Marriage as a Rat Race: Noisy Pre-Marital Investments with Assortative Matching[*]

V. Bhaskar[†]
Department of Economics
University College
London WC1E 6BT, UK

Ed Hopkins[‡]
School of Economics
University of Edinburgh
Edinburgh EH8 9JT, UK

November 2013

Abstract

We study pre-marital investments in a large frictionless marriage market with non-transferable utility. We assume stochastic returns to investment, which ensures uniqueness of equilibrium. The equilibrium in our continuum agent model is the limit of the equilibria of finite models, as the number of agents tends to infinity. Equilibrium investments are utilitarian efficient when the sexes are symmetric. However, if there is any asymmetry between the sexes, investments generically exceed the Pareto-efficient level. Girls will invest more than boys if their quality shocks are less variable than those for boys, or if they are the abundant sex.

Keywords: marriage, ex ante investments, gender differences, assortative matching tournament, sex ratio.
JEL codes: C72, C78, D62, H31, J12.

1   Introduction

We study the incentives of parents to invest in their children when these investments also improve their marriage prospects. We assume a frictionless marriage market with non-transferable utility. It has usually been thought that ex ante investments suffer from the hold-up problem, since a parent will not internalize the effects of such investments in her own child upon the welfare of the child's future spouse. However, Peters and Siow (2002) argue that in large marriage markets where the quality of one's match depends on the level of investment, a parent has an incentive to invest more in order to improve the match of her offspring. They argue that the resulting outcome will be Pareto efficient. This is a remarkable result, since they assume a marriage market without transferable utility. With transferable utility, Cole et al. (2001) show that in large markets, prices can provide incentives for efficient investment decisions.[1]

In this paper, we argue that the optimism of Peters and Siow (2002) must be somewhat tempered. When the return to investment is deterministic, we show that there is a very large set of equilibria. These include efficient outcomes, but also a continuum of inefficient ones. In order to overcome this embarrassment of riches, we propose a model where the returns to investment are stochastic. This is also realistic – talent risk is an important fact of life.[2] Equilibrium in this model is unique and we are therefore able to make determinate predictions. The model also allows us to address several questions of normative importance and social relevance. Are investments efficient, in the absence of prices? What are the implications of biological or social differences between the sexes for investment decision? What are the implications of sex ratio imbalances in countries such as China – Wei and Zhang (2011) argue that marriage market competition for scarce women underlies the high savings rate in China.

Our paper is related to the literature on matching tournaments or contests. This literature typically models a situation where there is a fixed set of prizes, and agents on the one side of the market compete by making investments, with prizes being allocated to agents according the rank order of their investments (see for example, Cole et al., 1992 and Hopkins and Kornienko, 2004, 2010). If the "prizes" derive no utility from these investments, e.g. when the prize is social status, then an agent's investment exerts a negative positional externality on the other side of the market, so that there is over-investment. On the other hand, if the "prizes" derive utility from these investments – for example, if men compete for a set of women with fixed qualities, or students compete for university places – then either over-investment or under-investment is possible, depending on how much these investments

---

[1] To appreciate the degree of transferability required, note that Mailath et al. (2010) show that one needs "personalized prices", which depend upon buyer characteristics as well as seller characteristics, in order ensure efficiency of investments. Felli and Roberts (2001) show that even in large finite markets, the hold-up problem may not disappear if the specificity of investments does not vanish.

[2] Recent studies of the inter-generational transmission of wealth, in the tradition of Becker and Tomes (1979), find an inter-generational wealth correlation of 0.4 in the United States, which is far from 1.

are valued (Cole et al., 2001; Hopkins, 2012).

In our context, investments are two-sided – the investments of men are valued by women and symmetrically, the investments of women are valued by men. Men do not care directly about how their investments are valued by women, they care only about the consequent improvement in match quality that they get. Women are in a similar situation, since they care only about the improvement in the quality of men that they might get. One might expect therefore, that this could give rise to under-investment or over-investment, depending on parameter values.

Surprisingly, our model yields clear conclusions. Under very special circumstances, when the sexes are completely symmetric, with identical distributions of shocks and a balanced sex ratio, investments will be efficient – not merely in the Pareto sense, but also from a utilitarian standpoint. However, if there are any differences between the sexes, whether it be differing returns to investments, different stochastic shocks or an unequal sex ratio, investments are generically excessive, as compared to Pareto-efficient investments. Since the intuition for the overinvestment result is somewhat subtle, and quite distinct from that in one-sided tournaments with positional externalities, we defer explaining this until the model is introduced.

The rest of the paper is set out as follows. Section 2 discusses the problems that arise in a model with deterministic returns, and other related literature. Section 3 sets out the model with noisy investments, and shows that a pure strategy equilibrium exists and is unique for general quality functions. We consider, in turn, additive and multiplicative shocks. Our main finding is that investments are generically excessive, relative to Pareto efficiency. We use our model to examine the observational implications of gender differences, and show that if talent shocks are more dispersed for boys than for girls, then girls will invest more than boys. We also examine the effects of sex ratio imbalances on investments, and show that the more abundant sex invests more (otherwise, in most of the paper, we focus on the case of a balanced sex ratio). Section 4 shows that when there are no gender differences, then investments are efficient, even if there is heterogeneity within each sex. Section 5 provides a finite agent justification for the continuum model that forms the bulk of the paper. We examine a model with finitely many agents, where there is uncertainty as to whether men will be excess, or women will be in excess. If the number of agents is large enough, there is a unique equilibrium that converges to the equilibrium of the continuum model. The final section concludes. The appendix contains proofs that are omitted in the text.

2   Motivation and Related Literature

The fundamental problem is the following: investment in a child benefits the child's future spouse, but the benefit to the spouse is not considered by the child's parents. There is

therefore a gap between the *privately optimal* investment in a child, which we denote $\bar{x}$, and the *socially optimal* level which is naturally greater. In the absence of prices, it is not clear that there are incentives for efficient investment. Peters and Siow (2002) (PS, henceforth) argue that nonetheless, equilibrium investments are socially efficient.

Let us consider the PS model, of investment with deterministic returns, but simplify by assuming that families are identical rather than differing in wealth. Assume a unit measure of boys, all of whom are ex ante identical, and an equal measure of girls, who are similarly ex ante identical. Assume that the quality of the child, as assessed by a partner in the marriage market, equals the level of parental investment, x. Suppose a boy is matched with a girl. The utility of the boy's parents is increasing in the investment level of the girl x_G, but they have to bear the cost of investment x_B in their son. Thus, if they choose x_B purely to maximize their utility, they would choose only the privately optimal investment $\bar{x}_B$. But the resulting investment levels $(\bar{x}_B, \bar{x}_G)$ are inefficient. Both families would be better off if they each raised their investments.

Suppose that the family of the boy believes that if they choose investment level x_B, the quality of their partner is given by a smooth, strictly increasing function, $\phi(x_B)$. They would choose investments to maximize their overall payoff, given the return function ϕ. Suppose also that the family of a girl believe that the match quality of their girl is also an increasing function of their own investment level x_G. Assume further that this return function equals $\phi^{-1}(x_G)$, the inverse of that for the boys.[3] Consider a profile of investments (x_B^{**}, x_G^{**}) such that x_B^{**} maximizes the payoffs of the boy's family given returns $\phi(x_B)$, x_G^{**} maximizes the payoffs of the girl's family given returns $\phi^{-1}(x_G)$, and $x_B^{**} = \phi^{-1}(x_G^{**})$, i.e. these expectations are actually realized. As PS argue, the profile (x_B^{**}, x_G^{**}) must be such that the indifference curves on the two sides of the market are mutually tangent, so that the investment profile must be Pareto-efficient.

A problem with this approach is that, while the expectations $\phi(x_B)$ are realized in equilibrium, they cannot be realized if the family of a boy chooses $x_B \neq x_B^{**}$. In particular, if a boy deviates and chooses $x_B < x_B^{**}$, the match $\phi(x_B) < x_G^{**}$ is not *feasible* since every girl in the market has quality x_G^{**}. In other words, while expectations are "rational" at the equilibrium, they are not so for any investment level that is not chosen in equilibrium.

We shall be explicit in this paper about the matching process that follows a profile of investments. Specifically, we will require the matching to be feasible, to be stable (in the sense of Gale and Shapley, 1962) and to be measure preserving. Despite these restrictions on matching off the equilibrium path, in terms of equilibria there is an embarrassment of riches, and a large set of equilibria. Let (x_B, x_G) be a pair of investments that are weakly greater

---

[3]If the return function for the girls is the inverse of that for the boys, then whenever a boy raises his investment from x_B to x_B' and finds that his partner quality rises from x_G to x_G', it must also be the case that when a girl raises her investment from x_G to x_G', her partner quality increases from x_B to x_B'. This property ensures efficiency of investments.

than the individually optimal investments $(\bar{x}_B, \bar{x}_G)$, and where the payoff of gender i from being matched with a partner with investment level x_j is weakly greater than the payoff from choosing the individually optimal investment level $\bar{x}_i$ and being unmatched. Any such pair can be supported as an equilibrium, by specifying that any agent who deviates to a lower investment level will be left unmatched. In the PS equilibrium, an individual making an efficient investment can expect to be matched with someone who invests similarly, there is in effect symmetry. However, here, if the parent of a boy deviates upwards, and chooses a higher level of investment, his son cannot realize a higher match quality, since all the girls are choosing x_G. We therefore have a "folk theorem"– any pair of investments satisfying the above conditions is an equilibrium. Efficient investments are an equilibrium, but so are inefficient ones.[4]

Turning now to the original PS environment where families differ in wealth, and thereby in their marginal costs of investment, we still find a continuum of inefficient equilibria. The equilibria we have constructed in the homogenous case are strict equilibria – any individual who invests differently does strictly worse. If we perturb wealth levels slightly, and wealth affects payoffs continuously, then these equilibria will continue to be strict. The only thing that is required is that the distribution of wealth is not too dispersed, so that there is a common level of investment that is not so low that it is below the richest family's privately optimal investment and not so high that the poorest family would prefer to deviate downwards and be unmatched. None of these equilibria are efficient. In fact, for all of them, a measure zero of agents make an efficient investment. If $\hat{x}$ is relatively low then all agents underinvest. If $\hat{x}$ is higher, some agents underinvest and some overinvest. One can also construct inefficient equilibria, with a heterogeneity of investment levels, even when wealth is more widely dispersed.[5]

Our paper shows that these problems can be resolved if we augment the model by adding an idiosyncratic element of match quality. This ensures that there is always an non-degenerate distribution of qualities on both sides of the market, thereby providing incentives to invest. Furthermore, equilibrium is unique, under some regularity conditions. Before proceeding to the model, we review some of the related literature not already discussed in the introduction.

Peters (2007) investigates two sided investments with finitely many agents. He assumes that individuals on the long side of the market may drop out of the market with some small probability,[6] and solves for an equilibrium in mixed strategies. Peters (2009) assumes that

---

[4]Equilibrium multiplicity also holds if the deviator is left unmatched with probability one-half, rather than for sure. This matching rule can be justified in a large finite model (see Section 5).

[5]For example, we may divide families into two groups, rich and poor, each of which groups has a common level of investment. The matching rule matches those families with sons who choose investment $\hat{x}_L$ to those families with daughters with the same investment, and matches those who choose $\hat{x}_H$ to daughters with the same investment.

[6]This works in a similar way to the uncertainty over participant numbers assumed in Section 5 of the current paper.

there is ex ante heterogeneity, rather than the noisy returns assumed in this paper. In both papers equilibrium investments are bounded away from the efficient level even as the number of participants goes to infinity.

Hoppe et al. (2009) analyze a signaling model of matching, where an agent cares about his or her match partner's underlying characteristic which is private information. Again there is ex ante heterogeneity rather than stochastic returns. Since investments are not directly valued, they are inherently wasteful, although they may improve allocative efficiency in the matching process. They also obtain a interesting comparative statics results, on gender differences and on the numbers of participants, that we relate to our own results. Hopkins (2012) finds that with one-sided investments, the level of investment can be inefficiently low.

Our approach differs from most of the theoretical literature on investments, which usually assumes ex ante heterogeneity or incomplete information (Hoppe et al., 2009; Peters, 2009; Hopkins, 2012). Agents are assumed to differ ex ante in terms of quality or wealth, giving rise to heterogeneity in investments. Instead, we build on the classic work of Lazear and Rosen (1981), who analyze a tournament where a finite number of identical workers compete for exogenously given prizes. By assuming that worker's output is noisy, they ensure that the optimization problem faced by the worker is smooth. Models with noisy returns face the difficulty that the optimization problem faced by the agent is not necessarily concave. By assuming that the noise is large enough, Lazear and Rosen ensure that the effort level that satisfies the first order condition is also globally optimal. With two-sided investments, our problems is somewhat more delicate, since it is the *relative dispersion* that matters. Increasing dispersion on one side, say men, increases the payoff to large deviations on the women's side. One of our contributions is to show how this analysis may be extended to two sided matching and investments, and to a situation where the number of agents is large.

Gall et al. (2009) also employ a model with noisy returns to investments, and examine investments, matching and affirmative action in a non-transferable utility setting. They consider a situation where efficiency requires negative assortative matching, but where stable matchings are positively assortative, providing a possible rationale for affirmative action. They allow for investments with stochastic returns and focus on the tradeoff between the positive role of affirmative action on match efficiency versus its possible negative effect upon investment incentives.

In empirical work on matching, typically the value of any match is assumed to have an idiosyncratic random element (Dagsvik, 2000; Choo and Siow, 2006). Our key finding here is that the structure of shocks not only affects the matching process, but is also critical for investment incentives.

The transferable utility model is an alternative paradigm that can be used to explain several empirical phenomena. Chiappori et al. (2009) use this model to explain the increasing education of women, while Iyigun and Walsh (2007) study the distributional consequences of institutional and gender differences for investments. While transferable utility models are

very useful, there are many reasons for the limited transferability of utility in the marriage context, including the inability to commit to future transfers at the time of marriage.[7] Finally, there is also work that considers investment incentives in the presence of search frictions – Acemoglu and Shimer (1999) study one-sided investments under transferable utility, while Burdett and Coles (2001) analyze a non-transferable utility model with two-sided investments.

3   A Matching Tournament with Noisy Investments

We now set out a model where the return to investments are stochastic. To simplify the analysis, we assume that there is no ex ante heterogeneity but there are stochastic returns so that there is heterogeneity ex post. Thus, all families are ex ante identical, save for the fact that some have boys and others have girls. We assume a balanced sex ratio and a continuum population, so that there are equal measures of boys and girls. Our main result in this section is to show that under these assumptions, and under certain technical conditions, a pure strategy equilibrium exists and is unique.

Assume that a parent i derives a direct private benefit $b_B(x)$ from an investment of x in the quality of a boy, and $b_G(x)$ in the quality of a girl and incurs a cost $\tilde{c}_B(x)$ and $\tilde{c}_G(x)$ respectively. Define the net cost of investment in a boy as $c_B(x) = \tilde{c}_B(x) - b_B(x)$, and assume that this is strictly convex and eventually increasing. Net costs for girls are similarly defined. The quality of a boy, q, depends upon the level of parental investment, x, and the realization of a random shock, ε, that is distributed with a density function $f(\varepsilon)$ and a cumulative distribution function $F(\varepsilon)$ on $[\underline{\varepsilon}, \bar{\varepsilon}]$. Let η denote the realization of the idiosyncratic quality component for a girl; this is distributed with a density function g and cdf G on $[\underline{\eta}, \bar{\eta}]$. We shall assume that the quality of a boy equals $q(x, \varepsilon)$, where q is continuous and strictly increasing in both arguments. Similarly, a girl's quality on the marriage market is given by $q(x, \eta)$. It is easy to see that our analysis also applies when partners valuation of quality is an increasing concave function of q. We assume that the "quality" from being unmatched, $\bar{u}$, is strictly less than the payoff from the lowest possible quality match. Note that matched or unmatched, the individual still pays the investment cost

Parents are altruistic and internalize the effects of their decisions on the utility of their own child, but not on the utility of their child's partner. Thus if a girl with parental investment x_G and shock η is matched with a boy whose parent has invested x_B, and who has shock realization ε, her payoff and that of her parents equals

$$U_G(x_G, x_B) = q(x_B, \varepsilon) + b_G(x_G) - \tilde{c}_G(x_G) = q(x_B, \varepsilon) - c_G(x_G). \tag{1}$$

---

[7]Gall et al. (2009) provide reasons why in many matching situations, including business partnerships, transfers between potential partners may not be fully flexible.

Similarly for a boy of type (x_B, ε) who is matched with a girl of type (x_G, η), his utility would be

$$U_B(x_B, x_G) = q(x_G, \eta) + b_B(x_B) - \tilde{c}_B(x_B) = q(x_G, \eta) - c_B(x_B). \tag{2}$$

The payoff for an unmatched individual is $\bar{u} - c_i(x_i)$ where $i \in \{G, B\}$.

Let $\bar{x}_B$ and $\bar{x}_G$ denote the *individually optimal* or Nash investments for boys and girls respectively. That is, for boys it is the investment that minimizes $c_B(x) = \tilde{c}_B(x) - b_B(x)$, or equivalently the investment such that $c'_B(\bar{x}_B) = 0$. The privately optimal investment for girls $\bar{x}_G$ is defined similarly.

Such investments are not Pareto efficient. Consider a social planner who chooses (x_B, x_G) to maximize

$$W(x_B, x_G) = \lambda \left[\int q(x_B, \varepsilon) f(\varepsilon) \, d\varepsilon - c_G(x_G) \right] + (1 - \lambda) \left[\int q(x_G, \eta) g(\eta) \, d\eta - c_B(x_B) \right], \tag{3}$$

for some $\lambda \in (0, 1)$, where λ is the relative weight placed on the welfare of girls. Differentiating with respect to x_B and x_G, setting to zero and rearranging, we obtain the first order conditions for Pareto efficiency,

$$\frac{c'_B(x_B)}{\int q_x(x_B, \varepsilon) f(\varepsilon) \, d\varepsilon} = \frac{\lambda}{1 - \lambda}. \tag{4}$$

$$\frac{c'_G(x_G)}{\int q_x(x_G, \eta) g(\eta) \, d\eta} = \frac{1 - \lambda}{\lambda}. \tag{5}$$

Rearranging the first order condition for welfare maximization, we obtain

$$c'_B(x_B) \times c'_G(x_G) = \int q_x(x_B, \varepsilon) f(\varepsilon) \, d\varepsilon \times \int q_x(x_G, \eta) g(\eta) \, d\eta. \tag{6}$$

In other words, any profile of Pareto-efficient investments satisfies this condition, irrespective of the value of λ. Pareto efficient investments always exceed the privately optimal level as the under the privately optimal investments, we have $c'_B(\bar{x}_B) = c'_G(\bar{x}_G) = 0$, whereas the right hand side of equation (6) is strictly positive. Of particular interest is the case where λ, the weight place on girls' welfare, is equal to their proportion in the population, one-half. Let x_B^{**}, x_G^{**} denote the efficient investments in this case. We shall call these the *utilitarian efficient* investments. These are the investments that parents would like the social planner to choose in the "original position", before the gender of their child is realized.

The Pareto efficiency condition (6) does not determine a unique investment level, but a continuous curve in (x_B, x_G) space. If a profile of investments (x_B, x_G) is such that the product of the marginal costs is strictly greater than the right hand side of equation (6),

and the marginal costs are positive,[8] then such a point lies above the Pareto efficiency curve. We say that have *overinvestment relative to Pareto-efficiency*, since it is possible to achieve Pareto efficiency by reducing either investment level.[9] Similarly, if the product of marginal costs is strictly less than the right hand side of the equation, we have *underinvestment relative to Pareto-efficiency.*

On the other hand, utilitarian efficiency determines a unique point in (x_B, x_G) space. Thus we may consider the investments of one side alone, say boys, and speak of underinvestment relative to the utilitarian level, without reference to the investments by girls. Thus, by the utilitarian criterion, one may have underinvestment by boys and overinvestment by girls. By the Pareto criterion, one can only have overinvestment or underinvestment, where this statement applies to the profile of investments, (x_B, x_G).

We shall focus upon pure strategy Nash equilibria where every parent on a given side of the marriage market chooses the same level of investment. Such an equilibrium will be called *quasi-symmetric* and consists of a pair (x_B^*, x_G^*). We require the matching to be stable and measure preserving. Given our specification of preferences, whereby all boys uniformly prefer girls of higher quality, and vice-versa, a stable measure preserving matching is essentially unique, and must be assortative. Since q is strictly increasing in the idiosyncratic shock, and since all agents on the same side of the market choose the same investment level, in equilibrium, there must be matching according to the idiosyncratic shocks alone. For a boy who has shock realization ε, let $\phi(\varepsilon)$ denote the value of η of his match. This satisfies

$$F(\varepsilon) = G(\phi(\varepsilon)), \tag{7}$$

or $\phi(\varepsilon) = G^{-1}(F(\varepsilon))$. That is, if boy is of rank z in the boy's distribution, he is matched with a girl of the same rank z in the girl's distribution. The non-degenerate distribution of qualities on both sides of the marriage market provides incentives of investment above the privately optimal level. If the parent of a boy invests a little more than x_B^*, he increases the boy's rank for any realization of ε. By doing so, he obtains a girl of higher rank. However, he is concerned not with the girl's rank but her quality.

One delicate issue concerns large deviations from the equilibrium, where the quality realization is outside the support of the equilibrium distribution of qualities. For example, if a boy deviates upwards and his quality exceeds $q(x_B^*, \bar{\varepsilon})$, stability implies that he will be matched with the best quality girl, of quality $q(x_G^*, \bar{\eta})$. If he deviates downwards and his quality is below $q(x_B^*, \underline{\varepsilon})$, then stability implies that he could be left unmatched (with

---

[8]No individual will choose investment below the individually optimal level, since higher investments can never reduce match quality. Thus we may restrict attention to investment levels such that marginal costs are non-negative.

[9]This follows from the concavity of $q(.)$ in x and the strict convexity of $c_B(.)$ and $c_G(.)$. If we reduce x_B then the left hand side of (6) decreases due to the strict convexity of the cost function, while the first term on the right-hand side increases, since $q_{xx}(x_B, \varepsilon) \leq 0$ by Assumption A2 in the Appendix. Since $c_B'(\bar{x}_B) = 0$, there exists a reduction in x_B such that (6) holds.

payoff $\bar{u}$) or matched with the lowest quality girl. We shall assume that both these outcomes have equal probability. Since we assume that being single has a low payoff, this deters large downward deviations. These assumptions are consistent with the requirement that the matching be stable and measure preserving. Moreover, the matching assumption can be justified as the limit of a model with a finite number of agents as the number of agents tends to infinity, as we show in Section 5, where we consider a model where the exact numbers of men and women is random: with probability one-half that there are slightly more men than women, and probability one-half the reverse is the case. Then a boy with the lowest quality is unmatched with probability one-half.

In the Appendix we show that the first order condition for the equilibrium investment in boys can be written as

$$\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} q_\eta(x_G^*, \phi(\varepsilon)) \frac{f(\varepsilon)}{g(\phi(\varepsilon))} \frac{q_x(x_B^*, \varepsilon)}{q_\varepsilon(x_B^*, \varepsilon)} f(\varepsilon) \, d\varepsilon = c_B'(x_B^*). \tag{8}$$

The intuition for the first order condition is that it balances the marginal cost c' of extra investment on the right hand side with its marginal benefit on the left hand side. The latter principally is determined by the possibility of an improved match from increased investment. Specifically, an increase in ε, a boy's shock, would improve his match, given the matching relation (7) at rate $\phi' = f/g$. Similarly, the first order condition for investment in girls is given by

$$\int_{\underline{\eta}}^{\bar{\eta}} q_\varepsilon(x_B^*, \phi^{-1}(\eta)) \frac{g(\eta)}{f(\phi^{-1}(\eta))} \frac{q_x(x_G^*, \eta)}{q_\eta(x_G^*, \eta)} g(\eta) \, d\eta = c_G'(x_G^*). \tag{9}$$

The match value of remaining single, $\bar{u}$, does not affect the first order condition for equilibrium investments, and thus does not affect the equilibrium level. This is so since in equilibrium, an individual is always matched with probability one when the sex ratio is balanced. Furthermore, we assume that the density function of shocks is zero at its lower bound, ensuring that $\bar{u}$ does not affect the derivative of the payoff function at equilibrium. However, the value of $\bar{u}$ does affect the payoff from large downward deviations, and we assume a "misery effect", i.e. that $\bar{u}$ is sufficiently small relative to the payoff from being matched.[10] This ensures that large downward deviations are not profitable.

We also have to ensure that large upward deviations are not profitable – this is not immediate, since the optimization problem faced by agents is not necessarily quasi-concave, just as in Lazear and Rosen (1981).[11] We assume throughout that the investment cost

---

[10]The total payoff to a boy from being single equals $\bar{u} - c_B(x_B)$. This incorporates the private benefit from investment, since $c_B(x_B) = \tilde{c}_B(x_B) - b_B(x_B)$ measures the cost net of the private benefit.

[11]In a one-sided tournament it suffices to assume that shocks affecting that side are sufficiently dispersed and that the cost function is convex. In a two-sided tournament, it is the *relative* dispersion on the two sides that matters, and one cannot make shocks relatively dispersed for both sides. See the discussion of the log-normal example in Subsection 3.1.3.

functions $c_B(x)$ and $c_G(x)$ are twice differentiable and strictly convex. In addition, we invoke, the following assumption:

Assumption A1: One of the following three conditions is satisfied:

a) F and G are distributions of the same type, i.e. $G(x) = F(ax + b)$.

b) $f(\varepsilon)$ and $g(\eta)$ are weakly increasing.

c) The second derivatives of the cost functions, $c_B''(x)$ and $c_G''(x)$ are non-decreasing, and satisfy $c_B''(x_B^*) > \gamma$ and $c_G''(x_G^*) > \gamma$, where γ is a constant that depends upon F and G.

A1 ensures that the overall payoff is strictly concave if an agent deviates upwards – the benefit function is concave if a) or b) hold, while c) ensures that the cost function is convex enough to overcome any non-concavity in the benefit function. We also require more standard technical assumptions, Assumption A2, on the concavity of the quality function, and upon the distributions of shocks. These are set out in the Appendix, and ensure existence of a quasi-symmetric equilibrium in pure strategies. For uniqueness of quasi-symmetric equilibrium, we invoke the following assumption:

Assumption A3: Quality is *generalized additive/multiplicative*:

$$q(x_B, \varepsilon) = \theta_B(x_B + \varepsilon) + (1 - \theta_B)\gamma(x_B)\varepsilon,$$

$$q(x_G, \eta) = \theta_G(x_G + \eta) + (1 - \theta_G)\gamma(x_G)\eta,$$

where $\theta_i \in [0, 1]$, $i \in \{G, B\}$. The function $\gamma(\cdot)$ is strictly increasing, twice differentiable and strictly concave, with $\gamma(0) = 0$.

Theorem 1 *Under Assumptions A1 and A2 and if the value to not being matched $(\bar{u})$ is sufficiently low, there exists a quasi-symmetric Nash equilibrium of the matching tournament. Under Assumption A3, the quasi-symmetric equilibrium is unique.*

Now that existence has been established, we can turn to some important questions about the qualitative nature of equilibrium behavior. Are investments Pareto-efficient? What are the factors that lead one sex to invest more than the other? In order to shed more light on these questions, we consider, in turn, different specifications of the quality function $q(\cdot)$.

3.1 Additive Shocks

We first analyze the case where $q(x, \varepsilon) = x + \varepsilon$. One interpretation is that investment or bequest are in the form of financial assets or real estate, while the shocks are to (permanent)
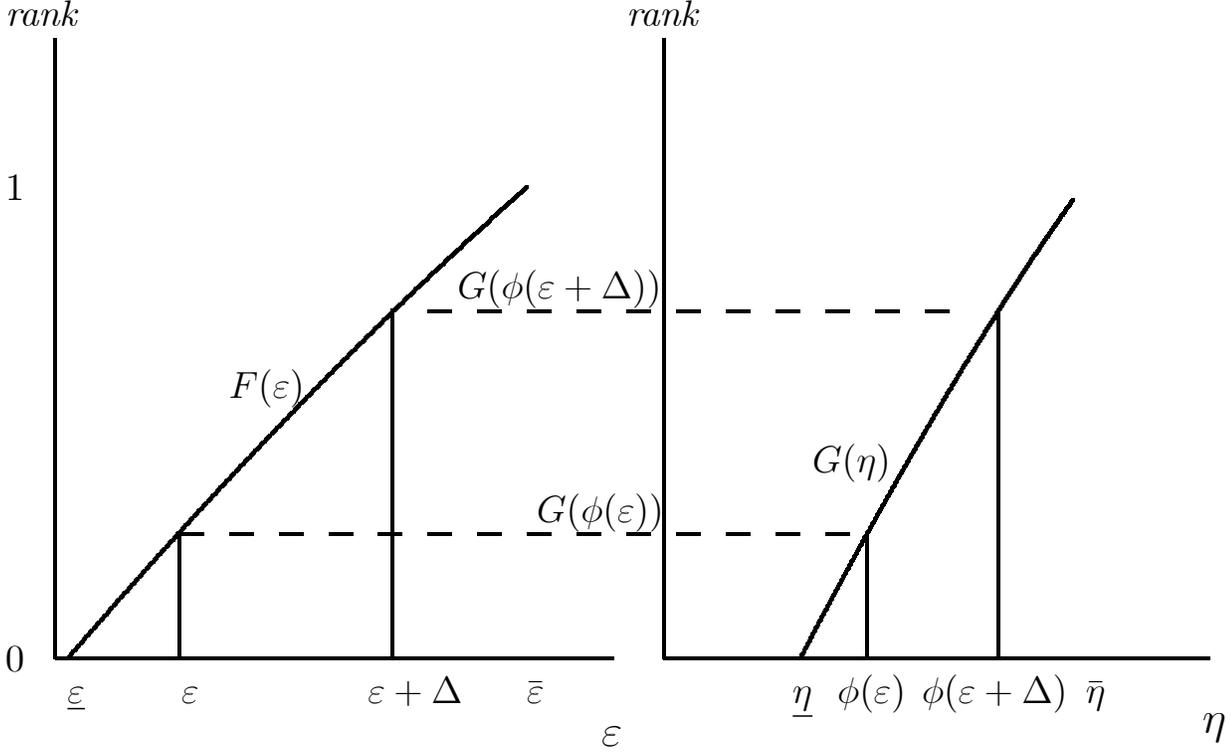
Figure 1: If a boy with shock realization ε increases investment by an amount Δ, he would overtake other boys with shock realizations between ε and $\varepsilon + \Delta$. The boy's match would improve from the girl with shock value $\phi(\varepsilon)$ to one at $\phi(\varepsilon + \Delta)$.

labor income of the child. The interpretation is that total household income is like a public good (as in Peters-Siow), which both partners share.

Consider a quasi-symmetric equilibrium where all boys invest x_B^* and all girls invest x_G^*. A boy with shock realization ε and of rank z in the distribution $F(.)$ will be matched with a girl of shock $\phi(\varepsilon)$ with same rank z in $G(.)$. Suppose that the parent of a boy invests a little more, $x_B^* + \Delta$, as in Figure 1. If his realized shock is ε, the improvement in the ranking of boys is approximately equal to $f(\varepsilon)\Delta$. The improvement in the quality of the matched girl, $\tilde{\Delta}$ must be such that $g(\eta)\tilde{\Delta} \approx f(\varepsilon)\Delta$, i.e. the improvement in the rank of his match must equal the improvement in his own rank. Thus the marginal return to investment in terms of match quality equals $\frac{f(\varepsilon)}{g(\phi(\varepsilon))}$ at any value of ε.

Integrating over all possible values of ε gives the first order condition for optimal investments in boys,

$$\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \frac{f(\varepsilon)}{g(\phi(\varepsilon))} f(\varepsilon) \, d\varepsilon = c_B'(x_B^*). \tag{10}$$

11

Similarly, the first order condition for investment in girls is

$$\int_{\underline{\eta}}^{\bar{\eta}} \frac{g(\eta)}{f(\phi^{-1}(\eta))} g(\eta) \, d\eta = c_G'(x_G^*). \tag{11}$$

The left hand side of the above equations (the marginal benefit) is constant, while the right hand side is strictly increasing in x_i, due to the convexity of the cost function. Thus, there is a unique solution to the first order conditions.

3.1.1 Efficiency

We now use the first order conditions (10) and (11) to examine the efficiency of investments. Under additive shocks, the marginal benefit to a girl from a boy's investment is one, regardless of the realization of the shock. Thus the Pareto efficiency conditions (4) and (5) reduce to $c_B'(x_B) = \frac{\lambda}{1-\lambda}$, and $c_G'(x_G) = \frac{1-\lambda}{\lambda}$. This implies that in any Pareto efficient allocation, $c_B'(x_B) \times c_G'(x_G) = 1$. The condition for utilitarian efficiency (where equal weight is placed on the welfare of boys and girls) is $c_B'(x_B) = c_G'(x_G) = 1$.

Suppose $F = G$, i.e. the distribution of shocks is the same. Thus, $\frac{f(\varepsilon)}{g(\phi(\varepsilon))} = 1$ for all values of ε, so that $c_B'(x_B^*) = c_G'(x_G^*) = 1$. Investments are utilitarian efficient even if the investment cost functions are different for the two sexes. As we shall see later, this is an example of a more general result – if there are no gender differences whatsoever, this ensures utilitarian efficiency. In general, if there are any differences between the sexes, $\frac{f(\varepsilon)}{g(\phi(\varepsilon))}$ will differ from 1, and so one cannot expect utilitarian efficiency. The following theorem sharpens this conclusion.

Theorem 2 *When noise is additive, in a quasi-symmetric equilibrium investments are generically excessive relative to Pareto efficiency.*

Proof. It is useful to make the following change in variables in the first order condition for the girls, (11). Since $\eta = \phi(\varepsilon)$,

$$d\eta = \phi'(\varepsilon) \, d\varepsilon = \frac{f(\varepsilon)}{g(\phi(\varepsilon))} \, d\varepsilon.$$

Thus the first order condition for girls is rewritten as

$$\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} g(\phi(\varepsilon)) \, d\varepsilon = c'(x_G^*).$$

12

Consider the product of the two first order conditions:

$$c'(x_B^*) \times c'(x_G^*) = \left(\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \frac{f(\varepsilon)}{g(\phi(\varepsilon))} f(\varepsilon)\, d\varepsilon \right) \left(\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} g(\phi(\varepsilon))\, d\varepsilon \right).$$

By the Cauchy-Schwarz inequality,

$$\left(\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \frac{f(\varepsilon)}{g(\phi(\varepsilon))} f(\varepsilon)\, d\varepsilon \right) \left(\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} g(\phi(\varepsilon))\, d\varepsilon \right) \geq \left[\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \left(\frac{f(\varepsilon)}{(g(\phi(\varepsilon)))^{1/2}} \right) (g(\phi(\varepsilon)))^{1/2}\, d\varepsilon \right]^2 = 1.$$

The Cauchy-Schwarz states that the inequality is strict if the two terms are linearly independent. Thus $c'(x_B^*) \times c'(x_G^*) > 1$ if $\frac{f(\varepsilon)}{\sqrt{g(\phi(\varepsilon))}}$ and $\sqrt{g(\phi(\varepsilon))}$ are linearly independent functions of ε. Since Pareto efficiency requires $c'(x_B^*) \times c'(x_G^*) = 1$, we have overinvestment generically if the distributions f and g differ. $\blacksquare$

Example 1 Let us assume that $F(\varepsilon) = \varepsilon$ on $[0,1]$, i.e. ε is uniformly distributed.[12] Assume that $G(\eta) = \eta^n$ on $[0,1]$. $F(\varepsilon) = G(\phi(\varepsilon))$ implies $\phi(\varepsilon) = \varepsilon^{\frac{1}{n}}$, $g(\phi(\varepsilon)) = n\varepsilon^{\frac{n-1}{n}}$. The equilibrium conditions are:

$$c'(x_B^*) = \int_0^1 \frac{f(\varepsilon)}{g(\phi(\varepsilon))} f(\varepsilon)\, d\varepsilon = \frac{1}{n} \int_0^1 \varepsilon^{\frac{1-n}{n}}\, d\varepsilon = 1,$$

$$c'(x_G^*) = \int_0^1 \frac{g(\eta)}{f(\phi^{-1}(\eta))} g(\eta)\, d\eta = n^2 \int_0^1 \eta^{2n-2}\, d\eta = \frac{n^2}{2n-1}.$$

The product of the marginal costs equals $\frac{n^2}{2n-1} > 1$ for $n > 1/2$ and $n \neq 1$. Efficiency requires that the product equals 1, which it only does for $n = 1$, i.e. when $f = g$.

The example provides additional intuition for the inefficiency result. Let $n = 2$, so that the density function for women, $g(\eta) = 2\eta$ on $[0,1]$. The incentive for investment for a man at any value of ε depends upon the ratio of the densities, $\frac{f(\varepsilon)}{g(\phi(\varepsilon))}$. This ratio exceeds one for low values of ε, but is less than one for high values of ε. Conversely, for women, the incentive to invest depends upon the inverse of this ratio, $\frac{g(\eta)}{f(\phi^{-1}(\eta))}$, which is low at low values of η but high at high values of η. In other words, the ratio of the densities plays opposite roles for the two sexes. However, the weights with which these ratios are aggregated differs between the sexes; high values of η are given relatively large weight in the case of women, since $g(\eta)$ is large in this case, while they are given relatively less weight in the case of men.

---

[12] The uniform distribution violates our Assumption A2, since $f(\underline{\varepsilon}) > 0$. This implies that the left hand derivative of a boy's payoffs with respect to investments is strictly greater than the right hand derivative – see footnote 19 in the Appendix. We focus on the equilibrium where the right hand derivative equals zero, i.e. the one with smallest investments. Any other equilibrium will have strictly larger investments.

3.1.2 Gender Differences

We use our model to examine a contentious issue – what are the implications of gender differences. Let us assume that the shocks to quality constitute talent shocks, and that quality is additive in talent and investment (our results in this section also apply when quality is multiplicative). One issue, that excites great controversy, is whether the distributions differ for men and women. For example, Baron-Cohen (2004) and Pinker (2003) argue that there are intrinsic gender differences that are rooted in biology, while Fine (2010) has attacked this view. In any case, in a study based on test scores of 15-year-olds from 41 OECD countries, Machin and Pekkarinen (2008) find that boys show greater variance than girls in both reading and mathematics test scores in most countries. We now explore the implications of differences in variability between the sexes.

Suppose that the shocks are more variable for men than for women. One way to formalize the idea of a distribution being more variable than another is the dispersive order. A distribution F is larger in the dispersive order than a distribution F, or $F \geq_d G$ if

$$g(G^{-1}(z)) \geq f(F^{-1}(z)) \text{ for all } z \in (0,1), \tag{12}$$

with the inequality being strict on a set of z values with positive measure (see Shaked and Shanthikumar (2007, pp148-9)). For example, if F and G are both uniform distributions, where the support of F is a longer interval than that of G, then $F \geq_d G$. A second example is two normal distributions – the one with the higher variance is larger in the dispersive order. These measures of dispersive order do not rely upon an equality of means (see Hopkins and Kornienko (2010) for further examples and discussion).

Suppose that $F \geq_d G$, so that $\frac{f(\varepsilon)}{g(\phi(\varepsilon))} \leq 1$ for all values of ε, and is strictly less on a set values of ε of positive measure. Thus the integral on the left-hand side of equation (10) is strictly less than one, and the integral on the left-hand side of equation (11) is strictly greater than one. As utilitarian efficiency requires $c'_B(x_B) = 1 = c'_G(x_G)$, boys under-invest and girls over- invest, relative to the utilitarian level. We therefore have the following proposition:

Proposition 1 *With additive shocks, if the distribution of shocks for boys is more dispersed than that for girls, that is $F \geq_d G$, then there is under-investment in boys and there is over-investment in girls relative to the utilitarian efficient level.*

The intuition for this result is as follows. If the distribution of shocks for boys is relatively dispersed, then at any realization of ε, a Δ increment in his investment only results in a small improvement in his rank, and thus of his partner. Since the quality of the girls is relatively compressed, this improvement in the rank of his partner only translates to a small increase in quality. In contrast, for a girl, an increment in investment results in a large improvement in her rank, and this improvement in the rank of her partner also translates to a large increase

in quality, given the higher dispersion in boy qualities. Therefore in equilibrium investment in girls is greater than in boys.

Empirically, the average performance of girls in school is often better than that of boys, especially in developed countries, where there is less discrimination. Our model provides a possible partial explanation for this – the incentives to invest for girls are greater, from marriage market matching considerations. While differences in the dispersion of shocks have strong implications, differences in the mean play no role in investment incentives. To see this, suppose that f is a translation of g, i.e. $f(\varepsilon) = g(\varepsilon + k)$ for some k. This implies that $\phi(\varepsilon) = \varepsilon + k$, so that $\frac{f(\varepsilon)}{g(\phi(\varepsilon))} = 1$ for every ε. Investments will be utilitarian efficient, and this difference in average quality has no implications for investment incentives.

3.1.3   Normal or Log-Normal Shocks

Suppose that the shocks are normally distributed, i.e. $\varepsilon \sim N(\mu_\varepsilon, \sigma_\varepsilon)$ and $\eta \sim N(\mu_\eta, \sigma_\eta)$. Thus $F(\varepsilon) = \Phi\left(\frac{\varepsilon - \mu_\varepsilon}{\sigma_\varepsilon}\right)$, and $G(\eta) = \Phi\left(\frac{\eta - \mu_\eta}{\sigma_\eta}\right)$, where Φ denotes the standard normal cdf. Thus the matching $\phi(\varepsilon)$ is linear, and $\frac{f(\varepsilon)}{g(\phi(\varepsilon))} = \frac{\sigma_\eta}{\sigma_\varepsilon}$ at all values of ε. Furthermore, linearity of the matching implies that $\phi'(\varepsilon + \Delta)$ is constant and equal to $\frac{\sigma_\eta}{\sigma_\varepsilon}$, implying that agents's optimization is strictly concave as long as the cost function is convex. Since shocks are unbounded, an agent is always matched even when he deviates downwards, and the misery effect plays no role in deterring downward deviations. The first order conditions for investment are

$$c_B'(x_B^*) = \frac{\sigma_\eta}{\sigma_\varepsilon}, \ \ c_G'(x_G^*) = \frac{\sigma_\varepsilon}{\sigma_\eta}.$$

Investments are always Pareto efficient but will not be utilitarian efficient if the variances differ. If one measures the degree of over or under investment relative to the utilitarian level by the associated marginal costs, and if $\sigma_\varepsilon > \sigma_\eta$, then the overinvestment by girls, relative to the utilitarian level, is proportional to the ratio of the standard deviations.

Our analysis can be extended to the case where an increasing function of the shocks is normally distributed. For example, consider log-normal shocks, so that $\ln(\varepsilon) \sim N(0, \sigma_\varepsilon)$ and $\ln(\eta) \sim N(0, \sigma_\eta)$. The first order conditions for investment (see the Appendix for derivation) are given by

$$\frac{\sigma_\eta}{\sigma_\varepsilon} \exp\left(\frac{1}{2}(\sigma_\eta - \sigma_\varepsilon)^2\right) = c_B'(x_B^*),$$

$$\frac{\sigma_\varepsilon}{\sigma_\eta} \exp\left(\frac{1}{2}(\sigma_\eta - \sigma_\varepsilon)^2\right) = c_G'(x_G^*).$$

As in the normal case, the ratio of the marginal costs is related to ratio of the variances. However, here we have

$$c_B'(x_B^*) \times c_G'(x_G^*) = \exp\left(\sigma_\eta - \sigma_\varepsilon\right)^2,$$

so that the outcome is neither utilitarian or even Pareto efficient unless $\sigma_\varepsilon = \sigma_\eta$. The extent of inefficiency is related to the difference in variances.

Finally, note that our general existence theorem does not apply to these examples since it assumes bounded shocks. In the normal case, linearity of the matching suffices to ensure that the maximization problem is strictly concave. The Appendix shows that large upward deviations are not profitable in the log-normal case under plausible assumptions on the ratio of variances and the convexity of costs. Thus our analysis can be extended more generally, beyond the class of distributions satisfying Assumption A1, if one assumes explicit forms for the distribution of shocks.

3.1.4 Sex Ratio Imbalances

Sex ratio imbalances are an important phenomenon in countries such as China and parts of India. These imbalances are extremely large in China, where it is estimated that one in five boys born in the 2000 census will be unable to find a marriage partner (see Bhaskar, 2011). Wei and Zhang (2011) argue that the high savings rate in China is partly attributable to the sex ratio imbalance. They argue that parents of boys feel compelled to invest more, in order to improve their chances of finding a partner, thus raising the overall savings rate. However, one might conjecture that this might be counter-balanced by the reduced pressure felt by the parents of girls. We therefore turn to our model to provide an answer to this question.

Assume that each sex is ex ante identical, and let the relative measure of girls equal $r < 1$. At the matching stage, since $r \leq 1$, all girls should be matched, and the highest quality boys should be matched. Since every girl is matched, the investment in her generates benefits for herself and for her partner (for sure). Thus the first best investment level in a girl, x_G^{**}, satisfies $c'(x_G^{**}) = 1$. Now consider investment in a boy. If we assume that the idiosyncratic component of match values is sufficiently small, then welfare optimality requires that only a fraction r of boys invest, and that their investments also satisfy $c'(\cdot) = 1$. However, if we restrict attention to symmetric investment strategies, then investment will take place in all boys, and since investment occurs before ε is realized, each boy has a probability r of being matched, and thus the utilitarian efficient level of investment in a boy, x_B^{**}, must satisfy $c'(x_B^{**}) = r$, i.e. the marginal cost must equal the expected marginal benefit. Similarly, the condition for Pareto efficiency, with arbitrary weights on the welfare of boys and girls is

$$c_B'(x_B^{**}) \times c_G'(x_G^{**}) = r. \tag{13}$$

We now turn to equilibrium. Suppose that all boys invest the same amount x_B^* and all girls invest the same amount x_G^*. Since only the top r fraction of boys will be matched, this corresponds to those having a realization of $\varepsilon \geq \tilde{\varepsilon}$ where $F(\tilde{\varepsilon}) = 1 - r$. In this case, a boy of

16

type $\varepsilon \geq \tilde{\varepsilon}$ will be matched with a girl of type $\phi(\varepsilon, r)$, where

$$1 - F(\varepsilon) = r[1 - G(\phi(\varepsilon, r))].$$

The derivative of this matching function is given by

$$\phi_\varepsilon(\varepsilon, r) = \frac{f(\varepsilon)}{rg(\phi(\varepsilon, r))}.$$

That is, an increase in ε increases a boy's match quality relatively more quickly, since the distribution of girls is relatively thinner, since $r < 1$.

Those boys with realizations below $\tilde{\varepsilon}$ will not be matched and receive a payoff $\bar{u} < \underline{\eta}$, what we have called the misery effect. As Hajnal (1982) has noted, in Asian societies such as China and India, marriage rates have historically been extremely high (over 99%, as compared to the traditional "European marriage pattern"with marriage rates around 90%). Thus the misery effect is likely to be large in Asian societies.

The first order condition for boys in an equilibrium where all boys invest the same amount x_B^*, while all girls invest the same amount x_G^* is given by

$$\frac{1}{r} \int_{\tilde{\varepsilon}(r)}^{\bar{\varepsilon}} \frac{f(\varepsilon)}{g(\phi(\varepsilon, r))} f(\varepsilon) \, d\varepsilon + f(\tilde{\varepsilon}(r))(\underline{\eta} + x_G^* - \bar{u}) = c_B'(x_B^*). \tag{14}$$

As compared to our previous analysis, we notice two differences. The first term is the improvement in match quality, and the sparseness of girls increases the investment incentives, due to the term in $1/r$. Additionally, an increment in investment raises the probability of one's son getting matched, at a rate $f(\tilde{\varepsilon})$, and the marginal payoff equals the difference between matching with worst quality girl and receiving $\underline{\eta} + x_G^*$, and not being matched and receiving $\bar{u}$. An unbalanced sex ratio tends to amplify investments in boys, for two reasons. First, a given increment in investment pushes boys more quickly up the distribution of girls, and second, there is an incentive to invest in order to increase the probability of match taking place at all, since there is discontinuous payoff loss from not being matched at $\tilde{\varepsilon}$, due to the misery effect.

Similarly, the first order condition for investment in girls is given by

$$r \int_{\underline{\eta}}^{\bar{\eta}} \frac{g(\eta)}{f(\phi^{-1}(\eta))} g(\eta) \, d\eta = c'(x_G^*). \tag{15}$$

Notice here that the role of $r < 1$ is to reduce investment incentives, since an increment in investment pushes a girl more slowly up the distribution of boy qualities. Furthermore, there is no counterpart to the misery effect for the scarcer sex, and the only reason to invest arises from the consequent improvement in match quality.

Since girls are in excess supply, a girl whose parents invest and whose quality realization is discretely lower than every other girl will still be able to find a partner. Such a girl will get a match payoff of $x_B^* + \tilde{\varepsilon}$, no matter how low her own quality. Thus, the conditions for the existence of a quasi-symmetric equilibrium are more stringent than in the balanced sex ratio case. Large downward deviations in investment will not be profitable provided that the dispersion in the qualities of boys is sufficiently large, and the cost function for girls is sufficiently convex.

Proposition 2 *If $r < 1$ and the noise is additive, there exists a unique quasi-symmetric equilibrium, provided that $f(.)$ is sufficiently dispersed and $c_B(.)$ is sufficiently convex. Investments are excessive relative to Pareto efficiency, for generic distributions of noise.*

It is worth pointing out that even absent the misery effect, there will be strictly excessive investments, even if the noise distributions are identical, unless they happen to be uniform. When $r < 1$, $g(\phi(\varepsilon))$ is not a linear transformation of $f(\varepsilon)$ unless f and g are uniform. Thus investment will be strictly greater than the efficient level.

As we have already noted, if F and G have the same distributions, and $r = 1$, investments are utilitarian efficient. Thus a balanced sex ratio is sufficient to ensure efficiency of investments in this case. This provides an additional argument for the optimality of a balanced sex ratio, over and above the congestion externality identified in Bhaskar (2011).

We may use our model to evaluate the theoretical basis of the empirical work by Wei and Zhang (2011), attributing the high savings rate in China to the sex ratio imbalance. Note first of all that, given the condition for utilitarian efficiency $c'(x_B^{**}) = r$, investment in boys should actually fall as r decrease below one. In a related signalling model, Hoppe et al. (2009) show that an increase in the number of men will increase total signalling by men; however, the effect on signalling by women is ambiguous and depends on the shape of the distribution of abilities amongst men. The shape of the distribution also matters for our model, but a further difficulty here is that investment by one side affects the incentives to invest by the other. In particular, the equilibrium choice of investment by girls enters the boys' first order condition (14). This potentially would also make the investment by boys respond ambiguously to the sex ratio becoming less equal.

As an example, consider the case where f and g are increasing and $f(\underline{\varepsilon}) = g(\underline{\eta}) = 0$:

$$r \int_{\underline{\eta}}^{\bar{\eta}} \frac{g(\eta)}{f(\phi^{-1}(\eta, r))} g(\eta) \, d\eta = \int_{\tilde{\varepsilon}(r)}^{\bar{\varepsilon}} g(\phi(\varepsilon, r)) \, d\varepsilon.$$

Since g is increasing and both ϕ and the range of integration $[\tilde{\varepsilon}(r), \bar{\varepsilon}]$ are increasing in r, it

follows that investment by girls is unambiguously increasing in r. Similarly,

$$\frac{1}{r} \int_{\tilde{\varepsilon}(r)}^{\bar{\varepsilon}} \frac{f(\varepsilon)}{g(\phi(\varepsilon, r))} f(\varepsilon) \, d\varepsilon = \int_{\underline{\eta}}^{\bar{\eta}} f(\phi^{-1}(\eta, r)) \, d\eta.$$

Since f is assumed increasing and ϕ^{-1} is decreasing in r, overall this expression in decreasing in r. That is the matching incentive for boys increases as the sex ratio becomes more uneven. However, the overall effect on boys' investments is ambiguous, as the LHS of (14) also depends on x_G^* which is increasing in r. One can at least conclude that for r close to one, so that $\tilde{\varepsilon}$ is close to $\underline{\varepsilon}$ and $f(\tilde{\varepsilon})$ is close to zero, x_B^* is decreasing in r.

In summary, under some assumptions, an uneven sex ratio can indeed increase investment incentives for men, but will also decrease incentives for women. The predicted effect on total investment is consequently ambiguous. An uneven gender ratio increases the relative weight of boys in the population, and their increased investment may be enough to increase the total, but this is not guaranteed.

3.2 Talent Shocks and Complementarities with Investment

Consider next the case where investment is in education and the uncertainty is talent risk. It is plausible that the return to investment depends upon the talent of the child. To model this, we suppose that quality is given by a multiplicative production function, $q(x, \varepsilon) = x\varepsilon$. With this production function, $q_x = \varepsilon$ and $q_\varepsilon = x$, so that the first order condition for equilibrium (8) and (9) can be re-written:

$$\frac{x_G^*}{x_B^*} \int \frac{f(\varepsilon)}{g(\phi(\varepsilon))} \varepsilon f(\varepsilon) \, d\varepsilon = c_B'(x_B^*), \tag{16}$$

$$\frac{x_B^*}{x_G^*} \int \frac{g(\eta)}{f(\phi^{-1}(\eta))} \eta g(\eta) \, d\eta = c_G'(x_G^*). \tag{17}$$

Unlike the additive case, the "reaction function" for the boys is upward sloping in the girl's investments, and vice versa. Intuitively, with if quality is multiplicative, an increase in the girl's investment level increases the dispersion in qualities on the girl's side, thereby increasing investment incentives for boys. Thus, with multiplicative shocks, one has interesting interaction effects between investments on the two sides of the market. First, we examine the implications of gender differences.

Proposition 3 *Assume that multiplicative shocks that have the same mean for boys and girls, and identical cost functions for the two sexes. If the distribution of shocks for boys is more dispersed than that for girls, that is $F \geq_d G$, then there is under-investment in boys*

and there is over-investment in girls relative to the utilitarian efficient level.

Proof. From the first order conditions, if $F \geq_d G$:

$$c'_B(x^*_B) = \frac{x^*_G}{x^*_B} \int \frac{f(\varepsilon)}{g(\phi(\varepsilon))} \varepsilon f(\varepsilon) \, d\varepsilon < \frac{x^*_G}{x^*_B} \mathbf{E}(\varepsilon),$$

$$c'_G(x^*_G) = \frac{x^*_B}{x^*_G} \int \frac{g(\eta)}{f(\phi^{-1}(\eta))} \eta g(\eta) \, d\eta > \frac{x^*_B}{x^*_G} \mathbf{E}(\eta).$$

Utilitarian investments x^{**}_B and x^{**}_G satisfy $c'_B(x^{**}_B) = \mathbf{E}(\varepsilon)$ and $c'_G(x^{**}_G) = \mathbf{E}(\eta)$. Since $\mathbf{E}(\varepsilon) = \mathbf{E}(\eta)$ and the cost functions are the same, $c'_B(x^*_B) < \mathbf{E}(\varepsilon)$ and $c'_G(x^*_G > \mathbf{E}(\eta)$. ∎

Multiplicative shocks have the interesting implication that an increase in investment by one side, say girls, also increases the dispersion in quality amongst girls, thereby providing more incentives to invest for the boys. So if shocks are relatively less variable amongst girls, the induced investments are such that the resulting difference in variability in *quality* between the two sides is less pronounced. Since girls invest more and boys invest less, this raises quality dispersion amongst girls, and reduces it amongst boys.

Both the additive model and the multiplicative model show that the side with more dispersed shocks has weaker incentives for investment. A similar argument is also found in Hoppe et al. (2009), where differences in the dispersion of exogenously given unobserved qualities affect signalling expenditures by men and women. The additive shocks model yields conclusions similar to Hoppe et al., since investments by one side do not affect investment incentives on the other side. The multiplicative model is richer, since increased investments by girls raise the incentives to invest for boys.

These interaction effects have interesting implications also when the mean value of shocks differs between the sexes. Suppose that $F(\varepsilon) = G(\varepsilon + k)$ for some $k > 0$, so that average quality is higher amongst the girls, but the distributions are of the same type. An example would be where shocks are normally distributed, with the girls having a higher mean. Since $\frac{f(\varepsilon)}{g(\phi(\varepsilon))} = 1$ for all values of ε the first order conditions reduce to

$$c'_B(x^*_B) = \frac{x^*_G}{x^*_B} \mathbf{E}(\varepsilon),$$

$$c'_G(x^*_G) = \frac{x^*_B}{x^*_G} \mathbf{E}(\eta).$$

From the first order conditions, we see that women invest more than men. However, the interaction effects imply that women invest less than the utilitarian efficient amount, and men invest more than the utilitarian level. Thus, in the multiplicative case, investment behavior partially offsets differences in mean quality.

Finally, the multiplicative model also shows us that facilitating investment by one side, say girls, also increases investment incentives for boys. In some developing countries such as India, government have sought to overcome discrimination against girls by subsidizing their education. Increased investment in girls raises the dispersion in their quality, thereby providing greater incentives to invest for the boys.

3.2.1 Investment under Traditional Gender Roles

Suppose that shocks are additive for women but multiplicative for men. One example is a traditional society, where women do not work, and so investment in them takes the form of a dowry; while parents invest in their sons' human capital. This interpretation fits our model if investments towards dowries must take place in advance, i.e. parents must forgo consumption in order to save for their daughter's dowry. The first order condition for investments in boys is

$$\frac{1}{x_B^*} \int \frac{f(\varepsilon)}{g(\phi(\varepsilon))} \varepsilon f(\varepsilon) \, d\varepsilon = c_B'(x_B^*). \tag{18}$$

Notice that this is independent of the investment level of the girls, since the quality of girls is additive. On the other hand, since the dispersion of qualities amongst the boys increases with their investment level, the investment by girls is increasing in boys' investments, as the girls' first order condition shows:

$$x_B^* \int \frac{g(\eta)}{f(\phi^{-1}(\eta))} g(\eta) \, d\eta = c_G'(x_G^*). \tag{19}$$

This mixed model can provide an explanation for why dowries may increase during the process of development, as for example in India.[13] Suppose that the marginal costs of investing in human capital fall, so that $c_B'(x_B)$ is reduced at any value of x_B. From (18), x_B^* must go up. From (19), x_G^* also increases. Intuitively, the increased investment by boys increases the variability in their quality, thereby increasing the incentives for the parents of girls for investing in their dowries. A similar argument can also be made if the return to human capital goes up – this will increase investment levels by boys, increasing the dispersion in their qualities. The increased dispersion in boy qualities increases the investment incentives for girls.

---

[13]Anderson (2007) surveys the evidence on dowries, while Anderson (2003) provides an alternative explanation for the increase in dowries during modernization, based on social stratification.

3.2.2 Efficiency in a Generalized Multiplicative/Additive Model

We now examine whether investments are Pareto efficient in a generalized model where quality has multilicative as well as additive components. Assume that quality satisfies assumption A3, set out on page 10. Utilitarian efficiency requires $c'_B(x_B^{**}) = \theta_B + (1-\theta_B)\gamma'(x_B^{**})\mathbf{E}(\varepsilon)$ and $c'_G(x_G^{**}) = \theta_G + (1-\theta_G)\gamma'(x_G^{**})\mathbf{E}(\eta)$. Pareto efficient investments are such that the product of the marginal costs equal $[\theta_B + (1-\theta_B)\gamma'(x_B^{**})\mathbf{E}(\varepsilon)][\theta_G + (1-\theta_G)\gamma'(x_G^{**})\mathbf{E}(\eta)]$. The following theorem shows that there will be generic overinvestment, if the densities are both symmetric, if they are different from each other.

Theorem 3 *Suppose that quality is generalized additive/multiplicative and that the distributions f and g are symmetric. Then in any quasi-symmetric equilibrium investments are generically excessive relative to Pareto efficiency.*

This overinvestment result applies both to the pure multiplicative model and to the one with traditional gender roles. The proof is based on on the Cauchy-Schwarz inequality and the result is a robust result. If f and g are symmetric and linearly independent, investment will be strictly too high. Now if we perturb the distributions so the $\tilde{f}$ is close to f and $\tilde{g}$ to g, then $c'(x_B^*) \times c'(x_G^*)$ will still be greater than those required for efficiency, since the integrals defining this are continuous in the distributions. In other words, we will have excessive investments even with asymmetric distributions as long as the asymmetries are not too large.

Why does the result require that the asymmetry not be too large? To provide some intuition for this, assume purely multiplicative quality, and return to the example where the distribution of shocks is uniform on $[0,1]$ for men and where the density function for women, $g(\eta) = 2\eta$ on $[0,1]$. Here again, the ratio of the densities that is relevant for men, $\frac{f(\varepsilon)}{g(\phi(\varepsilon))}$, is relatively large when ε is low. While these values of ε still have large weight (since $f(\varepsilon)$ is constant in ε), in the multiplicative case, the payoff to investment is low when ε is small. Under symmetry, neither particularly low values nor particularly high values of ε have any special weight and thus the inefficiency result applies.

4 No Gender Difference Implies Efficiency

We now consider the implications of ex-ante heterogeneity, where individuals differ even before shocks are realized, beginning with an illustrative example. Assume that the sex ratio is balanced. Suppose that we have two classes, H and L, with fractions θ_H and θ_L in the population. Assume that the marginal costs of investment are lower for the upper class, H. Let $c_H(.)$ and $c_L(.)$ be the cost functions, which depend upon class but not upon gender,

where $c'_H(x) < c'_L(x)$ for any x. Let $f_i(.)$ and $g_i(.), i \in \{H, L\}$ denote the density function of shocks for the boys from class i, and the girls from class i respectively. Assume that quality function is additive in the shocks and investment.

Consider a profile of investments $(x_{HB}, x_{LB}, x_{HG}, x_{LG})$ where each individuals who belong to the same class and same gender choose the same investment. This profile induces a distribution of qualities for the boys, $\tilde{F}(q)$, and of girls, $\tilde{G}(p)$. Since any stable measure preserving matching $\tilde{\phi}$ must be assortative, we must have $\tilde{F}(q) = \tilde{G}(\tilde{\phi}(p))$.

Let $x^*_{HB}, x^*_{LB}, x^*_{HG}, x^*_{LG}$ be the equilibrium investment levels. Suppose that the distribution of qualities of in both the sexes has a connected support, without any gaps. In class i, the first order condition for investment in boys is given by

$$\int_{q_{\min}}^{q_{\max}} \frac{\tilde{f}(q)}{\tilde{g}(\phi(q))} f(q - x^*_{iB}) \, dq = c'_B(x^*_{iB}). \tag{20}$$

The density function for boy's quality is given by (that for girls is analogous):

$$\tilde{f}(q) = \theta_H f_H(q - x^*_{HB}) + \theta_L f_L(q - x^*_{LB}).$$

Let the class differences be arbitrary, so that f_H can differ from f_L and g_H from g_L. However assume that there are no gender differences, so that $f_H = g_H$ and $f_L = g_L$. Consider a *gender neutral strategy* profile, where investments depend on class but not on gender, so that $x^*_{iB} = x^*_{iG}$ for $i \in \{H, L\}$. Since the shocks do not vary between the sexes, the induced distribution of qualities will be identical in the two sexes. That is, for any value q, $\tilde{F}(q) = \tilde{G}(q)$, implying that $\tilde{\phi}(q) = q$. This in turn implies that $\tilde{f}(q) = \tilde{g}(\phi(q))$. Therefore the left hand side of equation (20) equals one.

Consider a utilitarian social planner who puts equal weight on each type of individual, irrespective of gender or social class. Since the marginal benefit of additional investment in a boy is unity, for any girl who is matched with him, such a planner would set the marginal cost of investment to one. We conclude therefore that investment in boys is utilitarian efficient if there are no gender differences, even if there is large heterogeneity between classes. Similarly, investment in girls is utilitarian efficient.

This argument is very general – provided that there are no differences between the sexes, equilibrium investments will be utilitarian efficient even if there is wide heterogeneity within each sex. Assume that there is a finite set of types, indexed by $i \in \{1, 2, .., n\}$. Type i has a measure μ^i of boys and an equal measure of girls, with $\sum_{i=1}^n \mu^i = 1$. A boy or girl of type i has an idiosyncratic component of quality, ε, that is distributed with a density function $f_i(\varepsilon)$ and a cumulative distribution function $F_i(\varepsilon)$. We shall assume a general quality functions $q(x, \varepsilon)$, where q is continuous, strictly increasing in both arguments, and differentiable and concave in x, the investment. The cost of investment may also depend upon type, and is

denoted $c_i(x)$. We assume that Assumption A2 holds for each type, i.e. it holds for each cost function c_i and each density function f_i.

We assume: **No Gender Difference**: Men and women are symmetric with regards to costs of investment and the idiosyncratic component of quality. Specifically, for any type i: i) there are equal measures of men and women, ii) the investment cost functions and quality functions do not differ across the sexes, and iii) the idiosyncratic component of quality has the same distribution, f_i, that depends upon type but not on gender.

The assumption of no gender difference is strong, but there are reasonable conditions under which it is satisfied. Suppose that investment costs or the idiosyncratic component depend upon the "type" of the parent (e.g. parental wealth, human capital or social status), but not directly upon gender. If the gender of the child is randomly assigned, with boys and girls having equal probability, then no gender difference will be satisfied.

A utilitarian efficient profile of investments $(x_{Bi}^{**}, x_{Gi}^{**})_{i=1}^{n}$ is one that maximizes the sum of payoffs of all individuals, irrespective of type or gender. This satisfies the conditions below $\forall i$, i.e. the marginal social benefit from increased quality must equal the marginal cost to the individual:

$$c_i'(x_{Bi}^{**}) = \int q_x(x_{Bi}^{**}, \varepsilon) f_i(\varepsilon) \, d\varepsilon, \tag{21}$$

$$c_i'(x_{Gi}^{**}) = \int q_x(x_{Gi}^{**}, \eta) g_i(\eta) \, d\eta. \tag{22}$$

Under Assumption A2, the cost function c_i is strictly convex and $q_{xx}(.) \le 0$, so that the above conditions are sufficient for the profile to be utilitarian efficient. Assuming no gender difference, utilitarian efficiency requires that individuals of the same type choose the same investments even if they differ in gender, i.e. $x_{Bi}^{**} = x_{Gi}^{**} \equiv x_i^{**} \forall i$.

Consider now a quasi symmetric strategy profile $((x_{Bi}^*)_{i=1}^{n}, (x_{Gi}^*)_{i=1}^{n})$, specifying investment levels for each type of each gender. This profile, in conjunction with the realizations of idiosyncratic shocks, induces a cumulative distribution function of qualities, $\tilde{F}$, in the population of boys. Since ε is assumed to be atomless, and q is continuous, $\tilde{F}$ admits a density function $\tilde{f}$, although its support may not be connected if the investment levels of distinct types are sufficiently far apart (i.e. there may be gaps in the distribution of qualities). Similarly, let $\tilde{G}$ denote the cumulative distribution function of girl qualities, given $(x_{Gi})_{i=1}^{n}$. Since stable measure matchings will be assortative, i.e. if a boy of type q is matched to a girl of type $\phi(q)$ if and only if $\tilde{F}(q) = \tilde{G}(\phi(q))$. Thus the distributions $\tilde{F}$ and $\tilde{G}$ define the match payoffs associated with equilibrium investments for each type of boy and each type of girl. For the profile $((x_{Bi}^*)_{i=1}^{n}, (x_{Gi}^*)_{i=1}^{n})$ to be an equilibrium, it must satisfy the first order conditions for each type i, for the boys and girls respectively:

$$c_i'(x_{Bi}^*) = \int \tilde{\phi}'(.) q_x(x_{Bi}^*, \varepsilon) f_i(\varepsilon) \, d\varepsilon, \tag{23}$$

$$c_i'(x_{Gi}^*) = \int \left(\tilde{\phi}^{-1}\right)' (.) q_x(x_{Gi}^*, \eta) g_i(\eta) \, d\eta. \tag{24}$$

We shall call a strategy profile *gender neutral* if $x_{Bi} = x_{Gi} \forall i$, so each type of parent invests the same amount regardless of the gender of their child. Suppose that $((x_{Bi}^*)_{i=1}^n, (x_{Gi}^*)_{i=1}^n)$ is gender neutral and is an equilibrium. Under the assumption of no gender difference, the induced distributions of qualities are identical on the two sides, i.e. $\tilde{F}(.) = \tilde{G}(.)$. Thus $\tilde{\phi}(.)$ is the identity map on the support of $\tilde{F}(.)$. The first order conditions reduce to

$$c_i'(x_{Bi}^*) = \int q_x(x_{Bi}^*, \varepsilon) f_i(\varepsilon) \, d\varepsilon = c_i'(x_{Gi}^*)., \tag{25}$$

The first order conditions for an equilibrium that is gender neutral, (25) coincide with the first order conditions for utilitarian efficiency, (21) and (22). Thus if a gender neutral equilibrium exists, it must be utilitarian efficient. Also, if large deviations from the utilitarian profile are unprofitable for every type, then it is the unique gender neutral equilibrium.

A profile of investments, $((x_{Bi})_{i=1}^n, (x_{Gi})_{i=1}^n)$, has *no quality gaps* if the induced distributions of qualities, $\tilde{F}(q)$ and $\tilde{G}(p)$, have connected supports. Large deviations will be unprofitable as long as there are no quality gaps.

Theorem 4 *Suppose that there is no gender difference. The utilitarian efficient profile of investments is a gender neutral equilibrium if it has no quality gaps, if Assumption A2 is satisfied and $\bar{u}$ is sufficiently small. A gender neutral strategy profile that has no quality gaps is an equilibrium if and only if it is utilitarian efficient.*

The efficiency result applies plausibly to a non-marriage context. Consider a single-population matching model, where quality is a one-dimensional scalar variable. An example is partnership formation, e.g. firms consisting of groups of lawyers. Theorem 4 implies that one has efficient ex ante investments, even absent transferable utility. While the formal proof restricts attention to pair-wise matching, the extension to matches consisting of more than two partners is immediate.

The intuition for the efficiency result is as follows. Consider a gender neutral profile of investments, where there are no quality gaps. Then a boy of quality q will be matched with a girl of the same quality, i.e. $\tilde{\phi}(q) = q$. Thus the marginal return to investment equals the increment to his own quality, and thus private incentives and utilitarian welfare are perfectly aligned.

This result does require the no quality gap assumption, which will be satisfied if the support of the shocks is large enough, or if sufficient similarly between adjacent types so that their equilibrium investments are not too far apart. Interestingly, if there are quality gaps, then there is a tendency to overinvestment, rather than under investment. Let us

return to the two-class illustrative example at the beginning of this section, and suppose that the differences in utilitarian investments between the rich and the poor are so large that there is a quality gap. Suppose that an individual boy deviates from this profile and has a quality realization that is greater than the best poor boy, and smaller than the worst rich boy. Assume that the deviator is assigned either the match of the former or that of the latter, each with probability one-half. Under such a matching rule, the poor boys would have an incentive to deviate upwards – the rich boys have no incentive to deviate downwards.[14]

5   A Model with Finite Numbers

We now set out a model with finitely many boys and girls, where there is uncertainty as to whether there are slightly more boys than girls or the reverse.[15] Thus the lowest quality boy (or girl) will be unmatched with probability one-half. We show that the payoffs in this finite model converge to those in the continuum model as the number of participants becomes large. This provides a justification for our assumption in the continuum case, that a downward deviating agent, whose quality is below the support of the equilibrium distribution of qualities, is unmatched with probability one-half. We also show that if the number of agents is sufficiently large, there exists a unique quasi-symmetric equilibrium of the finite model, which converges to the equilibrium of the continuum model as the number of agents tends to infinity.

Assume that there are $2n + 1$ agents, with their sex being determined as follows: n of the agents are randomly chosen, (equi-probably, so that each agent has equal an chance of being chosen), and then a fair coin is tossed to determine whether these n chosen individuals are all male or all female. The $n + 1$ unchosen individuals are then specified to be of the opposite sex. We assume that at the time of investment, each individual knows their sex, but not whether they were among the n chosen individuals. Thus, at the investment stage, an individual does not know whether boys or girls are in excess.

Suppose that there are ℓ boys and m girls. For simplicity, we assume that quality is additive in investment and in the idiosyncratic shock. Shocks are i.i.d, and are drawn from F for the boys and G for the girls. Agents are matched assortatively in terms of quality, and if $\ell > m$, so that there are more boys than girls, then the bottom $\ell - m$ boys are left unmatched. Suppose all girls invest x_G and all boys invest x_B. Let $F_i^\ell(\varepsilon)$ denote the

---

[14]The matching rule can be justified as the limit of a model with a large but finite number of agents of each type, along the lines of argument in the next section. Efficiency can be obtained even with quality gaps if we modify the matching rule so that a deviator is always assigned the match of the next worst individual. This matching rule is formally correct in continuum model, but does not seem to correspond to the limit of a reasonable finite model.

[15]We thank Roger Myerson for suggesting this approach. See also Myerson (1998) for large games with a random set of players.

probability that a boy with shock ε is ranked i when there are ℓ boys. Write $E\eta_{(j,m)}$ as the expected value of the j-th order statistic for shocks for girls when there are m girls. The expected match of a boy when he invests $x_B + \Delta$, and his shock value is ε in the event that there are n boys and $n + 1$ girls is given by

$$\phi_{n,n+1}(\varepsilon + \Delta, x_G) = \sum_{i=1}^{n} F_i^n(\varepsilon + \Delta) \left[E\eta_{(i+1,n+1)} + x_G \right]. \tag{26}$$

His expected payoff when there are $n + 1$ boys and n girls is given by

$$\phi_{n+1,n}(\varepsilon + \Delta, x_G) = \sum_{i=2}^{n} F_i^n(\varepsilon + \Delta) \left[E\eta_{(i-1,n)} + x_G \right] + F_1^n(\varepsilon + \Delta)\bar{u}. \tag{27}$$

Assume that there are either n boys and $n+1$ girls or $n+1$ boys and n girls, with both these events having equal probability.[16] The benefit of a boy, when he invests $x_B + \Delta$, is given by

$$B_n(\Delta, x_G) = \frac{1}{2} \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \left[\phi_{n,n+1}(\varepsilon + \Delta, x_G) + \phi_{n+1,n}(\varepsilon + \Delta, x_G) \right] f(\varepsilon) \, d\varepsilon, \tag{28}$$

Since the benefit function $B_n(\Delta, x_G)$ is differentiable with respect to Δ, a quasi-symmetric equilibrium (x_{Bn}^*, x_{Gn}^*) must satisfy the first order conditions:

$$c_B'(x_{Bn}^*) = B_n'\left(0, x_{Gn}^*\right),$$

$$c_G'(x_{Gn}^*) = B_n'\left(0, x_{Bn}^*\right).$$

There is a unique solution to the first order conditions if n is sufficiently large. This defines an equilibrium, giving us the following theorem:

Theorem 5 *Suppose that quality is additive, and that Assumptions A1 and A2 are satisfied. For n sufficiently large, the game with uncertain finite numbers has a unique quasi-symmetric equilibrium with investments (x_{Bn}^*, x_{Gn}^*), provided that $\bar{u}$ is sufficiently small. Further, $\lim_{n\to\infty}(x_{Bn}^*, x_{Gn}^*) = (x_B^*, x_G^*)$, the equilibrium investments in the continuum model.*

We prove this result in the appendix, by showing that the benefit function in finite model and its derivatives converge to their counterparts in the continuum model. Since Theorem 2 shows that one has generic overinvestment in the continuum case, our main convergence result here, Theorem 5, implies that there is excessive investment when the number of agents is sufficiently large. However, exactly how investment incentives depend on population size is quite complex and we do not analyze it here.

---

[16]Conditional on an individual being a male, the probability that there are $n+1$ males is $\frac{n+1}{2n+1}$, since this is the probability that the individual is not among the n chosen in the first stage of the process determining the sexes. This converges to 0.5 as $n \to \infty$.

6   Conclusions

We examined a model of marriage with investments that have stochastic returns. This approach ensures the existence of a unique pure strategy equilibrium, in an area where models often have multiple equilibria or equilibria only in mixed strategies. Our main result is that generically investment is inefficiently high. The intuition for our inefficiency result is somewhat subtle – it is not due to the usual positional externality that arises in tournaments, since investments in our context are not wasteful. Investments by boys increase welfare for the girls, and vice-versa. Indeed, when the two sides or sexes are identical and there are no gender differences, one gets efficient investments. However, when there are differences between the sexes, there are some realizations of shocks where boys have a relatively higher incentive to invest, and other realizations where girls have a relatively higher incentives to invest. Each sex gives greater weight to those states where they have relatively greater incentives, giving rise to over-investment. When the sexes are identical, at every shock realization, both sexes have identical investment incentives, which ensures efficiency. Our model also has interesting observational implications. For example, if shocks are more variable for boys as compared to girls, boys invest less than girls. If there is an unbalanced sex ratio, the abundant sex invests more, while the scarcer one invests less.

While formal tests of the Pareto efficiency of investments at the aggregate level are yet to be developed, Wei and Zhang (2011) present evidence that the sex ratio imbalance drives higher savings in China, which is possibly inefficiently high. Similarly, the global boom in the higher education of women is arguably not explained by higher returns to education on the labor market (see Becker et al. (2010)). While this could be due to women having lower costs in the form of superior non-cognitive skills, our model suggest that matching considerations from the marriage market could explain why investment in education might differ by sex.

Appendix: Proofs

Assumption A2:

1. Assumptions on shocks: let $f(\varepsilon)$ and $g(\eta)$ be twice continuously differentiable with finite derivatives on their bounded supports $[\underline{\varepsilon}, \bar{\varepsilon}]$ and $[\underline{\eta}, \bar{\eta}]$ respectively. Further, we assume that $f(\underline{\varepsilon}) = 0$, $g(\underline{\eta}) = 0$ but $f(\varepsilon)$ and $g(\eta)$ are otherwise strictly positive on their supports, with $f'(\underline{\varepsilon}) > 0$, $g'(\underline{\eta}) > 0$.

2. Assumptions on costs: let $c_B(x_B)$ and $c_G(x_G)$ be continuously twice differentiable. $c_B(x_B)$ and $c_G(x_G)$ are eventually increasing. Further,

   (a) Convexity: $c_B''(x_B) > 0$ and $c_G''(x_G) > 0$.

(b) There exist privately optimal levels of investment $\bar{x}_B > 0$, $\bar{x}_G > 0$ such that $c'_B(\bar{x}_B) = 0$ and $c'_G(\bar{x}_G) = 0$.

(c) Bounded maximum rational investment: there are investment levels $\tilde{x}_B$, $\tilde{x}_G$ such that $\lim_{x \to \tilde{x}_B} c'_B(x) = \infty$ and $\lim_{x \to \tilde{x}_G} c'_G(x) = \infty$.

3. Assumptions on quality: let $q(x_B, \varepsilon)$ and $q(x_G, \eta)$ be strictly increasing and twice differentiable, with $q_{xx}(x_B, \varepsilon) \le 0$, $q_{\varepsilon\varepsilon}(x_B, \varepsilon) = 0$, $q_{xx}(x_G, \eta) \le 0$, $q_{\eta\eta}(x_G, \eta) = 0$.[17]

4. The value of not being matched is $\bar{u}$ which satisfies $\bar{u} < q(0, \underline{\varepsilon})$ and $\bar{u} < q(0, \underline{\eta})$. That is, a girl who invests x and who is not matched has total payoff $\bar{u} - c_G(x)$.

Proof of Theorem 1: We first derive the first order conditions as given in (8) and (9). We then show that these first order conditions are sufficient, and that no individual can benefit from large deviations. Next, we show that the first order conditions define a unique symmetric best response for the boys as a function of the investment level of the girls, and vice versa. These best response functions are continuous, establishing existence of equilibrium. Finally, we show that equilibrium is unique.

Consider a quasi-symmetric equilibrium where all boys invest x_B^* and all girls invest x_G^*. Suppose that a parent of a boy deviates from this equilibrium and invests $x_B^* + \Delta$ in his son. We may, without loss of generality, restrict Δ to lie in a compact interval $[a, b]$ where $a < 0 < b$.[18] If the realization of the shock for his son is ε, the son will hold the same rank in the population of boys as a boy with a shock level ξ, where $\xi(x, \varepsilon, \Delta)$ is defined by the equation

$$q(x_B^* + \Delta, \varepsilon) = q(x_B^*, \xi(x_B^*, \Delta, \varepsilon)).$$

For example, in the additive case $\xi(x_B^*, \Delta, \varepsilon) = \varepsilon + \Delta$. Given this deviation, the boy now holds rank $F(\xi)$ in the population of boys and can expect a match with a girl holding rank $G(\phi(\xi))$ in the population of girls. She would be of quality $q(x_G^*, \phi(\xi))$. It is easy to show that $\xi(x_B^*, \Delta, \varepsilon)$ has the following properties:

$$\xi_\Delta > 0, \xi_\varepsilon > 0, \xi_{\varepsilon\varepsilon} = 0, \xi_{\Delta\varepsilon} \ge 0, \xi_{\Delta x} \le 0, \xi_{\Delta\Delta} \le 0, \tag{29}$$

given our assumptions that $q_{xx} \le 0$, $q_{\varepsilon x} \ge 0$ and $q_{\varepsilon\varepsilon} = 0$.

For $\Delta > 0$, define $\hat{\varepsilon}(\Delta)$ by $\xi(x_B^*, \Delta, \hat{\varepsilon}) = \bar{\varepsilon}$. That is, given an upward deviation of Δ, $\hat{\varepsilon}(\Delta)$ is the shock value that results in exactly the same investment as the highest ranked non-deviating boy. Similarly for downward deviations, $\Delta < 0$, define $\tilde{\varepsilon}(\Delta)$ by the relation $\xi(x_B^*, \Delta, \tilde{\varepsilon}) = \underline{\varepsilon}$.

---

[17]That is, quality is assumed concave in investment (x) but linear in shocks (ε, η). These assumptions allow for the additive and multiplicative specifications, as well as many others, as special cases.

[18]No agent wants to deviate below the individually optimal investment level. Nor is there any advantage in deviating to $\Delta > \bar{\varepsilon} - \underline{\varepsilon}$, since one cannot do better than being matched with the best woman.

Importantly, if a deviating boy has a shock value that is greater than $\hat{\varepsilon}(\Delta)$, then he matches for sure with the highest ranking girl, that is the one with quality $q(x_G^*, \bar{\eta})$, where $\bar{\eta} = \phi(\bar{\varepsilon})$. Thus, if all other boys invest an amount x_B^* and all girls x_G^*, then the expected match quality or benefit $B(\Delta)$ of a boy investing $x_B^* + \Delta$, where $\Delta > 0$, is given by

$$B(\Delta) = \int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\Delta)} q(x_G^*, \phi(\xi(x_B^*, \Delta, \varepsilon))) f(\varepsilon) \, d\varepsilon + (1 - F(\hat{\varepsilon}(\Delta))) q(x_G^*, \bar{\eta}). \qquad (30)$$

The derivative of the expected match or benefit with respect to $\Delta > 0$ is then given by

$$B'(\Delta) = \int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\Delta)} q_\eta(x_G^*, \phi(\xi(x_B^*, \Delta, \varepsilon))) \phi'(\xi(x_B^*, \Delta, \varepsilon)) \xi_\Delta(x_B^*, \Delta, \varepsilon) f(\varepsilon) \, d\varepsilon. \qquad (31)$$

To evaluate this at $\Delta = 0$, note that ϕ' and ξ_Δ, evaluated at $\Delta = 0$, are given by

$$\phi'(\varepsilon) = \frac{f(\varepsilon)}{g(\phi(\varepsilon))}, \quad \xi_\Delta(x_B^*, \Delta, \varepsilon) = \frac{q_x(x_B^*, \varepsilon)}{q_\varepsilon(x_B^*, \varepsilon)}.$$

and $\xi(x_B^*, \Delta, \varepsilon)$ evaluated at $\Delta = 0$ is simply ε. Since $\xi = \varepsilon$ when $\Delta = 0$, $\hat{\varepsilon} = \bar{\varepsilon}$ when $\Delta = 0$. Thus, evaluating (31) at $\Delta = 0$ gives us

$$B'(\Delta)|_{\Delta=0} = \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} q_\eta(x_G^*, \phi(\varepsilon)) \frac{f(\varepsilon)}{g(\phi(\varepsilon))} \frac{q_x(x_B^*, \varepsilon)}{q_\varepsilon(x_B^*, \varepsilon)} f(\varepsilon) \, d\varepsilon. \qquad (32)$$

Thus the right hand derivative, evaluated at $\Delta = 0$, equals the left hand side of the first order condition (8).

Turning to downward deviations, $\Delta < 0$, $B(\Delta)$ is given by

$$B(\Delta) = \int_{\tilde{\varepsilon}(\Delta)}^{\bar{\varepsilon}} q(x_G^*, \phi(\xi(x_B^*, \Delta, \varepsilon))) f(\varepsilon) \, d\varepsilon + F(\tilde{\varepsilon}(\Delta)) \frac{\bar{u} + q(x_G^*, \underline{\eta})}{2}. \qquad (33)$$

The final term depends on the following crucial assumption. If the deviating boy has a low shock realization, in the interval $[\underline{\varepsilon}, \tilde{\varepsilon}]$, his quality will be less than the lowest quality in the equilibrium distribution of boys' qualities. With equal probability, he will be left unmatched and have utility $\bar{u}$ or match with the lowest quality girl. The derivative with respect to a downward deviation $\Delta < 0$ is

$$B'(\Delta) = \int_{\tilde{\varepsilon}(\Delta)}^{\bar{\varepsilon}} q_\eta(x_G^*, \phi(\xi(\cdot))) \phi'(\xi(\cdot)) \xi_\Delta(x_B^*, \Delta, \varepsilon) f(\varepsilon) \, d\varepsilon + f(\tilde{\varepsilon}(\Delta)) \frac{d\tilde{\varepsilon}}{d\Delta} \frac{(\bar{u} - q(x_G^*, \underline{\eta}))}{2} \qquad (34)$$

Since $\tilde{\varepsilon}$ is defined by the relation $q(x_B^* + \Delta, \tilde{\varepsilon}) = q(x_B^*, \underline{\varepsilon})$, $\tilde{\varepsilon} \to \underline{\varepsilon}$ as $\Delta \to 0$. Since $f(\underline{\varepsilon})$ is zero by A2, the above derivative approaches (32) as Δ goes to zero. Thus, the left and right derivatives exist and are equal at $\Delta = 0$ and give the first order condition (8). The first

order condition for girls (9) can similarly be derived.[19]

We now show that the integral defining $B'(\Delta)$ is well defined (even though $1/g(\phi(\varepsilon))$ is unbounded). As in the proof of Theorem 2, we make a change of variables, from ε to η:

$$B'(\Delta)\,|_{\Delta=0} \;=\; \int_{\underline{\eta}}^{\bar{\eta}} q_\eta(x_G^*,\eta)f(\phi^{-1}(\eta))\frac{q_x(x_B^*,\phi^{-1}(\eta))}{q_\varepsilon(x_B^*,\phi^{-1}(\eta))}d\eta.$$

Since $f(\phi^{-1}(\eta))$ is bounded, as is q_η and since $q_\varepsilon(x_B^*,\phi^{-1}(\eta))$ is bounded away from zero, the integral defining $B'(0)$ is well defined.

We now show that large deviations are not profitable, considering in turn upward deviations and downward deviations.

Upward deviations: Let $B''_+(\Delta)$ denote the derivative of $B'(\Delta)$ on $[0,b]$, where $B''_+(0) = \lim_{\Delta\uparrow 0} B''_+(\Delta)$, and equals the right hand derivative of $B'(\Delta)$ at $\Delta = 0$.[20] Let $U''_+(\Delta, x_B^*) = B''_+(\Delta) - c''(x_B^* + \Delta)$. We show that $U''_+(\Delta, x_B^*) < 0$ for $\Delta \in [0,b]$, which implies $B'(\Delta) - c'(x_B^* + \Delta) < B'(0) - c'(x_B^*)$, so that no upward deviation is profitable.

Suppressing most arguments, $B''_+(\Delta)$ can be written as

$$B''_+(\Delta) = \int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\Delta)} \left(q_{\eta\eta}(\phi'\xi_\Delta)^2 + q_\eta\phi''\xi_\Delta^2 + q_\eta\phi'\xi_{\Delta\Delta}\right) f(\varepsilon)\,d\varepsilon + \frac{d\hat{\varepsilon}}{d\Delta}q_\eta\phi'\xi_\Delta f(\hat{\varepsilon}). \qquad (35)$$

Thus $B''_+(\Delta) \leq 0$ if all the three functions $q_{\eta\eta}$, ϕ'' and $\xi_{\Delta\Delta}$ are (weakly) negative.[21] First, $q_{\eta\eta} = 0$ by assumption. Second, we have $\xi_{\Delta\Delta} \leq 0$ from (29). This leaves $\phi''(\cdot)$.

If a) F and G are of the same type (i.e. $G(x) = F(a + bx)$), then $\phi(\varepsilon)$ is linear and $\phi''(\cdot) = 0$. Thus, $B''_+(\Delta) \leq 0$, and $U''_+(.) < 0$ since $c''_B(.) > 0$.

We now consider b), $f'(\varepsilon) \geq 0$ and $g'(\eta) \geq 0$. The expression (35) can be written as

$$B''_+(\Delta) = \int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\Delta)} q_\eta\phi'\xi_{\Delta\Delta}f(\varepsilon)\,d\varepsilon + \int_{\underline{\varepsilon}}^{\hat{\varepsilon}(\Delta)} q_\eta\phi''\frac{\xi_\Delta^2}{\xi_\varepsilon}\xi_\varepsilon f(\varepsilon)\,d\varepsilon - q_\eta\phi'\frac{\xi_\Delta^2}{\xi_\varepsilon}f(\hat{\varepsilon}), \qquad (36)$$

---

[19]In some of our examples, e.g. those with a uniform distribution, the density f is non-zero at $\underline{\varepsilon}$, the lower bound of its support, and the left hand derivative of the benefit function, (32) is strictly greater than the right hand derivative, (34). Since optimality requires that the left hand derivative of benefits is greater than or equal to marginal costs, and that the right hand derivative is less than or equal, there is a continuum of equilibria in this case. We focus in these examples upon the equilibria with the smallest investments, i.e. where the right hand derivative equals marginal costs. Given that our results demonstrate overinvestment, investments will only be greater in any other equilibrium.

[20]$B'(\Delta)$ is not differentiable at 0.

[21]The integral defining the function $B''_+(\Delta)$ can be shown to be convergent using a similar argument as employed for $B'(\Delta)$.

31

given $q_{\eta\eta} = 0$ by assumption and $d\hat{\varepsilon}/\Delta = -\xi_\Delta/\xi_\varepsilon$. Note that the first integral is non-positive, due to our finding that $\xi_{\Delta\Delta} \leq 0$. The final term is negative. Turning to the second term, note that the ratio $\xi_\Delta^2/\xi_\varepsilon$ is increasing in ε because $q_{x\varepsilon} \geq 0$ by assumption, $f(\varepsilon)$ is also increasing by assumption, and their product is also increasing. Then, by the second mean value theorem for integrals,[22] there is a $c \in [\underline{\varepsilon}, \hat{\varepsilon}]$ such that the second integral is equal to

$$f(\hat{\varepsilon})\frac{\xi_\Delta^2(\hat{\varepsilon})}{\xi_\varepsilon(\hat{\varepsilon})}\int_c^{\hat{\varepsilon}(\Delta)} q_\eta \phi'' \xi_\varepsilon(\varepsilon)\, d\varepsilon = f(\hat{\varepsilon})\frac{\xi_\Delta^2(\hat{\varepsilon})}{\xi_\varepsilon(\hat{\varepsilon})}\left(q_\eta(x_G^*, \phi(\bar{\varepsilon}))\phi'(\bar{\varepsilon}) - q_\eta(x_G^*, \phi(\xi(c)))\phi'(\xi(c))\right),$$

where $\xi(c) = \xi(x_B^*, \Delta, c)$ and the arguments of ξ_Δ and ξ_ε are similarly abbreviated. That is, we can write the second and third terms of (36) together as

$$-q_\eta(x_G^*, \phi(c))\phi'(c)f(c)\frac{\xi_\Delta^2(x_B^*, \Delta, c)}{\xi_\varepsilon(x_B^*, \Delta, c)} < 0.$$

Thus $B_+''(\Delta) < 0$ if f is increasing.

Consider now case (c) in Assumption A1. By A2, $f'(.)$ and $g'(.)$ are continuous, implying that $B_+''(\Delta)$ is continuous on a compact domain, and thus bounded above by some γ, and we assume $c_B''(x_B^*) > \gamma$. Since $c_B''(.)$ is non-decreasing, this implies $U_+''(\Delta, x_B^*) < 0$ for $\Delta \in [0, b]$.

Downward deviations: We show that given $k > 0$, there exists $\bar{u}$ sufficiently small so that $B'(\Delta) - B'(0) > k\,|\Delta| \;\forall\, \Delta \in [a, 0]$.[23] This holds under Assumptions A2 alone, none of the stronger Assumptions A1 are necessary.

Write the integral from the expression for $B'(\Delta)$ in (34) as $\int \psi(x_G^*, x_B^*, \varepsilon, \Delta)f(\varepsilon)\, d\varepsilon$. Then consider the expression

$$h(\Delta) = \frac{\int_{\tilde{\varepsilon}(\Delta)}^{\bar{\varepsilon}} [\psi(x_G^*, x_B^*, \varepsilon, \Delta) - \psi(x_G^*, x_B^*, \varepsilon, 0)]f(\varepsilon)\, d\varepsilon - k\Delta}{f(\tilde{\varepsilon}(\Delta))},$$

which is well defined on $[a, 0)$ since $f(\tilde{\varepsilon}(\Delta)) > 0$ for $\Delta < 0$.[24] We need to show that $h(\Delta)$ is bounded on $[-a, 0)$. Since the numerator and the denominator in $h(\Delta)$ both converge to 0 as $\Delta \uparrow 0$, by L'Hopital's rule,

$$\lim_{\Delta \uparrow 0} h(\Delta) = \frac{\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \psi_\Delta(x_G^*, x_B^*, \varepsilon, \Delta)f(\varepsilon)\, d\varepsilon - k}{f'(\underline{\varepsilon})\frac{d\tilde{\varepsilon}}{d\Delta}}.$$

---

[22] More specifically, this is the special case known as Bonnet's Theorem, which considers the integral of the product of two functions where one (here $f(\varepsilon)\xi_\Delta^2/\xi_\varepsilon$) is non-negative and increasing. See for example, Bartle (2001, p194).

[23] We show this stronger result since we invoke this for the proof of Theorem 5.

[24] A deviation $\Delta < -(\bar{\varepsilon} - \underline{\varepsilon})$ results in being unmatched for sure, and thus if $\bar{u}$ is lower than the equilibrium payoff, such a deviation is unprofitable.

Since $\psi_\Delta(\varepsilon)f(\varepsilon)$ is a bounded function the integral in the numerator is bounded. Thus $\lim_{\Delta\uparrow 0} h(\Delta)$ exists since as $f'(\underline{\varepsilon}) > 0$ by Assumption A2. This establishes that $h(\Delta)$ is bounded below by some number L on $[a, 0)$. Choose $\bar{u}$ such that $q(x_G^*, \underline{\eta}) - \bar{u} > -2L$. Thus $B'(\Delta) - B'(0) > k\,|\Delta|\,\forall \Delta \in [a, 0)$. Since the cost function is strictly convex, this establishes that no downward deviation is profitable.

We now look at the reaction functions implied by the first order conditions. For any $x_G \in [\bar{x}_G, \tilde{x}_G]$, let $\hat{x}_B$ be a solution to the first order condition for the boys, i.e.

$$\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} q_\eta(x_G, \phi(\varepsilon)) \frac{f(\varepsilon)}{g(\phi(\varepsilon))} \frac{q_x(\hat{x}_B, \varepsilon)}{q_\varepsilon(\hat{x}_B, \varepsilon)} f(\varepsilon)\, d\varepsilon - c_B'(\hat{x}_B) = 0. \tag{37}$$

Let $h\,(x_B, x_G)$ denote the left hand side of the above equation, evaluated at an arbitrary x_B. Since $q_{\eta x} \leq 0$, $\frac{\partial h}{\partial x_G} \leq 0$. Since $q_{xx} \leq 0$ and $q_{\varepsilon x} \geq 0$ and $c_B'' > 0$, $\frac{\partial h}{\partial x_B} < 0$. Thus there is a unique value $\hat{x}_B(x_G)$ that solves the first order condition. By the implicit function theorem, the reaction function for the boys, $\hat{x}_B(x_G)$ is differentiable (and thus continuous), and weakly increasing, since $\frac{\partial h}{\partial x_G} \leq 0$ and $\frac{\partial h}{\partial \hat{x}_B} < 0$. An identical argument establishes that the reaction function for the girls, $\hat{x}_G(x_B)$, is continuous and increasing. Since the reaction functions $[\hat{x}_B(.), \hat{x}_G(.)]$ map $[\bar{x}_B, \tilde{x}_B] \times [\bar{x}_G, \tilde{x}_G]$ to itself, Brouwer's fixed point theorem ensures the existence of a quasi-symmetric equilibrium.

To show uniqueness, we use assumption A3 to re-write the equilibrium first order condition for boys (8) as

$$\frac{\theta_G + (1 - \theta_G)\gamma(x_G^*)}{\theta_B + (1 - \theta_B)\gamma(x_B^*)} \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} [\theta_B + (1 - \theta_B)\gamma'(x_B^*)\varepsilon] \frac{f(\varepsilon)}{g(\phi(\varepsilon))} f(\varepsilon)\, d\varepsilon = c_B'(x_B^*). \tag{38}$$

So if we take the product of the two first order conditions, we have $c_B'(x_B^*)c_G'(x_G^*)$ equal to

$$\int [\theta_G + (1 - \theta_G)\gamma'(x_G^*)\eta] \frac{g(\eta)}{f(\phi^{-1}(\eta))} g(\eta)\, d\eta \int [\theta_B + (1 - \theta_B)\gamma'(x_B^*)\varepsilon] \frac{f(\varepsilon)}{g(\phi(\varepsilon))} f(\varepsilon)\, d\varepsilon. \tag{39}$$

If we have two distinct equilibria, both x_B^* and x_G^* must be higher in one of the equilibria since the reaction functions are increasing. The product of the marginal benefits (39) is decreasing in x_B^* and x_G^* by the concavity of γ, and the product of the marginal costs $c_B'(x_B^*)c_G'(x_G^*)$ is increasing by the convexity of costs. So both equilibria cannot satisfy (39) and we have a contradiction. We have therefore proved the existence and uniqueness of quasi-symmetric equilibrium. ∎

Proof of the Log-Normal Example

Let quality be additive in x_B and ε, for the boys, and in x_G and η, for the girls. Let $h(\varepsilon)$ and $k(\eta)$ be a strictly increasing differentiable functions, such that $h(\varepsilon) \sim N(\mu_\varepsilon, \sigma_\varepsilon)$ and $k(\eta) \sim N(\mu_\eta, \sigma_\eta)$. Let $\tilde{f}$ and $\tilde{g}$ denote the density functions of the shocks, and let f denote

the density function of $h(\varepsilon)$, and g denote the density of $k(\eta)$. The matching $\phi(\varepsilon)$ is defined by

$$\tilde{F}(\varepsilon) = \tilde{G}(\phi(\varepsilon)) \Leftrightarrow F(h(\varepsilon)) = G(k(\phi(\varepsilon))).$$

This implies that

$$\phi'(\varepsilon) = \frac{f(h(\varepsilon))h'(\varepsilon)}{g(k(\phi(\varepsilon))k'(\phi(\varepsilon))}.$$

The first order condition for boys is given by

$$\int_0^\infty \phi'(\varepsilon)\tilde{f}(\varepsilon)\,d\varepsilon = c'_B(x_B^*).$$

Let us now make a change of variables, from ε to h. Since $\tilde{F}(\varepsilon) = F(h(\varepsilon))$, we also have that $\tilde{f}(\varepsilon) = f(h(\varepsilon)) \cdot h'(\varepsilon)$. Furthermore, $dh = h'(\varepsilon)d\varepsilon$. Thus

$$B'(0) = \int_{-\infty}^\infty \frac{f(h(\varepsilon))^2 h'(\varepsilon)}{g(k(\phi(\varepsilon)))k'(\phi(\varepsilon))}\,dh = \frac{\sigma_\eta}{\sigma_\varepsilon}\int_{-\infty}^\infty \frac{f(h(\varepsilon))h'(\varepsilon)}{k'(\phi(\varepsilon))}\,dh.$$

Changing variables again, back from h to ε:

$$B'(0) = \frac{\sigma_\eta}{\sigma_\varepsilon}\int_0^\infty \frac{\tilde{f}(\varepsilon)h'(\varepsilon)}{k'(\phi(\varepsilon))}d\varepsilon.$$

Specializing to the log normal case, suppose $h(\varepsilon) = \ln(\varepsilon)$ and $k(\eta) = \ln\eta$. Thus $h'(\varepsilon) = \frac{1}{\varepsilon}$ and

$$B'(0) = \frac{\sigma_\eta}{\sigma_\varepsilon}\int_0^\infty \frac{\tilde{f}(\varepsilon)\phi(\varepsilon)}{\varepsilon}d\varepsilon = c'_B(x_B^*).$$

Suppose that $\mu_\varepsilon = \mu_\eta = 0$. Since $F(h(\varepsilon)) = G(k(\eta))$, $\ln(\varepsilon)^{\frac{1}{\sigma_\varepsilon}} = \ln(\phi(\varepsilon))^{\frac{1}{\sigma_\eta}}$, so that $\phi(\varepsilon) = \varepsilon^{\frac{\sigma_\eta}{\sigma_\varepsilon}}$. The first order condition for the boys is

$$\frac{\sigma_\eta}{\sigma_\varepsilon}\int_0^\infty \tilde{f}(\varepsilon)\varepsilon^{\left(\frac{\sigma_\eta}{\sigma_\varepsilon}-1\right)}d\varepsilon = c'_B(x_B^*).$$

When ε is log-normally distributed, the expectation of ε^α equals $\exp(\alpha\mu_\varepsilon + \frac{1}{2}\alpha^2\sigma_\varepsilon^2)$. Since $\mu_\varepsilon = 0$,

$$\frac{\sigma_\eta}{\sigma_\varepsilon}\exp\left(\frac{1}{2}\left(\frac{\sigma_\eta - \sigma_\varepsilon}{\sigma_\varepsilon}\right)^2\sigma_\varepsilon^2\right) = \frac{\sigma_\eta}{\sigma_\varepsilon}\exp\left(\frac{1}{2}(\sigma_\eta - \sigma_\varepsilon)^2\right) = c'_B(x_B^*).$$

Let us now consider large deviations. For $\Delta < 0$, with probability $\frac{1}{2}\tilde{F}(\Delta)$ a downward deviation will result is being unmatched, and is unattractive if $\bar{u}$ is low enough. For $\Delta > 0$,

$$B'(\Delta) = \frac{\sigma_\eta}{\sigma_\varepsilon} \int_0^\infty \tilde{f}(\varepsilon)(\varepsilon + \Delta)^{(\frac{\sigma_\eta}{\sigma_\varepsilon}-1)} d\varepsilon.$$

Define $\alpha = \left(\frac{\sigma_\eta}{\sigma_\varepsilon} - 1\right)$. When $\alpha \leq 0$, $B'(\Delta)$ is clearly concave in Δ. So let us consider the case where $\alpha \in (0,1)$, so that the variance of the log of the shocks in the girls is less than twice that in the boys. A second order Taylor expansion of $(\varepsilon + \Delta)^\alpha$ around $\Delta = 0$ yields

$$(\varepsilon + \Delta)^\alpha = \varepsilon^\alpha + \alpha\varepsilon^{\alpha-1}\Delta + \alpha(\alpha-1)\varepsilon^{\alpha-2}\delta^2,$$

for some $\delta \in (0, \Delta)$. Thus,

$$\begin{aligned}
B'(\Delta) &= \frac{\sigma_\eta}{\sigma_\varepsilon}\left[\mathbf{E}(\varepsilon^\alpha) + \alpha\mathbf{E}(\varepsilon^{\alpha-1})\Delta + \alpha(\alpha-1)\mathbf{E}(\varepsilon^{\alpha-2})\delta^2\right], \delta \in [0, \Delta] \\
&\leq \frac{\sigma_\eta}{\sigma_\varepsilon}\left[\mathbf{E}(\varepsilon^\alpha) + \alpha\mathbf{E}(\varepsilon^{\alpha-1})\Delta\right].
\end{aligned}$$

and

$$c'(x_B^* + \Delta) = c'(x_B^*) + c''(x_B^*)\Delta + c'''(x_B^*)\tilde{\delta}^2, \tilde{\delta} \in [0, \Delta].$$

Assume that $c'''(x_B^*) \geq 0$. Thus the problem is concave in Δ if

$$\frac{\sigma_\eta}{\sigma_\varepsilon}\alpha\mathbf{E}(\varepsilon^{\alpha-1})\Delta - c''(x_B^*)\Delta \leq 0.$$

This is satisfied if

$$c''(x_B^*) \geq \frac{\sigma_\eta}{\sigma_\varepsilon}\left(\frac{\sigma_\eta - \sigma_\varepsilon}{\sigma_\varepsilon}\right)\exp\left(\frac{1}{2}(\sigma_\eta - 2\sigma_\varepsilon)^2\right).$$

In the boundary case where $\sigma_\eta = 2\sigma_\varepsilon$ the condition reduces to $c''(x_B^*) \geq 2$. Thus as long as the ratio of the variances is less than 2 and the second derivative of the cost function exceeds 2 at $x \geq x_B^*$, the problem is globally concave.

Proof of Proposition 2: Existence can be established by following the proof of Theorem 1 except in the case of downward deviations by girls. A girl choosing $x_g^* + \Delta$ for some $\Delta < 0$ faces a marginal benefit of

$$B'(\Delta) = \int_{\underline{\eta}}^{\bar{\eta}+\Delta} \frac{g(\eta)}{f(\phi^{-1}(\eta))}g(\eta - \Delta)\, d\eta$$

We can find a distribution $F(\varepsilon)$ sufficiently dispersed in the sense of the dispersion order (see (12) for a definition) so that $f(\cdot)$ is small enough to ensure that $B'(\Delta) > c_G'^*(x_G^* + \Delta)$ for $\Delta < 0$.

35

Turning to the efficiency of investments, we have $f(\tilde{\varepsilon})(\underline{\eta} + x_G^* - \bar{u}) > 0$, reflecting our assumption that the misery effect is strictly positive. Thus, combining the first order conditions (14) and (15), we have

$$c'(x_B^*) \times c'(x_G^*) > \left(\frac{1}{r} \int_{\tilde{\varepsilon}(r)} \frac{f(\varepsilon)}{g(\phi(\varepsilon))} f(\varepsilon)\, d\varepsilon \right) \left(r \int \frac{g(\eta)}{f(\phi^{-1}(\eta))} g(\eta)\, d\eta \right).$$

Making a change of variables, from η to ε results in

$$c'(x_B^*) \times c'(x_G^*) > \frac{1}{r} \left(\int_{\tilde{\varepsilon}(r)} \frac{f(\varepsilon)}{g(\phi(\varepsilon))} f(\varepsilon)\, d\varepsilon \right) \left(\int_{\tilde{\varepsilon}(r)} g(\phi(\varepsilon))\, d\varepsilon \right).$$

By the Cauchy-Schwarz inequality, it follows that

$$\frac{1}{r} \left(\int_{\tilde{\varepsilon}(r)} \frac{f(\varepsilon)}{g(\phi(\varepsilon))} f(\varepsilon)\, d\varepsilon \right) \left(\int_{\tilde{\varepsilon}(r)} g(\phi(\varepsilon))\, d\varepsilon \right) \geq \frac{1}{r} \left[\int_{\tilde{\varepsilon}(r)}^{\bar{\varepsilon}} \left(\frac{f(\varepsilon)}{(g(\phi(\varepsilon)))^{1/2}} \right) (g(\phi(\varepsilon)))^{1/2}\, d\varepsilon \right]^2 = \frac{r^2}{r}.$$

Thus, $c'(x_B^*) \times c'(x_G^*) > r$ while efficiency requires equality. ∎

Proof of Theorem 3: Let f and g be symmetric functions around their means, $\tilde{\varepsilon}$ and $\tilde{\eta}$. Symmetry implies that $f(\tilde{\varepsilon} - \Delta) = f(\tilde{\varepsilon} + \Delta)$ for any Δ. If f and g are both symmetric, then $\phi(\varepsilon)$ and $g(\phi(\varepsilon))$ are also symmetric around $\phi(\tilde{\varepsilon}) = \tilde{\eta}$ and $g(\phi(\tilde{\varepsilon}))$ respectively and $\frac{f(\varepsilon)}{g(\phi(\varepsilon))}$ is also symmetric around $\frac{f(\tilde{\varepsilon})}{g(\phi(\tilde{\varepsilon}))}$. Using these facts,

$$
\begin{aligned}
\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \varepsilon \frac{[f(\varepsilon)]^2}{g(\phi(\varepsilon))} d\varepsilon &= \int_{\underline{\varepsilon}-\tilde{\varepsilon}}^{0} (\tilde{\varepsilon} + \Delta) \frac{[f(\tilde{\varepsilon} + \Delta)]^2}{g(\phi(\tilde{\varepsilon} + \Delta))} d\Delta + \int_{\underline{\varepsilon}-\tilde{\varepsilon}}^{0} (\tilde{\varepsilon} - \Delta) \frac{[f(\tilde{\varepsilon} - \Delta)]^2}{g(\phi(\tilde{\varepsilon} - \Delta))} d\Delta \\
&= 2\tilde{\varepsilon} \int_{\underline{\varepsilon}-\tilde{\varepsilon}}^{0} \frac{[f(\tilde{\varepsilon} + \Delta)]^2}{g(\phi(\tilde{\varepsilon} + \Delta))} d\Delta \\
&= 2\tilde{\varepsilon} \int_{\underline{\varepsilon}}^{\tilde{\varepsilon}} \frac{[f(\varepsilon)]^2}{g(\phi(\varepsilon))} d\varepsilon.
\end{aligned}
$$

Similarly,

$$\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \frac{[f(\varepsilon)]^2}{g(\phi(\varepsilon))}\, d\varepsilon = 2 \int_{\underline{\varepsilon}}^{\tilde{\varepsilon}} \frac{[f(\varepsilon)]^2}{g(\phi(\varepsilon))}\, d\varepsilon.$$

A similar argument and a change of variables yields

$$\int_{\underline{\eta}}^{\bar{\eta}} \eta \frac{[g(\eta)]^2}{f(\phi^{-1}(\eta))}\, d\eta = \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \phi(\varepsilon) g(\phi(\varepsilon))\, d\varepsilon = 2\phi(\tilde{\varepsilon}) \int_{\underline{\varepsilon}}^{\tilde{\varepsilon}} g(\phi(\varepsilon))\, d\varepsilon.$$

36

We may therefore write the product of the equilibrium marginal costs as

$$\left(2\left[\theta_G + (1-\theta_G)\gamma'(x_G^*)\tilde{\eta}\right]\int_{\underline{\varepsilon}}^{\tilde{\varepsilon}} g(\phi(\varepsilon))\,d\varepsilon\right)\left(2\left[\theta_B + (1-\theta_B)\gamma'(x_B^*)\tilde{\varepsilon}\right]\int_{\underline{\varepsilon}}^{\tilde{\varepsilon}} \frac{[f(\varepsilon)]^2}{g(\phi(\varepsilon))}\,d\varepsilon\right). \quad (40)$$

By the Cauchy-Schwarz inequality, the above is weakly greater than

$$4\left[\theta_G + (1-\theta_G)\gamma'(x_G^*)\tilde{\eta}\right]\left[\theta_B + (1-\theta_B)\gamma'(x_B^*)\tilde{\varepsilon}\right]\left(\int_{\underline{\varepsilon}}^{\tilde{\varepsilon}} f(\varepsilon)\,d\varepsilon\right)^2,$$

which is equal to

$$\left[\theta_G + (1-\theta_G)\gamma'(x_G^*)\tilde{\eta}\right]\left[\theta_B + (1-\theta_B)\gamma'(x_B^*)\tilde{\varepsilon}\right].$$

Thus the product of marginal costs is strictly greater than for Pareto efficiency if $\frac{f(\varepsilon)}{\sqrt{g(\phi(\varepsilon))}}$ and $\sqrt{g(\phi(\varepsilon))}$ are linearly independent. ∎

Proof of Theorem 4: We have established that the first order conditions for utilitarian efficiency are identical to the first order conditions for an equilibrium that is gender neutral. Under Assumption A2, the second order conditions for payoff maximization are also satisfied at the utilitarian profile. The first order condition (23) for boys depends only on their investment x_{Bi}^*. Further, the left hand side is increasing and right hand side is decreasing in investment. Thus, there is clearly a unique solution to the first order condition that is independent of the girls' investment. Similarly girls' equilibrium investment does not depend on investment in boys. Thus if large deviations from the utilitarian profile are unprofitable, the utilitarian profile is the unique gender neutral equilibrium.[25]

Given the convexity of the cost function and $q_{xx} \leq 0$, the utilitarian investments x_i^{**} globally maximize the utilitarian payoffs, implying

$$\int q(x,\varepsilon)f_i(\varepsilon)\,d\varepsilon - c_i(x) \leq \int q(x_i^{**},\varepsilon)f_i(\varepsilon)\,d\varepsilon - c_i(x_i^{**})\forall x. \quad (41)$$

Since each type chooses x_i^{**} in any gender neutral equilibrium, the payoff of to any individual of type i equals the right hand side of (41). Thus if the payoff to the individual from deviating to any $x \neq x_i^{**}$ is less than or equal to the left-hand side of (41), no deviation is profitable. We show that this is the case if there are no quality gaps under the utilitarian profile, and if $\bar{u}$ is sufficiently small. Let $C(\tilde{F})$ denote the support of $\tilde{F}$ under the utilitarian profile, which is connected under the no quality gap assumption. If a deviating individual's $q(x,\varepsilon) \in C(\tilde{F})$, then his or her match payoff equals $q(x,\varepsilon)$. If $q(x,\varepsilon) > \max C(\tilde{F})$, then the match payoff equals $\max C(\tilde{F}) < q(x,\varepsilon)$. If $q(x,\varepsilon) < \min C(\tilde{F})$, then the match payoff

---

[25]Note that because of the strong assumption of symmetry between the sexes, we do not need Assumption A3 that was required to assure uniqueness in the general asymmetric case.

equals $\frac{\min C(\tilde{F})+\bar{u}}{2}$. Thus if $\bar{u}$ is sufficiently small, the match payoff is less than equal to $q(x,\varepsilon)$, and no deviation from x_i^{**} is profitable. ∎

Proofs for the Finite Case

Our strategy of proof for our main result, Theorem 5, that shows the existence of a quasi-symmetric equilibrium and its convergence to the quasi-symmetric equilibrium of the continuum model, is as follows. First, we consider the *decision problem* of a boy in a quasi-symmetric equilibrium. That is, we fix a configuration where every girl chooses investment x_G, where every other boy chooses investment x_B, and consider the benefit function of a boy who chooses $x_B + \Delta$. This payoff function is written $B_n(\Delta, x_G)$ (it does not depend upon x_B, but only on Δ). From the individual boy's point of view, x_G is exogenously given, while Δ is a choice variable. We show that this payoff function converges to the payoff function in the continuum case, uniformly in Δ for any given value of x_G. We also show convergence of the first derivative, and that the second derivative in the large finite case is uniformly negative if it is negative in the continuum model.

We then turn to the strategic problem and show that there is a unique solution to the first order conditions for boys and girls, (x_{Bn}^*, x_{Gn}^*), if n is sufficiently large. At this point we need to verify that no individual boy has an incentive to deviate from x_{Bn}^*. We use our uniform convergence results to show that in the resulting decision problem for the boy, $\Delta = 0$ is globally optimal provided that n is large enough.

It might be worthwhile, before proceeding to the proof, to explain some the issues involved in establishing uniform convergence. The matching function in the finite case, $\frac{1}{2}[\phi_{n,n+1}(\varepsilon + \Delta) + \phi_{n+1,n}(\varepsilon + \Delta)]$, converges pointwise to $\phi(\varepsilon + \Delta)$, the matching function in the continuum model (except at a single point, $\underline{\varepsilon}$). However, the finite agent matching functions (see (26) and (27)) are each continuous on their domain, while the continuum matching function is discontinuous to the left at $\underline{\varepsilon}$. Consequently, by Weirstrass's theorem, convergence *cannot* be uniform.[26] Fortunately, we do not require uniform convergence of the matching functions, but of the payoff function $B_n(.)$, which is the integral of the matching functions. The following lemma ensures that establishing pointwise convergence of the matching functions suffices for our purposes.

Let $\xi : \mathbb{R} \to \mathbb{R}$ equal zero except on a compact interval. Without loss of generality, let this interval be $[0,1]$, so that $\xi(x) = 0 \forall x \notin [0,1]$. Let $h : [-a, 1+a] \to \mathbb{R}$, and for $n \in \mathbb{N}$, let $h_n : [-a, 1+a] \to \mathbb{R}$. Assume that each of these functions ξ, h and h_n are measurable and bounded, and let $\bar{\xi} > 0$ be an upper bound for $|\xi|$. For $\Delta \in [-a, a]$, define $W_n(\Delta)$ and

---

[26]This does not preclude uniform convergence of the matching function on a restricted domain where the continuum matching function is continuous, as we show in lemma 3.

$W(\Delta)$ by

$$W_n(\Delta) = \int_0^1 h_n(x + \Delta)\xi(x)\, dx,$$

$$W(\Delta) = \int_0^1 h(x + \Delta)\xi(x)\, dx.$$

Lemma 1 *Assume that for almost any $z \in [-a, 1 + a]$, $h_n(z) \to h(z)$ as $n \to \infty$, i.e. h_n converges pointwise to h almost everywhere on $[-a, 1 + a]$. Assume also that $h_n(z) - h(z)$ is uniformly bounded by H, i.e. $|h_n(z) - h(z)| < H \;\forall n \in \mathbb{N}, \forall z \in [-a, 1 + a]$. Then, W_n converges to $W(\Delta)$ uniformly in Δ.*

Proof. By Egoroff's theorem, given any $\gamma > 0$, there exists a set $A \subset [-a, 1 + a]$ with Lebesgue measure greater than $1 + 2a - \gamma$ such that h_n converges uniformly to h on A. Thus, for any pair $\gamma, \epsilon > 0$, there exists $N : n > N \Rightarrow |h_n(z) - h(z)| < \frac{\epsilon}{\bar{\xi}}$ if $z \in A$.

Let ϵ be given. We show that we can find $N : n > N \Rightarrow |W_n(\Delta) - W(\Delta)| < 2\epsilon \forall \Delta \in [-a, a]$. Choose $\gamma < \frac{\epsilon}{H\bar{\xi}}$. Let A and N satisfy the conditions in the previous paragraph given these values of ϵ, γ. Then,

$$W_n(\Delta) - W(\Delta) = \int_{x+\Delta \in A} [h_n(x + \Delta) - h(x + \Delta)]\,\xi(x)\, dx$$
$$+ \int_{x+\Delta \notin A} [h_n(x + \Delta) - h(x + \Delta)]\,\xi(x)\, dx.$$

Thus,

$$|W_n(\Delta) - W(\Delta)| \leq \int_{x+\Delta \in A} |h_n(x + \Delta) - h(x + \Delta)|\,|\xi(x)|\, dx$$
$$+ \int_{x+\Delta \notin A} |h_n(x + \Delta) - h(x + \Delta)|\,|\xi(x)|\, dx.$$

Since ξ equals zero outside the interval $[0, 1]$ and its absolute value is bounded above by $\bar{\xi}$, $\int_{x+\Delta \in A} |\xi(x)|\, dx$ is bounded above by $\bar{\xi}$. Thus,

$$\int_{x+\Delta \in A} |h_n(x + \Delta) - h(x + \Delta)|\,|\xi(x)|\, dx < \frac{\epsilon}{\bar{\xi}}.\bar{\xi} = \epsilon.$$

Further,

$$\int_{x+\Delta \notin A} |h_n(x + \Delta) - h(x + \Delta)|\,|\xi(x)|\, dx \leq H\bar{\xi}\gamma < \epsilon.$$

This verifies that

$$|W_n(\Delta) - W(\Delta)| < 2\epsilon.$$

39

Since the choice of N and the set A is independent of the value of Δ, we have uniform convergence of $W_n(\Delta)$ to $W(\Delta)$. ∎

We now turn to the expected match value, $\phi_{n+1,n}(\varepsilon + \Delta, x_G)$, and show its convergence to $\phi(\varepsilon + \Delta, x_G)$, where

$$\phi(\varepsilon + \Delta, x_G) = \begin{cases} G^{-1}(F(\varepsilon + \Delta)) + x_G & \text{if } \varepsilon + \Delta \geq \underline{\varepsilon} \\ (\underline{\eta} + x_G + \bar{u})/2 & \text{if } \varepsilon + \Delta < \underline{\varepsilon} \end{cases}. \tag{42}$$

Let there be $n + 1$ boys and n girls. The quality shocks for each of these $2n + 1$ are independently drawn, from the distribution F for the boys and G for the girls. Our focus is on the random variable $\tilde{\phi}_{n+1,n}(\varepsilon + \Delta, x_G)$ that is the match value of a boy with shock realization $\varepsilon + \Delta$. The empirical distribution function corresponding to n iid draws from a distribution, say F, is denoted F_n, and is defined as follows. Let $(\varepsilon_1, \varepsilon_2, .., \varepsilon_n)$ denote a sample realization, and let $1_{(-\infty, t]}(x)$ denote an indicator variable that takes value 1 if $x \leq t$. $F_n(t)$, is defined by

$$F_n(t) = \frac{1}{n} \sum_{i=1}^{n} 1_{(-\infty, t]}(\varepsilon_i).$$

Thus if a boy has quality realization $\varepsilon + \Delta$, his rank in the empirical distribution equals $F_n(\varepsilon + \Delta)$.

Let G_n denote the empirical distribution function corresponding to the sample realizations $(\eta_1, \eta_2, .., \eta_n)$. The empirical quantile function, $G_n^{-1}(q)$ is defined for $q \in (0, 1]$ by

$$G_n^{-1}(q) = \inf\{t : G_n(t) \geq q\}.$$

Fix a type for the boy under consideration, $\varepsilon + \Delta$, and sample realizations $(\varepsilon_1, \varepsilon_2, ..., \varepsilon_n)$ and $(\eta_1, \eta_2, .., \eta_n)$ for the n other boys and the n girls respectively. The match value of this boy is given by

$$\tilde{\phi}_{n+1,n}(\varepsilon + \Delta, x_G) = \begin{cases} G_n^{-1}(F_n(\varepsilon + \Delta)) + x_G \text{ if } F_n(\varepsilon + \Delta) > 0 \\ \bar{u} \text{ if } F_n(\varepsilon + \Delta) = 0. \end{cases}$$

We now show that the random variable $\tilde{\phi}_{n+1,n}(\varepsilon + \Delta, x_G)$ converges weakly to the Dirac measure on $G^{-1}(F(\varepsilon + \Delta)) + x_G$ if $F(\varepsilon + \Delta) \in (0, 1)$. By the Glivenko-Cantelli theorem, $F_n(\varepsilon) \to F(\varepsilon)$ almost surely as $n \to \infty$, uniformly in $\varepsilon + \Delta$. By Lemma 1.2.9 in Reiss (1989, pp 19-20) G_n converges weakly to G if and only if $G_n^{-1}(q) \to G^{-1}(q)$ as $n \to \infty$ at every continuity point of G^{-1}. Since G is continuous and is strictly increasing on its support, $G^{-1}(q)$ is continuous on $(0, 1]$. Since the Glivenko-Cantelli theorem establishes that G_n converges to G almost surely, we conclude that $G_n^{-1}(q) \to G^{-1}(q)$ as $n \to \infty$ for every $q \in (0, 1]$.

Fix $\varepsilon + \Delta \in (\underline{\varepsilon}, \bar{\varepsilon}]$. By Glivenko-Cantelli,

$$\forall \epsilon, \delta_1 > 0, \exists N_1 : n > N_1 \Rightarrow \Pr \left\{ |F_n(\varepsilon + \Delta) - F(\varepsilon + \Delta)| > \delta_1 \right\} < \epsilon.$$

By Lemma 1.2.9 in Reiss, $\forall \epsilon, \delta_1, \delta_2 > 0, \exists N_2 : n > N_2 \Rightarrow$

$$\Pr \left\{ \left| G_n^{-1}(F(\varepsilon + \Delta) - \delta_1) - G^{-1}(F(\varepsilon + \Delta) - \delta_1) \right| > \delta_2 \right\} < \epsilon,$$

and

$$\Pr \left\{ \left| G_n^{-1}(F(\varepsilon + \Delta) + \delta_1) - G^{-1}(F(\varepsilon + \Delta) + \delta_2) \right| > \delta_2 \right\} < \epsilon.$$

Thus if $n > \max\{N_1, N_2\}$

$$\Pr \left\{ G^{-1}(F(\varepsilon + \Delta) - \delta_1) - \delta_2 \le G_n^{-1}(F_n(\varepsilon + \Delta)) \le G^{-1}(F(\varepsilon + \Delta) + \delta_1) + \delta_2 \right\} \ge (1 - \epsilon)^2.$$

This establishes that $\tilde{\phi}_{n,n+1}(\varepsilon + \Delta, x_G)$ converges weakly to a point mass at $G^{-1}(F(\varepsilon + \Delta)) + x_G$ as $n \to \infty$. Thus the expected match value, $\phi_{n,n+1}(\varepsilon + \Delta, x_G)$, converges to $G^{-1}(F(\varepsilon + \Delta)) + x_G$ as $n \to \infty$, for $\varepsilon + \Delta \in (\underline{\varepsilon}, \bar{\varepsilon}]$.

The above argument also applies when there are n boys and $n+1$ girls. The match value in this case is given by

$$\tilde{\phi}_{n,n+1}(\varepsilon + \Delta, x_G) = G_{n+1}^{-1} \left(\frac{(n-1) F_{n-1}(\varepsilon + \Delta) + 1}{n+1} \right) + x_G. \tag{43}$$

Since in $\frac{n-1}{n+1} \to 1$ and $\frac{1}{n+1} \to 0$ as $n \to \infty$, the above argument implies that $\phi_{n,n+1}(\varepsilon + \Delta, x_G)$ converges to $G^{-1}(F(\varepsilon + \Delta)) + x_G$ as $n \to \infty$, for $\varepsilon + \Delta \in (\underline{\varepsilon}, \bar{\varepsilon}]$. Similarly, when there are equal numbers of boys and girls (n),

$$\tilde{\phi}_{n,n}(\varepsilon + \Delta, x_G) = G_n^{-1} \left(\frac{(n-1) F_{n-1}(\varepsilon + \Delta) + 1}{n+1} \right) + x_G. \tag{44}$$

The same argument implies that $\phi_{n,n+1}(\varepsilon + \Delta, x_G)$ converges to $G^{-1}(F(\varepsilon + \Delta)) + x_G$ as $n \to \infty$, for $\varepsilon + \Delta \in (\underline{\varepsilon}, \bar{\varepsilon}]$. Indeed, convergence also obtains at $\underline{\varepsilon}$ since $\phi_{n,n}(\varepsilon + \Delta, x_G) = E\eta_{(1,n)} + x_G$, which converges to $\underline{\eta} + x_G$.

If $\varepsilon + \Delta \ge \bar{\varepsilon}$, then $\phi_{n+1,n}(\varepsilon + \Delta, x_G) = E\eta_{(n,n)} + x_G$ and $\phi_{n,n+1}(\varepsilon + \Delta, x_G) = E\eta_{(n+1,n+1)} + x_G$, both of which converge to $\bar{\eta} + x_G$ as $n \to \infty$ for $\varepsilon + \Delta \ge \bar{\varepsilon}$.

If $\varepsilon + \Delta \le \underline{\varepsilon}$, then $\phi_{n+1,n}(\varepsilon + \Delta, x_G) = \bar{u}$ and $\phi_{n,n+1}(\varepsilon + \Delta, x_G) = E\eta_{(2,n+1)} + x_G$, which converges to $\underline{\eta} + x_G$. Thus $\frac{1}{2}[\phi_{n+1,n}(\varepsilon + \Delta, x_G) + \phi_{n,n+1}(\varepsilon + \Delta, x_G)]$ converges to $(\underline{\eta} + x_G + \bar{u})/2$ if $\varepsilon + \Delta < \underline{\varepsilon}$.

We have therefore proved the following lemma:

Lemma 2 *Fix Δ in $[-a, a]$. Then, for any $\varepsilon \in [\underline{\varepsilon}, \bar{\varepsilon}]$ except $\varepsilon = \underline{\varepsilon} - \Delta$, $\frac{1}{2}[\phi_{n,n+1}(\varepsilon + \Delta, x_G) + \phi_{n+1,n}(\varepsilon + \Delta, x_G)]$ converges pointwise to $\phi(\varepsilon + \Delta, x_G)$ as $n \to \infty$.*

As we have already noted, while one cannot expect uniform convergence of the matching function in general, one can have uniform convergence on a restricted domain.

Lemma 3 *Fix Δ in $[0, a]$. Then $\phi_{n,n+1}(\varepsilon + \Delta, x_G)$ and $\phi_{n,n}(\varepsilon + \Delta, x_G)$ each converge to $\phi(\varepsilon + \Delta, x_G)$ as $n \to \infty$, uniformly in ε for $\varepsilon \in [\underline{\varepsilon}, \bar{\varepsilon}]$. Fix Δ in $[d, a]$ for some $d > 0$. Then $\phi_{n+1,n}(\varepsilon + \Delta, x_G)$ converges to $\phi(\varepsilon + \Delta, x_G)$ uniformly on $[\underline{\varepsilon}, \bar{\varepsilon}]$ as $n \to \infty$.*

Proof. Lemma 2 and its proof establish that $\phi_{n,n+1}(\varepsilon + \Delta, x_G)$ and $\phi_{n,n}(\varepsilon + \Delta, x_G)$ each converge to $\phi(\varepsilon + \Delta, x_G)$ at each value in $[\underline{\varepsilon}, \bar{\varepsilon} + a]$, while $\phi_{n+1,n}(\varepsilon + \Delta, x_G)$ converges to $\phi(\varepsilon + \Delta, x_G)$ at every point strictly greater than $\underline{\varepsilon}$. Each of these functions $\phi_{n,n+1}(\varepsilon + \Delta, x_G)$, $\phi_{n,n}(\varepsilon, x_G)$ and $\phi_{n+1,n}(\varepsilon + \Delta, x_G)$ are strictly increasing and continuous functions, converging to a continuous function on a compact domain. Polya's theorem implies that convergence is uniform (see Bartle (1976, Section 24)). ∎

We move on to the convergence of the benefit function B_n and its derivatives. Recall that $B_n(\Delta, x_G)$ and $B(\Delta, x_G)$ have been defined as:

$$B(\Delta, x_G) = \int_{\underline{\varepsilon} - \Delta}^{\bar{\varepsilon} - \Delta} \phi(\varepsilon + \Delta, x_G)) f(\varepsilon) \, d\varepsilon + [1 - F(\bar{\varepsilon} - \Delta)][\bar{\eta} + x_G] + F(\underline{\varepsilon} - \Delta)\frac{\eta + x_G + \bar{u}}{2}.$$

$$B_n(\Delta, x_G) = \frac{1}{2} \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \left[\phi_{n,n+1}(\varepsilon + \Delta, x_G) + \phi_{n+1,n}(\varepsilon + \Delta, x_G) \right] f(\varepsilon) \, d\varepsilon. \qquad (45)$$

The corresponding derivatives with respect to Δ, evaluated at $\Delta = 0$, equal

$$B'(0, x_G) = \int_{\underline{\varepsilon} - \Delta}^{\bar{\varepsilon}} \phi'(\varepsilon + \Delta, x_G)) f(\varepsilon) \, d\varepsilon + f(\underline{\varepsilon})\frac{\eta + x_G - \bar{u}}{2},$$

$$B'_n(0, x_G) = \frac{1}{2} \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \left[\phi'_{n+1,n}(\varepsilon + \Delta, x_G) + \phi'_{n,n+1}(\varepsilon + \Delta, x_G) \right] f(\varepsilon) \, d\varepsilon. \qquad (46)$$

Lemma 4 *$B_n(\Delta, x_G) \to B(\Delta, x_G)$ uniformly in Δ as $n \to \infty$. $B'_n(0, x_G) \to B'(0, x_G)$ as $n \to \infty$.*

Proof. The function $\frac{1}{2} \left[\phi_{n+1,n}(\varepsilon + \Delta, x_G) + \phi_{n,n+1}(\varepsilon + \Delta, x_G) \right]$ is uniformly bounded – below by $\bar{u}$ and above by $x_G + \bar{\eta}$. Thus Lemma 1 and Lemma 2 imply that $B_n(\Delta, x_G)$ converges to $B(\Delta, x_G)$ uniformly in Δ.

Use integration by parts to re-write the expression for $B'(0, x_G)$ as

$$B'(0, x_G) = -\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \phi(\varepsilon + \Delta) f'(\varepsilon)\, d\varepsilon + (\bar{\eta} + x_G) f(\bar{\varepsilon}) - \left(\frac{\eta + x_G}{2}\right) f(\underline{\varepsilon}). \qquad (47)$$

Similarly, using integration by parts, $B'_n(\Delta)$ can be re-written as

$$B'_n(0, x_G) = \begin{array}{l} -\frac{1}{2}\int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \left[\phi_{n,n+1}(\varepsilon, x_G) + \phi_{n+1,n}(\varepsilon, x_G)\right] f'(\varepsilon)\, d\varepsilon \\ +\frac{1}{2}\left[\phi_{n+1,n}(\bar{\varepsilon}, x_G) + \phi_{n,n+1}(\bar{\varepsilon}, x_G)\right] f(\bar{\varepsilon}) \\ -\frac{1}{2}\left[\phi_{n+1,n}(\underline{\varepsilon}, x_G) + \phi_{n,n+1}(\underline{\varepsilon}, x_G)\right] f(\underline{\varepsilon}). \end{array} \qquad (48)$$

Since $f'(\varepsilon)$ is bounded, Lemma 1 and Lemma 2 imply that the first term in (48) converges to the first term in (47). Lemma 2 implies that the second term in (48) converges to the second term term in (47). Since $f(\underline{\varepsilon}) = 0$, the third terms in (48) and (47) both equal zero. ∎

The second derivative on $(0, b]$ is given by:

$$B''_+(\Delta) = \int_{\underline{\varepsilon}-\Delta}^{\bar{\varepsilon}-\Delta} \phi''(\varepsilon + \Delta) f(\varepsilon)\, d\varepsilon - \phi'(\bar{\varepsilon}) f(\bar{\varepsilon} - \Delta). \qquad (49)$$

On $[-a, 0)$, it equals

$$B''_-(\Delta) = \int_{\underline{\varepsilon}-\Delta}^{\bar{\varepsilon}-\Delta} \phi''(\varepsilon + \Delta) f(\varepsilon)\, d\varepsilon + \phi'(\underline{\varepsilon}) f(\underline{\varepsilon} - \Delta) - f'(\underline{\varepsilon} - \Delta)\frac{x_G + \eta - \bar{u}}{2}. \qquad (50)$$

We define $B''_+(0)$ and $B''_-(0)$ by:

$$B''_+(0) = \lim_{\Delta \downarrow 0} B''_+(\Delta) = \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \phi''(\varepsilon) f(\varepsilon)\, d\varepsilon - \phi'(\bar{\varepsilon}) f(\bar{\varepsilon}), \qquad (51)$$

$$B''_-(0) = \lim_{\Delta \uparrow 0} B''_-(\Delta) = \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \phi''(\varepsilon) f(\varepsilon)\, d\varepsilon - f'(\underline{\varepsilon})\frac{x_G + \eta - \bar{u}}{2} \qquad (52)$$

Note that $B''_+(0)$ is not in general equal to $B''_-(0)$. The second derivatives in the finite case are

$$B''_{+n}(\Delta, x_G) = \frac{1}{2}\int_{\underline{\varepsilon}}^{\bar{\varepsilon}-\Delta} \left[\phi''_{n+1,n}(\varepsilon + \Delta, x_G) + \phi''_{n,n+1}(\varepsilon + \Delta, x_G)\right] f(\varepsilon)\, d\varepsilon, \Delta \geq 0. \qquad (53)$$

$$B''_{-n}(\Delta, x_G) = \frac{1}{2}\int_{\underline{\varepsilon}-\Delta}^{\bar{\varepsilon}} \left[\phi''_{n+1,n}(\varepsilon + \Delta, x_G) + \phi''_{n,n+1}(\varepsilon + \Delta, x_G)\right] f(\varepsilon)\, d\varepsilon, \Delta \leq 0. \qquad (54)$$

Our strategy is to show that if the global optimality conditions are satisfied in the continuum case, then they are also satisfied in the large finite case. For upward deviations, we show that the second derivative converges uniformly to the continuum derivative plus a negative term. For downward deviations in some interval, we do this by showing the uniform convergence of the second derivative to that in the continuum model. At this point, for the convergence of $B''_{-n}(\Delta, x_G)$, we invoke an additional assumption, $f(\bar{\varepsilon}) = 0$.

Lemma 5 *For any $k > 0, \exists N : n > N \Rightarrow B''_+(\Delta, x_G) < B''_+(\Delta, x_G) + k$ for any $\Delta \in (0, a]$. There exists $d > 0$ such that $B''_{-n}(\Delta, x_G) \to B''_-(\Delta, x_G)$ uniformly in Δ for $\Delta \in [-d, 0]$ as $n \to \infty$ if $f(\bar{\varepsilon}) = 0$.*

Proof. We show that $B''_{+n}(\Delta, x_G)$ can be written as the sum of two terms, where the first converges to $B''_+(\Delta, x_G)$ uniformly in Δ and the second is negative. By iterated integration by parts, we obtain

$$B''_+(\Delta) = \int_{\underline{\varepsilon}}^{\bar{\varepsilon} - \Delta} \phi(\varepsilon + \Delta) f''(\varepsilon) \, d\varepsilon - \phi(\bar{\varepsilon}) f'(\bar{\varepsilon} - \Delta) + \phi(\underline{\varepsilon} + \Delta, x_G) f'(\underline{\varepsilon}), \tag{55}$$

$$\begin{aligned} B''_{+n}(\Delta, x_G) &= \frac{1}{2} \int_{\underline{\varepsilon}}^{\bar{\varepsilon} - \Delta} \left[\phi_{n+1,n}(\varepsilon + \Delta, x_G) + \phi_{n,n+1}(\varepsilon + \Delta, x_G) \right] f''(\varepsilon) \, d\varepsilon \\ &\quad - \frac{1}{2} [\phi_{n+1,n}(\bar{\varepsilon}, x_G) + \phi_{n,n+1}(\bar{\varepsilon}, x_G)] f'(\bar{\varepsilon} - \Delta) \\ &\quad + \frac{1}{2} [\phi_{n+1,n}(\underline{\varepsilon} + \Delta, x_G) + \phi_{n,n+1}(\underline{\varepsilon} + \Delta, x_G)] f'(\underline{\varepsilon}). \end{aligned} \tag{56}$$

Since $f''(\varepsilon)$ is continuous on the support of f, it is bounded. Thus the first term of (56) (in the first line) converges to the first term in (55). Turning to the second terms in the two expressions, convergence follows since $\phi_{n+1,n}(\bar{\varepsilon}, x_G)$ and $\phi_{n,n+1}(\bar{\varepsilon}, x_G)$ each converge to $\phi(\bar{\varepsilon}, x_G)$. These terms are multiplied by $f'(\bar{\varepsilon} - \Delta)$, but since f' is continuous and therefore bounded, the convergence is also uniform in Δ.

This leaves the final term in (55), $\frac{1}{2}[\phi_{n+1,n}(\underline{\varepsilon} + \Delta, x_G) + \phi_{n,n+1}(\underline{\varepsilon} + \Delta, x_G)]$. This can be expressed as the sum of a term that converges to $\phi(\underline{\varepsilon} + \Delta, x_G)$ uniformly in Δ, and a second term that is negative. Rewrite $\phi_{n+1,n}(\underline{\varepsilon} + \Delta, x_G)$ as:

$$\phi_{n+1,n}(\underline{\varepsilon} + \Delta, x_G) = F(\varepsilon + \Delta) \phi_{n,n}(\underline{\varepsilon} + \Delta, x_G) + [1 - F(\varepsilon + \Delta)] \hat{\phi}_{n,n-1}(\underline{\varepsilon} + \Delta, x_G).$$

This expression is derived as follows. Fix a shock value $\underline{\varepsilon} + \Delta$ and an arbitrary boy among the *other* n boys. With probability $F(\underline{\varepsilon} + \Delta)$ this other boy has a shock realization below $\underline{\varepsilon} + \Delta$; in this event, the expected match value of type $\underline{\varepsilon} + \Delta$ is $\phi_{n,n}(\underline{\varepsilon} + \Delta, x_G)$, i.e. the expectation of $\tilde{\phi}_{n,n}(\underline{\varepsilon} + \Delta, x_G)$. With probability $1 - F(\underline{\varepsilon} + \Delta)$, the other boy has a shock

realization above $\underline{\varepsilon} + \Delta$. In this event, the random match value of type $\underline{\varepsilon} + \Delta$ is less than $\tilde{\phi}_{n,n}(\underline{\varepsilon}+\Delta, x_G)$ in the sense of first order stochastic dominance. Let $\hat{\phi}_{n,n-1}(\underline{\varepsilon}+\Delta, x_G)$ denote the expected match value in this case, which is (strictly) less than $\phi_{n,n}(\underline{\varepsilon} + \Delta, x_G)$. Using this, we rewrite

$$
\begin{aligned}
\phi_{n+1,n}(\underline{\varepsilon} + \Delta, x_G) + \phi_{n,n+1}(\underline{\varepsilon} + \Delta, x_G) &= \phi_{n+1,n}(\underline{\varepsilon} + \Delta, x_G) + \phi_{n,n}(\underline{\varepsilon} + \Delta, x_G) \\
&\quad -[1 - F(\underline{\varepsilon} + \Delta)][\phi_{n,n}(\underline{\varepsilon} + \Delta, x_G) - \hat{\phi}_{n,n-1}(\underline{\varepsilon} + \Delta, x_G)].
\end{aligned}
$$

By lemma 3, each of the terms in the first line converge uniformly to $\phi(\underline{\varepsilon}+\Delta, x_G)$, while the term in the second line is negative. We have shown that $B''_{+n}(\Delta, x_G)$ can be written as

$$
B''_{+n}(\Delta, x_G) = \tilde{B}''_{+n}(\Delta, x_G) - [1 - F(\underline{\varepsilon} + \Delta)][\phi_{n,n}(\underline{\varepsilon} + \Delta, x_G) - \hat{\phi}_{n,n-1}(\underline{\varepsilon} + \Delta, x_G)],
$$

where $\tilde{B}''_{+n}(\Delta, x_G)$ converges to $B''_+(\Delta, x_G)$ uniformly in Δ. Thus for any $k > 0$, $\exists N : n > N \Rightarrow \left| \tilde{B}''_{+n}(\Delta, x_G) - B''_+(\Delta, x_G) \right| < k$, implying that $B''_{n+}(\Delta, x_G) < B''_+(\Delta, x_G) + k$.

Turning to $B''_-(\Delta)$ and $B''_{-n}(\Delta, x_G)$, iterated integration by parts yields

$$
B''_-(\Delta) = \int_{\underline{\varepsilon}-\Delta}^{\bar{\varepsilon}} \phi(\varepsilon + \Delta) f''(\varepsilon) \, d\varepsilon + f'(\underline{\varepsilon} - \Delta) \left(\frac{x_G + \underline{\eta} + \bar{u}}{2} \right) - \phi(\bar{\varepsilon} + \Delta) f'(\bar{\varepsilon}) + \phi'(\bar{\varepsilon} + \Delta) f(\bar{\varepsilon}).
\tag{57}
$$

$$
\begin{aligned}
B''_{-n}(\Delta, x_G) &= \frac{1}{2} \int_{\underline{\varepsilon}-\Delta}^{\bar{\varepsilon}} \left[\phi_{n+1,n}(\varepsilon + \Delta, x_G) + \phi_{n,n+1}(\varepsilon + \Delta, x_G) \right] f''(\varepsilon) \, d\varepsilon \tag{58} \\
&\quad + \frac{1}{2} [\phi_{n+1,n}(\underline{\varepsilon}, x_G) + \phi_{n,n+1}(\underline{\varepsilon}, x_G)] f'(\underline{\varepsilon} - \Delta) \\
&\quad - \frac{1}{2} [\phi_{n+1,n}(\bar{\varepsilon} + \Delta, x_G) + \phi_{n,n+1}(\bar{\varepsilon} + \Delta, x_G)] f'(\bar{\varepsilon})) \\
&\quad + \frac{1}{2} [\phi'_{n+1,n}(\bar{\varepsilon} + \Delta, x_G) + \phi'_{n,n+1}(\bar{\varepsilon} + \Delta, x_G)] f(\bar{\varepsilon}). \tag{59}
\end{aligned}
$$

Uniform convergence of the first two terms in (58) to the first two terms in (57) is by the same argument as for $B''_+(.)$. Uniform convergence of the third term in (58) to the third term in (57) follows from lemma 3. Given the assumption that $f(\bar{\varepsilon}) = 0$, the fourth terms in (58) and (57) both equal zero. ∎

At this point, we could move directly to the proof of the main theorem if we assume that f and g equal zero at the upper bound of their supports. This assumption is not necessary, but relaxing it requires showing that $\phi'_{n,n+1}(\bar{\varepsilon} + \Delta, x_G)$ converges to $\phi'(\bar{\varepsilon} + \Delta, x_G)$ uniformly for Δ belonging to some interval $[-d, 0]$. This is shown in the following lemma.

Lemma 6 *There exists $d > 0$ such that for $\Delta \in [-d, 0]$, $\phi'_{n+1,n}(\bar{\varepsilon} + \Delta, x_G)$ and $\phi'_{n,n+1}(\bar{\varepsilon} +$*

$\Delta, x_G)$ *converge uniformly in Δ to $\phi'(\bar{\varepsilon} + \Delta, x_G)$ as $n \to \infty$.*

Proof. Let $d \in (0, \bar{\varepsilon} - \underline{\varepsilon})$, so that $\bar{\varepsilon} + \Delta \in [\underline{\varepsilon}, \bar{\varepsilon}]$. We focus on $\phi'_{n+1,n}(\bar{\varepsilon} + \Delta, x_G)$ since the argument for $\phi'_{n,n+1}(\bar{\varepsilon} + \Delta, x_G)$ is almost identical. This can be written as

$$
\begin{aligned}
\phi'_{n+1,n}(\bar{\varepsilon} + \Delta, x_G) &= f(\bar{\varepsilon} + \Delta) \sum_{i=2}^{n} F_i^n(\bar{\varepsilon} + \Delta) n \left[E\eta_{(i,n)} - E\eta_{(i-1,n)} \right] \qquad (60) \\
&+ f(\bar{\varepsilon} + \Delta) n (1 - F(\bar{\varepsilon} + \Delta))^{n-1} (E\eta_{(1,n)} + x_G - \bar{u}) > 0.
\end{aligned}
$$

Consider first the final term in the above expression. For any n, $n\left[1 - F(\bar{\varepsilon} + \Delta)\right]^n \leq n\left[1 - F(\bar{\varepsilon} - d)\right]^n$, and $n\left[1 - F(\bar{\varepsilon} + \Delta)\right]^n \to 0$ as $n \to \infty$ for every $\Delta \in [-d, 0]$. Since $\Delta = d$ provides an upper bound, the convergence is uniform in Δ.

Turning to the summation terms in (60), $f(\bar{\varepsilon} + \Delta)$ enters both $\phi'_{n+1,n}(\bar{\varepsilon} + \Delta)$ and $\phi'(\bar{\varepsilon} + \Delta)$, so it suffices to show that

$$
\sum_{i=2}^{n} F_i^n(\bar{\varepsilon} + \Delta) n \left[E\eta_{(i,n)} - E\eta_{(i-1,n)} \right] \to \frac{1}{g(G^{-1}(F(\bar{\varepsilon} + \Delta)))}
$$

as $n \to \infty$ uniformly in Δ.

Let $i(n) = (n+1)p$. We now show that the expression $n[E\eta_{(i(n),n)} - E\eta_{(i(n)-1,n)}]$ converges to $1/g(G^{-1}(p))$ as $n \to \infty$ uniformly in p for p belonging to an interval $[\tilde{p}, 1]$ where $g(G^{-1}(p))$ is bounded away from zero. Assume that $g(G^{-1}(1)) = g(\bar{\eta}) > 0$. Since g is continuous, there exists an interval $[\hat{p}, 1]$ such that $g(G^{-1}(p)) > \bar{g} > 0 \, \forall \, p \in [\hat{p}, 1]$, and let $\tilde{p} \in (\hat{p}, 1)$. From Arnold et al. (1992, p128) we have,

$$
E\eta_{i,n} = G^{-1}(p) + \frac{p(1-p)}{2(n+2)} \frac{d^2 G^{-1}(p)}{du^2} + O\left(\frac{1}{n^2}\right),
$$

where $dG^{-1}(p)/du$ is the derivative of $G^{-1}(u)$ evaluated at p. Furthermore, the terms that are $O\left(\frac{1}{n^2}\right)$ are so uniformly in p for $p \in [0,1]$. Now,

$$
\begin{aligned}
n \left[E\eta_{i,n} - E\eta_{i-1,n} \right] &= n \left(G^{-1}(p) - G^{-1}(p - \tfrac{1}{n+1}) \right) + n \frac{d^2 G^{-1}(p - \frac{1}{n+1})}{du^2} \left(\frac{\frac{(1-2p)}{n+1} - \left(\frac{1}{n+1}\right)^2}{2(n+2)} \right) \\
&+ n \frac{p(1-p)}{2(n+2)} \left(\frac{d^2 G^{-1}(p)}{du^2} - \frac{d^2 G^{-1}(p - \frac{1}{n+1})}{du^2} \right) + n O\left(\tfrac{1}{n^3}\right).
\end{aligned} \qquad (61)
$$

We first show that the first term on the right hand side converges uniformly to $1/g(G^{-1}(p))$ as $n \to \infty$ uniformly in p for $p \in [\tilde{p}, 1]$. By the mean value theorem, $\exists \, a \in (p - 1/(n+1), p)$ such that $[1/g(G^{-1}(a))]\frac{1}{n+1} = G^{-1}(p) - G^{-1}(p - 1/(n+1))$. Since $1/g(G^{-1}(.))$ is continuous and since $\frac{1}{n+1} \to 0$ as $n \to \infty$, $[1/g(G^{-1}(a))]\frac{n}{n+1} \to 1/g(G^{-1}(p))$ as $n \to \infty$. Convergence is uniform in p for $p \in [\tilde{p},]$ since $1/g(G^{-1}(.))$ is uniformly continuous on the compact interval

46

$[\hat{p}, 1]$.

Let us now turn to other terms in (61) to show that they go to zero as $n \to \infty$ uniformly in p, for $p \in (\tilde{p}, 1)$. This is true for $n \left(\frac{\frac{(1-2p)}{n+1} - \left(\frac{1}{n+1}\right)^2}{2(n+2)} \right)$ since it is of order $\frac{1}{n}$, and $p = 1$ provides an upper bound for its absolute value for large n. Turning to the third term, our lower bound on g implies that $\left(\frac{d^2 G^{-1}(p)}{du^2} - \frac{d^2 G^{-1}(p - \frac{1}{n+1})}{du^2} \right)$ is $O(1/n)$ uniformly in p, and thus this term also converges to zero as $n \to \infty$ uniformly in p, for $p \in (\tilde{p}, 1)$. The final term is of order $\frac{1}{n^3}$ since the difference between the two $O(\frac{1}{n^2})$ terms is $O\left(\frac{1}{n^3}\right)$.

Consider any $p \in (\tilde{p}, 1)$ and a sequence $(i(n), n)$, where $i(n) = [np] + 1$. We have verified that $n[E\eta_{i,n} - E\eta_{i-1,n}] \to \frac{1}{g(G^{-1}(p))}$ as $n \to \infty$, uniformly in p for $p \in (\tilde{p}, 1)$. Finally, by the Glivenko-Cantelli theorem, and as noted in the proof of Lemma 2, $F_n(\varepsilon + \Delta)$ converges to the Dirac measure on $F(\varepsilon + \Delta)$. Thus $\phi'_{n+1,n}(\bar{\varepsilon} + \Delta) \to f(\varepsilon + \Delta)/g(G^{-1}(F(\bar{\varepsilon} + \Delta)))$ uniformly in Δ for $\Delta \in [-d, 0]$. ∎

Proof of Theorem 5. We show first that there exists a unique solution to the first order conditions if n is sufficiently large. Define $\hat{x}_{Bn}(x_G)$ and $\hat{x}_{Gn}(x_B)$ by

$$c'_B(\hat{x}_{Bn}) = B'_n(0, x_G),$$

$$c'_G(\hat{x}_{Gn}) = \tilde{B}'_n(0, x_B),$$

where $\tilde{B}'_n(0, x_B)$ denotes the marginal benefit function for the girls. Since $c'(\bar{x}_B) = 0$ and $c'(x_B) \to \infty$ as $x_B \to \tilde{x}_B$, and $B'_n(0, x_G) > 0$, $\hat{x}_{Bn}(x_{Gn})$ is uniquely defined. Its derivative is

$$\frac{d\hat{x}_{Bn}}{dx_G} = \frac{\partial B'_n(0, x_G)}{\partial x_G} \frac{1}{c''_B(\hat{x}_{Bn})}.$$

Since $c''_B(.) > 0$ and is continuous, it is bounded below on $[\bar{x}_B, \tilde{x}_B]$ by some $\gamma_B > 0$. Similarly, $c''_G(.)$ is bounded below on $[\bar{x}_G, \tilde{x}_G]$ by $\gamma_G > 0$. Let $\gamma = \min\{\gamma_B, \gamma_G\}$. By differentiating (46) with respect to x_G, we obtain

$$\frac{\partial B'_n(0, x_G)}{\partial x_G} = n \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} f^2(\varepsilon)(1 - F(\varepsilon))^{n-1} \, d\varepsilon.$$

Define $y(\varepsilon) = -f(\varepsilon)(1 - F(\varepsilon))^n$, so that

$$\frac{dy}{d\varepsilon} = nf^2(\varepsilon)(1 - F(\varepsilon))^{n-1} - f'(\varepsilon)(1 - F(\varepsilon))^n.$$

47

Thus,

$$
\begin{aligned}
\frac{\partial B'_n(0, x_G)}{\partial x_G} &= \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} \frac{dy}{d\varepsilon} \, d\varepsilon + \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} f'(\varepsilon)(1 - F(\varepsilon))^n \, d\varepsilon \\
&= [-f(\varepsilon)(1 - F(\varepsilon))^n]_{\underline{\varepsilon}}^{\bar{\varepsilon}} + \int_{\underline{\varepsilon}}^{\bar{\varepsilon}} f'(\varepsilon)(1 - F(\varepsilon))^n \, d\varepsilon.
\end{aligned}
$$

The first term equals zero since $f(\underline{\varepsilon}) = 0$. In the second term, the function being integrated is uniformly bounded, and converges point-wise to zero as $n \to \infty$, and thus the integral also converges to zero. Thus $\frac{\partial B'_n(0, x_G)}{\partial x_G} \to 0$ as $n \to \infty$, and there exists N_B such that $n > N_B \Rightarrow \frac{\partial B'_n(0, x_G)}{\partial x_G} < \frac{\gamma}{2}$.

Since the argument is identical for the girls, there exists $N_G : n > N_G \Rightarrow \frac{\partial \tilde{B}'_n(0, x_B)}{\partial x_B} < \frac{\gamma}{2}$. Consider the space $[\bar{x}_G, \tilde{x}_G] \times [\bar{x}_B, \tilde{x}_B] \subset \mathbb{R}^2$, equipped with the Euclidean metric. If $n > \max\{N_G, N_B\}$, the mapping $(\hat{x}_{Bn}, \hat{x}_{Gn})$ from $[\bar{x}_B, \tilde{x}_B] \times [\bar{x}_G, \tilde{x}_G]$ to itself is a contraction. Thus there exists a unique fixed point, (x^*_{Bn}, x^*_{Gn}), where the first order conditions are satisfied.

Consider the sequence $(x^*_{Bn}, x^*_{Gn})_{n=\max\{N_G, N_B\}+1}$. Since $B'_n(0, x^*_{Gn}) \to B'(0)$ as $n \to \infty$, $x^*_{Bn} \to x^*_B$ as $n \to \infty$. Similarly, $x^*_{Gn} \to x^*_G$ as $n \to \infty$.

To complete the proof, we now show that no deviation from x^*_{Bn} is profitable for n sufficiently large. Recall that U''_+ is defined by $U''_+(\Delta, x^*_B) = B''_+(\Delta) - c''_B(x^*_B + \Delta)$, and that $\gamma > 0$ is a lower bound for c''_B on $[\bar{x}_B, \tilde{x}_B]$. In the proof of Theorem 1, we established that under Assumptions A1 and A2, $\exists \rho > 0 : B''_+(\Delta, x^*_B) - \gamma < -2\rho$ if $\Delta \in [0, b]$. Recall from the proof of lemma 5 that $B''_{+n}(\Delta) < \tilde{B}''_{+n}(\Delta)$ and that $\tilde{B}''_{+n}(\Delta)$ converges uniformly to $B''_+(\Delta)$. Let n be sufficiently large that $\left| B''_+(\Delta) - \tilde{B}''_{+n}(\Delta, x^*_{Gn}) \right| < \rho$, consider a $\tilde{\Delta} \in [0, b]$ deviation from x^*_{Bn}. Since

$$
[B'_n(\tilde{\Delta}, x^*_{Gn}) - c'(x^*_{Bn} + \tilde{\Delta})] - [B'_n(0, x^*_{Gn}) - c'(x^*_{Bn})] \leq \int_0^{\tilde{\Delta}} \left(B''_{+n}(\Delta, x^*_{Gn}) - \gamma_B \right) d\Delta < -\rho\tilde{\Delta},
$$

no upward deviation is profitable.

We now turn to downward deviations. As in the proof of Theorem 1 we can choose $\bar{u}$ sufficiently small such that $B'(\Delta, x_G) - B'(0) > k |\Delta|$ for some $k > 0$. Given this, we show that $B'_n(\tilde{\Delta}, x^*_{Gn}) - B'_n(0, x^*_{Gn}) > 0$ for $\tilde{\Delta} \in [-d, 0)$, which establishes the unprofitability of downward deviations in this set, due the convexity of the cost function.

$$
B'_n(\tilde{\Delta}, x^*_{Gn}) - B'_n(0, x^*_{Gn}) = \int_0^{\tilde{\Delta}} B''_{-n}(\Delta, x^*_{Gn}) d\Delta.
$$

Since $B''_{-n}(\Delta, x^*_{Gn})$ converges uniformly to $B''_-(\Delta, x^*_{Gn})$, we may choose $N_1 : n > N_1 \Rightarrow$

$\left| B''_{-n}(\Delta, x^*_{Gn}) - B''_-(\Delta, x^*_{Gn}) \right| < \frac{k}{4}$. Since $x^*_{Gn} \to x^*_G$ and $n \to \infty$, and since $B''_-(\Delta, x_G)$ is uniformly continuous in x_G on the compact interval $[-d, 0]$, we may therefore choose $N_2 : n > N_2 \Rightarrow \left| B''_-(\Delta, x^*_{Gn}) - B''_-(\Delta, x^*_G) \right| < \frac{k}{4}$. Thus, if $n > \max\{N_1, N_2\}$,

$$
\begin{aligned}
\int_0^{\tilde{\Delta}} B''_{-n}(\Delta, x^*_{Gn}) d\Delta &= \int_0^{\tilde{\Delta}} \left[B''_{-n}(\Delta, x^*_{Gn}) - B''_-(\Delta, x^*_{Gn}) \right] d\Delta \\
&\quad - \int_0^{\tilde{\Delta}} \left[B''_-(\Delta, x^*_G) - B''_-(\Delta, x^*_{Gn}) \right] d\Delta + \int_0^{\tilde{\Delta}} B''_-(\Delta, x^*_G) \, d\Delta \\
&\geq -\frac{k}{4} \left| \tilde{\Delta} \right| - \frac{k}{4} \left| \tilde{\Delta} \right| + k \left| \tilde{\Delta} \right| > 0.
\end{aligned}
$$

Thus, $B'_n(\tilde{\Delta}, x^*_{Gn}) > B'_n(0, x^*_{Gn})$ and no small downward deviation in the set $[-d, 0)$ is profitable.

The final step is to show that no large downward deviation, $\Delta < -d$, is profitable. This follows from the uniform convergence of the payoff function. In the proof of Theorem 1, we established that in the continuum model, the loss in payoff from a downward deviation $\Delta < -d$ is bounded away from zero if $\bar{u}$ is sufficiently small. Since $B_n(.)$ converges uniformly to $B(.)$, since $(x^*_{Gn}, x^*_{Bn}) \to (x^*_G, x^*_B)$ as $n \to \infty$, and since $B_n(\Delta, x_G)$ is continuous in x_G, $\exists N : n > N \Rightarrow B_n(\Delta, x^*_{Gn}) - c(x^*_{Bn} + \Delta) < B_n(0, x^*_{Gn}) - c(x^*_{Bn})$ for $\Delta < -d$. This concludes the proof of the Theorem. ∎

References

Acemoglu, Daron, and Robert Shimer (1999) "Holdup and efficiency with search frictions", *International Economic Review*, 40, 827-849.

Anderson, Siwan (2003) "Why dowry payments declined with modernization in Europe but are rising in India", *Journal of Political Economy*, 111-2, 269-310.

Anderson, Siwan (2007) "Economics of brideprice and dowry", *Journal of Economic Perspectives*, 21 (4), 151-174.

Arnold, Barry C., N. Balakrishnan and H.N. Nagaraja (1992) *A First Course in Order Statistics*, Wiley.

Baron-Cohen, S. (2003) *The Essential Difference: the Truth about the Male and Female Brain*, Basic Books.

Bartle, Robert G. (2001) *A Modern Theory of Integration.* American Mathematical Society.

Becker, Gary S. (1973) "A theory of marriage: part 1", *Journal of Political Economy*, 81, 813-846.

Becker, Gary S., Hubbard, William H.J. and Murphy, Kevin M. (2010) "Explaining the Worldwide Boom in Higher Education of Women", *Journal of Human Capital*, 4(3), 203-241.

Becker, Gary S. and Nigel Tomes (1979) "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility", *Journal of Political Economy*, 87 (6), 1153-1189.

Bhaskar, V. (2011) "Sex Selection and Gender Balance", *American Economic Journal: Microeconomics* , 3(1): 214-44.

Burdett, Kenneth, and Melvyn Coles (2002) "Transplants and Implants: The Economics of Self Improvement", *International Economic Review*, 42 (3), 597-616.

Chiappori, Pierre-André, Murat Iyigun and Yoram Weiss (2009) "Investment in Schooling and the Marriage Market", *American Economic Review*, 99, 1689-1713.

Choo, Eugene and Aloysius Siow (2006) "Who Marries Whom and Why", *Journal of Political Economy*, 114, 175-201.

Cole, Harold L., George J. Mailath and Andrew Postlewaite (1992) "Social norms, savings behavior, and growth", *Journal of Political Economy*, 100 (6), 1092-1125.

Cole, Harold L., George J. Mailath and Andrew Postlewaite (2001) "Efficient non-contractible investments in large economies", *Journal of Economic Theory*, 101, 333-373.

Dagsvik, John K. (2000) "Aggregation In Matching Markets", *International Economic Review*, 41, 27-57.

Felli, Leonardo and Kevin W.S. Roberts (2001) "Does competition solve the hold-up problem?", working paper.

Fine, Cordelia (2008) "Will Working Mothers' Brains Explode? The Popular New Genre of Neurosexism", *Neuroethics*, 1, 69-72.

Gall, Thomas, Patrick Legros and Andrew Newman (2009) "Mis-match, Re-match, and Investment", working paper.

Hajnal, John (1982) "Two Kinds of Preindustrial Household Formation System", *Population and Development Review*, 8 (3), 449-494.

Hopkins, Ed (2012) "Job Market Signalling of Relative Position or Becker Married to Spence", *Journal of the European Economic Association*, 10(2), 290-322.

Hopkins, Ed and Tatiana Kornienko (2004), "Running to Keep in the Same Place: Consumer Choice as a Game of Status", *American Economic Review*, 94 (4), 1085-1107.

Hopkins, Ed and Tatiana Kornienko (2010) "Which Inequality? The Inequality of Endowments Versus the Inequality of Rewards", *American Economic Journal: Microeconomics*, 2(3), 106-137.

Hoppe, Heidrun C., Benny Moldavanu and Aner Sela (2009) "The theory of assortative matching based on costly signals", *Review of Economic Studies*, 76, 253-281.

Iyigun, Murat and Randall P. Walsh (2007) "Building the Family Nest: Premarital Investments, Marriage Markets, and Spousal Allocations", *Review of Economic Studies* 74, 507-535.

Lazear, Edward P. and Sherwin Rosen (1981) "Rank-order tournaments as optimum labor contracts", *Journal of Political Economy*, 89 (5), 841-864.

Machin, Stephen and Tuomas Pekkarinen (2008) "Global Sex Differences in Test Score Variability", *Science*, 322, 1331-1332.

Mailath, George J., Andrew Postlewaite and Larry Samuelson (2010) "Pricing and Investments in Matching Markets", forthcoming *Theoretical Economics*.

Myerson, Roger (1998) "Population Uncertainty and Poisson Games", *International Journal of Game Theory*, 3, 375-392.

Peters, Michael (2007) "The pre-marital investment game", *Journal of Economic Theory*, 137, 186-213.

Peters, Michael (2009) "Truncated Hedonic Equilibrium", working paper.

Peters, Michael and Aloysius Siow (2002) "Competing Premarital Investments", *Journal of Political Economy* , 110, 592-608.

Pinker, Susan (2008) *The Sexual Paradox: Men, Women and the Real Gender Gap.* Scribner.

Shaked, Moshe and J. George Shanthikumar (2007) *Stochastic Orders*, New York: Springer.

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