Institut d'Etudes Politiques de Paris

Master in Economics

Master's Thesis

Returns to College on the Marriage Market: a Simple Roy Model with Perfect Foresight

Author: Edoardo Ciscato
Supervisor: Pr. Jean-Marc Robin:

SciencesPo.

Abstract

Education is recognized to be a valuable asset on the marriage market: more educated people are likely to have higher wages, more cultural and social capital, higher opportunities of meeting people of the same educational level and greater bargaining power within the household. Therefore, college education produces important returns on the marriage market. In this paper, I make a first attempt to combine the self-selection approach to the educational choice started by Willis and Rosen (1979) with the idea that schooling returns from the marriage market are crucial in explaining graduation patterns. In doing so, I turn to the family economics literature on intra-household dynamics in order to determine the surplus shares enjoyed by each partner for a given income distribution across the couple (mainly inspired by Chen and Woolley, 2001). Once obtained the distribution of utility levels, I set a fully parametric Roy model with perfect foresight to analyze self-selection patterns. The paper contains many oversimplifications: there is no space for uncertainty and neither for a real competitive marriage market. In spite of this, the model simply tries to convey the idea that there is ground for future research in this direction. In addition, I propose an empirical application with 2008-2012 ACS data for California on married and single women: results provide (weak) evidence that hierarchical sorting on labor-market skills is offset by the distribution of unobservable marriage-market-related skills, i.e. women seem to pursue a degree to increase their competitiveness on the marriage market rather than to exploit their relative advantage in labor-market skills.

Acknowledgements

I would like to thank my supervisor Pr. Jean-Marc Robin, as well as Pr. Denis Fougère, Pr. Alfred Galichon, Pr. Marion Goussé and PhD candidate Simon Weber for extremely helpful discussion about my master's thesis's topic. I would also like to thank my parents for their support over my five years of college.

1 Introduction.

In the last decades, the educational gender gap reduced in many countries and, in some cases, it reversed[1]. Despite labor earnings tend to be lower for women than for men in absolute value, the returns to schooling seem to be higher[2]: college education, in fact, not only ensures better job opportunities because of higher productivity due to specialized skills and ability signaling, but it also allows women to escape discrimination (Dougherty, 2005; Chiappori, Iyigun, and Weiss, 2009). Nevertheless, this does not seem sufficient to provide an exhaustive explanation for recent trends, which is why analyzing marriage patterns may provide further insights.

Investments in schooling typically occur before individuals decide whether and who to marry. An investment in education is a valuable asset on the marriage market not only because high income is a desirable trait, but also because of the acquired cultural and social capital and potential advantages in childrearing: in his early theory of family economics, Becker (1973) identifies complementarities in non-labor market as a source of gains from marriage. In addition, college campus is an important marriage market, where relatively homogamous young men and women have many opportunities to meet. Finally, higher education - and possibly the consequent higher wage-earning potential - increases the bargaining power of the spouse within the household. It follows that college education produces returns also with respect to the marriage market, since higher education can lead to higher marriage surplus.

A growing literature deals with the problem of schooling investment decisions before the opening of the marriage market: among the most important, Peters and Siow (2002), Iyigun and Walsh (2007), Chiappori, Iyigun, and Weiss (2009) and the recent working paper Chiappori, Salanié, and Weiss (2015). Whereas the first three papers are mainly theoretical and focus on equilibrium sorting patterns with uncoordinated investments, only recently Chiappori, Salanié, and Weiss (2015) have brought this class of models to the data, in an extension of the seminal model by Choo and Siow (2006) on frictionless marriage markets. Their work focuses on the estimation of college premium on the marriage market, finding that White women's premium has increased in the last decades, whereas White men's premium has been stable over time, and providing further evidence that returns to schooling from the marriage market are crucial in explaining attendance and graduation patterns.

On the other hand, a large literature has analyzed how and why individuals decide whether to go to college or not as a self-selection problem. The final choice is indeed the result of a rational decision process, in which the agent considers his potential gains after taking into account his own personal traits, skills and costs that each option involves. Therefore, individuals self-select according to their own characteristics. Typically, the econometrician observes neither the counterfactual outcomes, nor the whole set of characteristics and skills. The self-selection problem has been formally described by Roy (1951), who was interested in workers' sectoral choice and wage premia. Many empirical applications tackling diverse issues have followed: among the best-known, women's participation to

[1] Goldin, Katz, and Kuziemko (2006) analyze the trend of educational gender ratio in United States across the 20th century: the ratio went from 2.3 graduate men to for each graduate woman in 1947 to parity in 1980.

[2] There is a large literature measuring and comparing returns to schooling across genders. See Trostel, Walker, and Woolley (2002) and Dougherty (2005) for cross-country analysis and detailed literature review. Most researchers seem to agree on the fact that returns to schooling are higher for women in most industrialized countries.

labor market (Heckman, 1974), workers' registration to unions (Lee, 1978), college enrollment (Willis and Rosen, 1979), participation to training programs (Ashenfelter and Card, 1985), migration (Borjas, 1987).

In particular, Willis and Rosen (1979) were among the first to cope with self-selection in schooling and set a benchmark for college attendance models. They assume that agents act on the discounted value of lifetime earnings: under perfect foresight, this simplifies to sectoral initial earnings wage growth rate. In introducing their exclusion restrictions, they argue that family background characteristics only affect earnings only through schooling choices and thus they include parents' traits in the choice equation while excluding them from outcome equations. They conclude that cross-sectoral differential in wages and growth rates, financial constraints and family background all matter for the choice of attending college.

Since their work, the related theoretical and applied literature has expanded. Of particular relevance is the contribution of Heckman and Honore (1990), who thoroughly discuss the contents of the original Roy model and the robustness of its implications in nonnormal cases, dealing with non-parameteric identification issues. Their work has served as a reference for many of the following extensions of the classic Roy model. More recent works have extended self-selection models to multiple schooling level (Dahl, 2002), uncertainty (Heckman, Carneiro, and Hansen, 2003; Cunha, Heckman, and Navarro, 2005; Cunha and Heckman, 2007), non-pecuniary gains (Arcidiacono, 2004; Beffy, Fougere, and Maurel, 2012; D'Haultfœuille and Maurel, 2013), introducing semi-parametric specifications with an increasing degree of sophistication and realism in the description of schooling decisions. The nature of these works is oriented to empirical application and aims to assess returns to schooling with a high degree of precision.

In this paper, I make a first attempt to combine the self-selection approach with the idea that schooling returns from the marriage market are crucial in explaining graduation patterns. This is something new in the literature and could potentially open doors for further research that tries to combine two different fields.

This void in the literature is easily explainable, since the assessment of marriage surplus shares is an extremely hard task. As a first experiment, I propose to use a simple household model similar to the one by Chen and Woolley (2001) and Browning, Chiappori, and Lechene (2010) in order to determine resource allocation within the sample. Using an intra-household model instead of a marriage market model, I abstract from the endogenous choice of whether and who to marry: individuals know that if they graduate they will have a partner with a certain income level - intuition suggests probably higher - whereas if they do not they will find another partner with another income level - probably lower. Once recovered the future payoffs, I focus on the choice of whether or not to pursue a degree and I set a fully parametric Roy model in the spirit of Willis and Rosen (1979), with heterogeneity in sectoral outcomes and costs of graduating. Agents have perfect foresight and forecast both their own and their partner's income with absolute precision in both possible states.

Despite the evident oversimplification in the specification of the model, I push it to its limit by estimating it with 2008-2012 American Community Survey data for California and try to provide a simple interpretation of the results. I focus on 25-to-50-year-old married women to exploit an interesting non-standard feature of the model: I can estimate the distribution of returns to schooling also including women that are outside of labor force, which constitute a non-negligible part of the sample (almost one third). I also provide a brief comparison with a subsample of 25-to-50-year-old single women. The estimated self-selection patterns seem to suggest that there is weak hierarchical sorting on unobservable skills among singles: self-selection for singles is interpreted as driven by labor market skills. The situation radically changes in the case of married women: graduates seem to choose to attain a degree because of their poor unobservable skills when they need to compete on the non-graduate labor *and* marriage market with those that do not hold a degree. Non-graduates, instead, seem to be aware that their skills allow them to achieve a sufficiently high living standard without incurring the cost of going through college. I explain this reversal of fortune arguing that unobservables skills on the marriage market are important determinants of selection.

The paper is organized as follows: in section 2 I introduce and solve the household problem, also discussing its implications; in section 3 I set the Roy model for the schooling choice, explain the estimation technique and the determination of utility levels and preference parameters; in section 4 I describe the sample, present the results for married women and set the comparison with singles.

2 The Household Problem.

The model represents the choice of whether to attain a college degree or not, given the outcomes the agent would obtain in each possible future state. The model differs from the classic case of Willis and Rosen (1979) in that agents take into account the surplus they can obtain from the marriage market. Thus, the model is structured in two parts: first, agents decide whether to continue to study, then they make their consumption choices within the newly formed household, together with their respective partners. In order to analyze the educational choice, it is necessary to solve the model backwards to understand what the agent expects from his future household life. In this section, I propose a non-cooperative household model based on Chen and Woolley (2001) and Browning, Chiappori, and Lechene (2010). Solving for household consumption demands, I obtain agents' indirect utility functions. In this way, with data on individual income and marriage patterns at hand, it is possible to compute utility levels and use them for the binary choice model described in section 3.

2.1 A Non-cooperative Household Model.

What agents strive for since the beginning of the decision process is the utility they obtain from within the households they form. Once married, agents need to decide how to spend their labor income in consumption goods jointly with their respective partners. I assume that every spouse $i \in (m, f)$ draws his utility from consumption of both a private good q_i and a public good Q. Every spouse receives a given wage $y_i \geq 0$, without ruling out the possibility that one of the partners (but not both) has zero income. Although there is no time-intensive home production technology, I allow, at least at the beginning, for the possibility that the relative price p_i of the public good may differ across sexes[3]. The

[3]This means that, with exactly the same amount of money, one partner is actually able to purchase a larger amount of public good. One could think of this as a difference in the capacity of choosing the products from the shelves in a store. Alternatively, it could stand for the difference in the ability of applying or activating a product, e.g. changing a light bulb or hanging a painting with a nail.

price of the private good is instead normalized to one without loss of generality.

Most importantly, I assume that the spouses adopt a non-cooperative behavior. This assumption is common to many models of the household since the seminal contributions by Manser and Brown (1980) and McElroy and Horney (1981). The two papers were among the first to address the following issue: how can the partners find an enforceable agreement on the allocation of time and money if cooperation is not assumed? Their approach consists in the repartition of the marriage surplus according to a Nash bargaining process with divorce as disagreement points[4]. Making the bargaining process explicit, they depart from unitary models inspired by the work of Samuelson (1956), who made the simplifying assumption that one individual is in charge of choosing an allocation that maximizes a weighted welfare function of the family. Afterwards, starting from Chiappori (1988, 1992), the collective approach allowed analysts to focus on the impact of *distribution factors* (incomes, prices, sex-ratios, etc.) on the household equilibrium under the general assumption that such equilibrium is Pareto-efficient, regardless of which bargaining process is at work.

In this paper, I introduce a representation of the household that is simple enough to recover indirect utility functions with a manageable functional form. In doing so, I draw inspiration from the works by Chen and Woolley (2001) and Browning, Chiappori, and Lechene (2010) in what follows. As anticipated, agents consume a public good which is the primary source of gains from marriage. Although marriage allows agents to push their Pareto frontiers further, non-cooperative behavior causes underprovision of the public good. However, I also introduce *caring preferences*, i.e. I assume that every agent's utility function contains the partner's utility function. This is the secondary source of gains from marriage and, as I will show later, limits the relevance of public good underprovision.

Finally, I allow for money transfers between partners. There is no restriction on the sign of the *net* transfer π: a positive value implies that the husband is making a transfer, whereas a negative sign implies the opposite. At this stage, Chen and Woolley (2001) propose two possible ways to determine how agents choose the direction and the size of the transfer. One solution is that the partners reach an agreement through a generalized Nash-bargaining process, with the no-transfer equilibrium as a disagreement point. Alternatively, any of the two partners may independently and voluntarily decide to make a money transfer if he/she wishes to do so. I assume that we find ourselves in this second case. I discuss in section 2.3 the implications, the advantages and the drawbacks of this choice.

2.2 The Cournot-Nash Equilibrium.

In this section, I provide a formal exposition of the model. Individual preferences for goods (q, Q) are represented by a log-additive utility function. Caring implies the presence of the partner's utility function which enters additively multiplied by a parameter $s \in [0, 1]$. Thus, for $i, j \in \{m, f\}$ with $i \neq j$ and for $\alpha > 0$:

$$W_i(q_i, Q) = \alpha \log q_i + \log Q + s(\alpha \log q_j + \log Q). \tag{2.1}$$

Logarithmic form ensures that the assumptions by Chen and Woolley (2001) on the utility function

[4] Actually, Manser and Brown (1980) expand their framework to a broader set of bargaining processes. In addition, they discuss two kinds of reservation utilities corresponding to the disagreement points: one is singlehood, while the other is marriage with the best alternative partner on the marriage market.

$u'_{q_i}(q_i,.) > 0$, $u''_{q_i}(q_i,.) < 0$, $u'_Q(.,Q) > 0$, $u'_Q(.,Q) < 0$ $u'_{q_i}(0,.) = \infty$ and $u'_Q(.,0) = \infty$ are respected. In particular, the two last imply that agents need to consume a positive quantity of both goods. I assume that men and women do not differ in preference parameters, i.e. α and s are equal for both sexes.

The available budget for a man m is equal to $y_m - \pi$, whereas it is $y_f + \pi$ for a woman. In this section, I first treat π as given: one could think of it either as set exogenously or even equal to zero. In this way, we can focus on what happens when the equilibrium is only determined by the strategic interaction that arises with the presence of a public good. Now, given prices and income, one can solve the following optimization problem for a man m:

$$\begin{aligned}
&\max_{q_m, Q_m} u(q_m, Q) \\
&\quad s.t.\ q_m + Q \leq y_m - \pi && (2.2)\\
&\quad Q = Q_m + Q_f && (2.3)\\
&\quad q_m \geq 0 && (2.4)\\
&\quad Q_m \geq 0. && (2.5)
\end{aligned}$$

It is easy to see how budget constraint (2.2) is always binding because of individual rationality, whereas (2.4) is never binding as long as the individual has positive income, since $u'_{q_i}(0,.) = \infty$. On the contrary, the constraint (2.5) can be binding, since the total amount of public good is given by the sum (2.3): if the agent m knows that the partner is providing some public good, he might decide to free-ride and set $Q_m = 0$. Any woman f faces a symmetric problem.

In order to assess the amount of public good provided, it is necessary to solve for the Cournot-Nash equilibrium by stating the First Order Conditions and finding the Best Response functions. I first deal with interior solutions: I assume that both partners supply a positive amount of public good and then I check under which conditions an equilibrium of this kind is feasible. Denote λ the Lagrangian multiplier for (2.2) and recall that the problem for f is symmetric to the one for m:

$$\begin{aligned}
FOC:\ & \frac{\alpha}{q_m} = \lambda && (2.6)\\
& \frac{1+s}{Q_m + Q_f} = p\lambda && (2.7)\\
& q_m + Q = y_m - \pi
\end{aligned}$$

$$\begin{cases}
Q_m^{BR}(Q_f) = \dfrac{1+s}{1+\alpha+s}\dfrac{y_m - \pi}{p_m} - \dfrac{\alpha}{1+\alpha+s}Q_f & (2.8)\\
Q_f^{BR}(Q_m) = \dfrac{1+s}{1+\alpha+s}\dfrac{y_f + \pi}{p_f} - \dfrac{\alpha}{1+\alpha+s}Q_m. & (2.9)
\end{cases}$$

First note that Best Response functions are linear in the quantity provided by the other player. In addition, since $\alpha/(1+\alpha+s)$ is smaller than one, when the wife increases Q_f by one unit, the husband reacts by reducing his supply Q_m by less than one unit (and vice versa). Because of this, if constraints on Q_i are not binding for both the husband and the wife, there always exists a unique interior

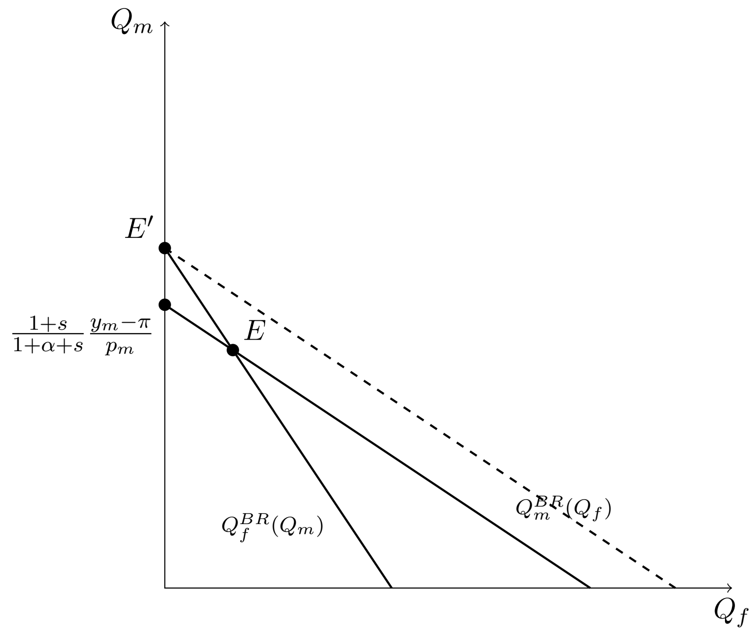


Figure 2.1: Cournot-Nash equilibrium

solution. Agents determine their provision also according to their respective incomes: if y_m increases, Q_m increases too for any Q_f. From (2.8) and (2.9) we can recover the Cournot-Nash equilibrium quantities:

$$\begin{cases} Q_m^* = \dfrac{(1+\alpha+s)(y_m-\pi)p_f - \alpha(y_f+\pi)p_m}{(1+2\alpha+s)p_m p_f} & (2.10) \\ Q_f^* = \dfrac{(1+\alpha+s)(y_f+\pi)p_m - \alpha(y_m-\pi)p_f}{(1+2\alpha+s)p_m p_f}. & (2.11) \end{cases}$$

The conditions of existence of an interior solution can be found by solving a system of inequalities once simultaneously set $Q_m^{CN} > 0$ and $Q_f^{CN} > 0$. But first it is useful to locate the equilibrium on a graph (Figure 2.1). A unique interior solution E exists as long as the two functions have their intersection in the first quadrant. However, whenever the income of one partner becomes too high relatively to the other's, there is no interior solution. If the wife's income y_f is kept constant, it is possible to find the threshold income $y_m^{E'}$ - implicitly defined by the dashed line - for which the equilibrium is exactly E'. This implies that, for $y_m \geq y_m^{E'}$, we have a corner solution with $Q_f = 0$. Conversely, a high y_f relatively to y_m leads to a corner solution with $Q_m = 0$.

A formal description of this result is provided by Chen and Woolley (2001). I provide a reformulation of their theorems for the specific case of log-additive utility. Also remember once again that π is given in this section and could be normalized to zero.

Proposition 1 *Assuming that the utility function is equal to (2.1), the pre-transfer Nash equilibrium:*

1. *is a unique interior solution if:*

$$\frac{\alpha}{1+\alpha+s}\frac{p_m}{p_f} < \frac{y_m-\pi}{y_f+\pi} < \frac{1+\alpha+s}{\alpha}\frac{p_m}{p_f}. \tag{2.12}$$

2. *is a unique corner solution with $Q_m > 0$ and $Q_f = 0$ if:*

$$\frac{y_m - \pi}{y_f + \pi} \geq \frac{1 + \alpha + s}{\alpha} \frac{p_m}{p_f} \tag{2.13}$$

3. *is a unique corner solution with $Q_m = 0$ and $Q_f > 0$ if:*

$$\frac{y_m - \pi}{y_f + \pi} \leq \frac{\alpha}{1 + \alpha + s} \frac{p_m}{p_f}. \tag{2.14}$$

Consider the case of an interior solution. From equilibrium quantities (2.10) and (2.11), it is possible to derive the relevant quantities, all denoted with * as superscript:

$$Q^* = \frac{1+s}{1+2\alpha+s} \frac{(y_m - \pi)p_f + (y_f + \pi)p_m}{(1+2\alpha)p_m p_f} \tag{2.15}$$

$$q_m^* = \frac{\alpha}{1+2\alpha+s} \frac{(y_m - \pi)p_f + (y_f + \pi)p_m}{p_f} \tag{2.16}$$

$$q_f^* = \frac{\alpha}{1+2\alpha+s} \frac{(y_m - \pi)p_f + (y_f + \pi)p_m}{p_m}. \tag{2.17}$$

We can notice that an increase in y_m surely has a positive impact on the welfare of the family, because Q^*, q_m^* and q_f^* all increase. However, it changes the way public good is provided, since Q_m^* increases while Q_f^* decreases. The inverse happens when y_f grows. As one may expect, for higher values of α, both Q_m^* and Q_f^* are unambiguously lower, whereas q_m^* and q_f^* are higher. On the contrary, when the parameter s for caring preferences is higher, both Q_m^* and Q_f^* are higher, whereas q_m^* and q_f^* are lower. Caring, in fact, limits the inefficiency caused by Cournot competition in the provision of the public good: if s were equal to 1, then the wife's and the husband's problem would both coincide with the maximization of the household's welfare, underprovision would disappear and the equilibrium would be fully efficient.

Finally, look at prices p_m and p_f: a man m decreases his provision Q_m^* when p_m grows, although he increases it when p_f grows[5]. The inverse happens for a woman f. The equilibrium quantity Q^* negatively depends on both p_m and p_f, whereas a man's private consumption q_m increases with p_m but decreases with p_f, since the man needs to cope with a lower provision from his wife. The difference in public good prices gives rise to a puzzle for the family: clearly, it would be efficient to entrust to the more productive spouse the whole production of public good[6]. The literature defines it as a situation of non-neutrality (Lundberg and Pollak, 1993; Konrad and Lommerud, 1995; Chen and Woolley, 2001). If one spouse were in charge of solving the household problem, he would assign to the most productive member of the family sufficient resources to take care both of his own consumption

[5] Clearly, one could define a relative price p_m/p_f to analyze the equilibrium levels of public good. Nonetheless, even doing so, the price levels p_m and p_f still matter as the private good q is the numéraire.

[6] It is also interesting to remark that it is efficient to have only the more productive spouse producing the public good because the total amount Q is a simple sum of the two purchased quantities. With more complex home production technologies - notably time-intensive ones - one could require that the contribution of both partners is essential. For instance, a Cobb-Douglas function with the time dedicated to housework of both partners as inputs would prevent complete specialization, as both of them need to participate to home production. This consideration is crucial if one wants to focus on home production and consumption complementarities.

and of the public good provision for the whole family. This would be the case if the model belonged to the *unitary* class, in the spirit of Samuelson (1956) and Becker (1974). However, if prices p_m and p_f are equal, it does not matter who is actually purchasing the public good. Since the following section, I will mainly assume that $p_m = p_f$: this will lead us back to a situation of neutrality and has important implications on the equilibrium quantities.

Before moving on, it is still necessary to analyze what happens in corner solutions. Once assessed that under conditions (2.13) or (2.14) one of the partners will not purchase any public good, the other partner will reply by changing his behavior consistently. Now suppose that (2.13) holds, i.e the wife's income is low with respect to the husband's: given that in the problem of the wife the constraint binds at $Q_f = 0$, she spends all her income on private consumption ($q_f = y_f + \pi$). The problem of the husband does not change: his optimality conditions are always given by (2.6) and (2.7). Although he is now the only responsible for public good provision, he has no control over the wife's (small) share of the family budget and her private consumption: therefore, the situation of public good underprovision persists. Denote with ** the equilibrium quantities under (2.13):

$$Q_f^{**} = 0, \quad q_f^{**} = y_f + \pi$$

$$Q_m^{**} = Q^{**} = \frac{1+s}{1+\alpha+s}\frac{y_m - \pi}{p_m} \tag{2.18}$$

$$q_m^{**} = \frac{\alpha}{1+\alpha+s}(y_m - \pi). \tag{2.19}$$

Note that, as long as (2.13) holds, an increase y_f is only beneficial to the wife, while an increase in y_m generates both higher Q^{**} and q_m^{**}. As concern preference parameters, a higher α leads to a higher q_m^{**} but a lower Q_m^{**}, whereas a higher s generates an opposite effect. Finally, note that now only p_m matters, being negatively correlated with Q^{**}. It is also true that, when p_m/p_f grows high, it is more likely that the conditions (2.12) for the interior solution to hold are respected. In fact, if p_m/p_f is high enough to escape the corner solution, the wife decisively contributes to home production thanks to high productivity and despite her low income and low expenditure in the public good.

2.3 The Voluntary Transfer.

Following Chen and Woolley (2001), I assume that agents can make voluntary transfers to their respective partners. Clearly, every spouse can only make a transfer *in favor* of his partner, although, as already anticipated, the *net* transfer π can be either positive or negative and cannot be larger than the spouses' respective incomes, i.e. $\pi \in [-y_f, y_m]$. Agents choose the amount they wish to transfer in their maximization program. Nevertheless, the new First Order Condition $\partial W_i(.,.)/\partial \pi = 0$ leaves the old (2.6) and (2.7) unchanged (since I already wrote them for a fixed π).

Once again, I deal with the case where the positivity constraint (2.5) does not bind and the one where it binds separately. When there exists an interior solution and both partners contribute to public good provision, one can still recover the relationships (2.10) and (2.11) from the FOCs, while the equations (2.15), (2.16) and (2.17) follow. For these last three equilibrium quantities depend on π only through the factor $(y_m - \pi)p_f + (y_f + \pi)p_m$, the indirect utility function increases at any point when such factor grows higher. Since the agent still needs to determine π, he wants to maximize

$y_m p_f + y_f p_m + (p_m - p_f)\pi$: as predicted in section 2.2, whenever p_m is higher than p_f, the husband would like to make a transfer to the wife so that she purchases the public good at lower price. This is a way to introduce a simple specialization process in the model, although the lack of coordination may prevent partners from completely entrusting public good provision to only one of them. Henceforth, I assume that $p_m = p_f = p$, which implies that the equilibrium quantities Q^*, q_m^* and q_f^*, as well as the indirect utility functions, are unaffected by the net transfer π. In fact, even if the husband transfers a positive amount π to the wife, she would only employ it in the provision of public good (Q_f^* rises), letting the partner decrease his contribution (Q_m^* falls) and increase his private consumption. To conclude, as long as conditions (2.12) hold, equilibrium quantities are the same as in the pre-transfer case for any feasible level π. The net transfer only affects the way public good is provided.

Now consider the case when the condition (2.13) for a corner solution holds (an equivalent reasoning applies when condition (2.14) holds). In this case, the wife is stuck at $Q_f = 0$ and thus uses her income (if she has any) to fund private consumption. However, even if she spends all her disposable income, she would like to purchase even more q_f: by writing the full Kuhn-Tucker conditions with the binding constraint $Q_f = 0$, one can see that the marginal utility from private consumption is higher than in the case of the interior solution, i.e. the wife's FOC (2.6) does not hold any more. On the other hand, the husband keeps the same FOCs (2.6) and (2.7), although now, observing the pre-transfer equilibrium quantities given by (2.18) and (2.19), it is clear that q_m and $Q_m = Q$ - together with the indirect utility function - depend on π in a nontrivial way. On the one hand, in fact, the husband may dislike the idea of a transfer because it subtracts resources from his expenditure in q_m and Q. On the other hand, because of caring preferences, he does not want his wife to have very low private consumption. It is thus necessary to add the last First Order Condition $\partial W_i(.,.)/\partial\pi = 0$ and solve for optimal π. Substituting for $q_f = y_f + \pi$, the program can be rewritten as:

$$\max_{q_m, Q, \pi} \alpha \log q_m + \alpha s \log(y_f + \pi) + (1+s)\log Q \quad s.t. \quad q_m + Q \leq y_m - \pi$$
$$Q \geq 0, \quad \pi \geq 0.$$

Actually, if one substitutes back $q_f = y_f + \pi$ in the utility function and in the budget constraint, then it becomes clear that choosing π is equivalent to choosing q_f (up to y_f). The fact that the husband essentially takes control over the household decision-making process in a non-cooperative framework may particularly remind of the *Rotten Kid Theorem* (Becker, 1974), a household where the head maximizes the family welfare function and compensates any imbalance among the members by restoring the efficient allocation. The case of voluntary transfers leads to similar results, although it presents a fundamental difference: there is no family head *ex-ante*, although in some particular cases there is one spouse acting as if he were the head. In addition, the high-income spouse's welfare function perfectly coincides to the household welfare function only when $s = 1$.

Once set the program, I first assume that the constraint $\pi \geq 0$ is not binding. The FOCs are given by equations (2.6) and (2.7) together with a new condition on π. It is then possible to compute the transfer equilibrium quantities, which bear superscript T:

$$\frac{\partial W_i(.,.)}{\partial \pi} = 0 \Leftrightarrow \frac{\alpha}{y_m - \pi - pQ_m} = \frac{\alpha s}{y_f + \pi} \tag{2.20}$$

$$
\begin{aligned}
\pi^T &= \frac{\alpha s}{(1+\alpha)(1+s)} y_m - \frac{1+\alpha+s}{(1+\alpha)(1+s)} y_f \qquad (2.21)\\
Q^T_m = Q^T &= \frac{1}{1+\alpha}\frac{y_m+y_f}{p}\\
q^T_m &= \frac{1}{1+s}\frac{\alpha}{1+\alpha}(y_m+y_f)\\
q^T_f &= \frac{s}{1+s}\frac{\alpha}{1+\alpha}(y_m+y_f)
\end{aligned}
$$

Comparative statics suggests that an increase in s causes π^T and q^T_f to be larger, while q^T_m falls and Q^T stays unchanged. In particular, when $s = 1$ the husband maximizes household welfare, and thus $q^T_m = q^T_f$. However, the quantity Q^T never changes: when the husband takes over public good purchase, free-riding and underprovision disappear and Q^T is always at the efficient level, independently of caring preferences. The parameter s only determines how the remaining part of the high-income partner's budget is shared between the wife's and the husband's private consumption.

From equation (2.21), one can immediately recover the condition for which the optimal π^T is actually positive. Whenever the constraint is binding and $\pi^T = 0$, there is no change with respect to the corner solution described in section 2.2, since the husband decides not to transfer any money. On the contrary, when the wife's income is sufficiently low, the husband enjoys a high marginal utility from the transfer.

$$
\pi^T > 0 \Leftrightarrow \frac{y_m}{y_f} > \frac{1+\alpha+s}{\alpha s} \qquad (2.22)
$$

Therefore, the income ratio not only determines whether the equilibrium is an interior or a corner solution relatively to the positivity constraint on Q_f (or Q_m), but also whether there is a positive (or negative) net transfer π or not. Note that the threshold income ratio for a positive π defined by (2.22) is necessarily larger than the threshold defined by condition (2.13) (and by the dashed line in Figure 2.1) for a positive Q_f. The two thresholds coincide only when $s = 1$. This is rather intuitive: if the husband has a relatively large income, he takes control over public good provision and let the wife spend her income on private goods; however, if a caring husband's income is much larger than the wife's, he also funds her private consumption through a transfer. This leads to the following proposition that concludes the formal analysis:

Proposition 2 *Assuming that the utility function is equal to (2.1), the Nash equilibrium with voluntary transfers is characterized by:*

1. *no net transfer π if:*
$$\frac{\alpha s}{1+\alpha+s} \leq \frac{y_m}{y_f} \leq \frac{1+\alpha+s}{\alpha s}$$
2. *a positive net transfer π if:*
$$\frac{y_m}{y_f} > \frac{1+\alpha+s}{\alpha s} \geq \frac{1+\alpha+s}{\alpha}$$
3. *a negative net transfer π if:*
$$\frac{y_m}{y_f} < \frac{\alpha s}{1+\alpha+s} \leq \frac{\alpha}{1+\alpha+s}.$$

One last observation concerns the number of different equilibria that could possibly arise, according to which constraints are binding or not. Combining Propositions 1 and 2, it is straightforward to notice that there are five different types of equilibria, of which: one is the interior solution; two are corner solutions with respect to individual public good provision but with no voluntary transfers; two are always corner solutions but with positive transfers.

2.4 Implications of the Model.

I conclude section 2 by briefly discussing the implications of this non-cooperative household model for the empirical analysis of the college graduation choice. Before anything else, it is important to remark that the model is fully explicit on how the spouses decide to use their respective incomes: the theory provides a foundation for household demand as a result of a joint utility maximization problem solved by the spouses. Although nobody would believe that people actually compute best response functions when making actual decisions, it is crucial to have a theory that at least tries to clarify the roles of different variables in the decision process. Chen and Woolley (2001) stress that the non-cooperative approach can be considered as complementary to the collective approach, since the latter, despite being an extremely useful tool for empirical analysis, does not explain *how* the decision making occurs. Nonetheless, the non-cooperative model adopted here is also very demanding in terms of assumptions, especially as concerns functional forms of both preferences and public good provision system. Since the model is later used for empirical analysis, this lack of flexibility is a potential weakness.

In sections 2.2 and 2.3, I showed the relevance of ex-ante income distribution across the couple and of preference structures, both by solving the model and through some simple comparative statics. The importance of ex-ante relative income is stressed by Lundberg and Pollak (1993) and Browning, Chiappori, and Lechene (2010), because it justifies policies that target one partner in the provision of subsidies (especially if one believes that partners have different preferences) altering the income distribution within the household. This observation reveals the juxtaposition with *unitary models*, where the family allocation is chosen as if there were a head of the household that maximizes global welfare: in this latter case, ex-ante income distribution would not have any relevance.

On the other hand, although in this non-cooperative model relative income matters for determining intervals for the existence of corner solutions and the positivity of transfers, in some cases small variations in the income sources are actually irrelevant. The model, in fact, implies *local income pooling*: at the interior solution and at corner solutions with positive transfers, the demands for goods q and Q only depend on the total household income $(y_m + y_f)$. This is crucial, because for three out of five possible equilibria indirect utility also depends on overall household income only. Empirical evidence seems to reject income pooling[7]: although here the property holds only locally and ex-ante income distribution does matter for determining the allocation, an improved version of this model should make the consumption demand sensitive to any small change in relative income.

[7] Income pooling tests aim to show if a variation in relative income of the couple has an impact on household demand. Thomas (1990), Browning, Bourguignon, Chiappori, and Lechene (1994) and Browning and Chiappori (1998) are among the best-known. Studies by Lundberg, Pollak, and Wales (1997) and Duflo (2003) analyze the impact of public policies on household demand to disentangle the effect of the variation in relative income from possible confounding factors. Browning, Chiappori, and Weiss (2014) contains an exhaustive literature review on the topic.

The comparison with unitary and collective models also reveals that the non-cooperative approach has another important implication. The equilibrium allocation, in fact, is not necessarily efficient: in presence of a public good, the Cournot game leads to underprovision and inefficiency, at least as long as there is no transfer between partners. Critics of this approach point at the fact that family members interact on a regular basis, observe each other's actions and have a good knowledge of each other's preferences: because of repeated interaction and symmetric information, most of the recent economic literature on family argues that the resource allocation should be efficient (see, for instance, Chiappori, 1988, 1992; Browning, Chiappori, and Weiss, 2014). In this paper, the decision process only leads to an efficient allocation when there is a corner solution with positive transfer, because of the *local family head* property. Although inefficiency may be a drawback, on the other hand the solution to the household problem is characterized as a stable Nash equilibrium that arises from the optimality conditions of the family members: this is an advantageous representation, because it addresses commitment issues without requiring any assumption on stability.

More than efficiency, the real drawback of voluntary transfers is the lack of a representation of the bargaining power distribution. Depending on preference parameters, in fact the model may have either an interior solution where both partners have identical consumption levels or a corner solution when one spouse maximizes a household welfare function weighted according to his own tastes. If the transfer were discussed instead than being chosen individually, the solution would change and relative bargaining power would lead to different equilibria than those discussed in sections 2.2 and 2.3. In addition, the bargaining process requires an analysis of disagreement points that also imply best outside options for each spouses. For instance, a partner with low income but high attractiveness on the marriage market may succeed in extracting a higher share of surplus from the bargaining process.

Finally, another limitation is the lack of both time-use patterns and non-monetary surplus from marriage. Household specialization is often regarded as a source of gains and, although its importance may have slightly decreased (see Greenwood, Seshadri, and Yorukoglu, 2005, among the others), simple descriptive statistics seem to suggest that married women still reduce their labor supply with respect to single women (see also section 4.1 for some descriptive statistics). In this model, since the only input to public good provision is income, low-wage individuals concentrate their poor resources on private consumption but are excluded from the public good production process. Non-monetary gains and consumption complementarities are other important sources of surplus from marriage: in a refined version of a household model, they would depend on individual traits as in the early model for the marriage market by Becker (1973). Yet, caring preferences are the simplest way to make the low-income individual count in the generation of surplus: although they may be an oversimplification, they still are a manageable expedient to introduce supplementary gains from marriage.

3 A Roy Model for Education.

In this section, I introduce a classic binary choice model similar to Willis and Rosen (1979) and Borjas (1987) and I expose the estimation procedure. Finally, I conclude the section briefly discussing the identification and the choice of utility levels and preference parameters, which represent the payoffs of agents' decision.

3.1 The Roy Model.

In this section, I introduce a simple binary choice model with perfect foresight similar to Willis and Rosen (1979). Outcomes denoted with v_{ki}, $k \in \{0,1\}$, are the utility levels the agent i attains if he earns a college degree ($k = 1$) or if he does not ($k = 0$). Utility levels are explained by a set of personal characteristics x_i and unobserved heterogeneity $(\varepsilon_{1i}, \varepsilon_{0i})$. In the following exposition, I do not include sector-specific skills, i.e. I do not assume that there is at least one covariate x_{ij} such that $\beta_{kj} = 0$ and $\beta_{k'j} \neq 0$, with $k, k' \in \{0,1\}$ and $k \neq k'$. I rule out this possibility only because in the subsequent empirical application I will not use any sector-specific skill, although the model could easily be adapted to include them. On the other hand, the agent i faces an individual cost C_i when he decides to go to college and complete his studies: C_i depends on personal traits x_i, on unobserved heterogeneity ε_{ci} and on a set z_i of covariates that are not included in the outcome equations, i.e. for any z_{ij} I assume $\beta_{kj} = \beta_{k'j} = 0$. This condition acts as an exclusion restriction for the model: in this section and in the following (3.2), I clarify the role of this restriction for the identification of the linear parameters in the main equations and in the covariance structure.

First of all, outcomes and costs are described by the following equations:

$$\begin{cases} v_{1i} = x_i'\beta_1 + \varepsilon_{1i} & (3.1) \\ v_{0i} = x_i'\beta_0 + \varepsilon_{0i}i & (3.2) \\ C_i = x_i'\delta_x + z_i'\delta_z + \varepsilon_{ci}. & (3.3) \end{cases}$$

In addition, as in the early literature on Roy models (e.g. Willis and Rosen, 1979; Borjas, 1987), I assume that $(\varepsilon_{1i}, \varepsilon_{0i}, \varepsilon_{ci})$ are jointly normal:

$$\begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{0i} \\ \varepsilon_{ci} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} \sigma_1^2 & & \\ \sigma_{01} & \sigma_0^2 & \\ \sigma_{1c} & \sigma_{0c} & \sigma_c^2 \end{pmatrix} \right).$$

Always following the literature on self-selection, I make the usual (but crucial) independence assumption $(\varepsilon_{1i}, \varepsilon_{0i}, \varepsilon_{ci}) \perp (x, z)$. Finally, the choice equation with the binary dependent variable $d_i = k$ embodies the condition under which the agent decides to attain the degree ($d_i = 1$):

$$d_i = \mathbb{1}\{v_{1i} - C > v_{0i}\} = \mathbb{1}\{\varepsilon_{0i} - \varepsilon_{1i} + \varepsilon_{ci} < x_i'(\beta_1 - \beta_0 - \delta_x) - z_i'\delta_z\} = \mathbb{1}\{\eta_i < \omega_i'\gamma\} \qquad (3.4)$$

where the last equality defines the random component $\eta_i = \varepsilon_{0i} - \varepsilon_{1i} + \varepsilon_{ci}$ with zero mean and variance σ_η^2, as well as the parameters $\gamma = \beta_1 - \beta_0 - \delta$ for $\omega_i = (x_i, z_i)$. Note that, because of the exclusion restriction, $\gamma_z = -\delta_z$.

The nature of the problem allows to observe only $v_i = dv_{1i} + (1 - d_i)v_{0i}$: because of self-selection, the distribution of outcomes for those who actually graduated is different than the distribution of v_{1i} over the population. In particular, the expected outcome for the two subpopulations can be written as follows:

$$E[v_{1i}|x_i, z_i, d_i = 1] = x_i'\beta_1 + E[\varepsilon_{1i}|x_i, z_i, d_i = 1] \tag{3.5}$$

$$E[v_{0i}|x_i, z_i, d_i = 0] = x_i'\beta_0 + E[\varepsilon_{0i}|x_i, z_i, d_i = 0] \tag{3.6}$$

where $E[\varepsilon_k|x_i, z_i, d_i = k] \neq E[\varepsilon_k|x_i, z_i] = 0$ since the random component η_i that determines the choice d_i is correlated with $(\varepsilon_{1i}, \varepsilon_{0i})$ by construction. Because of the normality assumptions on the structure of unobserved heterogeneity, it is possible to write these expected values conditional on the choice d_i in a clear form:

$$E[\varepsilon_{1i}|x_i, z_i, d_i = 1] = E[\varepsilon_{1i}|\frac{\eta_i}{\sigma_\eta} < \omega_i'\gamma^*] = -\rho_{1\eta}\sigma_1\frac{\phi(\omega_i'\gamma^*)}{\Phi(\omega_i'\gamma^*)} = -\rho_{1\eta}\sigma_1\lambda_1(\omega_i'\gamma^*) \tag{3.7}$$

$$E[\varepsilon_{0i}|x_i, z_i, d_i = 0] = E[\varepsilon_{0i}|\frac{\eta_i}{\sigma_\eta} > \omega_i'\gamma^*] = \rho_{0\eta}\sigma_0\frac{\psi(\omega_i'\gamma^*)}{1 - \Psi(\omega_i'\gamma^*)} = \rho_{0\eta}\sigma_0\lambda_0(\omega_i'\gamma^*). \tag{3.8}$$

Note that substituting these two conditional means in equations (3.5) and (3.6), it is now possible to write the conditional expected value for v_{ki} so that it is linear in x_i and $\lambda_{ki}(\omega_i'\gamma^*)$. Since λ_{ki} is a non-linear transformation, there cannot be perfect multicollinearity between x_i and $\lambda(.)_{ki}$ and the parameters of the linear model are formally identified even without any exclusion restriction. Nonetheless, in the case $\omega_i = x_i$, collinearity problems are likely to emerge both because there can be very strong correlation between x_i and $\lambda_{ki}(x_i'\gamma^*)$ and because, depending on the properties of the support of x_i (especially if small), the transformation λ may be almost linear. In general, it is better to include an exclusion restriction, such as a variable present in ω_i but not in x_i, so that the relationship between x_i and $\lambda(.)_k i$ does not depend on the form of λ or the support of x_i but rather on the information contained in the control function $\lambda(.)_{ki}$ (see Heckman, Carneiro, and Hansen, 2003, for formal exposition).

Equations (3.7) and (3.8) implicitly define $\gamma^* = \gamma/\sigma_\eta$, correlation rates $\rho_{k\eta} = Corr(\varepsilon_{ki}, \eta_i)$ and inverse Mills ratios $\lambda_k(\omega_i'\gamma^*)$. In addition, they provide a starting point to study the structure of the covariance. Now, in fact, it is possible to write the residuals $\zeta_{ki} = v_{ki} - E[v_{ki}|x_i, z_i, d_i = k]$, which constitute the residual variation of v_{ki} once conditioned on x_i, z_i and $d_i = k$. The analysis of such residuals allows to write the following relationships, crucial for the identification of σ_0^2 and σ_1^2:

$$Var[\zeta_{1i}] = E[\varepsilon_{1i}^2|\frac{\eta_i}{\sigma_\eta} < \omega_i'\gamma*] = \sigma_1^2 - (\rho_{1\eta}\sigma_1)^2(\omega_i'\gamma*)\lambda_1(\omega_i'\gamma*) \tag{3.9}$$

$$Var[\zeta_{0i}] = E[\varepsilon_{0i}^2|\frac{\eta_i}{\sigma_\eta} < \omega_i'\gamma*] = \sigma_0^2 + (\rho_{0\eta}\sigma_0)^2(\omega_i'\gamma*)\lambda_0(\omega_i'\gamma*). \tag{3.10}$$

Going back to the choice equation (3.4), the identification of the structural parameters γ and consequently of the cost function parameters δ is only possible if there is a way to separately identify σ_η^2. One way to recover these parameters is to introduce another exclusion restriction: under the assumption that there is one x_{ij} so that $\delta_j = 0$ and for given β_1 and β_0, it is possible to identify σ_η^2 and consequently γ and δ, in a way that will be clear in the discussion of estimation in section 3.2.

Now that the theoretical structure of the problem has been explored, it is possible to say more on the parameters that should be targeted in the analysis. Thanks to equations (3.8) and (3.7), it is now simple to recover the expected outcome from securing a college degree for those who did obtain it and the expected outcome from *not* having a degree for those who did not obtain it. Nonetheless, in order to assess the returns to college education for both subgroups, it is necessary to build counterfactuals to estimate what their respective outcomes would have been if they had made the opposite decision. The goal is to retrieve, for $k = \{0, 1\}$, the distribution of the following marginal treatment effect:

$$E[v_{1i} - v_{0i}|x_i, z_i, d_i = k] = x'(\beta_1 - \beta_0) + E[\varepsilon_{1i} - \varepsilon_{0i}|x_i, z_i, d_i = k] \tag{3.11}$$

The way to obtain *net* marginal treatment effect - i.e. costs of graduating included - is slightly different. Consider the exemplifying case where $k = 1$:

$$\begin{aligned} E[v_{1i} - v_{0i} - C_i|x_i, z_i, d_i = 1] &= x'(\beta_1 - \beta_0 - \delta) + E[\varepsilon_{1i} - \varepsilon_{0i} - \varepsilon_{ci}|x_i, z_i, d_i = 1] = \\ &= x'(\beta_1 - \beta_0 - \delta) - E[\eta_i|\frac{\eta_i}{\sigma_\eta} < \omega_i'\gamma^*] = x'(\beta_1 - \beta_0 - \delta) + \sigma_\eta \lambda_1(\omega_i'\gamma^*) \end{aligned} \tag{3.12}$$

Finally, it would be insightful to obtain information on the covariance structure of $(\varepsilon_{1i}, \varepsilon_{0i}, \varepsilon_{ci})$. Some of these parameters are readily identified, while some others require more attention. While I have already discussed some critical points in this section, all remaining issues are clarified with the exposition of the estimation technique in section 3.2.

3.2 Estimation.

As discussed at the beginning of section 3, the identification of the Roy model with normal random components has been widely studied (e.g. Heckman and Honore, 1990). The estimation follows the method proposed by Heckman (1976, 1979) for selection correction. The procedure is based on multiple steps for the estimation of choice and outcome equations. Subsequently, it is possible to obtain estimates for the cost function parameters and for (part of) the covariance matrix of $(\varepsilon_{1i}, \varepsilon_{0i}, \varepsilon_{ci})$. In what follows, I report a brief exposition of the different steps.

1. The first step consists in the estimation of the choice equation (3.4). I first deal with the reduced form of the choice equation: since $\frac{\eta_i}{\sigma_\eta} \sim \mathcal{N}(0, 1)$, it is possible to run a Probit regression and obtain consistent estimates $\hat{\gamma}^*$. Recall that $\gamma^* = (\beta_1 - \beta_0 - \delta)/\sigma_\eta$: for now, it is not possible to estimate the structural parameters γ directly.

2. The second step consists in the estimation of outcome equations (3.5) and (3.6). The parameters $(\beta_1, \beta_0, \rho_{1\eta}\sigma_1, \rho_{0\eta}\sigma_0)$ can be estimated by Ordinary Least Squares after rewriting the error terms as in equations (3.8) and (3.7) and computing the estimated mills ratios $\lambda_0(\omega_i'\hat{\gamma}*)$ and $\lambda_1(\omega_i'\hat{\gamma}*)$. Hence, one can compute the residuals $\hat{\zeta}_{0i}$ and $\hat{\zeta}_{1i}$ from the linear regressions, so that it is possible to obtain consistent estimates for σ_0^2 and σ_1^2 from the empirical equivalents of equations (3.9) and (3.10).

3. The estimation process can be extended to include a third step that corrects for heteroskedasticity with a Feasible Generalized Least Squares estimator for β_1 and β_0 (and σ_1^2, σ_0^2). Note from equations (3.9) and (3.10) that, in the outcome equations, the variance of the error term ζ_{ki} differs

across individuals because of selection correction. Since residuals $\hat{\zeta}_{0i}$ and $\hat{\zeta}_{1i}$ are available from OLS regressions, it is possible to construct an $N \times N$ diagonal matrix $\hat{\Omega}_k = diag(\hat{\zeta}_{k1}^2, \hat{\zeta}_{k2}^2, ..., \hat{\zeta}_{kN}^2)$. The matrix can then be used to obtain Weighted Least Squares estimators for each outcome equation.

Once recovered the main parameters of interest throughout the main estimation process, it is straightforward to check which remaining parameters are identified. Note that the coefficients γ are identified up to a factor σ_η and the estimates $\hat{\gamma}^*$ are already available from the first step. Since I required an exclusion restriction so that there is at least one x_{ij} for which $\delta_j = 0$, it is possible to estimate the structural parameters of the choice equation as follows. Call x_i^R the covariates for which the aforementioned restriction holds, while x_i^U are the others. Using $\hat{\beta}_1$ and $\hat{\beta}_0$, one can single out the covariate $x_i'(\hat{\beta}_1 - \hat{\beta}_0)$ so that it is linear in $1/\sigma_\eta$ in the latent inequality. In brief, this means running the following Probit regression:

$$d_i = \mathbb{1}\{v_{1i} - C > v_{0i}\} = \mathbb{1}\{\eta_i < [x_i'(\hat{\beta}_1 - \hat{\beta}_0)]\frac{1}{\sigma_\eta} - z_i'\frac{\delta_z}{\sigma_\eta} - (x_i^U)'\frac{\delta_x}{\sigma_\eta}\}.$$

Note that it is only thanks to the exclusion restriction that avoiding perfect multicollinearity is possible, since $x_i^U \neq x_i$. It follows that an estimate for σ_η^2 and consequently those for parameters δ are readily obtainable. Estimators for the structural γ parameters can now be easily computed by rescaling the reduced form $\hat{\gamma}^*$ estimated at the first step.

As concerns the covariance structure of random components, the theory provides a few restrictions on the relationships between its components. It is straightforward to compute $\rho_{k\eta}$, $k = \{0, 1\}$, since at this stage the estimates for σ_k^2 and $\rho_{k\eta}\sigma_k$ are already available. In addition, σ_η^2 is known and it amounts to $Var[\varepsilon_i] = Var[\varepsilon_{0i} - \varepsilon_{1i} + \varepsilon_{ci}]$. However, the number of restrictions is not sufficient to solve for all the components of the covariance. The following system, in fact, still has four unknowns $(\sigma_c^2, \sigma_{1c}, \sigma_{0c}, \sigma_{01})$:

$$\begin{cases} \rho_{1\eta}\sigma_1 = Cov(\frac{\eta}{\sigma_\eta}, \varepsilon_1) = \frac{\sigma_{01} + \sigma_{1c} - \sigma_1^2}{\sigma_\eta} & (3.13) \\ \rho_{0\eta}\sigma_0 = Cov(\frac{\eta}{\sigma_\eta}, \varepsilon_0) = \frac{\sigma_0^2 - \sigma_{01} + \sigma_{0c}}{\sigma_\eta} & (3.14) \\ \sigma_\eta^2 = \sigma_0^2 + \sigma_1^2 + \sigma_c^2 - 2\sigma_{01} + 2\sigma_{0c} - 2\sigma_{1c}. & (3.15) \end{cases}$$

Because of this, the empirical analysis of the covariance is limited to $\rho_{1\eta}$, $\rho_{0\eta}$, σ_1^2 and σ_0^2. Nonetheless, these parameters are sufficiently insightful in order to discuss the relevance of self-selection, which crucially depends on the sign of $\rho_{1\eta}$, $\rho_{0\eta}$ as discussed in section 4.3. Another strategy would be to make an assumption on σ_c^2, such as to set it at the same order of magnitude of σ_1^2 and σ_0^2. In this way, one could learn something on the sign of the covariance terms. In particular, it is possible to identify a threshold value for σ_c^2 at which the covariance σ_{01} changes sign. From the system above, it follows:

$$\sigma_{01} > 0 \Leftrightarrow \sigma_c^2 < \sigma_1^2 + \sigma_0^2 + \sigma_\eta(\rho_{1\eta}\sigma_1 - \rho_{0\eta}\sigma_0 + \sigma_\eta) \tag{3.16}$$

which means that if $\rho_{1\eta}\sigma_1 - \rho_{0\eta}\sigma_0 + 1 > 0$ the variance σ_c^2 of unobserved heterogeneity in costs must not be too high to have a positive correlation between "skills" in the two sectors. On the contrary, if

$\rho_{1\eta}\sigma_1 - \rho_{0\eta}\sigma_0 + 1$ is negative and exceeds the sum $\sigma_1^2 + \sigma_0^2$, σ_{01} would necessarily be negative because σ_c^2 is always positive, which is potentially a useful restriction. The interpretation of these relationships will be clarified in light of the results in section 4.3.

3.3 Utility Levels and Preference Parameters.

The main peculiarity of this paper is the idea of considering utility levels derived from the solution of the household problem as the outcomes for the binary choice. Since there is no uncertainty in the model, agents know the exact payoffs v_{i1} and v_{i0} they can reap for both available paths. This implies that each agent knows his/her partner in advance, or at least the partner's income. None of the two interpretations sounds natural. In the first place, it is extremely unlikely that every agent can predict the partner's identity by the end of the high school. But, even in the realistic case where agents do not know their future spouses, they would still make a reasonable forecast about their future household life, i.e. whether they will be married and how their partner might look like. Ideally, this would require a more sophisticated model of the marriage market that explains how individuals form rational expectations on their marital status and potential partner, although, unfortunately, this task falls outside of the scope of this work, as discussed later in section 4.4 and in the conclusion. In this paper, I only focus on the case where perfect foresight extends to the partner's income and thus agents know their future household income with pinpoint accuracy.

In the rest of the section, I briefly discuss the identification of the parameters of the household problem and how in practice I compute the utility levels v_{i1} and v_{i0} to use in the Roy model. The set of parameters is given by (α, s, p). In principle, with data on prices at hand, one could calibrate p and identify α and s from the First Order Conditions of the household problem. However, this requires detailed data on family consumption patterns, in order to measure the household demand for public good and individual demands for private goods. Consider the relationship between the amount of public good Q and of private good q_i demanded by each individual:

$$\frac{q_i}{Q} = \frac{\alpha}{1+s}p.$$

From this equation, one cannot separately identify α and s. Nonetheless, assuming that preferences do not vary after marriage, it is possible to recover α from singles' First Order Conditions (where $s = 0$) and then go back to the married agents' problem to retrieve s. Note that this does not require the assumption that preferences are identical across genders. GMM estimation in the spirit of Hansen and Singleton (1982) could be explored in order to estimate (s, α). Although estimation is possible, there would be several issues that would threaten the reliability of a procedure relying on the identification discussed here. For instance, the difficulty in assessing what and to which extent goods are public and the possible distortion due to lack of information on home-produced goods (especially related to housework and caring). Therefore, in this work, I choose (α, s) arbitrarily and let them vary to see how they impact the results. In appendix B, I present the distribution of equilibrium utility levels for different picks of the preference parameters.

Now consider the structure of indirect utility functions. Regardless of which equilibrium prevails, the wife's and husband's indirect utility functions depend on both incomes y_m and y_f and price levels p, given preference parameters (s, α). This means that, as long as data on marriage patterns are available and for each married individual i the partner's income is known, it is possible to compute utility levels $v_i = v(y_m, y_f; p, \alpha, s)$. The function $v(.)$ is not smooth, since it takes a different form according to relative income y_m/y_f, although one can verify that at thresholds between intervals the function has kinks but no jump. Moreover, the function needs to be defined separately for husbands and wives since for all non-interior solutions utility levels are different across the couple. Therefore, using the equilibrium quantities found in section 2, I write the function $v_f(.)$ over its whole support for any married woman f. To find $v_m(.)$ for her husband, it actually suffices to switch the intervals.

$$v(y_m, y_f; p, \alpha, s) =$$

$$\begin{cases} (1+\alpha)(1+s)\log(y_m+y_f) + \alpha \log \frac{\alpha}{1+\alpha}\frac{s}{1+s} + \alpha s \log \frac{\alpha}{1+\alpha}\frac{1}{1+s} + (1+s)\log \frac{1}{1+\alpha}\frac{1}{p} & if \ \frac{y_f}{y_m} < \frac{\alpha s}{1+\alpha+s} \\ \alpha \log y_f + (1+s+\alpha s)\log y_m + \alpha s \log \frac{\alpha}{1+\alpha+s} + (1+s)\log \frac{1+s}{1+\alpha+s}\frac{1}{p} & if \ \frac{\alpha s}{1+\alpha+s} < \frac{y_f}{y_m} < \frac{\alpha}{1+\alpha+s} \\ (1+\alpha)(1+s)\log(y_m+y_f) + \alpha(1+s)\log \frac{\alpha}{1+2\alpha+s} + (1+s)\log \frac{1+s}{1+2\alpha+s}\frac{1}{p} & if \ \frac{\alpha}{1+\alpha+s} < \frac{y_f}{y_m} < \frac{1+\alpha+s}{\alpha} \\ (1+\alpha+s)\log y_f + \alpha s \log y_m + \alpha \log \frac{\alpha}{1+\alpha+s} + (1+s)\log \frac{1+s}{1+\alpha+s}\frac{1}{p} & if \ \frac{1+\alpha+s}{\alpha} < \frac{y_f}{y_m} < \frac{1+\alpha+s}{\alpha s} \\ (1+\alpha)(1+s)\log(y_m+y_f) + \alpha \log \frac{\alpha}{1+\alpha}\frac{1}{1+s} + \alpha s \log \frac{\alpha}{1+\alpha}\frac{s}{1+s} + (1+s)\log \frac{1}{1+\alpha}\frac{1}{p} & if \ \frac{y_f}{y_m} > \frac{1+\alpha+s}{\alpha s} \end{cases}$$

Despite looking complicated, writing down the functional form of $v(y_m, y_f; p, \alpha, s)$ is helpful. In fact, it is possible to apply an important transformation: subtracting $-(1+s)\log p$, which is present for every interval, the price level disappears. In practice, this means that the $-(1+s)\log p$ is absorbed by the intercept in the linear model described by equations (3.5) and (3.6). Therefore, there is no need to calibrate the relative price p in the estimation.

From the structure of the indirect utility function, it is straightforward to see how variations in preference parameters affect utility levels, as well as how local income pooling occurs in three out of five intervals. Changes in parameters (α, s) have two effects: they both shift utility levels up and down through the additive constants and magnify or reduce the marginal effect of (individual or household) income. The appendix B shows how the utility level distribution for married women in the sample described at section 4.1 varies with the parameters. For higher levels of s and α, the whole distribution shifts to the right and has fatter tails. For low s and α, the peak is extremely pronounced - especially in the graduate subsample - and the distribution becomes more symmetric, with a strong concentration around the mode.

Finally, in the empirical analysis I estimate the Roy model for single women using the indirect utility function retrieved by solving the trivial household problem for singles. The function takes the following form for any man or woman i:

$$v(y_i; p, \alpha) = (1+\alpha)\log y_i + \alpha \log \frac{\alpha}{1+\alpha} + \log \frac{1}{1+\alpha}\frac{1}{p}$$

Note that it is still possible to drop the term $\log p$ and let it be absorbed by the constant of the linear model.

4 Empirical Analysis.

In this section, I present the results from the estimation of the model with 2008-2012 data for California, after a brief discussion of the exclusion restrictions. I focus on women's college graduation since the model allows to estimate returns to schooling for women that are not in the labor force, which is a non-standard feature of this model. The estimation aims to shed some light on the parameters of choice and outcome equations, on the structure of the variance and on the distribution of the expected outcomes and returns.

4.1 Sample.

The data source for the empirical analysis is the 2008-2012 Public Use Microdata Sample for the state of California from the United States Census Bureau. Tables and graphs are available in appendix A. The database is extremely large and provides detailed information on family structure, housing, income and main individual traits. On the other hand, it provides little or no information on college studies,working career, parents' characteristics and psychophysical traits. However, thanks to the large size, it is possible to select subsamples of interest without being forced to work with small numbers, a characteristic that makes the database a suitable starting point to see how the model performs. In this case, I create four subsamples: husbands, wives, single men and single women. First, I select married couples where both partners are aged between 25 and 50. Then, I select single men and women in the same age bracket. This sample selection excludes cohabitating and same-sex couples, who both have distinct patterns and would deserve separate analysis.

In addition, I also exclude households with an unemployed member: this means that I ignore both families when any of the two partners is currently not working but looks for a job and unemployed singles. 5.37% of married men and 5.75% of married women declare themselves unemployed, while among singles the rates are 7.52% for men and 7.66% for women. Since I am interested in the wage-earning potential of individuals, unemployment is a hindrance because it is likely to be a temporary - and possibly unexpected - status that prevents the observation of such potential. On the contrary, I include people outside of labor force, who amount to 30.59% of married women, 4.43% of married men, 13.38% of single women and 8.65% of single men. These statistics on participation suggest that married women are much more likely to quit the labor market than singles, while the contrary is true for men. This is a possible clue for household specialization. Finally, I also include people that have not finished their studies yet, who are a negligible percentage because the 25-50 age bracket typically does not include schooling age. At this stage, the sample is composed of 112,043 couples, 49,340 single women and 37,869 single men.

I only consider two personal traits: age and ethnicity. Tables A.1, A.2, A.3 and A.4 provide a description of age and ethnic patterns in the sample. California is characterized by a pronounced ethnic diversity, especially because of the large Hispanic and Asian populations. Note that the population of singles is younger than the married one and that some ethnic groups - Hispanic and Asian - are overrepresented among married, while others - White and Black - are underrepresented. Table A.5 provides an overview of ethnic homogamy rates in the married population. Homogamy rates are computed as the ratio between the observed number of couples of a certain ethnic mix and a counterfactual density computed as if couples were formed randomly with respect to ethnic background (i.e. using the product of ethnic groups' marginal densities). As one could expect, values on the diagonal are far higher than those off the diagonal, which suggests that people sort into couples according to their ethnicity.

In this study, I only consider annual labor income, given as the sum of salary and self-employment earnings. Table A.6 provides an overview of labor income distributions for the different subsamples. Interestingly, single men seem to earn less than married, while the inverse is true for women, which is another clue for household specialization: in fact, this may be due not only to higher wage rates but also to a higher fraction of time dedicated to work. In addition, note that correlation between partners' income is positive but weak for married couples (0.1406). Finally, in the empirical application I exclude 1,072 households that still have no labor income after selecting on employment, and I trim outliers of the household income distribution (1% at the top, 1% at the bottom). Similarly, I exclude singles with zero income: in the case of single women, the cut is more significant, since 5,698 out of 49,340 need to be excluded from the sample. Also in the case of singles, I apply trimming in an analogous way.

Another variable I look at in the analysis is the rate of college attendance by cohort. Not every person that goes to college succeeds in obtaining a degree, therefore this rate is considerably higher than the graduation rate. For instance, over the whole 25-50 year old non-student population in the main sample, 21% of the individuals attended college without completing it, whereas 35.41% graduated. As shown by figure A.1, the enrollment rate continued to increase - of about 7 percentage points - for cohorts of women between 1962 and 1987, whereas it is more oscillating around the same value - about 52% - for men. As concerns completion rates for the subsamples considered in this analysis, 41.10% of married women, 39.49% of married men, 40.01% of single women and 42.98% of single men graduated. Note that several studies (e.g Bound, Lovenheim, and Turner, 2010; Light and Strayer, 2000) report a mismatch between the trends of enrollment and completion, with the former steadily increasing and the latter stagnating or declining in the very last decades.

4.2 Exclusion Restrictions.

Exclusion restrictions are imposed on attendance by cohort and age. The percentage of women subscribing to college is assumed to matter only for the determination of the costs, and not directly of the outcomes, i.e. $\beta_{k,\%attend} = 0$. A higher attendance among women is supposed to reduce costs of completing graduate studies because it generates positive externalities for the newcomers, such as a decrease in cases of discrimination in class and on campus, larger friendship networks, growing number of specific student associations, etc. The same may be true for men, especially as concerns sports and other leisure activities. In addition, increased college attendance of students of the same gender is likely to generate peer pressure in the choice of what to do after high school. On the other hand, considering colleges as marriage markets, higher enrollment may imply fiercer competition. However, consider that the variable in question is *not* equivalent to gender ratio: an increase in women's attendance does not necessarily coincide with an unfavorable decrease in the ratio men/women, although figure A.1 shows that men's attendance is relatively stable over time. In general, I expect attendance rates to increase the cost of *not enrolling* to college and to decrease the cost of *completing* graduate studies.

One of the main threats to the validity of the exclusion restriction is the possibility that changes in the percentage of graduates affects equilibrium wages on the labor market. Nonetheless, what should matter the most for wage determination for graduate and non-graduate workers is the *overall* percentage of graduates, and not the percentage by gender. The problem partly persists if the latter still matters in particular sectors, i.e. if gender specific skills are relevant in wage determination. However, what endangers the identification strategy the most is the distribution of the excluded variable (shown in table A.1): ideally, college attendance should not be correlated with age too strongly and should display sufficient variability. If the instrument does not have these characteristics, it adds little information and collinearity persists. This is especially risky in the case of men, because variability of college attendance is relatively limited.

In addition, I impose another exclusion restriction on age, so that $\delta_{age} = \delta_{age^2} = 0$. Willis and Rosen (1979) make a similar assumption for working experience, which unfortunately is not available in the ACS data. In the outcome equations age should primarily capture the positive effect of working experience, and possibly the likelihood of having a wealthier partner. The main concern may be the presence of psychological and monetary costs of graduating at a slightly older age. Nonetheless, with the set of variables at disposal, age remains the main candidate for an exclusion restriction on δ. Finally, note that another problem might arise if the difference of the outcome equation coefficients $\beta_1 - \beta_0$ of the excluded variable were close to zero, since the identification of σ_η^2 would fail. In this analysis, this does not seem to be the case, since the marginal effect of age on utility levels differ across graduate and non-graduates.

Because of all the critical points stressed in this section, the results must be interpreted with extreme caution. As explained in section 3.2, although one could hope to handle a Roy model without any of them, valid exclusion restrictions are still the best way to ensure that identification is successful. If such restrictions are weak, then results are much less credible.

4.3 Results for Married Women.

In this section I first focus on the results for married women: with 30.59% of them out of the labor force, a correct estimation of their returns to college must consider what share of surplus they obtain from marriage. Their actual resources and consumption patterns, in fact, depend on the dynamics of the household and, in particular, the model set in section 3 predicts that their utility also depends on the partner's income, and not only on personal income.

I estimate the model for several values of (α, s), with the three-step estimation procedure with heteroskedasticity correction. Tables with the estimated coefficients can be found in appendix C, whereas graphs with the distributions of the returns to college for graduates and non-graduates can be found in appendix D. The estimates for the coefficients of the model do not vary significantly with (α, s): signs and significance levels are the same for every considered pair (α, s). I will go back to this later, while I first report results for $\alpha = 1.2$ and $s = 0.5$ in table 4.1, which is arbitrarily chosen as a benchmark case. The first two columns report estimates for the outcome equations. The estimated coefficients suggest that utility levels are concave in age for graduates. This should not be surprising since utility levels depend on the incomes of both partners: concavity is likely to reflect both the shape of the labor earnings curve over time and the increase over time of the probability that the husband is already an experienced worker. Concavity is instead absent in the outcome for non-graduates. Note that, by sampling only individuals aged between 25-50, concavity is likely to be underestimated because the sample contains no information about wage curves after 50. Ethnic minorities mainly suffer from utility losses (except for Hispanic graduates), which are stronger for non-graduates. This may due to a penalty both on the labor market - because of discrimination - and on the marriage market, in that non-graduate women from ethnic minorities may be more likely to marry poorer men. Homogamy rates shown in table A.5 seem to provide support for this statement, since income distributions by ethnic group differ from each other in a non-negligible way: median income for White married men is the highest (76,924), while it is progressively lower for Asian (70,592), Black (53,496) and Hispanic (36,383). Therefore, a White woman may not only have an advantage over ethnic minorities in her working career, but she is also more likely to have a high-income partner.

	β_1	$SE(\beta_1)$	β_0	$SE(\beta_0)$	γ	$SE(\gamma)$	δ	$SE(\delta)$
Graduates	0.000	0.000	0.000	0.000	6.928***	1.087	-6.928	0.298
Constant	28.418***	1.046	30.089***	0.291	-8.439***	0.868	6.777***	0.241
Age	0.339***	0.042	0.093***	0.029	0.246***	0.014	0.000	0.000
Age^2	-0.004***	0.001	-0.001	0.000	-0.003***	0.000	0.000	0.000
Black	-0.333***	0.135	-0.319***	0.102	-0.568***	0.042	0.557***	0.042
Hispanic	-0.523	0.392	-1.312***	0.215	-1.737***	0.015	2.529***	0.041
Asian	-0.262***	0.079	-1.031***	0.076	0.418***	0.016	0.350***	0.040
Inv. Mills Ratio	-0.467	0.481	1.083***	0.372	-	-	-	-

$\sigma_0^2 = 5.0550,\ \sigma_1^2 = 4.0183,\ \rho_{0\eta} = 0.4799,\ \rho_{1\eta} = 0.2337$

$N = 108,751,\ N_1 = 44,692,\ \alpha = 1.2,\ s = 0.50.$

Statistical significance is assessed with two-tailed tests: *** implies significance at 1%∗, ** at 5% and * at 10%.

Table 4.1: Married women ($\alpha = 1.2, s = 0.50$)

In the last column of table 4.1, the estimates for the cost function are all significant and the direction of marginal effects are rather intuitive. Note that a positive δ implies a positive marginal effect on the costs of graduating, whereas a positive γ indicates a positive marginal effect in the choice equation estimated with the Probit regressions: I focus on δ because the interpretation is simpler. First of all, the positive constant represents a lump sum cost for completing college, whereas all ethnic minorities seem to suffer from an additional cost with respect to the White reference group, which is consistent with heterogeneity in primary and high schools formation, theories of racial discrimination and ethnic identity in schools[8].

The percentage of graduate women by cohort calls for particular attention: its coefficient has a significant and negative impact on costs as predicted in section 4.2. However, note that the coefficient is large only because of the unit of measure: an increase of one percentage point in female enrollment causes costs to decrease by about 0.07, which means that, for instance, an Asian woman would need a 5% increase in enrollment to compensate her penalty due to ethnic background. Although the instrument is always significant at 1% level, figure A.1 shows that attendance spans over a 7-percentage-point interval: the validity of the instrument is probably at the limit. This suggests

[8]See for instance Akerlof and Kranton (2002) for an analysis of social norms and ethnic identity in schools and Fryer Jr and Levitt (2004) for an analysis of the possible explanations of why children from ethnic minorities (here of Black origins) lose ground in the first years of school.

that the same instrument would not work if I were to estimate the model on the men's sample: since the variability of male enrollment is extremely limited, the validity of the instrument would be compromised.

Going back to outcome equations, now consider the estimates of the linear coefficients for the inverse Mills ratios. A positive coefficient indicates that the conditional expected unobserved component has a higher mean for the subsample observed in that state with respect to the whole sample mean for unobserved skills in the same state, i.e. people who selected into that "sector" - to use Roy's original terminology - have unobserved skills that help them to perform well in it. In the case of non-graduate married women, I find that the linear coefficient for $\lambda_0(\omega_i'\hat{\gamma}^*)$ is significant and positive (1.083): it thus seems that those who do not choose to attain a degree belong to the upper distribution of the unobserved component ε_{0i}.

More intuitively, this implies that, because of some unobserved characteristics, non-graduate women are successful even without a degree: at their place, those women who actually graduated would not perform equally well. Thinking of the labor market, the canonical interpretation is that non-graduate women have some sector-specific skills that allow them to attain a sufficiently high wage and consequent desirable living standards without paying the cost of going to college. On the contrary, graduate women prefer to escape from the "unskilled" sector and go through college before entering the labor market. Nonetheless, since in this model payoffs also depend on husbands' income, this result suggests that non-graduate women are also aware that they can find a partner with a sufficiently high income without attending - or without completing - college. On the contrary, women that chose to attain a degree forecast, given their unobserved characteristics, that they would not have been able to find a "good" partner without the help of their studies. Hence, they choose to attain a college degree, which can be intuitively explained by two non-mutually exhaustive reasons: first, since they have higher wages (as shown in figure A.2), they are also more attractive on the marriage market[9]; second, on the college campus it may be easier to find a wealthier husband.

At the same time, the linear coefficient for $\lambda_1(\omega_i'\hat{\gamma}^*)$ for the married graduates subsample has negative sign but no statistical significance. This may be quite surprising because one would expect to find that those who obtained a degree were the best suited to work in the skilled sector (and to find the highest-income partners). Nevertheless, the result suggests that their unobserved skills do not provide them with any particular advantage over the rest of the population. The sign of the non-significant estimate even seems to suggest that non-graduates could actually be able to outperform those who have a degree even in the "skilled" sector. In any case, the coefficient being negative or null, this relative lack of advantage in the "skilled" sector is by far less important than the strong disadvantage that graduates would experience in their counterfactual state. In fact, their choice to graduate is motivated more by the will of *not* staying in the "unskilled" sector rather than that of moving to the "skilled" sector, as clarified in the former paragraph.

Now consider gross and net returns to college defined by equations (3.11) and (3.12). Figures 4.1 and 4.2 show the distribution of gross and net returns to college in the benchmark case ($\alpha = 1.2$,

[9] Clearly, at college they could also earn other desirable "skills and assets" - such as cultural knowledge, social networks, upper-class behavior, etc. - that allow them to be more competitive on the marriage market. In general, homogamy with respect to education is a widely-studied phenomenon, whose reasons are multiple and complex. See Kalmijn (1998) and Schwartz and Mare (2005) for a complete overview of demographic trends and possible explanations.

$s = 0.50$). In black, there are the marginal returns for those who actually graduated, in red there are the marginal returns for those who did not graduate. The difference in the distribution between the two subsamples is due to two reasons. The first is the diversity in the composition of the graduate and non-graduate subsample with respect to observed characteristics x_i: for instance, Hispanic women have an additional incentive to graduate with respect to White ones, since $\beta_{1,Hispanic} - \beta_{0,Hispanic} = 0.7890$ is positive, but this is washed out by the high additional cost $\delta_{Hispanic} = 2.529$, so that their net marginal returns to college are smaller than those for the White reference group. Since only 11.69% of married women actually graduated and Hispanic constitute a large part of the sample (see A.2), they will be relatively more relevant in determining the distribution of net returns among non-graduates, making it shift to the left. The second reason is the above-mentioned self-selection process and concerns unobserved heterogeneity. Graduate women have higher marginal returns to college because their unobserved characteristics would make them perform poorly in the unskilled sector. In fact, note that the results show that the component of the marginal returns for graduates in equation (3.11) accounting for self-selection $E[\varepsilon_{1i} - \varepsilon_{0i}|x_i, z_i, d_i = 1] = -(\rho_{1\eta}\sigma_1 - \rho_{0\eta}\sigma_0)\lambda_1(\omega_i'\gamma^*)$ is positive, because $-\rho_1\hat{\eta}\sigma_1 + \rho_0\hat{\eta}\sigma_0$ is greater than zero (0.616), which certainly explains part of the distance between the distributions. Finally, note that, as one would expect, net marginal returns are lower - but still mostly positive also for non-graduates - for both subsamples and that the distributions are more concentrated than those of gross returns.

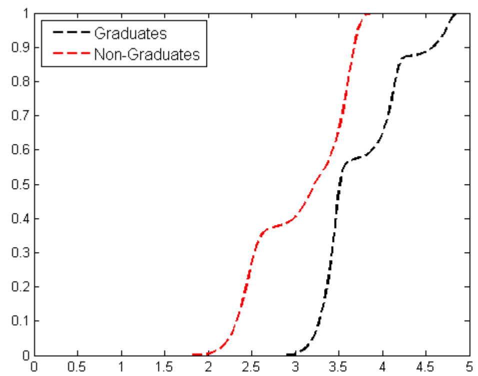


Figure 4.1: Gross marginal returns for married women ($\alpha = 1.2, s = 0.50$)

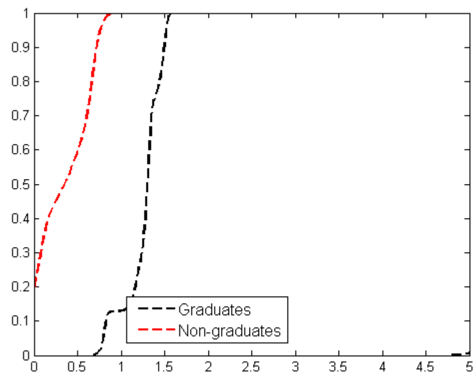


Figure 4.2: Net marginal returns for married women ($\alpha = 1.2, s = 0.50$)

Other figures displaying gross marginal returns for different picks of α and s can be found in section D. A comparison is needed to understand the role of preference parameters and household dynamics. Note first that the distributions of gross marginal returns is slightly flatter and shifts to the right for higher s and α. Most importantly, for higher levels of s there seems to be a greater distance between the distribution of expected returns for graduates and the one for non-graduates (see figures D.2 and D.4 with respect to figures D.1 and D.3). Remember that an increase in the relevance of caring preferences implies three things: first, when there is no positive transfer, public good provision grows (and approaches the efficiency level); second, when there is a positive transfer from the husband to the wife, the transfer becomes larger; third, for any equilibrium, the "sympathy" for partner's consumption grows. Therefore, gains from increased caring are systematically higher for *any* married woman in *any* counterfactual state. Nonetheless, it seems that the increase in s exacerbates self-selection: by looking carefully at the tables in appendix C, one can notice that the differences $\hat{\beta}_1 - \hat{\beta}_0$ are relatively stable when varying the preference parameters[10], which means that the component of gross returns due to observables most likely cannot account for the whole change. It follows that a part of the change must be explained by an increase in the linear coefficients of the selection correction terms and in the distribution of the inverse Mills ratios. Because of the self-selection process previously described, an explanation that could justify this shift in the distance between distributions points at the fact that the partner's income becomes more important when caring preferences matter more. Consequently, if a woman knows that her unobserved characteristics do not allow to find a desirable partner without a college degree, she wants to "escape" from the non-graduate population even more than before.

In order to exhaust the analysis on married women's educational choice, consider the information obtained on the covariance. For instance, take $\hat{\rho}_{0\eta} = 0.4799$, which is larger than $\hat{\rho}_{1\eta} = 0.2337$: a relatively high correlation with η_i implies that, for higher value of the specific unobservable factor ε_{0i}, the probability of not attaining a degree increases fast, as evident from the choice equation 3.4. It is indeed the relationship between the two $\rho_{k\eta}$ that capture the effects of self selection. The observed high correlation $\rho_{0\eta}$ may imply a relative high covariance σ_{0c} between unobservables in the non-graduate state and unobservables in costs, or a low covariance σ_{01} between unobservable skills in both states. On the contrary, a relatively low but positive $\rho_{1\eta}$ implies the inequality $\sigma_{01} + \sigma_{1c} - \sigma_1^2 < 0$, which is not of particular help in the analysis of signs. The condition (3.16), instead, may provide some clue: the threshold for a positive covariance σ_{01} is given by $\sigma_c^2 < 10.5032$: unless $\sigma_1^2 = 4.0183$ and $\sigma_0^2 = 5.0550$ are less than half the variance of unobserved heterogeneity in costs, it seems that σ_{01} can be positive. Actually, note that the condition (3.16) does not hold - i.e. unobserved skills are negatively correlated across sectors - only when there are strong positive sorting patterns on unobservables, which means that every agent goes to the sector where he is most productive and which implies that both linear coefficients for selection correction terms are strong and positive. This is not the case for married women, since apparently non-graduates outperform graduates in the "unskilled" sector, whereas the reverse is not true.

4.4 Married vs Singles.

In this section, I introduce a brief comparison between the results presented in section 4.3 and those obtainable from the estimation of the model with a subsample of single women. The comparison aims to show the structural differences in the process of self-selection and in the outcome distributions. Before moving on, I must caution the reader that the model is not sophisticated enough to establish a relationship between utility levels of singles and married women. The indirect utility function of single women, in fact, only depends on personal income and utility levels are systematically lower than those for married women. However, this is not due to the fact that public good is provided jointly and through two sources of earnings (or directly household income, when local income pooling holds),

[10]Consider tables C.1 and C.2, where for given $\alpha = 1.2$, s shifts from 0.50 to 0.75: the absolute values of the differences between the difference $\hat{\beta}_1 - \hat{\beta}_0$ is $\{0.663, 0.056, 0.001, 0.024, 0.001, 0.158\}$, where the order of the labels is the same as in the tables. The most relevant change concerns the constant term, for which $\hat{\beta}_1 - \hat{\beta}_0$ grows with s by 0.663. At the same time, also the difference in coefficients for the selection correction term grows significantly (by 0.175) and relatively more than the other linear coefficients, consistently with what explained in the main text. Similar differences can be computed between tables C.3 and C.4.

but rather to the magnifying effect that caring preferences have on utility. On the contrary, singles do not enjoy any "prize" for their status. This is of course an oversimplification due to the fact that the model does not treat the choice of marital status and the one of the partner as endogenous. In more complex frameworks, agents choose whether and who to marry and have a reservation utility for staying single that may depend on preferences and opportunities to date other potential partners in the future. However, the model abstracts from this aspect: this, among the several implications, prevents the comparison between singles' and married women's utility levels.

Another brief reminder is due: as anticipated in section 4.1, I exclude those observations whose income is zero, because consumption would also be zero and the marginal utilities for both goods would be infinite. Contrarily to the case of married women, the cut in the sample is more relatively large. The lack of zero-income observations may distort the distribution of outcomes, especially because of the 5,698 excluded observations only 754 have a college degree: thus, the outcome distribution for non-graduates may be inflated.

In light of these limitations, I present the results from the estimation of the Roy model on 25-to-50-year-old single women. Table 4.2 contains the estimated coefficients of the structural equations for $\alpha = 1.2$, the same value as the benchmark case. As concerns outcome equations, the coefficients β_1 and β_0 there are some differences to discuss. The evolution of utility level over time is slightly slower, which may be due to the fact that there is a unique income source. Estimates for ethnic minorities suggest that discrimination persists for Black, whereas it increases for Hispanic and decreases for Asian. Recalling the homogamy channel briefly described in section 4.3, this difference could be explain by the fact that Hispanic women succeed in avoiding the negative impact of their ethnic background through marriage, whereas Asian women perform better - at least as well as White - on the labor market but suffer from a penalty compared to other groups when the partner's income plays a role. The cost structure is similar, although the effect of female college attendance is now doubled: could it be that singles enjoyed the company of other women during college? Although this may be true, as reminded in section 4.3 the size of the effect is still small because of the unit of measure.

	β_1	$SE(\beta_1)$	β_0	$SE(\beta_0)$	γ	$SE(\gamma)$	δ	$SE(\delta)$
Graduates	0.000	0.000	0.000	0.000	22.690***	3.627	-22.690***	0.920
Constant	16.810***	0.370	20.422***	0.979	-16.455***	2.809	12.831***	0.522
Age	0.264***	0.026	0.052**	0.026	0.213***	0.040	0.000	0.000
Age^2	-0.003***	0.000	-0.000	0.000	-0.003***	0.000	0.000	0.000
Black	-0.527**	0.299	-0.605***	0.192	-1.765***	0.070	1.845***	0.071
Hispanic	-1.017**	0.543	-1.191***	0.297	-3.082***	0.050	3.254***	0.057
Asian	0.209	0.134	-0.123	0.125	1.238***	0.062	-0.908***	0.078
Inv. Mills Ratio	0.924	0.753	-0.690	0.568	-	-	-	-

$\sigma_0^2 = 4.5219$, $\sigma_1^2 = 2.8461$, $\rho_{0\eta} = -0.3280$, $\rho_{1\eta} = -0.5507$

$N = 42,769$, $N_1 = 17,030$, $\alpha = 1.2$.

Statistical significance is assessed with two-tailed tests: *** implies significance at 1%*, ** at 5% and * at 10%.

Table 4.2: Single women ($\alpha = 1.2$)

Most importantly, for the single women's subsample the self-selection process seems to have changed radically. Unfortunately, estimates for linear coefficients of the inverse Mills ratios are not significant and have relatively high variance considered the sample size. Their signs indicate that, if there is any effect, the conditional mean for the unobserved heterogeneity component in the "skilled" sector is higher for those who actually hold a degree. On the contrary, those who do not have a degree do not outperform the graduates in the "unskilled" sector. The situation appears as the opposite than what occurs for married women: those who attained a degree selected into college because they have suitable characteristics, whereas those who did not graduate stayed in the "unskilled" sector only because they are *less worse* and thus their returns to college are lower. If these patterns are driven by labor market skills, then what occurs is a sort of hierarchical sorting according to unobservable abilities.

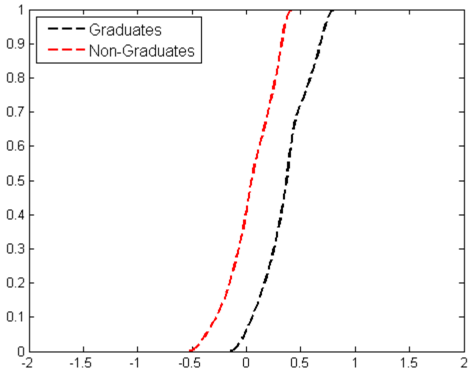


Figure 4.3: Gross marginal returns for single women ($\alpha = 1.2$)

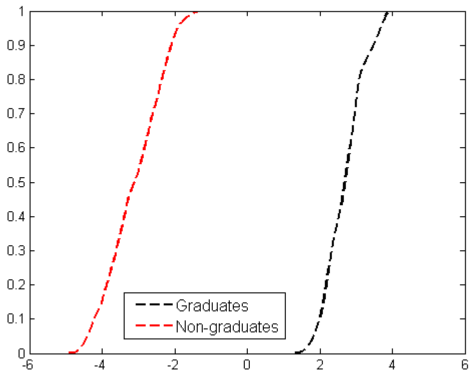


Figure 4.4: Net marginal returns for single women ($\alpha = 1.2$)

The substantial difference between the patterns observed for married women is of particular interest. One possible explanation is that the presence of a second source of income in the household and, more in general, the possibility to enjoy marriage surplus turn the tide in favor of non-graduate women when the partner is considered. Non-graduates, in fact, seem to have sufficient advantage over graduates on the marriage market not to feel forced to pursue a degree: the effect of hierarchical sorting according to labor market skills is then offset by the distribution of unobservable factors that matter on the marriage market. Finally, note that, since the estimates do not provide enough statistical evidence for self-selection in the case of singles, one could also argue that self-selection is completely absent in their case: this would still be an important different with respect to the married women's subsample, who do experience self-selection. The reasoning is similar, although hierarchical sorting on the labor market should be considered as a weak phenomenon, attenuated or inexistent.

5 Conclusion and Perspectives.

This paper explores the possibility of combining the self-selection approach to educational choice with the assessment of marriage surplus shares proposed by family economics theories. As concerns both the household problem and the binary choice, I use simple modeling specifications that belong to the two separate literatures: the non-cooperative household models with voluntary transfers by Chen and Woolley (2001) and Browning, Chiappori, and Lechene (2010) inspire the family dynamics described in this analysis, whereas the classical parametric model with perfect foresight by Willis and Rosen (1979) is the benchmark for the Roy model adopted here. Despite the oversimplification, the empirical application on the 2008-2012 ACS sample for California in section 4 leads to interesting results: weak hierarchical sorting on labor market skills occurs in the subsample of singles, whereas in the case of married women higher returns to college for graduates are driven by poor unobserved abilities in the unskilled labor market and in the marriage market.

The framework set in this paper can be a starting point for further research, which could develop it in several directions. As regards the Roy model, the introduction of uncertainty and non-pecuniary gains - with all the consequent technical sophistications - seems to be the primary way to follow (see the introduction and Cunha and Heckman, 2007, for an overview of the topic). Uncertainty has mainly concerned the future series of earnings streams, with education as a potential insurance to limit the negative impact of undesired variability (e.g. Blundell, Pistaferri, and Preston, 2008). However, uncertainty is another key aspect of the literature on matching models with uncoordinated investments: young students cannot know whether they will be willing to marry or who their partner will be, and thus they form rational expectations about their future marital status. In addition, note that theoretical and applied models of matching markets rely on different forms of conditional independence assumptions between observable and unobservable traits: adapting them to a dynamic context and deriving an individual expected utility function that depends on both sets of characteristics represents another challenging issue. Finally, a further complication is the difficulty in determining marriage gains throughout the whole marital life cycle, especially because singlehood, divorce and remarriage trends are becoming progressively more complicated. This suggests that it may be wise to turn to a dynamic representation of the life-cycle and tackle the problem from a different point of view, following the approach of Shimer and Smith (2000).

Hopefully, the paper conveys the idea that the analysis of returns to education on the marriage market is crucial to improve the comprehension of schooling and marriage patterns and that combining different strands of literature can be an insightful exercise. College attendance and the choice of the partner are two important - and strongly interconnected - decisions: analyzing them allows to better understand the medium- and long-term trends in the distributions of household income and labor supply.

References

Akerlof, G. A., and R. E. Kranton (2002): "Identity and schooling: some lessons for the economics of education," *Journal of Economic Literature*, 40(4), 1167–1201.

Arcidiacono, P. (2004): "Ability sorting and the returns to college major," *Journal of Econometrics*, 121(1), 343–375.

Ashenfelter, O., and D. Card (1985): "Using the longitudinal structure of earnings to estimate the effect of training programs," *The Review of Economics and Statistics*, 67(4), 648–660.

Becker, G. S. (1973): "A theory of marriage: part I," *Journal of Political Economy*, 81(4), 813–846.

Becker, G. S. (1974): "A theory of social interactions," *Journal of Political Economy*, 82(6), 1063–93.

Beffy, M., D. Fougere, and A. Maurel (2012): "Choosing the field of study in postsecondary education: Do expected earnings matter?," *Review of Economics and Statistics*, 94(1), 334–347.

Blundell, R., L. Pistaferri, and I. Preston (2008): "Consumption inequality and partial insurance," *American Economic Review*, 98(5), 1887–1921.

Borjas, G. J. (1987): "Self-selection and the earnings of immigrants," *American Economic Review*, 77(4), 531–53.

Bound, J., M. F. Lovenheim, and S. Turner (2010): "Why have college completion rates declined? An analysis of changing student preparation and collegiate resources," *American Economic Journal: Applied Economics*, 2(3), 129–157.

Browning, M., F. Bourguignon, P.-A. Chiappori, and V. Lechene (1994): "Income and outcomes: a structural model of intrahousehold allocation," *Journal of Political Economy*, 102(6), 1067–96.

Browning, M., and P.-A. Chiappori (1998): "Efficient intra-household allocations: a general characterization and empirical tests," *Econometrica*, 66(6), 1241–78.

Browning, M., P.-A. Chiappori, and V. Lechene (2010): "Distributional effects in household models: separate spheres and income pooling*," *The Economic Journal*, 120(545), 786–799.

Browning, M., P.-A. Chiappori, and Y. Weiss (2014): *Economics of the Family*. Cambridge University Press.

Chen, Z., and F. Woolley (2001): "A Cournot–Nash model of family decision making," *The Economic Journal*, 111(474), 722–748.

Chiappori, P. (1988): "Rational household labor supply," *Econometrica*, 56(1), 63–90.

——— (1992): "Collective labor supply and welfare," *Journal of Political Economy*, 100(3), 437–67.

Chiappori, P.-A., M. Iyigun, and Y. Weiss (2009): "Investment in schooling and the marriage market," *American Economic Review*, 99(5), 1689–1713.

Chiappori, P.-A., B. Salanié, and Y. Weiss (2015): "Partner choice and the marital college premium: analyzing marital patterns over several decades," .

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

Cunha, F., J. Heckman, and S. Navarro (2005): "Separating uncertainty from heterogeneity in life cycle earnings," *Oxford Economic Papers*, 57(2), 191–261.

Cunha, F., and J. J. Heckman (2007): "Identifying and estimating the distributions of ex post and ex ante returns to schooling," *Labour Economics*, 14(6), 870–893.

Dahl, G. B. (2002): "Mobility and the return to education: testing a Roy model with multiple markets," *Econometrica*, 70(6), 2367–2420.

D'Haultfœuille, X., and A. Maurel (2013): "Inference on an extended Roy model, with an application to schooling decisions in France," *Journal of Econometrics*, 174(2), 95–106.

Dougherty, C. (2005): "Why are the returns to schooling higher for women than for men?," *Journal of Human Resources*, 40(4), 969–988.

Duflo, E. (2003): "Grandmothers and granddaughters: old-age pensions and intrahousehold allocation in South Africa," *The World Bank Economic Review*, 17(1), 1–25.

Fryer Jr, R. G., and S. D. Levitt (2004): "Understanding the black-white test score gap in the first two years of school," *Review of Economics and Statistics*, 86(2), 447–464.

Goldin, C., L. F. Katz, and I. Kuziemko (2006): "The homecoming of American college women: the reversal of the college gender gap," *Journal of Economic Perspectives*, 20(4), 133–156.

Greenwood, J., A. Seshadri, and M. Yorukoglu (2005): "Engines of liberation," *The Review of Economic Studies*, 72(1), 109–133.

Hansen, L. P., and K. J. Singleton (1982): "Generalized instrumental variables estimation of nonlinear rational expectations models," *Econometrica*, 50(5), 1269–1286.

Heckman, J., P. Carneiro, and K. Hansen (2003): "Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice," *International Economic Review*, 44(2), 361–422.

Heckman, J. J. (1974): "Shadow prices, market wages, and labor supply," *Econometrica*, 42(4), 679–94.

——— (1976): "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models," in *Annals of Economic and Social Measurement*, vol. 5, pp. 475–492. NBER.

——— (1979): "Sample selection bias as a specification error," *Econometrica*, 47(1), 153–61.

Heckman, J. J., and B. E. Honore (1990): "The empirical content of the Roy model," *Econometrica*, 58(5), 1121–1149.

Iyigun, M., and P. R. Walsh (2007): "Building the family nest: premarital investments, marriage markets, and spousal allocations," *Review of Economic Studies*, 74(2), 507–535.

Kalmijn, M. (1998): "Intermarriage and homogamy: causes, patterns, trends," *Annual Review of Sociology*, 24(1), 395–421.

Konrad, K. A., and K. E. Lommerud (1995): "Family policy with non-cooperative families," *Scandinavian Journal of Economics*, 97(4), 581–601.

Lee, L.-F. (1978): "Unionism and wage rates: a simultaneous equations model with qualitative and limited dependent variables," *International Economic Review*, 19(2), 415–33.

Light, A., and W. Strayer (2000): "Determinants of college completion: school quality or student ability?," *Journal of Human Resources*, 35(2), 299–332.

Lundberg, S., and R. A. Pollak (1993): "Separate spheres bargaining and the marriage market," *Journal of Political Economy*, 101(6), 988–1010.

Lundberg, S. J., R. A. Pollak, and T. J. Wales (1997): "Do husbands and wives pool their resources? Evidence from the United Kingdom Child Benefit," *Journal of Human Resources*, 32(3), 463–480.

Manser, M., and M. Brown (1980): "Marriage and household decision-making: a bargaining analysis," *International Economic Review*, 21(1), 31–44.

McElroy, M. B., and M. J. Horney (1981): "Nash-bargained household decisions: toward a generalization of the theory of demand," *International Economic Review*, 22(2), 333–49.

Peters, M., and A. Siow (2002): "Competing premarital investments," *Journal of Political Economy*, 110(3), 592–608.

Roy, A. D. (1951): "Some thoughts on the distribution of earnings," *Oxford Economic Papers*, 3(2), 135–146.

Samuelson, P. A. (1956): "Social indifference curves," *The Quarterly Journal of Economics*, 70(1), 1–22.

Schwartz, C. R., and R. D. Mare (2005): "Trends in educational assortative marriage from 1940 to 2003," *Demography*, 42(4), 621–646.

Shimer, R., and L. Smith (2000): "Assortative matching and search," *Econometrica*, 68(2), 343–369.

Thomas, D. (1990): "Intra-household resource allocation an inferential approach," *The Journal of Human Resources*, 25(4), 635–664.

Trostel, P., I. Walker, and P. Woolley (2002): "Estimates of the economic return to schooling for 28 countries," *Labour Economics*, 9(1), 1–16.

Willis, R., and S. Rosen (1979): "Education and self-selection," *Journal of Political Economy*, 87(5), S7–S36.

A Descriptive Statistics.

	Age Brackets					
	25-30	31-35	36-40	41-45	46-50	
White	42.9	41.2	41.7	44.3	49.6	44.4
Black	3.1	2.8	2.7	3.0	3.3	3.0
Others	0.6	0.6	0.5	0.6	0.6	0.6
Asian	10.6	16.6	17.9	17.2	17.5	16.8
Hispanic	42.8	38.9	37.2	34.9	29.0	35.3
Marginals	8.6	16.9	23.2	25.9	25.4	

Table A.1: Age and ethnic patterns for 25-50 year-old married men

	Age Brackets					
	25-30	31-35	36-40	41-45	46-50	
White	39.2	37.7	39.9	43.8	49.4	41.8
Black	2.1	2.1	2.1	2.4	2.9	2.3
Others	0.6	0.7	0.6	0.5	0.7	0.6
Asian	16.0	20.5	20.3	19.3	17.5	19.1
Hispanic	42.2	39.0	37.1	33.9	29.5	36.2
Marginals	14.2	20.7	25.0	23.9	16.3	

Table A.2: Age and ethnic patterns for 25-50 year-old married women

	Age Brackets					
	25-30	31-35	36-40	41-45	46-50	
White	49.2	47.1	47.7	53.4	59.8	52.0
Black	5.8	5.8	6.7	7.5	8.5	6.9
Others	0.7	0.7	1.1	1.1	1.1	0.9
Asian	16.5	17.7	15.0	11.0	9.4	13.6
Hispanic	27.8	28.7	29.5	27.0	21.3	26.5
Marginals	20.3	17.1	17.9	20.2	24.5	

Table A.3: Age and ethnic patterns for 25-50 year-old single men

	Age Brackets					
	25-30	31-35	36-40	41-45	46-50	
White	43.8	38.4	38.1	43.1	49.5	43.2
Black	9.9	10.7	11.6	11.5	11.7	11.2
Others	0.7	0.9	0.9	0.9	1.0	0.9
Asian	13.6	13.6	13.3	12.3	11.3	12.7
Hispanic	31.9	36.3	36.1	32.2	26.4	32.0
Marginals	16.9	16.1	18.7	21.6	26.8	

Table A.4: Age and ethnic patterns for 25-50 year-old single women

	Husbands				
Wives	**White**	**Black**	**Others**	**Asian**	**Hispanic**
White	1.922	0.471	0.776	0.136	0.258
Black	0.323	27.557	0.525	0.067	0.168
Others	0.878	1.356	67.055	0.277	0.437
Asian	0.338	0.314	0.412	4.716	0.086
Hispanic	0.290	0.370	0.492	0.066	2.468

Table A.5: Homogamy rates for married couples in the sample

	Income Percentiles								
	0.01	**0.10**	**0.25**	**0.40**	**0.50**	**0.60**	**0.75**	**0.90**	**0.99**
Married Women	0	0	0	10,395.08	20,204.14	31,223.278	51,975.4	86,881.92	206,344.6
Single Women	0	0	13,683.9	25,678.08	34,347.04	43,440.96	62,370.48	93,555.72	203,284.80
Husbands	0	14,855.69	30,772.51	45,613.01	56,705.76	70,686.55	98,432.64	152,043.40	432,247.70
Single Men	0	5,305.605	23,344.66	37,422.29	47,076.48	57,172.94	80,042.12	119,462.60	407,113.40

Table A.6: Labor income distributions by category (in dollars)

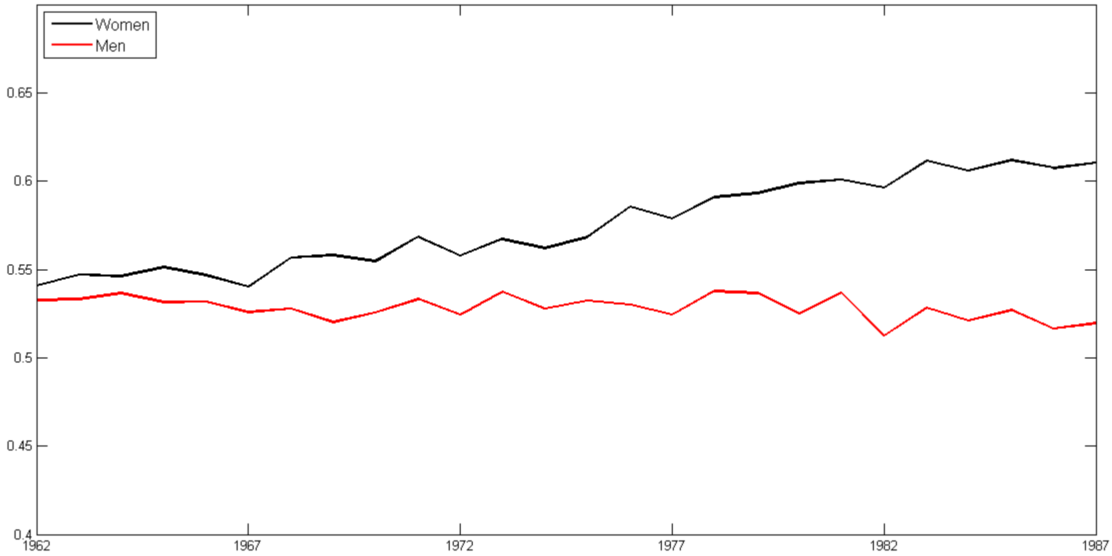


Figure A.1: Enrollment trends for 25-50 year-old population of California

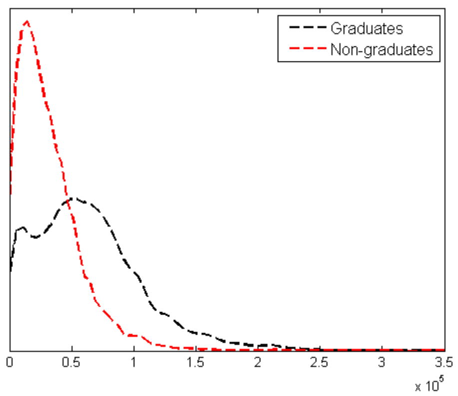


Figure A.2: Income distribution of working married women

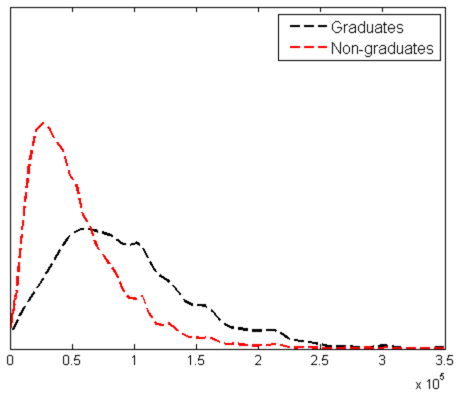
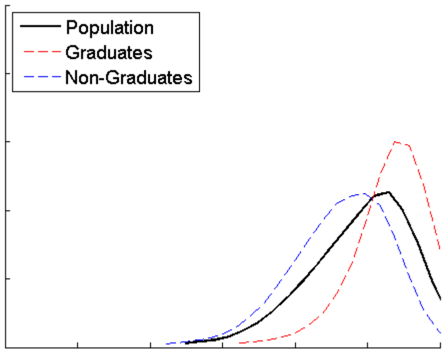


Figure A.3: Income distribution of working married men (by educational category of the wife)

B Outcome Distributions.



$s = 1, \alpha = 1.2$

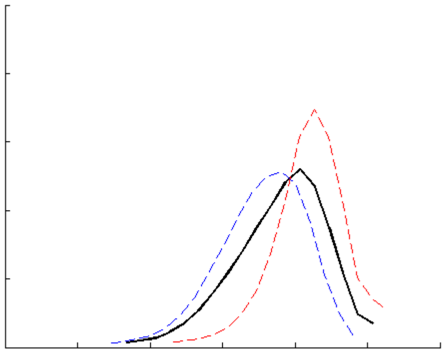


Figure B.2: $s = 0.75, \alpha = 1.2$

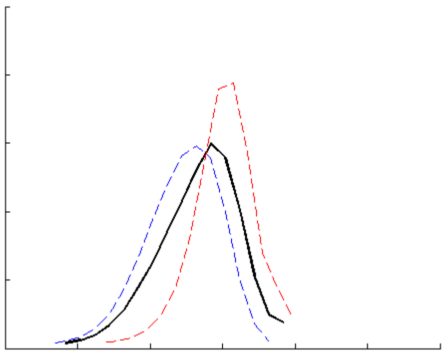


Figure B.3: $s = 0.5, \alpha = 1.2$

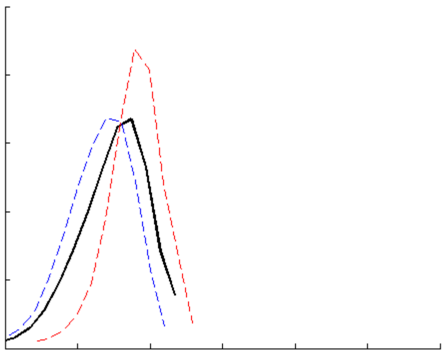


Figure B.4: $s = 0.25, \alpha = 1.2$

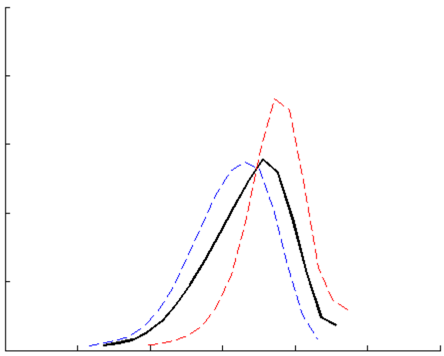


Figure B.5: $s = 1, \alpha = 0.8$

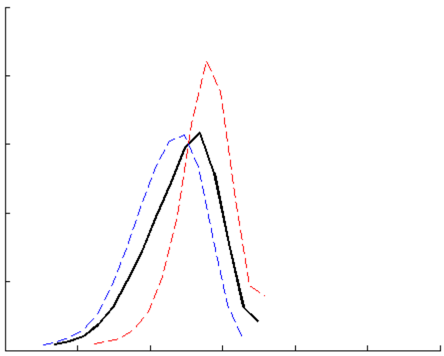
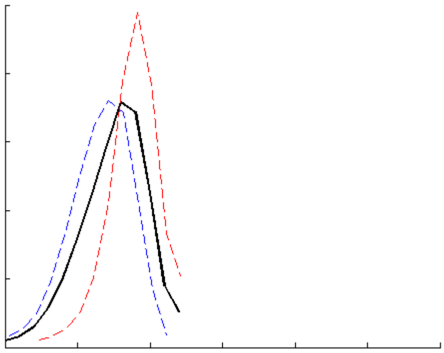


Figure B.6: $s = 0.75, \alpha = 0.8$



$s = 0.5, \alpha = 0.8$

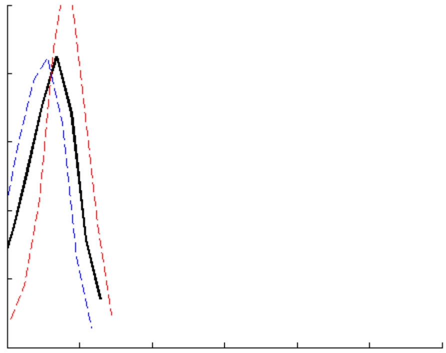


Figure B.8: $s = 0.25, \alpha = 0.8$

C Additional Tables of Estimates.

	β_1	$SE(\beta_1)$	β_0	$SE(\beta_0)$	γ	$SE(\gamma)$	δ	$SE(\delta)$
Graduates	0.000	0.000	0.000	0.000	6.928***	1.087	-6.928	0.298
Constant	28.418***	1.046	30.089***	0.291	-8.439***	0.868	6.777***	0.241
Age	0.339***	0.042	0.093***	0.029	0.246***	0.014	0.000	0.000
Age^2	-0.004***	0.001	-0.001	0.000	-0.003***	0.000	0.000	0.000
Black	-0.333***	0.135	-0.319***	0.102	-0.568***	0.042	0.557***	0.042
Hispanic	-0.523	0.392	-1.312***	0.215	-1.737***	0.015	2.529***	0.041
Asian	-0.262***	0.079	-1.031***	0.076	0.418***	0.016	0.350***	0.040
Inv. Mills Ratio	-0.467	0.481	1.083***	0.372	-	-	-	-

$\sigma_0^2 = 5.0550,\ \sigma_1^2 = 4.0183,\ \rho_{0\eta} = 0.4799,\ \rho_{1\eta} = 0.2337$

$N = 108,751,\ N_1 = 44,692,\ \alpha = 1.2,\ s = 0.50.$

Statistical significance is assessed with two-tailed tests: *** implies significance at 1%$*$, ** at 5% and * at 10%.

Table C.1: Married women -benchmark case ($\alpha = 1.2, s = 0.50$)

	β_1	$SE(\beta_1)$	β_0	$SE(\beta_0)$	γ	$SE(\gamma)$	δ	$SE(\delta)$
Graduates	0.000	0.000	0.000	0.000	8.722***	1.325	-8.722***	0.366
Constant	32.914***	1.214	35.248***	0.336	-10.449***	1.057	8.178***	0.291
Age	0.411***	0.049	0.109***	0.033	0.300***	0.017	0.000	0.000
Age^2	-0.005***	0.001	-0.001	0.000	-0.004***	0.000	0.000	0.000
Black	-0.461***	0.156	-0.423***	0.118	-0.691***	0.051	0.658***	0.051
Hispanic	-0.718	0.454	-1.506***	0.249	-2.099***	0.018	2.908***	0.042
Asian	-0.296***	0.092	-1.223***	0.088	0.506***	0.020	0.416***	0.049
Inv. Mills Ratio	-0.440	0.558	1.285***	0.431	-	-	-	-

$\sigma_0^2 = 6.7631,\ \sigma_1^2 = 5.4126,\ \rho_{0\eta} = 0.4937,\ \rho_{1\eta} = 0.1999$

$N = 108,751,\ N_1 = 44,692,\ \alpha = 1.2,\ s = 0.75.$

Statistical significance is assessed with two-tailed tests: *** implies significance at 1%$*$, ** at 5% and * at 10%.

Table C.2: Married women ($\alpha = 1.2, s = 0.75$)

	β_1	$SE(\beta_1)$	β_0	$SE(\beta_0)$	γ	$SE(\gamma)$	δ	$SE(\delta)$
Graduates	0.000	0.000	0.000	0.000	5.736***	0.894	-5.736***	0.245
Constant	23.440***	0.855	24.841***	0.238	-6.993***	0.713	5.576***	0.198
Age	0.278***	0.035	0.075***	0.023	0.204***	0.011	0.000	0.000
Age^2	-0.003***	0.000	-0.000	0.000	-0.003***	0.000	0.000	0.000
Black	-0.287***	0.110	-0.272***	0.083	-0.470***	0.034	0.455***	0.034
Hispanic	-0.444	0.320	-1.066***	0.176	-1.433***	0.01	2.054***	0.032
Asian	-0.214***	0.065	-0.850***	0.062	0.344***	0.013	0.291***	0.033
Inv. Mills Ratio	-0.369	0.393	0.897***	0.304	-	-	-	-

$\sigma_0^2 = 3.3733,\ \sigma_1^2 = 2.6767,\ \rho_{0\eta} = 0.4866,\ \rho_{1\eta} = 0.1569$

$N = 108,751,\ N_1 = 44,692,\ \alpha = 0.8,\ s = 0.50.$

Statistical significance is assessed with two-tailed tests: *** implies significance at 1%∗, ** at 5% and * at 10%.

Table C.3: Married women ($\alpha = 0.8, s = 0.50$)

	β_1	$SE(\beta_1)$	β_0	$SE(\beta_0)$	γ	$SE(\gamma)$	δ	$SE(\delta)$
Graduates	0.000	0.000	0.000	0.000	7.125***	1.075	-7.125***	0.298
Constant	27.198***	0.993	29.073***	0.274	-8.558***	0.858	6.693***	0.238
Age	0.335***	0.040	0.089***	0.027	0.246***	0.014	0.000	0.000
Age^2	-0.004***	0.001	-0.001	0.000	-0.003***	0.000	0.000	0.000
Black	-0.375***	0.128	-0.348***	0.096	-0.566***	0.041	0.538***	0.041
Hispanic	-0.579	0.372	-1.233***	0.204	-1.717***	0.015	2.374***	0.035
Asian	-0.244***	0.075	-1.001***	0.072	0.413***	0.016	0.342***	0.040
Inv. Mills Ratio	-0.372	0.456	1.051***	0.352	-	-	-	-

$\sigma_0^2 = 4.5262,\ \sigma_1^2 = 3.6224,\ \rho_{0\eta} = 0.4947,\ \rho_{1\eta} = 0.1985$

$N = 108,751,\ N_1 = 44,692,\ \alpha = 0.8,\ s = 0.75.$

Statistical significance is assessed with two-tailed tests: *** implies significance at 1%∗, ** at 5% and * at 10%.

Table C.4: Married women ($\alpha = 0.8, s = 0.75$)

D Additional Marginal Treatment Distributions.

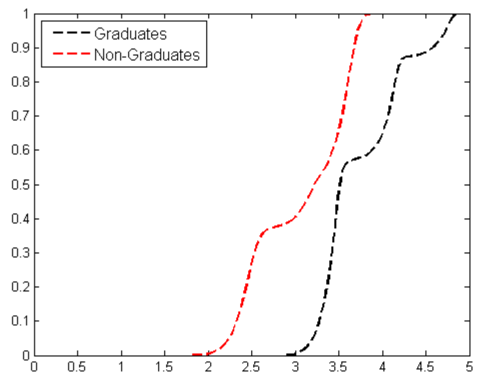


Figure D.1: Married women -benchmark case ($\alpha = 1.2, s = 0.50$)

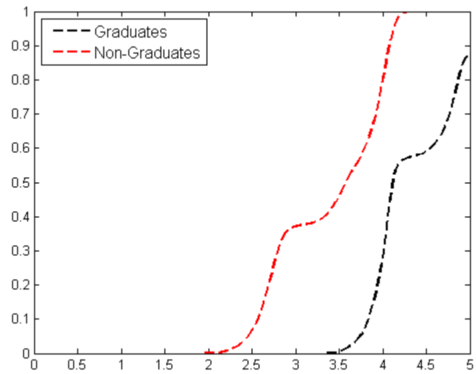


Figure D.2: Married women ($\alpha = 1.2, s = 0.75$)

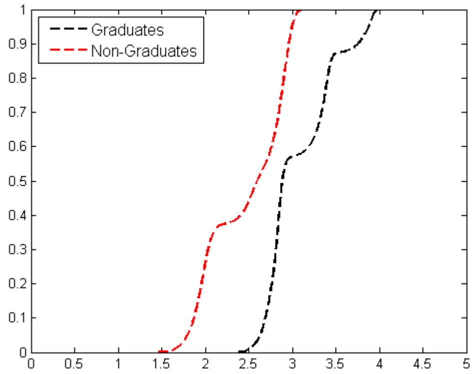


Figure D.3: Married women ($\alpha = 0.8, s = 0.50$)

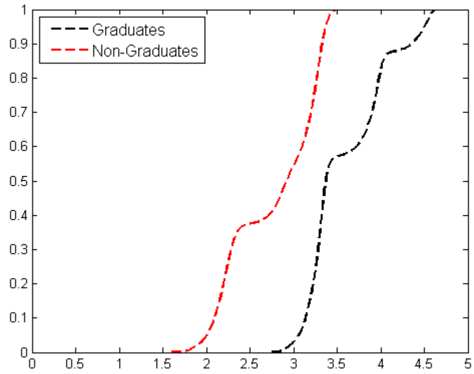


Figure D.4: Married women ($\alpha = 0.8, s = 0.75$)