginal distributions of male educational 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 positive assortative matching choices[8]. As educational attainment rose among the general populace, the possibility of increase positive assortative matching would have increased among men with higher educational attainment. This upward trend in educational attainment is however quite separate from the effect of the OCP on familial investments in children, nor can it be attributed to the economic reforms since the agents were born of parents in an era prior to 1979. Nonetheless to account for the changes across time, the relative changes amongst these three cohorts of individuals is examined to ascertain the effect of the OCP. It should also be noted that since both the OCP and Economic Reforms had differential impacts provincially, it would not be surprising to see inter-provincial differences in matching patterns, since it would largely depend upon the relative strengths of the policies (the OCP or Economic Reforms).

For each province, comparisons of the change in matching between the cohorts (1940s versus 1960s, 1940s versus 1950s and 1950s versus 1960s) allows us to examine the trends in matching. This is done through the examination of the overlap between the empirical density matrix to that expected under positive (negative) assortative matching. It is the differences between these asymptotically normally distributed scalar overlap measures between cohorts which facilitate understanding of the evolution of the matching process. In the absence of any trends towards positive assortative matching (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. However, should there be a "linear" trend towards positive assortative matching, the effect that is due to the OCP or the economic reform can be gleaned from examining the difference in the measures from two comparisons, 1940s versus 1960s, and 1950s versus 1960s, which is similar to a difference-in-difference analysis. As was noted in the introduction, the sorting attribute examined is educational attainment whose classification 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 institution of the new educational laws [9].

[8]Due to a lack of data, we are also unable to discern if the individuals were married in rural towns prior to being observed within the urban context.

[9]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 5: Marginals of Married Individuals

		Shaanxi		Jilin		Hubei		Sichuan		Guangdong		Shandong	
		Males	Females	Males	Females	Males	Females	Males	Females	Males	Females	Males	Females
40s Cohort	1	0.0625	0.1124	0.0848	0.1317	0.0718	0.1893	0.1134	0.2065	0.1397	0.2210	0.0843	0.1917
	2	0.3239	0.4438	0.3336	0.4359	0.3448	0.4136	0.3113	0.3868	0.2995	0.3683	0.3363	0.4005
	3	0.1720	0.1713	0.1668	0.1374	0.1698	0.1467	0.1491	0.1393	0.1757	0.1811	0.1730	0.1469
	4	0.1899	0.1929	0.1275	0.1577	0.1519	0.1560	0.1923	0.1803	0.1364	0.1255	0.1745	0.1603
	5	0.2517	0.0797	0.2873	0.1374	0.2617	0.0944	0.2340	0.0871	0.2488	0.1042	0.2319	0.1007
Number of Obs.		1343	1343	1427	1427	1949	1949	2663	2663	1833	1833	1341	1341
50s Cohort	1	0.0307	0.0449	0.0123	0.0235	0.0279	0.0370	0.0764	0.0823	0.0589	0.0644	0.0269	0.0441
	2	0.3120	0.2994	0.3371	0.3427	0.3086	0.3435	0.4040	0.4058	0.2430	0.2593	0.3081	0.3727
	3	0.2852	0.4559	0.3356	0.3591	0.3053	0.3843	0.1978	0.2878	0.3548	0.4344	0.2421	0.2997
	4	0.0947	0.0870	0.0907	0.1295	0.1258	0.1164	0.1157	0.1322	0.1007	0.1252	0.1519	0.1633
	5	0.2775	0.1128	0.2243	0.1452	0.2324	0.1188	0.2062	0.0919	0.2426	0.1167	0.2710	0.1202
Number of Obs.		1827	1827	2679	2679	3266	3266	4409	4409	2700	2700	2970	2970
60s Cohort	1	0.0074	0.0180	0.0117	0.0098	0.0045	0.0103	0.0174	0.0106	0.0019	0.0140	0.0027	0.0060
	2	0.1605	0.2302	0.2227	0.2411	0.1207	0.2187	0.1522	0.2256	0.1043	0.1909	0.1388	0.2358
	3	0.3538	0.4234	0.3454	0.3822	0.3252	0.3736	0.3164	0.3512	0.4535	0.4218	0.2699	0.3382
	4	0.1288	0.1668	0.1380	0.1515	0.1271	0.1536	0.1324	0.1551	0.0847	0.1462	0.1599	0.2033
	5	0.3495	0.1616	0.2822	0.2153	0.4226	0.2439	0.3816	0.2575	0.3557	0.2272	0.4287	0.2168
Number of Obs.		947	947	1630	1630	1550	1550	2070	2070	1074	1074	1845	1845

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

Overlap amongst the male and female attainment distributions (the sum of the minimums of the male and female proportions in each attainment category) is a measure of the degree to which exact positive assortative matching is *feasible* (i.e. all males of type 1 matches with female of type 1, all males of type 2 match with female of type 2, etc.). Examining this stylized measure yields an indication of the *degree* to which positive assortative matching is *feasible*. Such an overlap measure for each province and the three cohorts are respectively; Jilin was 0.8146, 0.9204 and 0.9320; Shandong was 0.8333, 0.8494 and 0.7907; Hubei was 0.8069, 0.8775 and 0.8210; Guangdong was 0.8447, 0.8740 and 0.8403; Sichuan was 0.8278, 0.8865 and 0.8705; and Shaanxi was 0.8241, 0.8149 and 0.8109. Comparing the potential for assortative pairing between the cohorts born in the 1940s and 1950s, there was a general increase (the exception being Shaanxi). On the other hand, when comparing between 1950s and 1960s, there is infact a decrease in potential with the exception of Jilin. This suggests that, should a significant increase in assortative pairing between the 1950s versus the 1960s cohorts be found, it is very possible that it is a result of the OCP, without regard to trends towards positive assortative matching.

4 Testing the Matching Hypotheses

The empirical joint densities of the data are reported in Table 6. First, note that the diagonal probabilities of the joint density provide some evidence of increased assortative pairing between the cohorts born in the 1940s and 1950s which is not surprising given the capacity for assortative matching has increased between the two cohorts (In other words, the comparison between this two cohorts is akin to examining the marital effects due to the cultural revolution between 1966 and 1969.). What is interesting is that this was also true among provinces where capacity for positive assortative pairing for the 1960s cohort decreased. Closer inspection of the marginal densities in table 3 reveals that the fall in capacity is largely due to a decreased proportion of low educational attainment individuals, while the increases in positive asortative pairing among the 1960s cohorts are among individuals with higher educational attainment realizations. Further note, as predicted by the model, the lower rates of matching among individuals of low attainment, that is individuals with elementary, and middle school education. Nonetheless, this evidence is suggestive, and will serve only as a guide in the subsequent analysis.

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

			Jilin Females					Shandong Females					Hubei Females				
			1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
40s Cohort	Males	1	0.0545	0.0255	0.0021	0.0021	0.0000	0.0545	0.0174	0.0076	0.0061	0.0000	0.0384	0.0254	0.0036	0.0047	0.0005
		2	0.0326	0.2258	0.0290	0.0333	0.0064	0.0742	0.1591	0.0409	0.0500	0.0136	0.0737	0.1755	0.0337	0.0472	0.0099
		3	0.0198	0.0651	0.0566	0.0149	0.0149	0.0212	0.0826	0.0318	0.0242	0.0136	0.0213	0.0722	0.0389	0.0270	0.0093
		4	0.0205	0.0453	0.0212	0.0283	0.0149	0.0273	0.0803	0.0265	0.0318	0.0053	0.0332	0.0592	0.0213	0.0244	0.0161
		5	0.0113	0.0679	0.0276	0.0800	0.1005	0.0167	0.0568	0.0402	0.0492	0.0689	0.0270	0.0774	0.0462	0.0535	0.0602
50s Cohort	Males	1	0.0033	0.0056	0.0033	0.0011	0.0000	0.0131	0.0104	0.0027	0.0010	0.0007	0.0076	0.0152	0.0052	0.0000	0.0000
		2	0.0115	0.2514	0.0502	0.0152	0.0115	0.0138	0.1806	0.0642	0.0316	0.0171	0.0137	0.1868	0.0844	0.0213	0.0064
		3	0.0048	0.0536	0.2157	0.0379	0.0216	0.0067	0.0854	0.1100	0.0279	0.0108	0.0079	0.0719	0.1780	0.0244	0.0213
		4	0.0007	0.0205	0.0320	0.0253	0.0130	0.0034	0.0501	0.0370	0.0427	0.0215	0.0034	0.0369	0.0491	0.0277	0.0085
		5	0.0037	0.0167	0.0558	0.0495	0.0959	0.0084	0.0491	0.0837	0.0595	0.0686	0.0040	0.0360	0.0667	0.0415	0.0820
60s Cohort	Males	1	0.0055	0.0049	0.0012	0.0000	0.0000	0.0000	0.0011	0.0011	0.0000	0.0005	0.0000	0.0019	0.0025	0.0000	0.0000
		2	0.0036	0.1518	0.0559	0.0055	0.0030	0.0027	0.0726	0.0409	0.0178	0.0086	0.0051	0.0682	0.0331	0.0057	0.0083
		3	0.0012	0.0625	0.2198	0.0273	0.0346	0.0027	0.0850	0.1318	0.0360	0.0151	0.0064	0.0860	0.1739	0.0325	0.0280
		4	0.0000	0.0079	0.0474	0.0589	0.0237	0.0011	0.0328	0.0425	0.0597	0.0194	0.0000	0.0287	0.0433	0.0325	0.0223
		5	0.0000	0.0115	0.0577	0.0589	0.1573	0.0000	0.0446	0.1216	0.0866	0.1759	0.0000	0.0350	0.1178	0.0847	0.1841
			Guangdong Females					Sichuan Females					Shaanxi Females				
			1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
40s Cohort	Males	1	0.0703	0.0463	0.0153	0.0060	0.0005	0.0561	0.0455	0.0046	0.0076	0.0008	0.0178	0.0304	0.0037	0.0074	0.0015
		2	0.0785	0.1417	0.0327	0.0365	0.0098	0.0751	0.1495	0.0319	0.0406	0.0102	0.0475	0.1930	0.0386	0.0319	0.0134
		3	0.0283	0.0703	0.0507	0.0147	0.0114	0.0209	0.0649	0.0383	0.0209	0.0046	0.0148	0.0594	0.0572	0.0327	0.0082
		4	0.0251	0.0490	0.0278	0.0256	0.0098	0.0341	0.0653	0.0277	0.0542	0.0118	0.0163	0.0742	0.0275	0.0690	0.0022
		5	0.0191	0.0605	0.0529	0.0431	0.0741	0.0247	0.0580	0.0357	0.0588	0.0584	0.0178	0.0869	0.0431	0.0512	0.0542
50s Cohort	Males	1	0.0178	0.0271	0.0134	0.0007	0.0004	0.0298	0.0356	0.0081	0.0009	0.0020	0.0144	0.0083	0.0094	0.0000	0.0000
		2	0.0271	0.1060	0.0833	0.0227	0.0048	0.0327	0.2301	0.0984	0.0363	0.0081	0.0211	0.1713	0.1048	0.0133	0.0033
		3	0.0175	0.0781	0.2052	0.0331	0.0186	0.0102	0.0541	0.0979	0.0174	0.0160	0.0067	0.0543	0.1951	0.0094	0.0161
		4	0.0015	0.0212	0.0424	0.0216	0.0141	0.0059	0.0408	0.0305	0.0300	0.0092	0.0000	0.0249	0.0394	0.0266	0.0055
		5	0.0019	0.0279	0.0870	0.0472	0.0796	0.0032	0.0469	0.0512	0.0476	0.0571	0.0050	0.0438	0.1031	0.0371	0.0870
60s Cohort	Males	1	0.0000	0.0009	0.0000	0.0000	0.0000	0.0034	0.0120	0.0010	0.0000	0.0000	0.0041	0.0031	0.0000	0.0000	0.0000
		2	0.0046	0.0557	0.0399	0.0056	0.0000	0.0034	0.0814	0.0457	0.0096	0.0116	0.0041	0.0966	0.0586	0.0021	0.0010
		3	0.0093	0.0984	0.2674	0.0427	0.0418	0.0024	0.0948	0.1680	0.0299	0.0231	0.0072	0.0771	0.2148	0.0360	0.0195
		4	0.0000	0.0167	0.0325	0.0232	0.0130	0.0005	0.0169	0.0380	0.0448	0.0323	0.0000	0.0164	0.0524	0.0432	0.0134
		5	0.0000	0.0186	0.0901	0.0761	0.1634	0.0019	0.0202	0.0997	0.0698	0.1897	0.0021	0.0298	0.1069	0.0843	0.1274

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

The corresponding indices and tests for positive and negative assortative matching using the overlap measure 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 assortative pairing by the 1960s cohort over the other two cohorts.

Increased positive assortative matching is examined via:

$$vs.\begin{array}{rl} H_0: & \Delta\mathbf{OV}_p > 0 \\ H_1: & \Delta\mathbf{OV}_p \leq 0 \end{array}$$

and decreased negative assortative matching via:

$$vs.\begin{array}{rl} H_0: & \Delta\mathbf{OV}_n < 0 \\ H_1: & \Delta\mathbf{OV}_n \geq 0 \end{array}$$

From table 7 note that in all instances, the overlap measures are all statistically significantly different from complete overlap, and that the empirical joint density is a closer match to the positive assortative joint density matrix than that generated by negative assortative matching. Next, examining the change in assortative matching between the 1940s and 1950s cohort, note that the hypothesis of increased positive assortative matching, and decreased negative assortative matching cannot be rejected for Shandong, Hubei, Guangdong, and Sichuan. For Jilin and Shaanxi, it seems there is an increase in both positive and negative assortative matching, noting that the empirical joint density is closer to positive than negative assortative matching. Considering the fact that the capacity for positive assortative matching rose between the two cohorts, 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 positive assortative matching. However, comparing the 1950s and 1960s cohorts, note the significant increase in positive assortative matching but statistically insignificant change in negative assortative matching for Jilin, Hubei, Sichuan and Shaanxi. For Shandong, there is a significant increase in positive assortative matching and a significant decrease in negative assortative matching, while Guangdong recorded a significant increase in both positive and negative assortative matching, with the overlap with positive assortative matching joint density being higher. Similar conclusions can be made when comparing the 1960s and 1940s cohorts. This then suggests that there was a significant increase in positive assortative matching in the 1960s which, negating considerations of trends and coupled with the decreased capacity for positive assortative matching among the members of the 1960s cohort, suggests that this is a consequence of the OCP (or that the OCP effects dominate that due to the Economics Reforms).

To control for the effects of trends from increased preference for positive assortative matching in urban China, a difference-in-difference analysis is performed by examining the relative change in overlap measure between two comparisons, the results of which are reported in table 7. Given that with the exception of Shaanxi, all other provinces had experienced an increase in capacity for positive assortative matching from the 1940s to the 1950s cohort, we can test whether the increase in positive assortative matching between the 1950s and 1960s cohort is significantly greater than that between the 1940s and 1950s which would control for trends towards increased preference for positive assortative matching. This comparison is reported in the first comparison of the final panel of *Difference-in-Difference*. Note that in this comparison, Shandong, Hubei and Sichuan all experienced a significantly higher rate of increase in positive assortative matching between the 1950s and 1960s cohorts. The results continue to suggests that the increase in positive assortative matching is a result of the OCP.

On the other hand, the results for Guangdong and Shaanxi suggests that positive assortative matching has slowed down significantly. One possible reason is principally due to the lower capacity for positive assortative matching for Guangdong and Shaanxi among the 1960s cohort. The result for Guangdong could be the result of the strong economic growth which would have particularly affected the 1950s and 1960s cohorts, suggesting that the fall in positive assortative matching there might be tempered by the income effect suggested in the model of section 2. For Jilin, the only province in our sample that exhibited continued increase in positive assortative matching capacity, much of the increase in positive assortative matching seem to have been exhausted by the arrival of the 1950s cohort such that the increase in positive assortative matching is lower than that observed between the 1940s and 1950s cohort. This suggests that for Shaanxi and Jilin, the dominating effect that drove the change in positive assortative matching is simply due to trends (in post cultural revolution China), or urbanization.

Table 7: Matching by Birth Cohort (All Years: 1989, 1991 to 2001)

	Province	Jilin		Shandong		Hubei		Guangdong		Sichuan		Shaanxi	
		Positive	Negative	Positive	Negative	Positive	Negative	Positive	Negative	Positive	Negative	Positive	Negative
	Overlap Statistic	0.6645	0.3149	0.5788	0.5000	0.6002	0.5026	0.6093	0.4164	0.6100	0.4431	0.5939	0.3809
40s Cohort		(0.0126)	(0.0124)	(0.0136)	(0.0138)	(0.0112)	(0.0114)	(0.0114)	(0.0115)	(0.0095)	(0.0097)	(0.0134)	(0.0132)
	Number of Obs.	1413	1413	1320	1320	1926	1926	1835	1835	2636	2636	1347	1347
	Overlap Statistic	0.7382	0.3871	0.6106	0.3682	0.6583	0.4389	0.6903	0.4097	0.6098	0.3878	0.7228	0.4296
50s Cohort		(0.0085)	(0.0094)	(0.0089)	(0.0088)	(0.0083)	(0.0087)	(0.0089)	(0.0095)	(0.0073)	(0.0073)	(0.0105)	(0.0117)
	Number of Obs.	2689	2689	2974	2974	3281	3281	2690	2690	4433	4433	1804	1804
	Overlap Statistic	0.7669	0.3722	0.7187	0.3287	0.7631	0.4389	0.7215	0.4930	0.7020	0.3924	0.7677	0.4450
60 Cohort		(0.0104)	(0.0119)	(0.0104)	(0.0109)	(0.0107)	(0.0125)	(0.0137)	(0.0152)	(0.0100)	(0.0107)	(0.0135)	(0.0159)
	Number of Obs.	1647	1647	1859	1859	1570	1570	1077	1077	2077	2077	973	973
	50s - 40s	0.0737	0.0722	0.0318	-0.1318	0.0581	-0.0637	0.0811	-0.0067	-0.0003	-0.0553	0.1289	0.0488
		(0.0152)	(0.0155)	(0.0163)	(0.0164)	(0.0139)	(0.0143)	(0.0145)	(0.0149)	(0.0120)	(0.0121)	(0.0170)	(0.0176)
		[1.0000]	[1.0000]	[0.9748]	[0.0000]	[1.0000]	[0.0000]	[1.0000]	[0.3271]	[0.4910]	[0.0000]	[1.0000]	[0.9972]
Change in	60s - 40s	0.1023	0.0573	0.1399	-0.1713	0.1629	-0.0638	0.1122	0.0767	0.0920	-0.0507	0.1738	0.0642
Assortative		(0.0163)	(0.0172)	(0.0171)	(0.0176)	(0.0155)	(0.0169)	(0.0178)	(0.0191)	(0.0138)	(0.0144)	(0.0190)	(0.0207)
Mating		[1.0000]	[0.9996]	[1.0000]	[0.0000]	[1.0000]	[0.0001]	[1.0000]	[1.0000]	[1.0000]	[0.0002]	[1.0000]	[0.9990]
	60s - 50s	0.0287	-0.0149	0.1080	-0.0395	0.1047	0.0000	0.0311	0.0834	0.0922	0.0046	0.0449	0.0154
		(0.0134)	(0.0152)	(0.0137)	(0.0140)	(0.0136)	(0.0152)	(0.0163)	(0.0179)	(0.0124)	(0.0130)	(0.0172)	(0.0197)
		[0.9836]	[0.1623]	[1.0000]	[0.0024]	[1.0000]	[0.4990]	[0.9718]	[1.0000]	[1.0000]	[0.6391]	[0.9956]	[0.7826]
	(50s-40s)-(60s-50s)	0.0450	0.0871	-0.0762	-0.0923	-0.0466	-0.0637	0.0500	-0.0901	-0.0925	-0.0599	0.0840	0.0333
		(0.0203)	(0.0217)	(0.0213)	(0.0216)	(0.0194)	(0.0209)	(0.0218)	(0.0233)	(0.0173)	(0.0178)	(0.0242)	(0.0265)
		[0.9868]	[1.0000]	[0.0002]	[0.0000]	[0.0082]	[0.0012]	[0.9890]	[0.0001]	[0.0000]	[0.0004]	[0.9997]	[0.8960]
Difference-	(60s-40s)-(60s-50s)	0.0737	0.0722	0.0318	-0.1318	0.0581	-0.0637	0.0811	-0.0067	-0.0003	-0.0553	0.1289	0.0488
in-		(0.0211)	(0.0229)	(0.0220)	(0.0225)	(0.0206)	(0.0228)	(0.0241)	(0.0262)	(0.0186)	(0.0194)	(0.0256)	(0.0286)
Difference		[0.9998]	[0.9992]	[0.9265]	[0.0000]	[0.9976]	[0.0026]	[0.9996]	[0.3994]	[0.4942]	[0.0022]	[1.0000]	[0.9558]
	(50s-40s)-(60s-40s)	0.0287	-0.0149	0.1080	-0.0395	0.1047	0.0000	0.0311	0.0834	0.0922	0.0046	0.0449	0.0154
		(0.0223)	(0.0231)	(0.0236)	(0.0240)	(0.0208)	(0.0222)	(0.0229)	(0.0242)	(0.0183)	(0.0189)	(0.0255)	(0.0272)
		[0.9009]	[0.2593]	[1.0000]	[0.0498]	[1.0000]	[0.4993]	[0.9127]	[0.9997]	[1.0000]	[0.5968]	[0.9606]	[0.7146]

Standard errors are in parenthesis and $\Pr(X \leq x)$ are in brackets.

It is of interest to compare the difference in overlap between the 1960s versus the 1940s cohort, and 1960s versus the 1950s cohort. We know that for Shandong, Guangdong, and Shaanxi there was a decline in capacity for positive assortative matching between the 1940s and 1960s cohort. While Jilin, Sichuan and Hubei saw increases in capacity for positive assortative matching, the change in capacity was far larger than the difference between the 1960s and 1950s cohort. This means that we should expect the rate of change in positive assortative matching to be greater for all the provinces in the 1940s versus 1960s comparison than the 1960s versus the 1950s comparison. This result is the second comparison of the final panel. Note that for all the provinces unanimously, the increase in positive assortative matching over 1940s and 1960s was indeed larger than that exhibited between the 1950s and 1960s.

The last comparison of the final panel reports the difference between the change between the 1950s versus 1940s, and that of the 1960s versus the 1940s. Given the change in capacity, there is little to be gleaned regarding trends, and the results accords with expectations that the latter difference in overlap is smaller since the effect is dominated by the change in capacity over the 1940s and 1950s. Note however that the difference is not significant at the 5% level for Jilin, Guangdong, and Shaanxi, further suggesting the likelihood that the change in matching behavior in the 1960s cohort is in fact dominated by the effects of the OCP. On the aggregate, the evidence substantially supports the hypothesis that the OCP altered the individual's spousal choice.

Finally, if the OCP induced an increase in positive assortative matching, for higher type individuals, this increase would reduce the likelihood of the individual choosing the "lowest" type individuals as partners, which in turn implies that there may be a stochastic dominant shift in the cumulative distribution of spouses across the cohorts[10]. For the lower educational attainment realizations, the stochastic dominance relationship is an empirical question since it involves a shrinkage of the range of prospective spousal type. The former prediction was likewise affirmed by the data[11].

[10] The authors thank Aloysius Siow for suggesting this alternative approach.

[11] A complete discussion of the implications of the model regarding stochastic dominance, and the stochastic dominance test results are available from the authors upon request.

5 Conclusion

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 One Child Policy on the Chinese populace in 1979 have been examined in terms of the urban populations in six provinces. As a guide to the analysis, a simple model of family formation was developed, which generated predictions regarding the direction of impact the introduction of a binding constraint on family size could have had on a household. It predicts an increase in the marginal benefits and consequently incidence of positive assortative matching 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 positive assortative matches with economic growth.

The matching predictions were empirically examined via bi-annual samples of urban households in six Chinese provinces taken from 1987 through 2001. An index was developed for measuring the intensity of positive assortative matching. Based upon the degree of overlap between the hypothetical joint density of matches posited by a particular matching scheme and the empirical joint density of matches that actually occurred, the index turned out to be conveniently asymptotically normally distributed, thus permitting simple comparisons between pre- and post-OCP matching patterns. 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. The model predicted that increases in income would engender reductions in the intensity of matching, whereas the OCP would engender increases in the intensity of matching. Thus, given the increase in incomes over the period of the introduction of the OCP, any increase in the intensity of matching could be attributed (at least in direction if not in magnitude) to 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[12], the intensity of positive and

[12] A bi-product of this analysis was evidence that the OCP suppressed the extent to which gender

negative assortative matching was examined. The index indicated significant increases (decreases) in the intensity of positive (negative) assortative matching and this was accompanied by a significant reduction in the incidence of marriage, all of which accorded with the predictions of the model. Thus the evidence here suggests that the One Child Policy may have precipitated an increase (decrease) in positive (negative) assortative matching.

selection of children occurred, though there was evidence that the probability of having an "accident" after a first born that was female was greater than the probability of an "accident" after a male firstborn.

A Appendix

A.1 Proof of Propositions

Proof. 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 (4),

$$\frac{\partial k'}{\partial \widetilde{n}} = \frac{q_n \widetilde{n} + q + q_n k' - q_{kn}\left(yx - \widetilde{n}k'\right)}{q_{kk}\left(yx - \widetilde{n}k' - q_k \widetilde{n} - q_n \widetilde{n}\right)} \leq 0 \tag{A-1}$$

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. Proof of Proposition 2: Differentiating (3) and (4) with respect to y respectively gives,

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

$$\frac{\partial k^*}{\partial y} = -\frac{q_k x}{\left(q_{kk}\left(yx - n^* k^*\right) - 2q_k k^*\right)} \geq 0 \tag{A-3}$$

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

Proof. Proof of Proposition 3: For the proof of point 1, differentiating $\underline{t_w^R}$ in (7) with respect to the number of children $\widetilde{n}$,

$$\frac{\partial \underline{t_w^R}}{\partial \widetilde{n}} = \frac{qk' - q_{\widetilde{n}}(yx - \widetilde{n}k')}{q_{\underline{t_w^R}}\left(yx - \widetilde{n}k'\right) + qyx_{\underline{t_w^R}} - yv_{\underline{t_w^R}}} \leq 0 \tag{A-4}$$

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 3, 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 (8) with respect to $\widetilde{n}$ as above.

$$\frac{\partial \overline{t_w^R}}{\partial \widetilde{n}} = \frac{qk'' - q_{\widetilde{n}}\left(yx - \widetilde{n}k''\right)}{q_{\overline{t_w^R}}\left(yx - \widetilde{n}k''\right) + qyx_{\overline{t_w^R}} - yv_{\overline{t_w^R}}} \geq 0 \tag{A-5}$$

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 3, 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

$$\Pr(\underline{t_w^R} \leq t_w^* \leq \overline{t_w^R}) = 1 \tag{A-6}$$

It follows that,

$$\int_{\underline{t_w^R}}^{\overline{t_w^R}} f(t_w^*|t_h)dt_w^* = \frac{1}{f(t_h)} \int_{\underline{t_w^R}}^{\overline{t_w^R}} g(t_w^*, t_h)dt_w^* = \frac{1}{f(t_h)} \left[G(\overline{t_w^R}, t_h) - G(\underline{t_w^R}, t_h)\right] = 1 \tag{A-7}$$

where $g(.)$ and $G(.)$ are respectively the joint density and joint distribution functions. The total differential of (A-7) with respect to $\widetilde{n}$ may be written as,

$$\frac{1}{f(t_h)} \left[\frac{\partial G(\overline{t_w^R}, t_h)}{\partial \overline{t_w^R}} \frac{\partial \overline{t_w^R}}{\partial \widetilde{n}} - \frac{\partial G(\underline{t_w^R}, t_h)}{\partial \underline{t_w^R}} \frac{\partial \underline{t_w^R}}{\partial \widetilde{n}}\right] d\widetilde{n} + \frac{1}{f(t_h)} \frac{\partial \left[G(\overline{t_w^R}, t_h) - G(\underline{t_w^R}, t_h)\right]}{\partial \widetilde{n}} d\widetilde{n} = 0 \tag{A-8}$$

Since

$$\frac{1}{f(t_h)} > 0, \frac{\partial G(t_w, t_h)}{\partial t_w} > 0, \frac{\partial \overline{t_w^R}}{\partial \widetilde{n}} > 0, \frac{\partial \underline{t_w^R}}{\partial \widetilde{n}} < 0 \tag{A-9}$$

It may be observed that

$$\frac{\partial \left[G(\overline{t_w^R}, t_h) - G(\underline{t_w^R}, t_h)\right]}{\partial \widetilde{n}} < 0 \tag{A-10}$$

■

Proof. Proof of Proposition 4: As in the proof of proposition 3, differentiate $\underline{t_w^R}$ and $\overline{t_w^R}$ in (7) and (8) with respect to y respectively.

$$\frac{\partial \underline{t_w^R}}{\partial y} = \frac{-qx + v(\underline{t_w^R}) + v(t_h)}{q_{\underline{t_w^R}}(yx - \underline{n}k) + qyx_{\underline{t_w^R}} - yv_{\underline{t_w^R}}} \leq 0 \tag{A-11}$$

First note that by assumption 3, 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, by assumption 5, the numerator is negative, and the inequality follows.

$$\frac{\partial \overline{t_w^R}}{\partial y} = \frac{-qx + v(\overline{t_w^R}) + v(t_h)}{q_{\overline{t_w^R}}(yx - \overline{n}k) + qyx_{\overline{t_w^R}} - yv_{\overline{t_w^R}}} \geq 0 \tag{A-12}$$

By assumption 3, and $\overline{t_w^R} \geq t_h$, the denominator is negative. By assumption 5, the numerator is negative, and the inequality follows. The rest of the arguments are similar to proposition 3. ■

A.2 A Brief Discussion about the Overlap Measure

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

$$\mathbf{V} = \sqrt{n}\begin{bmatrix} \frac{j_{1,1}-\pi_{1,1}}{\sqrt{\pi_{1,1}}} & \frac{j_{1,2}-\pi_{1,2}}{\sqrt{\pi_{1,2}}} & \cdots & \frac{j_{1,N}-\pi_{1,N}}{\sqrt{\pi_{1,N}}} \\ \frac{j_{2,1}-\pi_{2,1}}{\sqrt{\pi_{2,1}}} & \frac{j_{2,2}-\pi_{2,2}}{\sqrt{\pi_{2,2}}} & \cdots & \frac{j_{2,N}-\pi_{2,N}}{\sqrt{\pi_{2,N}}} \\ \vdots & \vdots & \vdots\vdots\vdots & \vdots \\ \frac{j_{M,1}-\pi_{M,1}}{\sqrt{\pi_{M,1}}} & \frac{j_{M,2}-\pi_{M,2}}{\sqrt{\pi_{M,2}}} & \cdots & \frac{j_{M,N}-\pi_{M,N}}{\sqrt{\pi_{M,N}}} \end{bmatrix} \tag{A-13}$$

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

$$\mathbb{v}' = \left(\sqrt{\pi_{1,1}}, ..., \sqrt{\pi_{1,N}}, ..., \sqrt{\pi_{M,1}}, ..., \sqrt{\pi_{M,N}}\right) \tag{A-14}$$

and

$$\Omega = \mathbf{I} - \mathbb{v}\mathbb{v}' \tag{A-15}$$

Then by the results in Rao (1973) pages 383 and 391, and Anderson et al. (2009), we have

$$\mathbb{V} \xrightarrow{a} N_{MN}\left(\mathbf{0}, \Omega\right) \tag{A-16}$$

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,

$$\mathbf{j} \xrightarrow{a} N_{MN}\left(\pi, \frac{1}{n}(\text{dg}(\mathbb{v}))\Omega(\text{dg}(\mathbb{v}))'\right) \tag{A-17}$$

$$\Rightarrow \mathbf{i}'\mathbf{j} \xrightarrow{a} N\left(\mathbf{i}'\pi, \frac{1}{n}\mathbf{i}'(\text{dg}(\mathbb{v}))\Omega(\text{dg}(\mathbb{v}))'\mathbf{i}\right) \tag{A-18}$$

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 vecΠ^p and vecΠ^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-18),

$$\mathbf{OV}_p := \mathbf{i}'\mathbf{j}^{\min} \xrightarrow{a} N\left(\mathbf{i}'\pi^{\min}, \frac{1}{n}\mathbf{i}'(\text{dg}(\mathbb{v}^{\min}))\Omega^{\min}(\text{dg}(\mathbb{v}^{\min}))'\mathbf{i}\right) \tag{A-19}$$

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

$$\mathbb{v}^{\min\,\prime} = \left(\sqrt{\pi_{1,1}^{\min\,\prime}}, ..., \sqrt{\pi_{1,N}^{\min\,\prime}}, \sqrt{\pi_{2,1}^{\min\,\prime}}, ..., \sqrt{\pi_{2,N}^{\min\,\prime}}, \sqrt{\pi_{3,1}^{\min\,\prime}}, ..., \sqrt{\pi_{M,N}^{\min\,\prime}}\right) \tag{A-20}$$

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

References

Anderson, Gordon, "Nonparametric Tests of Stochastic Dominance in Income Distributions," *Econometrica*, 1996, *64*, 1183–1193.

—— **and Teng Wah Leo**, "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*, 2009, *55* (s1), 607–629.

—— **and Ying Ge**, "The size distribution of Chinese cities," *Regional Science and Urban Economics*, 2005, *35* (6), 756–776.

—— **,** —— **, and Teng Wah Leo**, "Distributional Overlap: Simple, Multivariate, Parametric and Non-Parametric Tests for Alienation, Convergence and General Distributional Difference Issues," 2009. Forthcoming in *Econometric Reviews*.

Atkinson, Anthony, "On the Measurement of Inequality," *Journal of Economic Theory*, 1970, *2*, 244–263.

Becker, Gary, *Treatise on the Family* 1993. Harvard University Press.

Bourguignon, Francois and Gary Fields, "Discontinuous Losses from Poverty, Generalized Measures, and Optimal Transfers to the Poor," *Journal of Public Economics*, 1997, *63*, 155–175.

Browning, Martin, "Children and Household Economic Behavior," *Journal of Economic Literature*, 1992, *30*, 1434–1475.

Caldwell, John, *Theory of Fertility Decline* 1982. London Academic Press.

—— **, P.H Reddy, and Pat Caldwell**, "The Causes of Demographic Change in Rural South India: A Micro Approach," *Population and Development Review*, 1982, *8*, 718.

Choo, Eugene and Aloysius Siow, "Who marries whom and why," *Journal of Political Economy*, 2006, *114* (1), 175–201.

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

Dardanoni, Valentino, "On Measuring Social Mobility," *Journal of Economic Theory*, 1993, *61*, 372–394.

Davidson, Russell and Jean-Yves Duclos, "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality," *Econometrica*, 2000, *68*, 1435–1464.

Deaton, Angus, "Theoretical and Empirical Approaches to Consumer Demand Under Rationing," 1981. In *Essays in the Theory and Measurement of Consumer Behaviour: In Honour of Sir Richard Stone*, edited by Angus Deaton. Cambridge University Press.

Knuth, Donald, *Marriages Stables*, Montreal: Les Presses de l'Universite de Montreal, 1976.

Maasoumi, Esfandiar, "On Mobility," 1996. In *Handbook of Applied Economic Statistics*, edited by D. Giles and A. Ullah. Marcel Dekker.

Neary, Peter and Kevin Roberts, "The Theory of Household Behaviour Under Rationing," *European Economic Review*, 1980, *13*, 25–42.

Peng, Xizhe and Zhigang Guo, *The Changing Population of China* 2000. Blackwell Publishing.

Peters, Michael and Aloysius Siow, "Competing Pre-Marital Investments," *Journal of Political Economy*, 2002, *110*, 592–608.

Quah, Danny, "Empirics for Economic Growth and Convergence," *European Economic Review*, 1996, *40*, 1353–1375.

Rao, C. Radhakrishna, *Linear Statistical Inference and Its Applications* 1973. John Wiley & Sons.

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

Shorrocks, Anthony, "Income Mobility and the Markov Assumption," *Economic Journal*, 1976, *86*, 566–577.

——— , "The Measurement of Mobility," *Econometrica*, 1978, *46*, 1013–1024.

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

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