DISCUSSION PAPER SERIES

IZA DP No. 6533

Becker Meets Ricardo: Multisector Matching with Social and Cognitive Skills

Robert J. McCann
Xianwen Shi
Aloysius Siow
Ronald Wolthoff

April 2012

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Becker Meets Ricardo: Multisector Matching with Social and Cognitive Skills

Robert J. McCann
University of Toronto

Xianwen Shi
University of Toronto

Aloysius Siow
University of Toronto
and IZA

Ronald Wolthoff
University of Toronto
and IZA

Discussion Paper No. 6533
April 2012

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 6533
April 2012

ABSTRACT

Becker Meets Ricardo: Multisector Matching with Social and Cognitive Skills*

This paper presents a tractable framework for studying frictionless matching in school, work, and marriage when individuals have heterogeneous social and cognitive skills. In the model, there are gains to specialization and team production, but specialization requires communication and coordination between team members, and individuals with more social skills communicate and coordinate at lower resource cost. The theory delivers full task specialization in the labor and education markets, but incomplete specialization in marriage. It also captures well-known matching patterns in each of these sectors, including the commonly observed many-to-one matches in firms and schools. Equilibrium is equivalent to the solution of a utilitarian social planner solving a linear programming problem.

JEL Classification: E24, J12, J24, J31

Keywords: social skill, cognitive skill, matching, sorting, education, labor, marriage, social welfare, linear programming

Corresponding author:

Ronald Wolthoff
University of Toronto
Department of Economics
150 St. George Street
Toronto, ON M5S 3G7
Canada
E-mail: ronald.wolthoff@utoronto.ca

* McCann thanks the University of Nice Sophia-Antipolis and the University of Chicago's Becker-Friedman Institute for Economic Research and Stevanovich Center for Financial Mathematics for their hospitality and support during various stages of this work. Shi and Siow thank SSHRC for financial support. We also thank Jan Eeckhout, Jim Heckman, Bernard Salanie, Michael Waldman and various seminar and conference participants for helpful comments.

1 Introduction

Over their lives, most individuals will participate in schooling, work and marriage. Activities in these three sectors occur within teams. Students match with teachers, workers match with managers, and husbands match with wives. Within each sector, there is substantial variation in outcomes. For each individual, outcomes are correlated across sectors/markets.

A large empirical psychology literature and a small one in economics have shown that social skills affect how individuals perform in these sectors. For example, labor market participants cite social skills as an important factor in hiring decisions and labor market success (Posner, 1981), communication between spouses is an important determinant of marital success (Cleek and Pearson, 1985), and colleges screen applicants for leadership skills (Bruggink and Gambhir, 1996).[1]

Why are social skills valued? In order to answer this question, we need a theory of social skills that describes social interactions among individuals with heterogeneous social skills. Moreover, from a positive perspective, the theory of social skills should differentiate the observable effects of social skills from that of cognitive skills. Unlike the theory of human capital, the theory of social skills is less developed.

This paper develops a theory of social and cognitive skills in team production and frictionless matching (hereafter SC model). Our theory of cognitive skills follows Becker (1973, 1974) where the cognitive skills of teammates are complements in the production of team output in each sector. As is well known, this complementarity will result in positive assortative matching (PAM) by cognitive skills in equilibrium.

Our theory of social skills builds on two classic ideas in team production: specialization and task assignment. Adam Smith argued that when workers specialize in different tasks, they can produce more output per worker than if each worker does every task. Smith did not discuss which worker should work on which task. Subsequently, Ricardo argued that the principle of comparative advantage should determine optimal task assignment.

Smith also noted that there are limits to specialization. In order to realize the potential

[1] Generally, psychologists regard social skills as multidimensional. Several of the factors in their standard model of personality traits, the Big Five model, affect the social skill factor as used in this paper. Some researchers subscribe to the concept of "emotional intelligence" which is close to the one social skill factor discussed in this paper. Almlund, et. al. (2011) is a summary of the empirical psychology literature written for economists.

gains to specialization, individuals who specialize and engage in team production have to incur communication and collaboration costs. Trading off the gains from specialization against communication and collaborative costs, individuals form teams and choose tasks in each of the sectors. As market participants recognize and the large empirical literature shows, some individuals are better at communication and collaboration, and team members prefer to work with them.[2] Hence, we analyze a model in which individuals are heterogenous not only in cognitive skills but also in social skills.

Consider a production process in which two tasks, I and C, have to be completed to produce output.[3] Without specialization, one individual does both tasks sequentially. That is, the individual finishes one task and then does the other. With specialization, one individual may do task I while the other will do task C. So both individuals will become more efficient in doing their specialized task. In addition, both tasks can be done simultaneously which may also increase output significantly. For example, a dentist can work along side several dental hygienists tending to several patients simultaneously. A dentist working alone will have to do dentistry and teeth cleaning one patient at a time.[4] There is often an efficiency gain from simultaneous specialization in team production even if every individual has the same productivity in a specific task.

Under specialization, the teammates have to coordinate their actions. These communication activities take time away from production. Following Garicano (2000), we assume that the communication cost is one-sided and is only borne by the team member in task C. Furthermore, we assume that an individual with more social skill, when assigned to task C, uses less time to communicate with his or her teammate, and therefore has more time for production in task C.

With every sector having qualitatively the same production function, we study how the education, labor and marriage markets operate. In the labor market, we obtain full specialization in tasks, that is, an individual with high social skill specializes in task C while an individual with low social skill specializes in task I. Within a team in the labor market, only

[2] See, for example, Goleman (1997), Kuhn and Weinberger (2005), Rivera (2011), and the references therein.

[3] Here I stands for "individualistic", while C stands for "collaborative".

[4] As another example, two single parents, each acting alone, cannot pick up their own children and send them to after school activities, and simultaneously cook dinner. On the other hand, two parents, with two children, can divide up the tasks and do them simultaneously.

one individual will do task C and all the remaining teammates do task I. We refer to the former individual as the manager, and refer to the latter ones as workers. Therefore, our model generates many-to-one teams, a commonly observed organizational form in the labor market. Moreover, managers who have more social skills and thus lower communication and coordination cost will manage larger teams (a larger span of control), while the social skills of workers are irrelevant. Hence, cognitive and social skills are not isomorphic.[5]

In contrast, full specialization is not optimal in a monogamous marriage market where team size is fixed at two, that is, the division of labor in marriage is limited by monogamy. Each spouse will do both task I and C. Hence, individuals will care about both social and cognitive skills of their spouses, leading to PAM along both skills.

When young, individuals augment their cognitive skills in the education market. Social skill is assumed to be constant throughout life. Each school in the education market consists of one teacher (who does task C) and several students (who do task I). A teacher with higher social skill can manage a larger class, and thus can command a higher wage. A student who has strong social skill will expect to become a manager or a teacher later in life. Since there is cognitive skill complementarity between a manager and their workers or between a teacher and their students, such a student will invest more in cognitive skill compared with another student who will become a worker. Moreover, a future manager/teacher with more social skill will have a larger span of control, so the return to cognitive skill investment is even higher.[6] Therefore, students with high social skills will compete with students who have higher cognitive ability but lower social skill to match with teachers with high cognitive skill.[7] The model generates an endogenous positive correlation between social and cognitive skills of adults.

We simulate the equilibrium of the model with a bivariate uniform distribution of initial cognitive abilities and social skills. The simulation generates an income distribution which qualitatively consistent with a log normal distribution. In addition, as we argue above, there

[5] In one-factor (cognitive) models of achievement, Larry Summers would have been the president of the United States and George Bush Jr. would have been a nobody. Most observers underestimated George Bush Jr.'s significant social skills (Kristof, 2000).

[6] Smeets and Warzynski (2008) finds a positive relationship between managers' wages and the number of individuals they supervise.

[7] Unlike popular belief, our model suggests that George Bush Jr. went to Yale to augment his cognitive skill rather than to build his social network.

is an educational gap between workers and managers/teachers.

Finally, the paper also makes a methodological contribution. We show that the equilibrium of our multi-factor multi-market matching model with endogenous occupational choice is equivalent to the solution of a utilitarian social planner solving a linear programming problem, substantially extending the existing results. It also facilitates the proof for the existence of equilibrium and makes it easy to numerically compute the equilibrium.

Most components of our SC model have been studied previously in isolation. We build on and integrate their insights. So it will be convenient to defer our review of the literature until the end of the paper. For now, we acknowledge our intellectual debt to Garicano (2000), Garicano and Rossi-Hansberg (2004, 2006), whose work on communication costs, task assignments, organizational design and equilibrium one factor (cognitive skill) matching in the labor market, inspired our work. One way to view our contribution is that we extend their concerns to allow individuals to have heterogenous communication costs. We study the implications of this additional dimension of heterogeneity for schooling, labor and marriage markets.

A caveat is in order. In this paper, the role of social skill is kept simple so that we can include multiple sectors in our analysis. Thus this model should be viewed as a first pass theory of frictionless matching with social and cognitive skills.

This paper proceeds as follows. Section 2 introduces the model, discusses the micro-foundation of the production technology, and describes the equilibrium in each market. Section 3 formally discusses the linear programming problem and proves existence of the equilibrium. In section 4, we illustrate the properties of the equilibrium by simulation. Section 5 discusses related literature and section 6 concludes.

2 The Model

Consider a society where risk-neutral individuals live for two periods. First, individuals enter the education market as students. Subsequently, as adults, they work in the labor market and participate in the marriage market. All three markets are perfectly competitive and individuals behave as price-takers. We will also assume a discount rate of zero and a perfect capital market.

Each individual in the society is born with two skills, a gross social skill η, with $\eta \in \left[\underline{\eta}, \overline{\eta}\right]$,

$0 < \underline{\eta} < \overline{\eta} \leq 1$, and initial cognitive ability a, with $a \in [\underline{a}, \overline{a}]$, $0 < \underline{a} < \overline{a}$. The gross social skill η is fixed for the entire life of the individual and cannot be changed.[8] Adult cognitive skill is augmentable by attending school in the first period. Individuals are heterogeneous in their skills. So the bivariate distribution of a and η is not degenerate.

Education takes place in schools consisting of one teacher and a certain number of students. A school charges each of its students a tuition fee τ which it uses to hire the teacher. Schools can enter the education market freely, so they make zero profit in equilibrium, i.e. the tuition fees exactly cover the teachers' wage ω. The initial cognitive ability a of a student and his school choice determine the final cognitive skill of this student upon graduation, which we denote by k, with $k \in \left[\underline{k}, \overline{k}\right]$, $0 < \underline{k} < \overline{k}$.

After graduating from school, students become adults in the second period and enter the labor market. Each adult can become either a teacher in one of the schools or an employee in a firm. Each firm employs a certain number of individuals to produce output which is sold at a normalized price of one. Firms can freely enter the labor market. Therefore, each firm makes zero profit in equilibrium, and its output exactly covers the wages ω of its employees.

Simultaneously, each adult enters the marriage market and look for a spouse. Schools and firms can potentially be large. The marriage market is monogamous and each marriage consists of one husband and one wife. For pedagogical convenience, we assume that the sex ratio, the ratio of men to women by type, in this society is one for each type.[9] In any marriage, the two spouses produce marital output which is divided between them. Let h be the marital payoff for an adult.

Each adult consumes his or her own labor earnings and there is no complementarity in consumption from labor earnings and marital consumption. So individuals' lifetime payoffs are equal to the sum of their labor market earnings and marital payoffs minus tuition costs, $\omega + h - \tau$.[10] Individuals choose who to match with in the three sectors to maximize their net payoffs.

The main innovation of this paper is to introduce a team production function and apply it to all three sectors – education, labor and marriage. The next subsection will describe how this technology works.

[8] Social and other non-cognitive skills can be acquired when young (E.g. Heckman, et. al. (2011)). Our framework can accommodate this acquisition but we ignore it here to keep the paper length manageable.

[9] Our model can deal with an uneven sex ratio, but this is beyond the scope of the current paper.

[10] The payoffs ω, h, and τ are equilibrium objects. We discuss their determinants in more detail below.

2.1 Team Production Technology

In our model, output is produced by completion of two tasks, I and C. Both tasks need to be completed successfully for production to take place. Each individual is endowed with one unit of time in each market – school, work and marriage – and is not allowed to substitute time across markets.

Suppose individual i of type (η_i, k_i) spends $\theta_i^I \in [0,1]$ unit of time on task I and $\theta_i^C \in [0,1]$ unit of time on task C, then output is given by the following Leontief production function:

$$Y\left(\theta_i^I, \theta_i^C; \eta_i, k_i\right) = \beta k_i \min\left\{\theta_i^I, \gamma\theta_i^C\right\} \tag{1}$$

subject to the time constraint $\theta_i^I + \theta_i^C \leq 1$. We assume $\gamma > 1$ so that task C takes less time to do than task I, and assume $\beta \in (0,1)$ to reflect the loss of insufficient specialization and/or lack of team production. Therefore, if individual i of type (η_i, k_i) works alone – performs both task I and C herself – she will choose $\theta_i^I = \gamma/(1+\gamma)$ and obtain output $\beta k_i \gamma/(1+\gamma)$. Since she is working alone, she does each task sequentially which is implicit in her time budget constraint, and there is no need for coordination. As a result, individual i's gross social skill η_i does not enter the production function.

Now consider two individuals, i of type (η_i, k_i) and j of type $\left(\eta_j, k_j\right)$, forming a team to produce output. Within a team, one individual can specialize in task I while the other can specialize in task C or they can each do both tasks. As discussed in the introduction, there are two types of potential gains from specialization. First, a worker who specializes in one task becomes more efficient in doing it. Second, both tasks can be done simultaneously.

While there are potential gains from specialization and team production, specialized team members need to coordinate their activities to efficiently produce output. Following Garicano (2000), we assume that the communication and coordination cost is one-sided: the cost to coordinate the activities of two team members will be borne by one of them. We assume that the person assigned to task C bears the coordination cost. Coordination is costly because it takes time away from producing C. How much time is taken away depends on the gross social skill of the worker assigned to task C. Suppose person j of gross social skill η_j spends a total of θ_j^C units of time on task C (producing and coordinating), then a fraction $(1-\eta_j)$ of θ_j^C is spent on communication and coordination, and the remaining fraction η_j of θ_j^C is used to produce C. Therefore, a more socially skilled worker in task C, who has a higher η_j, can spend less time in coordinating with his teammates, and therefore produces more of

task C.

Our one-sided communication cost assumption implies that the social skill of the teammate in task I does not affect team output. We motive this behavioral assumption by the following thought experiment. Consider an existing team. Replace one teammate with another individual with more social skill than the other members of the team. This new teammate will be able to better communicate and coordinate activities of the team to increase output and so will be assigned to task C. Now consider replacing a second teammate with another individual again with significantly higher social skill. Since the first new teammate is already coordinating work, will this second new team mate clash with the first? Put another way, does a team need more than one leader? If one team leader is sufficient in many team environments, then the second new teammate will not significantly increase output. So formally we assume that the social skill of individuals in task I do not affect team output.

We now describe the production function with team production. Let i spend $\theta_i^I \leq 1$ units of time on task I and j spend $\theta_j^C \leq 1$ units of time on task C. Total team output is given by

$$Y(\theta_i^I, \eta_i, k_i, \theta_j^C, \eta_j, k_j) = \sqrt{k_i k_j} \min\left\{\theta_i^I, \gamma\eta_j\theta_j^C\right\}. \tag{2}$$

There are three main differences between the production technologies for working alone, (1) and working in a team, (2). First, the gross social skill η_j of j working in task C affects output. Since $\eta_j \leq 1$, the need for coordination lowers output. Second, we drop $\beta < 1$ in the team production function to capture the gains to specialization and joint production relative to working alone. As long as β is sufficiently small, output under joint production will exceed that of the two persons each working alone. Finally, following Becker, we assume that cognitive skills of co-workers in team production are complementary, which is captured by $\sqrt{k_i k_j}$, leading to PAM with respect to cognitive skills in co-workers.[11]

The production technology will generate teams of many-to-one matching in a competitive labor market. First, the technology generates specialization in production where one employee will be assigned exclusively to task I and the other to task C. As long as $\gamma\eta_j \geq 1$, the employee assigned to task C will be left with unused time when the other employee's time is completely exhausted. The employee in task C can work with other employees in task I. Thus, an optimal team will have one employee in task C and many in task I as with

[11] If $\overline{k} \leq 1$, one can, as in Kremer (1993), interpret $\sqrt{k_i}$ as the probability for individual i to succeed in performing task I while $\sqrt{k_j}$ the probability for individual j to succeed in performing task C.

one manager for several workers. Also note that when $\gamma\eta_j \geq 1$ and β is sufficiently small, specialization and joint production is more productive than working alone.

Since the coordination cost is assumed to be one-sided, individual i in task I does not incur any coordination cost. This is reflected in (2) that the gross social skill of i, η_i, who is assigned to task I, does not enter the production function. This restriction greatly improves tractability and has behavioral implications which we will discuss later.

In contrast to the labor market, the number of spouses in a marriage in a monogamous marriage market is fixed at two. Since both spouses have one unit of time in marriage, they have to spend all their time with each other in producing marital output. As we will show later, full task specialization in marriage is no longer optimal. Each spouse will do both task I and task C.

We now turn to the education market. In this market, let individual i be a student and individual j be a teacher. The inputs in the production function are the initial cognitive ability a_i of student i and the adult cognitive skill k_j of the teacher j, while output is the adult cognitive skill k_i of individual i. That is, the production function (3) can be adapted to the education market as follows

$$k_i = \sqrt{a_i k_j} \min\left\{\theta_i^I, \gamma\eta_j\theta_j^C\right\}.$$

In what follows, we will analyze the implications of the above team production technology for the education, labor and marriage markets. For expositional convenience, we define social skill $n_j \equiv \gamma\eta_j$, $n_j \in [\underline{n}, \overline{n}]$. In terms of social skill n, the team production function becomes

$$Y(\theta_i^I, n_i, k_i, \theta_j^C, n_j, k_j) = \sqrt{k_i k_j} \min\left\{\theta_i^I, n_j\theta_j^C\right\} \tag{3}$$

Comparing (3) with (1), the tradeoff between team production and single agent production is clear. Since $n_j < \gamma$, team production is costly due to communication and coordination costs. This cost is reduced when the social skill of individual j, n_j, increases. On the other hand, $\beta < 1$ in (1) reflects the cost of producing alone. To ease exposition, we assume throughout that

$$\underline{n} \geq 2 \text{ and } \beta \leq 1/2 \tag{4}$$

so that team production is always superior to working alone.[12]

There are several different ways to motivate the team production function (3). In Appendix A, we briefly sketch two alternative micro-foundations.

[12]To see why condition (4) is sufficient, note that a type-(n,k) individual produces output $\beta k\gamma/(1+\gamma)$

2.2 Labor Market

It is convenient for us to first analyze the labor market. An adult in the labor market is characterized by his social and cognitive skills, (n, k). Here we take the distribution $\alpha(n, k)$ of adult skills as given, but we will endogenize it when we study the education market.

Let $\omega(n, k)$ denote the equilibrium wage for a type-(n, k) adult. Each adult can either work for a firm as an employee or work for a school as a teacher. If a firm hires the adult for $\theta \in [0, 1]$ units of time, then the firm will pay the adult $\theta\omega(n, k)$. The adult will supply $(1 - \theta)$ units of time to other firms and earn $(1 - \theta)\omega(n, k)$ from them. If a school hires a type (n, k) adult as a teacher, it also has to pay $\omega(n, k)$ per unit of time.

There is a perfectly elastic supply of firms and schools in the labor market. So in equilibrium, all firms and schools make zero profit, and the wage function $\omega(n, k)$ must satisfy the demand of firms and schools for adults. For now, we focus only on the problem that firms solve.

A firm is a collection of teams. Given our production function (3), there is no interaction across teams, so we can study the problem of one team in a firm. A team is a collection of employees chosen by the firm who assigns tasks to them and pays them market wages. We first establish that our production function (3) implies full specialization in task assignment in the labor market.

Proposition 1 *Each team's profit is maximized by allocating every employee to a specific task, either I or C, for the entire length of the production process.*

Proof. Consider a team which employs a type-(n, k) adult for a short time interval Δ. The firm can allocate the employee to either task I or task C.

If employee (n, k) is allocated to task I during the time interval Δ, then the firm has to hire another adult (n', k') from the labor market to perform task C for $\Delta(n')^{-1}$ units of time in order to produce output $\sqrt{kk'}\Delta$. Choosing (n', k') optimally, the firm's profits of having (n, k) in task I for Δ time interval is given by

$$\pi^I(n, k, \Delta) = \max_{(n',k')} \sqrt{kk'}\Delta - \omega(n', k')(n')^{-1}\Delta - \omega(n, k)\Delta. \tag{5}$$

by working alone. Now suppose he forms a three-person team with two other type-(n, k) individuals: two of them do I, and the third does C. If $n \geq 2$, this team produces at least $2k$ according to our team production function (3). Therefore, team production is better than working alone as long as (4) holds.

If employee (n,k) is instead allocated to task C in period Δ, then the firm needs to hire another adult (n'',k'') to do task I for $n\Delta$ units of time to produce output $\sqrt{kk''}n\Delta$. The associated profits $\pi^C(n,k,\Delta)$ are

$$\pi^C(n,k,\Delta) = \max_{(k'',n'')} \sqrt{kk''}n\Delta - \omega(n'',k'')\,n\Delta - \omega(n,k)\,\Delta. \tag{6}$$

Therefore, the firm would assign employee (n,k) to task I if and only if $\pi^I(n,k,\Delta) - \pi^C(n,k,\Delta) \geq 0$. The sign of $\pi^I(n,k) - \pi^C(n,k)$ is independent of Δ, the length of time that is available for production. Consequently, the firm's profits are maximized by allocating employee (n,k) to either I or C for the entire duration of the production process. ■

Proposition 1 indicates that within a firm each employee will specialize in performing either task I or task C. Moreover, since all firms have access to the same production technology, Proposition 1 also implies that task assignments must be the same across firms: if one firm strongly (weakly) prefers to assign an employee of type (n,k) to task I (C), all other firms will do the same. As a result, it is without loss of generality to assume that each team has only one team member performing task C.[13]

This proposition also implies that task assignment within a team is determined by comparative advantage.

Corollary 1 *Consider any two members with types (n_i,k_i) and (n_j,k_j) in a team. Member i will be assigned to task I if*

$$\omega(n_i,k_i)(1-n_i^{-1}) < \omega(n_j,k_j)(1-n_j^{-1})$$

The above corollary is obtained by comparing the profits from assigning member i to task I and j to task C to produce a fixed amount of output, versus the reverse assignment. The insight of the above corollary is known since Ricardo.

Finally, for fixed cognitive skill k, the task assignment is sorted according to individuals' social skills.

Proposition 2 *For each cognitive skill level k, there exists a cutoff value $\widehat{n}(k) \in [\underline{n},\overline{n}]$ such that individuals with social skill $n < \widehat{n}(k)$ perform task I, and individuals with social skill $n \geq \widehat{n}(k)$ perform task C.*

[13] If there are $q > 1$ team members performing task C in a particular team, the firm hiring these team members can split and re-organize the team such that each (new) team has only one member performing task C, without lowering profits.

Proof. Applying the envelope theorem to (5) and (6) yields

$$\frac{d}{dn}\left(\pi^I(n,k,\Delta)-\pi^C(n,k,\Delta)\right)=-\left[\sqrt{kk''}+\omega(n'',k'')\right]\Delta<0.$$

Therefore, the value of $\pi^I(n,k,\Delta)-\pi^C(n,k,\Delta)$ crosses zero only once and from above. ■

Let the employee (n,k) be optimally assigned to task I, that is, $\pi^I(n,k,\Delta)\geq\pi^C(n,k,\Delta)$. We will call these employees workers and denote their occupation by w. Note that the amount of team output produced in Δ time interval is $\sqrt{kk'}\Delta$ which is independent of n. Put another way, the firm does not value a worker's social skill and thus will not be willing to pay for it. Instead, it only pays attention to the worker's cognitive skill k. Therefore, the equilibrium wage of workers of skill (n,k), which we denote by $\omega_w(n,k)$, is independent of n. To simplify notation, we will write $\omega(k)\equiv\omega_w(n,k)$.

On the other hand, if the employee (n,k) is assigned to task C, the profit $\pi^C(n,k,\Delta)$ from hiring (n,k) depends on n. We call these employees managers and denote their occupation by m. Their wages will depend on both n and k and are denoted by $\omega_m(n,k)$.

Consider a team with a manager of type (n,k) where $n\geq 2$. According to Proposition 1, this manager is only matched with other employees who perform task I, i.e., workers. Hence, all teams in the labor market consist of many-to-one matchings. The number of workers that a manager supervises, which can be interpreted as the span of control or the capacity of the manager, is exactly equal to the manager's social skill.

Let the team with a type-(n,k) manager choose n workers with respective types $(k_1,...,k_n)$ in order to maximizes its profits. The team solves the following maximization problem:[14]

$$\max_{(k_1,...,k_n)}\sum_{i=1}^{n}\left[\sqrt{kk_i}-\omega(k_i)\right]-\omega_m(n,k).$$

Given the additive separability of the total output, the optimal choice of workers satisfies $k_1^*=...=k_n^*=\mu(k)$ with

$$\mu(k)\in\arg\max_{k'}\sqrt{kk'}-\omega(k'). \tag{7}$$

Therefore, we have proved the following result.

Lemma 1 *In equilibrium, it is optimal for a team to hire workers with the same cognitive skill.*

[14] For expositional purposes, we treat n as an integer here. More generally, we can write the maximization problem as follows: $\max_{\{k_i\}_{i\in[0,n]}}\int_0^n\left[\sqrt{k\cdot k_i}-\omega(k_i)\right]di-\omega_m(n,k)$.

The function $\mu(k)$ determines the worker type matched to a type-(n,k) manager. It depends on the manager's cognitive skill k, but not on his or her social skill n, and fully captures the sorting between workers and managers in the labor market. Hence, we call $\mu(k)$ the equilibrium matching function in labor market. Given Lemma 1, we can rewrite the profits of the team with manager (n,k) as follows:

$$n\left[\sqrt{k\mu(k)}-\omega(\mu(k))\right]-\omega_m(n,k).$$

The free-entry condition for firms implies that the above expression must be zero. Therefore, the manager's wage is given by

$$\omega_m(n,k)=n\left[\sqrt{k\mu(k)}-\omega(\mu(k))\right].$$

Define $\phi(k)$ as

$$\phi(k)\equiv\sqrt{k\mu(k)}-\omega(\mu(k))=\max_{k'}\sqrt{kk'}-\omega(k'). \tag{8}$$

We can interpret $\phi(k)$ as the profits per worker generated by a type-k manager. The equilibrium wage for the manager (n,k) can be rewritten as

$$\omega_m(n,k)=n\phi(k). \tag{9}$$

Next we address the issue of sorting in labor market. Both workers and managers are heterogeneous in their cognitive skill, so an important question is which worker types work for which manager. Applying the envelope theorem to equation (7), we obtain:

Lemma 2 *The equilibrium matching function $\mu(k)$ is strictly increasing.*

Given that $\mu(k)$ is weakly increasing, we can define the generalized inverse function $\mu^{-1}(\cdot)$ of $\mu(\cdot)$ as

$$\mu^{-1}(k)=\min\{k':\mu(k')=k\}.$$

That is, $\mu^{-1}(k)$ is the lowest cognitive skill among managers hiring type-k workers. Now we can link the equilibrium wage $\omega(k)$ and $\phi(k)$ with the equilibrium matching function $\mu(k)$.

Lemma 3 *Given an equilibrium matching function $\mu(k)$, wages $\phi(k)$ and $\omega(k)$ are given by*

$$\begin{aligned}\phi(k) &= \phi\left(\overline{k}\right)-\frac{1}{2}\int_k^{\overline{k}}\sqrt{\frac{\mu(x)}{x}}dx,\\ \omega(k) &= \omega\left(\overline{k}\right)-\frac{1}{2}\int_k^{\overline{k}}\sqrt{\frac{\mu^{-1}(x)}{x}}dx.\end{aligned}$$

Proof. We can apply the envelope theorem to (8) and obtain that

$$\frac{d\phi(k)}{dk} = \frac{1}{2}\sqrt{\frac{\mu(k)}{k}}. \tag{10}$$

Furthermore, the necessary first-order condition of the maximization problem (8) is

$$\frac{d\omega(k')}{dk'}|_{k'=\mu(k)} = \frac{1}{2}\sqrt{\frac{k}{k'}},$$

which can be rewritten as

$$\frac{d\omega(k)}{dk} = \frac{1}{2}\sqrt{\frac{\mu^{-1}(k)}{k}}. \tag{11}$$

The claims then follow immediately. ■

To conclude the analysis of labor market, we should also characterize the equilibrium occupation choice. However, since schools also compete in the labor market for teachers, we will carry out the analysis of occupation choice when we investigate the education market.

2.3 Marriage Market

A crucial feature that distinguishes the marriage market from the labor market and the education market is monogamy, i.e. all marital matches are bilateral. We formalize monogamy as each spouse in a marriage devoting all their time in the marriage market with each other. We show that the full specialization solution we found earlier for the labor market cannot be optimal for the marriage market.

Consider a marriage between two individuals (n_i, k_i) and (n_j, k_j). Suppose individual i is fully specialized and spends his entire time unit on task I. Individual j spends $\theta_j^C = 1/n_j$ units of time on task C to create $\sqrt{k_i k_j}$ of output, and spends the remaining time $(n_j - 1)/n_j$ working alone. The total output in this household is given by

$$\sqrt{k_i k_j} + \left(1 - \frac{1}{n_j}\right)\beta k_j \frac{\gamma}{1+\gamma}. \tag{12}$$

Consider the alternative arrangement where both individuals spend time on both tasks. The marital output for such a marriage can be written as

$$Y^M(\theta_i^I, n_i, k_i, \theta_j^I, n_j, k_j) = \sqrt{k_i k_j}(\min(\theta_i^I, n_j(1-\theta_j^I)) + \min(\theta_j^I, n_i(1-\theta_i^I))).$$

A household will choose the time allocation $\left(\theta_i^I, \theta_j^I\right)$ to maximize the marital output. It is easy to verify that the optimal solution is given by

$$\theta_i^I = \frac{n_i n_j - n_j}{n_i n_j - 1}, \text{ and } \theta_j^I = \frac{n_i n_j - n_i}{n_i n_j - 1}.$$

Therefore, optimal marital output equals

$$Y^M(n_i, k_i, n_j, k_j) = \sqrt{k_i k_j}\frac{2n_i n_j - n_j - n_i}{n_i n_j - 1}. \tag{13}$$

We now argue that full specialization is not optimal in marriage. Suppose, by contradiction, that full specialization by one spouse is optimal, so that the household output is given by (12). Take two such households and re-arrange the marriage so that two type-(n_i, k_i) individuals marry to each other, and two type-(n_j, k_j) individuals marry to each other. Moreover, within each new marriage, each spouse spend time on both tasks. Then according to (13), the total output for the two new marriages is

$$\frac{2n_i}{n_i + 1}k_i + \frac{2n_j}{n_j + 1}k_j.$$

Therefore, full specialization is not optimal in marriage if

$$\frac{2n_i}{n_i + 1}k_i + \frac{2n_j}{n_j + 1}k_j \geq 2\sqrt{k_i k_j} + 2\left(1 - \frac{1}{n_j}\right)\beta k_j \frac{\gamma}{1+\gamma}.$$

It is straightforward to check that the above inequality holds under condition (4), i.e., $\underline{n} \geq 2$ and $\beta \leq 1/2$.

Let $h(n_i, k_i)$ denote the equilibrium payoff that an individual of type (n_i, k_i) obtains from marriage. If this individual has a spouse of type (n_j, k_j), then his payoff must satisfy $h(n_i, k_i) = Y^M(n_i, k_i, n_j, k_j) - h(n_j, k_j)$. Recall that the sex ratio, ratio of men to women, by type is assumed to be unity. The proposition below characterizes both the equilibrium marriage pattern and an individual's marital payoff $h(n_i, k_i)$:

Proposition 3 *In the marriage market equilibrium, individuals only marry within their own type. The equilibrium marriage income for a type-(n, k) individual equals one half of the marital output of a type-(n, k) man marrying a type-(n, k) woman, that is, $h(n, k) = nk/(n + 1)$.*

Proof. See Appendix B, where it is verified that $2Y^M(n, k, n', k') \leq Y^M(n, k, n, k) + Y^M(n', k', n', k')$. ■

We conclude this subsection by pointing out how the different sorting patterns in the labor and marriage markets affect individuals with low social skills. In the labor market, a team member with poor social skill but high cognitive skill may still have significant labor earnings, because managers and workers match assortatively only by cognitive skills. That

is, the labor market is able to mitigate an individual's lack of social skill. On the other hand, monogamy restricts specialization in marriage. As a result, spouses match by both social and cognitive skills. So an individual with low social skill has lower marital output. Compared with the labor market, individuals are more strongly penalized for low social skills in the marriage market.

2.4 Education Market

In the education market, students need to decide which school they wish to attend. Given our production function, a type-(n_t, k_t) teacher can manage n_t students. Moreover, the cognitive skill that a student accumulates depends only on the student's initial cognitive ability a_s and the cognitive skill of her teacher k_t. Therefore, the tuition charged by a school with teacher (n_t, k_t) does not depend on the social skill of the teacher, n_t, and we can write tuition as $\tau(k_t)$. The profit for a school with a type-(n_t, k_t) teacher and n_t students is $n_t\tau(k_t) - \omega_t(n_t, k_t)$, so the free entry condition for schools implies that the teacher's wage must equal

$$\omega_t(n_t, k_t) = n_t\tau(k_t). \tag{14}$$

A simple arbitrage argument then shows that $\tau(k_t)$ must be increasing in k_t.

In order to study the education choice of students, we need to compute the return to schooling which depends on labor market earnings. Since equilibrium wages in the labor market vary across occupations, it is useful to first discuss the equilibrium occupation choice in the labor market. Note that a type-(n, k) adult can choose to become a worker, a manager, or a teacher, so his wage must be

$$\omega(n, k) = \max\{\omega_w(n, sk), \omega_m(n, k), \omega_t(n, k)\},$$

which can be rewritten as

$$\omega(n, k) = \max\{\omega(k), n\phi(k), n\tau(k)\}.$$

For a given cognitive skill level k, there exists a cutoff $\widehat{n}(k) \in [\underline{n}, \overline{n}]$ such that individuals with social skill $n < \widehat{n}(k)$ become workers, and individuals with social skill $n \geq \widehat{n}(k)$ become either managers or teachers. In particular, if $\widehat{n}(k)$ is interior, then a type-$(\widehat{n}(k), k)$ adult will be either indifferent between becoming a worker and becoming a manager, or indifferent

between becoming a worker and becoming a teacher. For the former case, we must have

$$\widehat{n}(k) = \omega(k)/\phi(k), \tag{15}$$

while for the latter, we must have

$$\widehat{n}(k) = \omega(k)/\tau(k). \tag{16}$$

Note that if, for a particular value of k, there are both managers and teachers, then we must have $\phi(k) = \tau(k)$. In this case, we cannot separate managers from teachers in terms of social skill n, because a manager with social skill $(n_1 + n_2)$ can switch his position with two teachers with respective social skill n_1 and n_2, and vice versa. Therefore, the equilibrium masses of managers and teachers are indeterminate.

Now consider the schooling decision of a type-(n_s, a_s) student. He chooses the type of school k_t that maximizes his lifetime income:

$$\begin{aligned} & \max_{k_t} \left\{ \omega\left(n_s, \sqrt{a_s k_t}\right) + h\left(n_s, \sqrt{a_s k_t}\right) - \tau(k_t) \right\} \\ = & \max_{k_t} \left\{ \max\left\{ \omega\left(\sqrt{a_s k_t}\right), n_s \phi\left(\sqrt{a_s k_t}\right), n_s \tau\left(\sqrt{a_s k_t}\right) \right\} + \frac{n_s}{n_s + 1}\sqrt{a_s k_t} - \tau(k_t) \right\}. \end{aligned}$$

Going to a school with a better teacher k_t results in a higher value of cognitive skill and thus a higher payoff in the labor and marriage markets, but comes at a higher cost of tuition. Let $\rho(n_s, a_s)$ denote the equilibrium school choice of a type-(n_s, a_s) student. Formally, $\rho(n_s, a_s)$ is defined as

$$\rho(n_s, a_s) \in \arg\max_{k_t} \left[\max\left\{ \omega\left(\sqrt{a_s k_t}\right), n_s \phi\left(\sqrt{a_s k_t}\right), n_s \tau\left(\sqrt{a_s k_t}\right) \right\} + \frac{n_s}{n_s + 1}\sqrt{a_s k_t} - \tau(k_t) \right]$$

The next lemma addresses sorting in the education market, that is, how the equilibrium school choice $\rho(n_s, a_s)$ varies with respect to a student's characteristics.

Lemma 4 *Given n_s and conditional on being a worker or a manager, a student with higher initial cognitive skill a_s will choose a teacher with higher k_t:*

$$\frac{\partial \rho(n_s, a_s)}{\partial a_s}\Big|_{occupation} \geq 0.$$

Given a_s and conditional on being a worker or a manager, a student with higher social skill n_s will choose a teacher with higher k_t:

$$\frac{\partial \rho(n_s, a_s)}{\partial n_s}\Big|_{occupation} \geq 0.$$

Proof. See Appendix B. ■

Therefore, conditional on occupation choice, students with higher initial cognitive ability a or higher social skills n will choose to attend schools of higher quality. In equilibrium, students with different cognitive abilities may choose the same school because of compensating differences in their social skill.

Finally, we want to highlight one interesting feature of the equilibrium education choice in our model. Specifically, we show that on the margin students who become managers or teachers invest discretely more in schooling than students who become workers.[15] Intuitively, with endogenous occupation choice, the return to schooling has a kink for the marginal student type who is indifferent between two occupations, creating a wedge between the optimal education choice for students with social skill just below and above this marginal type.

Proposition 4 *Let $k_t = \rho(n_s, a_s)$ denote the equilibrium education choice of a type-(n_s, a_s) student. If $\widehat{n}'(k) \neq 0$ at $k = k_t$, then $\rho(n_s, a_s)$ is discontinuous at $n_s = \widehat{n}(\sqrt{a_s k_t})$.*

Proof. See Appendix B. ■

The above idea of polarization of education choice is illustrated in Figure 1. The x-axis represents the students' eventual human capital $k_s = \sqrt{a_s k_t}$, and $c(k_s)$ represents the education cost for a student type (a_s, n_s). For illustration purposes, let us ignore the marriage market for the moment. Then both $c(k_s)$ and $\omega(k_s)$ do not depend the students' social skills n_s, while the wage for managers increases in n_s. Fix the student's ability a_s. Then the two curves $c(k_s)$ and $\omega(k_s)$ are fixed, and thus the optimal education choice k_s is determined by maximizing the distance $\omega(k_s) - c(k_s)$. Let k_s^* denote the optimal eventual education choice for student a_s if he aims to be a worker. By varying n_s, there exist one n_s^* such that the wage curve for manager type (n_s^*, k_s^*) passes through the point $(k_s^*, \omega(k_s^*))$. If in equilibrium there exist adults of type (n_s^*, k_s^*), then k_s^* must also be optimal choice for students of type (a_s, n_s^*) who want to be a manager. This is generically impossible, as illustrated in the figure: the distance $n\phi(k_s) - c(k_s)$ will not be maximized at k_s^*.

[15] There is evidence that managers have a higher rate of return to schooling than non-managers (e.g., Hirsch 1978). Building on the Roy model, Keane and Wolpin (1997) showed that the rate of return to schooling is higher for white collar workers compared with blue collar workers after controlling for a one factor unobserved ability of the individual.

[FIGURE 1 AROUND HERE]

2.5 Equilibrium

We can now formally define the equilibrium in our model. Recall that, although our model has an overlapping-generation structure, we focus on the steady state equilibrium by treating it as a two-period model.

Definition 1 *A (steady state) equilibrium consists of wages $\omega(n,k)$, tuition $\tau(k)$, marital payoffs $h(n,k)$, matching functions $\mu(k)$, $\rho(n,k)$ and $\nu(n,k)$ for the labor market, education market and marriage market, respectively, and a distribution $\alpha(n,k)$ of adult types, such that*

1. *Profit maximization: firms and schools choose the number and types of individuals to form teams to maximize their profits, given wages and tuition.*
2. *Free entry: the number of firms and schools is such that each firm and school earns zero profits.*
3. *Utility maximization: individuals choose who to match with in each sector and how to divide tasks to maximize their lifetime payoff, given wages, tuition and marital payoffs.*
4. *Market clearing: wages, tuition and marital payoffs are such that demand equals supply for each type of adult and/or student in each of the three sectors.*
5. *Consistency: the distribution of adult types is consistent with educational choices and the distribution of student types.*

In the previous subsections, we have already derived various properties of the equilibrium. A more formal characterization of the equilibrium will be provided in the next section via linear programming.

3 Linear Programming Approach

In this section, we will use linear programming techniques to analyze our model. This analysis serves three purposes. First, we show that the planner's optimization problem is a linear program, which greatly facilitates our later simulation exercises. Since markets are

competitive and there are no externalities, the social planner's solution can be implemented in decentralized markets. The equilibrium wages and utilities can be described as Lagrange multipliers attached to the constraints in the planner's problem, and they must solve the dual to the planner's problem. Second, we establish the existence of the primal solution, which is non-trivial given our continuous type space and endogenous education choice. In a companion manuscript (McCann et al., 2012), we show the maximum of the primal problem is also attained by a pair of wage functions which solve the dual problem derived below and are twice differentiable almost everywhere, thus justifying the foregoing differentiation-based analysis. Third, we also sketch how to derive our earlier characterization systematically through linear programming.

In what follows, we will formulate the linear programming problem and report the results. The details of the analysis and proofs for the results stated in this section are contained in McCann et al. (2012).

Let $A \equiv [\underline{n}, \overline{n}] \times [\underline{k}, \overline{k}]$ denote the type space for adults. Let $S \equiv [\underline{n}, \overline{n}] \times [\underline{a}, \overline{a}]$ denote the type space for students. The probability measure of student types $\sigma(n_s, a_s) \geq 0$ is exogenously given, absolutely continuous, and $\sigma(S) = 1$.

In the education market, we want to find a joint measure $\varepsilon \geq 0$ on $S \times A$ of many-to-one student-teacher pairings. The supply and demand constraint in the education market requires that the total number of type (n_s, a_s) students in all schools cannot exceed the total supply of type (n_s, a_s) students. Since a teacher of type (n_t, k_t) can mentor n_t students, we must have, for all (n_s, a_s),

$$\int_{(n_t,k_t)\in A} n_t \varepsilon(n_s, a_s; dn_t, dk_t) \leq \sigma(n_s, a_s). \tag{17}$$

Similarly, in the labor market, we want to find a joint measure $\lambda \geq 0$ on $A \times A$ of many-to-one pairings of workers to managers. Given our production technology, the output Y^L for a team consisting of a type (n_m, k_m) manager and n_m type (n_w, k_w) workers is given by

$$Y^L(n_w, k_w; n_m, k_m) = n_m \sqrt{k_w k_m},$$

independent of the workers' social skill n_w. As in the education market, the total demand of type (n, k) workers, type (n, k) managers, and type (n, k) teachers must not exceed the total supply of type (n, k) adults, for all (n, k). Since a manager of type (n_m, k_m) has the

capacity to supervise up to n_m workers, we must have, for all (n,k),

$$\begin{aligned} & \int_{(n_m,k_m)\in A} n_m \lambda(n,k;dn_m,dk_m) + \int_{(n_w,k_w)\in A} \lambda(dn_w,dk_w;n,k) + \int_{(a_s,n_s)\in S} \varepsilon(dn_s,da_s;n,k) \\ \leq & \int_{(n_t,k_t)\in A} \frac{2k}{k_t} n_t \varepsilon(n,k^2/k_t;dn_t,dk_t). \end{aligned} \tag{18}$$

where the term $2k/k_t$ on the right hand side is due to a change of variable $a_s = k^2/k_t$.

Finally, in the marriage market, we have perfect assortative matching. The equilibrium payoff from marriage for a type (n,k) adult is

$$h(n,k) = \frac{n}{n+1}k.$$

Thus the planner's primal linear program is given by

$$\sup_{\varepsilon,\lambda} \int_{A\times A} n_m \sqrt{k_w k_m} \lambda(dn_w,dk_w;dn_m,dk_m) + \int_{S\times A} n_t h(n_s,\sqrt{a_s k_t}) \varepsilon(dn_s,da_s;dn_t,dk_t) \tag{19}$$

given the constraints $\varepsilon \geq 0$, $\lambda \geq 0$, (17), and (18). We prove in McCann et al. (2012) that a solution to the planner's problem exists, using the Banach-Alaoglu Theorem and the Riesz Representation Theorem.

We now move to the dual program of the primal problem (19), the solution to which will be the equilibrium wages in the decentralized markets. The dual program can be derived heuristically as follows. Let $u : S \to R$ and $v : A \to R$ denote the Lagrange multipliers conjugate to the constraints (17) and (18), respectively. The Lagrangian function for the primal is given by

$$\begin{aligned} & \mathcal{L}(\varepsilon,\lambda;u,v) \\ = & \int_{A\times A} [n_m\sqrt{k_w k_m} - n_m v(n_w,k_w) - v(n_m,k_m)] \lambda(dn_w,dk_w;dn_m,dk_m) \\ & + \int_{S\times A} [n_t h(n_s,\sqrt{a_s k_t}) + v(n_s,\sqrt{a_s k_t}) n_t - v(n_t,k_t) - n_t u(n_s,a_s)] \varepsilon(dn_s,da_s;dn_t,dk_t) \\ & + \int_S u(n_s,a_s)\sigma(dn_s,da_s). \end{aligned}$$

Let $\mathcal{Q}$ denote the set of (u,v) satisfying the following two constraints: first, for all $(n_w,k_w),(n_m,k_m) \in A$,

$$n_m v(n_w,k_w) + v(n_m,k_m) \geq n_m \sqrt{k_w k_m}, \tag{20}$$

and second, for all $(n_s, a_s) \in S, (n_t, k_t) \in A$,

$$v(n_t, k_t) + n_t u(n_s, a_s) \geq n_t \left(\frac{n_s}{n_s + 1} \sqrt{a_s k_t} \right) + n_t v(n_s, \sqrt{a_s k_t}). \tag{21}$$

Then by the duality principle, the constrained maximum value of the primal problem must be equal to the unconstrained minimax

$$\inf_{u,v} \sup_{\varepsilon,\lambda \geq 0} \mathcal{L}(\varepsilon, \lambda; u, v) = \inf_{(u,v) \in \mathcal{Q}} \sup_{\varepsilon,\lambda \geq 0} \mathcal{L}(\varepsilon, \lambda; u, v).$$

The equality follows because if $(u, v) \notin \mathcal{Q}$, we can always find ε and λ such that the supremum is unbounded. Therefore, we can rewrite the dual program as

$$\inf_{u,v} \int_{(n_s,a_s) \in S} u(n_s, a_s) \sigma(dn_s, da_s), \tag{22}$$

subject to (20) and (21).

The Lagrange multipliers $u(n, a)$ and $v(n, k)$ can be interpreted as the indirect utility (or wages) that students and adults of various types derive from their position in these markets. They must be nonnegative. To see this, note that, by setting $n_w = n_m = n$ and $k_w = k_m = k$ in (20), we obtain $v\,(n, k) \geq nk \geq 0$ for all n and k. Similarly, by setting $n_t = n_s = n$ and $k_t = a_s = a$ in (21), we obtain

$$u(n, a) \geq \frac{n}{n+1} a + \frac{n-1}{n} v(n, a).$$

Since $n \geq 2$, we have $u(n, a) \geq 0$ for all n and a.

If both type spaces A and S are discrete, the standard dual principle immediately implies that there is no gap between the primal and the dual, and that the dual is obtained. With continuous type spaces, the existence proof of the dual problem is more subtle and involved.

As a corollary to the existence proof, we can obtain the differentiability of the wage functions, $v\,(n, k)$ and $u\,(n, a)$, up to the second order. Two of the relevant tools are Rademacher's Theorem and Alexandrov's Theorem.

Next, we sketch how we can characterize our equilibrium matching through condition (20) and (21). Condition (20) must hold with equality for λ-a.e. worker-manager pair and condition (21) also holds with equality for ε-a.e. teacher-student pair. These equalities reflect the fact that the equilibrium permits no opportunities for arbitrage.

Given differentiability of function v, we deduce the following first-order conditions from (20): for λ-a.e. pair of worker (n_w, k_w) and manager (n_m, k_m),

$$\frac{\partial v(n_w, k_w)}{\partial k_w} = \frac{1}{2}\sqrt{\frac{k_m}{k_w}} \tag{23}$$

$$\frac{\partial v(n_w, k_w)}{\partial n_w} = 0 \tag{24}$$

$$\frac{\partial v(n_m, k_m)}{\partial k_m} = \frac{1}{2} n_m \sqrt{\frac{k_w}{k_m}} \tag{25}$$

$$v(n_w, k_w) + \frac{\partial v(n_m, k_m)}{\partial n_m} = \sqrt{k_w k_m} \tag{26}$$

It follows from (24) that the worker types enjoy wage $v(n_w, k_w) = \omega(k_w)$ independent of n_w. We also see from (23) that each worker's cognitive skill k_w determines the cognitive skill of his manager

$$\frac{1}{2}\sqrt{\frac{k_m}{k_w}} = \frac{d\omega(k_w)}{dk_w}$$

The wage constraint (20) now implies $v(n_m, k_m) = n_m \phi(k_m)$ holds for a.e. type (n_m, k_m) of manager, where $\phi(k_m)$ is defined as

$$\phi(k_m) = \sup_k \sqrt{k k_m} - \omega(k).$$

The formula $v(n_m, k_m) = n_m \phi(k_m)$ implies the worker and manager regions must be disjoint. It is easy to verify that the last two first-order conditions (25) and (26) are automatically satisfied. Condition (25) implies that

$$\frac{1}{2}\sqrt{\frac{k_w}{k_m}} = \frac{d\phi(k_m)}{dk_m}$$

so the cognitive skill of each workers k_w is determined by the human-capital of his manager k_m. Continuity of v across the worker/manager interface $\widehat{n}(k)$ forces $\omega(k) = \widehat{n}(k)\phi(k)$.

By following a similar procedure, we can derive the equilibrium matching pattern in the education market by operating on condition (21).

4 Simulation

We illustrate the properties of the equilibrium by simulating the model. The simulation is computationally straightforward because of the equivalence between the market equilibrium and the solution of the planner's problem, which is a linear program as shown in the previous

section. We solve the primal problem (19), abstracting from the marriage market since it does not change any of the outcomes qualitatively but obscures some of the effects that we want to highlight.[16] Note that solving (19) gives us the equilibrium wages as the multipliers.

For the simulation, we consider a mass 1 of students with uniformly distributed skills. The support of the distribution of social skill n and initial cognitive skill a is assumed to be equal to $[\underline{n}, \overline{n}] \times [\underline{a}, \overline{a}] = [2, 10] \times [0.1, 1]$. The distribution of adult cognitive skill k and its support $\left[\underline{k}, \overline{k}\right]$ are endogenously determined through the students' education choices. We discretize the support of both student and adult skill for computational reasons, by using 19 grid points for n and 91 grid points for a and k.

[FIGURE 2 AROUND HERE]

The results of the simulations are presented in figure 2-5. Figure 2 displays the education choice $\rho(n_s, a_s)$ i.e. the cognitive skill k_t of the teacher chosen by students with social skill n_s and initial cognitive skill a_s. Students who become workers later in life are displayed in yellow, while students who will be managers swmor teachers are shown in brown. The figure confirms several of the equilibrium properties. For example, students with higher values of a match with teachers with higher values of k, as we described in Lemma 4. Given the absence of the marriage market, the value of n has no effect on the education choice of students who will become workers, but does affect the school attended by future managers and teachers. Figure 2 also clearly shows the educational gap between workers and managers/teachers that we described in Proposition 4.

[FIGURE 3 AROUND HERE]

Figure 3 shows the occupational choice of the adults. Most adults become workers, which are displayed in yellow. In order to become a manager or a teacher, the adult requires a sufficiently high value of both n and k. Note that in this equilibrium, managers (black) and teachers (light-brown) are almost entirely separated in terms of cognitive skill k. Relative to managers, teachers have either lower or higher cognitive skill. Only for a few values of k we find pooling of managers and teachers (indicated by dark-brown in the figure). Teachers with lower cognitive skills are matched with students who will become future workers, while teachers with higher cognitive skills are matched with students who will become future

[16] To be precise, it slightly reduces the magnitude of the educational gap.

managers or teachers. The gap reflects the discontinuity in the demand for teachers with different cognitive skills: students who are future managers and teachers invest discretely more in education than students who are future workers.[17]

Figure 3 further shows that certain combinations of n and k are absent in the labor market. For example, no adult has a very low value of k since everyone invests at least a certain amount in education. Similarly, low values of n combined with very high values of k do not arise, since students with low values of n realize that they will never become managers or teachers in the second period. Therefore, they are generally not willing to attend very high levels of education. Finally, the figure clearly shows the educational gap between workers and managers/teachers as an empty band between the two corresponding sets of adult types.

[FIGURE 4 AROUND HERE]

Figure 4 displays the wages earned in the labor market. More precisely, it shows the wage $\omega(n,k)$ for an adult of type-(n,k), again using yellow to indicate workers and brown to indicate managers / teachers. As discussed in section 2, a worker's wage does not vary with n, while the wages of managers and teachers are linearly increasing in n. Note further that the worker with the highest wages may earn more than the least-paid managers and teachers.

[FIGURE 5 AROUND HERE]

Figure 5 shows the cross-sectional wage density. One feature stands out: in line with empirical evidence from real labor markets, the simulated wage density is asymmetric with a short left tail and a long right tail. It is important to stress that this occurs even though we assumed a bivariate uniform distribution for a and n. Introducing any asymmetry in this distribution would only further strengthen this result. Note that workers, even with very high cognitive skills, do not receive the highest wages. The highest wages go to managers/teachers with strong social and cognitive skills. It turns out that heterogeneity in social skills is important for generating the right skewness in the wage distribution.[18]

[17] A university degree generally requires fours years of schooling. Most professors also have PhDs which require significantly more schooling. Although Spence style signalling models also imply bunching of educational attainment, those models do not imply that teachers of high schooling attainment students should have significantly higher schooling themselves.

[18] For example, we simulated a model with a fixed $n = 5$ for all individuals and did not obtain a long right tail in the wage distribution.

Finally, we simulated one comparative static with the model. We extend the support of social skill by increasing $\overline{n}$[19]. As a result of this change, the equilibrium number of managers and teachers falls and the skewness of log wages increases. This result is reminiscent of the Rosen's (1981) superstar phenomenon as well as Garicano and Rossi-Hansberg's (2006) concern about advances in communication technology.

5 Related Literature

Our SC model builds on several classical ideas in economics, including matching, comparative advantage, and task assignment. We also draw insights from a more recent literature on the role of non-cognitive skills in individual behavior. We review the relevant literature and relate our findings in this section.

A large empirical literature in psychology and a smaller one in economics document the quantitative importance of non-cognitive skills in affecting individual behavior and outcomes.[20] For example, Heckman et al. (2006) show that non-cognitive skill strongly influences schooling decisions, occupational choice, and wages. For many labor market outcomes, the effect of a change in non-cognitive skill is found to be comparable to or greater than a corresponding change in cognitive skill.

Researchers have considered a variety of non-cognitive skills. Among psychologists, the Big Five model of personality traits is particularly popular (see e.g. McCrae and Costa, 1999, and John et al., 2008). Heckman et al. (2011) shows that a three-factor structure fits the data well. Two of these correspond to our cognitive and social factors. Their third factor is related to the marginal utility of leisure which we ignored in this paper. Consistent with our SC model, empirical researchers often find that the empirical cognitive and non-cognitive factors which they recover are not orthogonal to each other.[21]

Adding social skill as a second dimension of heterogeneity to a standard matching model leads to new questions regarding task assignment and sorting. In particular, task assignment and sorting patterns will no longer be based on cognitive skill alone.

The literature on marriage matching with transferable utility was initiated by Becker

[19] An alternative but similar exercise would be to increase γ, reducing the time needed to do task C relative to task I

[20] Borghans et al. (2008) and Almlund et al. (2011) review this literature for economists.

[21] See e.g., Markon, et al. (2005) and DeYoung (2006).

(1973; 1974),[22] and was subsequently extended in various ways (see e.g. Shimer and Smith, 2000; Legros and Newman, 2002). The assumption of one-dimensional heterogeneity is maintained in most of this literature. Notable exceptions include Anderson (2003) and Chiappori, Oreffice and Quintana-Domeque (2010). In the former paper, individuals are heterogeneous in both social status and wealth, while in the latter they are heterogeneous in educational attainment and smoking behavior. Our work differs in various ways. First, individuals in our model are heterogeneous in social and cognitive skills, and we consider matching in various sectors. Further, we allow for endogenous occupation choice in the labor market and endogenous human capital accumulation in the education market.

Task assignment on the basis of comparative advantage dates back to Ricardo. Roy (1951) was the first to apply this concept to occupation choice based on occupation-specific skills. These ideas have been formalized and extended by various authors, with Sattinger (1975) being an early example.[23] In Sattinger's model, however, the two sides of the market are exogenously determined. In that sense, our model is closer to contributions by Lucas (1978), Rosen (1978, 1982), Garicano and Rossi-Hansberg (2004, 2006), in which agents endogenously choose their roles within the firm.[24]

By embedding task assignment, occupation choice and schooling choice in a multisector matching framework, our model with a common production technology in all three sectors is able to explain different matching patterns in the three sectors and various empirical observations.

First, our model delivers many-to-one matching and hierarchy structure in the labor market, similar to Lucas (1978), Rosen (1978, 1982), and Garicano (2000). Furthermore, in our model, the individual assignment to different positions depends on not only the cognitive skills but also the social skills. In particular, individuals with higher social skills are more likely to be managers, which is consistent with empirical observations documented in Smeets and Warzynski (2008). In our model, there is positive assortative matching along the cognitive skill dimension between managers and workers, but equilibrium occupation choices are

[22] The empirical relevance of this model has been demonstrated by Choo and Siow (2006), Brandt, Siow and Vogel (2011), Siow (2010), Chiappori, Salanie and Weiss (2010).

[23] For applications of the Roy model, see e.g. Heckman and Honore (1990), Heckman and Sedlacek (1985).

[24] In a model where workers are heterogenous in *one* dimensional skills and firms are heterogenous in *one* dimensional quality of endowment, Eeckhout and Kircher (2011) study how firms solve the tradeoffs of span of control over more versus better workers.

based on both dimensions of skills.[25]

Second, our equilibrium in the marriage market exhibits PAM as in Becker (1973, 1974), but it sorts along both dimensions of skills. Moreover, full specialization is not optimal in marriage. Our mechanism for incomplete specialization in monogamy recalls Adam Smith's argument on limits to the division of labor. Becker (1991) also studies the division of labor within the household, and argues that at least one spouse will be fully specialized in a task. Pollak (2011) argues that Becker's full specialization is a special case based on concerns different from those discussed here.

Third, as in the labor market, equilibrium matchings in the education market are many-to-one. Students sort in both dimensions of skills, so students with different combination of social skill and initial cognitive skill are enrolled in the same school. Thus, our model provides an explanation to the puzzle why high and low cognitive ability students are in the same school.[26]

While positing a different production technology than Garicano (2000) in the labor market, we follow his concern about the importance of communication cost in limiting team size. Garicano and Rossi-Hansberg (2004, 2006) used his technology to study a one factor (cognitive skill) matching model of the labor market with organizational design, occupational choice and worker investment. One way to view our contribution is that we extend their concerns to allow individuals to have heterogenous communication costs and we study what this additional dimension of heterogeneity implies for the schooling, labor and marriage markets.

Our theory also delivers predictions consistent with various empirical observations. First, conditional on cognitive ability, our model predicts that managers' wages are increasing in their social skills. The social skills in our model can be interpreted as leadership or span of control. Kuhn and Weinberger (2005) find that individuals who occupy leadership posi-

[25] There are mixed evidences of positive assorting in the labor market, see for example, Abowd, Kramarz and Margolis (1999). Lise, Meghir and Robin (2011) develop and estimate an equilibrium wage determination model that also incorporates positive assortative matching between workers and jobs. They allow for search frictions, but agents in their model are heterogenous in one dimension and they focus on the labor market only.

[26] A common answer in the literature is peer effects. In their surveys on peer effects among college students, however, Epple and Romano (2011) and Sacerdote (2011) do not find quantitatively large academic student peer effects for students with low cognitive skills.

tions in high school as students are more likely to become managers as adults. Moreover, conditional on cognitive skills, these individuals earn wage premium up to 33%. Interestingly, leadership skills command a higher wage premium within managerial occupations than elsewhere.

The empirical literature on the span of control also points to a positive relationship between managers' wages and the number of individuals they supervise. Smeets and Warzynski (2008) show on the basis of a survey data that wages and bonuses are increasing in the span of control. Furthermore, using survey data from a panel of more than 300 large U.S. companies over the period 1986-1999, Rajan and Wulf (2006) find a simultaneous increase in the span of control of CEOs and wage inequality.

As shown in the simulation, our model can generate a non-monotonic wage distribution that is skewed to the right, which is consistent with empirical regularities summarized in Neal and Rosen (2000). Sattinger (1975), Rosen (1978, 1982) and Waldman (1984) have used one-factor models to show that task assignment can generate right tail skewness in the earnings distribution.[27] These papers, however, do not explore whether the predicted earnings distributions are qualitatively consistent with the entire earnings distribution (roughly lognormal with a fat right tail).[28] Acemoglu and Autor (2011) argues that task assignments and skill biased technical change are needed to explain the recent evolution of the US labor market.

On the technical side, Chiappori, McCann, Nesheim (2010) show that the frictionless multifactor marriage matching model is equivalent to a utilitarian social planner's linear programming problem.[29] We extend this equivalence to a frictionless multisector multifactor many-to-many matching framework with endogenous occupational choice.

[27] See Gibbons and Waldman (1999) for a survey, and Gabaix and Landier (2008) for a more recent contribution.

[28] Li (2012) shows how a model of task assignment can approximate the fat right tail and the evolution of empirical wage distribution by introducing Pareto learning.

[29] See also Li and Suen (2001), and McCann and Trokhimtchouk (2010) for a single sector, unidimensional model with endogenous occupation choice.

6 Conclusion

Motivated by the empirical literature on the importance of social skills for lifetime outcomes, this paper presents a social and cognitive skills model (SC model) of human capabilities. The model generates different matching patterns for the education, labor, and marriage market and a large number of other empirical predictions. The new results in this paper are primarily due to our new team production function, which allows individuals to specialize on the basis of comparative advantage, but introduces the need for costly coordination.

In the paper, we make several simplifying assumptions to keep the analysis as transparent and tractable as possible. Our model is therefore best viewed as an initial exploration of how cognitive and social skills may work in the various environments, and it provides several avenues for further research.

First, the theoretical model can be extended by relaxing some of our assumptions. For example, we aggregate all non-cognitive factors into one social factor. Multiple factors could be considered. Further, one could allow individuals to decide on the time allocation between the labor and marriage market, or allow different individuals to value marital output and labor market income differently. The introduction of two-sided communication costs or interaction among workers may help generate richer predictions regarding matching in the labor market. In addition, one could introduce gender differences in the distributions of cognitive and non-cognitive skills,[30] or allow individuals to accumulate not only cognitive skills but also social skills.

It would also be interesting to consider more heterogeneity on the firm side, e.g. with respect to firm resources or price of output. Such heterogeneity may interact with the social skill of a manager, generating different spans of control for managers with the same n. Finally, more elaborate task structures could be considered, e.g. with a larger number of tasks, with an endogenous division of tasks to workers, or with a multi-level hierarchy of tasks. The last extension may yield multi-level hierarchies among a firm's employees as well.

This paper also generates many empirical research questions. A first order problem is to separately identify social and cognitive skills in different sectors. For example, what are the testable implications of social skills regarding separations in the labor and marriage markets? Motivated by this paper, ongoing research by Kambourov, Siow and Turner (2012) shows

[30] See e.g. Siow (1998) for an analysis of gender roles.

that individuals with more previous divorces are more likely to separate from their current employer, and individuals with more previous separations from employers are more likely to divorce their current spouse.

Another avenue for empirical research is to study non-cognitive factors which affect the schooling decision and subsequent occupational choices. In particular, the model predicts that the rate of return to schooling is increasing in the span of a manager. In their survey of the recent empirical literature on the determinants of increasing earnings inequality by cognitive skills, Acemoglu and Autor (2011) argued that a task-assignment model with occupational choice is needed to fit some of the finer details of the phenomenon. It will be interesting to see if our model can be used to shed light on this empirical literature.

Finally, there is need to integrate the large non-behavioral empirical literature on cognitive and non-cognitive skills, and the recent empirical behavioral models on cognitive and non-cognitive skill acquisitions with the equilibrium matching concerns studied here.

Appendix A: Alternative Micro-foundation for Our Technology

There are several different ways to motivate our team production technology. In this appendix, we sketch out two alternatives. To keep the exposition brief, we only highlight the differences compared to our main motivation.

Alternative I

Output is produced by successful completion of two productive tasks, A and B. There is return to specialization in these tasks, which provides individuals with an incentive to form teams. However, team production introduces a need for coordination, i.e. task C. We model the coordination task as follows. Consider a team member i who specializes in task A. He has to frequently check with another team member j who has knowledge about the production of task B to make sure that his production follows the team plan. The cost of coordination is again incurred by the receiver (i.e. member j) and depends on his social skill n_j. Similarly, a team member who specializes in task B needs to frequently check with a member who has knowledge about the production of task A.

Consider first an individual i of type (n_i, k_i) who works alone. If he spends $\theta_i^A \in [0,1]$ units of time on task A and $\theta_i^B \in [0,1]$ units of time on task B, then output is given by the following Leontief production function:

$$Y\left(\theta_i^A, \theta_i^B; n_i, k_i\right) = \beta \min\left\{k_i\theta_i^A, k_i\theta_i^B\right\}, \tag{27}$$

subject to the time constraint $\theta_i^A + \theta_i^B \leq 1$. The parameter $\beta \in (0,1)$ captures the loss in output due to insufficient specialization. .

Consider next a two-member team with individual i of type (n_i, k_i) and individual j of type (n_j, k_j). Without loss of generality, let i specialize in task A and j in task B. Suppose i spends $\theta_i^A \in [0,1]$ units of time on task A, and j spends $\theta_j^B \in [0,1]$ units of time on task B. Successful coordination then requires that member i and j respectively spend $\theta_i^C = \theta_j^B/n_i$ and $\theta_j^C = \theta_i^A/n_j$ units of time on coordination. Total team output is given by

$$\min\left\{\sqrt{k_ik_j}\theta_i^A, \sqrt{k_ik_j}\theta_j^B\right\} \tag{28}$$

subject to $\theta_i^A + \theta_i^C \leq 1$ and $\theta_j^A + \theta_j^C \leq 1$. Here we omit the parameter β to reflect the benefit of full specialization with respect to A and B. In addition, we assume that the cognitive skills of the two team members are complementary in producing output, which is captured by the term $\sqrt{k_i k_j}$.

In a two-member team, both individuals need to spend some time on coordination, that is, there is no full specialization with respect to task C. Therefore, it may be advantageous to extend a two-member team with a third member (say member m) who can coordinate both task A and B. The output of three-member team is given by

$$\min \left\{ \sqrt{k_i k_m}\theta_i^A, \sqrt{k_j k_m}\theta_j^B \right\} \tag{29}$$

with $\theta_i^A, \theta_j^B, \theta_m^C \in [0,1]$ and subject to $\theta_i^A + \theta_j^B = n_m \theta_m^C$. Note that if member m has a high value of social skill n_m, she may have extra time left which she can use to coordinate more worker pairs.

In education market, there is only one productive task (task A) representing learning by students. The coordinative task C represents teacher instruction and coordination. Successful accumulation of human capital requires completion of both tasks. The inputs in the production function are the initial cognitive skill a_i of student i and the adult cognitive skill k_j of the teacher j, while output is the adult cognitive skill k_i of individual i, determined by the production function

$$k_i = \sqrt{a_i k_j}\theta_i^A,$$

subject to $\theta_i^A, \theta_j^C \in [0,1]$ and $\theta_i^A = n_j \theta_j^C$.

Alternative II

Output is produced by completion of one productive task. Individuals who work in teams can ask each other for help when they encounter problems in production. However, discussing problems and solutions is time consuming. We assume the time cost of communication is fully incurred by the receiver.

Consider first an individual i of type (n_i, k_i) who works alone. The output that he produces is proportional to his cognitive skill k_i:

$$Y(\theta_i; n_i, k_i) = \beta k_i \theta_i$$

where $\theta_i \in [0,1]$ is the amount of time that he spends on production, and $\beta \in (0,1)$ captures the idea that producing alone may be less effective because one cannot ask for help when problems arise.

Consider next a two-member team with individual i of type (n_i, k_i) and individual j of type (n_j, k_j). Suppose i spends θ_i units of time on production, and j spends θ_j units of time on production. Then individual i has to spend $\psi_i = \theta_j/n_i$ in helping individual j, and individual j has to spend $\psi_j = \theta_i/n_j$ in helping individual i. The total team output is given by

$$Y(\theta_i, \theta_j; n_i, k_i, n_j, k_j) = \sqrt{k_i k_j}\,(\theta_i + \theta_j)$$

subject to $\theta_i + \psi_i \leq 1$ and $\theta_j + \psi_j \leq 1$. Here we omit the parameter β to reflect the benefit of team production with communication. In addition, the output produced now depends on the cognitive skill of both team members which are assumed to be complementary.

In the two-member team, both individuals need to spend some time on communication, that is, there is no full specialization in communication. It may be advantageous to extend a two-member team with a third member (say member m) who can specialize in communication. The output of three-member team is given by

$$\sqrt{k_i k_m}\theta_i + \sqrt{k_j k_m}\theta_j$$

subject to $\theta_i + \theta_j = n_m \psi_m$ with $\theta_i, \theta_j, \psi_m \in [0,1]$. Note that if member m has a high value of social skill n_m, she may have extra time left which she can use to help more workers.

In the education market, student i performs the productive task, while teacher j provides help and instruction when needed. The inputs in the production function are the initial cognitive skill a_i of student i and the adult cognitive skill k_j of the teacher j, while output is the adult cognitive skill k_i of individual i, determined by the production function

$$k_i = \sqrt{a_i k_j}\theta_i$$

subject to $\theta_i = n_j \psi_j$ with $\theta_i, \psi_j \in [0,1]$.

Appendix B: Omitted Proofs

Proof of Proposition 3. We first show that equilibrium a type-(n,k) individual must match with his/her own type. We assume by contradiction that in equilibrium a type-(n,k)

individual marries a type-$(n+\Delta_n, k+\Delta_k)$ individual with $\max\{|\Delta_n|, |\Delta_k|\} > 0$. Now consider a switch of two such pairs by matching each individual with their own type. The gain from such a switch is given by

$$\begin{aligned} & Y^M(n,k;n,k) + Y^M(n+\Delta_n, k+\Delta_k; n+\Delta_n, k+\Delta_k) - 2Y^M(n,k;n+\Delta_n,k+\Delta_k) \\ = & \frac{2n}{n+1}k + \frac{2(n+\Delta_n)}{(n+\Delta_n)+1}(k+\Delta_k) - 2\sqrt{k(k+\Delta_k)}\frac{2n(n+\Delta_n)-2n-\Delta_n}{n(n+\Delta_n)-1} \\ \geq & \left(\sqrt{\frac{2n}{n+1}\frac{2(n+\Delta_n)}{(n+\Delta_n)+1}} - \frac{2n(n+\Delta_n)-2n-\Delta_n}{n(n+\Delta_n)-1}\right) 2\sqrt{k(k+\Delta_k)}. \end{aligned}$$

If $\Delta_n = 0$, then $\Delta_k \neq 0$, which implies that the above inequality is strict and the last expression is zero, so we have the desired contradiction. In order to generate a contradiction for the case of $\Delta_n \neq 0$, it is sufficient to show

$$\sqrt{\frac{2n}{n+1}\frac{2(n+\Delta_n)}{(n+\Delta_n)+1}} - \frac{2n(n+\Delta_n)-2n-\Delta_n}{n(n+\Delta_n)-1} > 0 \tag{30}$$

for $\Delta_n \neq 0$. First observe that we can without loss to assume $\Delta_n > 0$, because if $\Delta_n < 0$, we can always obtain the same condition by redefining

$$\Delta_n' = -\Delta_n,\ n' = n+\Delta_n, n'+\Delta_n' = n.$$

Next, we reformulate (30) as

$$\frac{4n(n+\Delta_n)(n(n+\Delta_n)-1)^2 - (n+1)(n+\Delta_n+1)(2n(n+\Delta_n)-2n-\Delta_n)^2}{(n+1)(n+\Delta_n+1)(n(n+\Delta_n)-1)^2} > 0$$

Since the denominator is positive, it is sufficient to show numerator is positive. Let $f(n,\Delta_n)$ denote the numerator. Since

$$f(1,\Delta_n) = 2\Delta_n^3 > 0$$

for all $\Delta_n > 0$, it is sufficient to show that $f(n,\Delta_n)$ is increasing in n for fixed Δ_n. With some algebra, we can show that

$$\frac{\partial f(n,\Delta_n)}{\partial n} = \Delta_n^2(6n+3\Delta_n-2)$$

which is positive for all $\Delta_n \geq 0$ and $n \geq 2$. This completes the proof that equilibrium must exhibit PAM. The claim that $h(n,k) = nk/(n+1)$ is then immediate. ■

Proof of Lemma 4. Note that we can write

$$\phi(k) = \max_{k'} \sqrt{kk'} - \omega(k').$$

The necessary first-order and second-order conditions are

$$\frac{1}{2}\sqrt{\frac{k}{k'}} - \omega'(k') = 0 \text{ and } -\frac{1}{4k'}\sqrt{\frac{k}{k'}} - \omega''(k') \leq 0,$$

which imply that

$$\omega''(k') \geq -\frac{1}{2k'}\omega'(k').$$

Now first suppose student (n_s, a_s) will become a worker. Define

$$\Pi(n_s, a_s; n_t, k_t) \equiv \omega\left(\sqrt{a_s k_t}\right) + \frac{n_s}{n_s+1}\sqrt{a_s k_t} - \tau(k_t).$$

Then we have

$$\begin{aligned}
\frac{\partial^2 \Pi(n_s, a_s; n_t, k_t)}{\partial a_s \partial k_t} &= \frac{1}{4}\omega''\left(\sqrt{a_s k_t}\right) + \frac{1}{4\sqrt{a_s k_t}}\omega'\left(\sqrt{a_s k_t}\right) + \frac{1}{4\sqrt{a_s k_t}}\frac{n_s}{n_s+1} \\
&\geq -\frac{1}{4}\frac{1}{2\sqrt{a_s k_t}}\omega'\left(\sqrt{a_s k_t}\right) + \frac{1}{4\sqrt{a_s k_t}}\omega'\left(\sqrt{a_s k_t}\right) + \frac{1}{4\sqrt{a_s k_t}}\frac{n_s}{n_s+1} \\
&= \frac{1}{8\sqrt{a_s k_t}}\omega'\left(\sqrt{a_s k_t}\right) + \frac{1}{4\sqrt{a_s k_t}}\frac{n_s}{n_s+1} \\
&> 0
\end{aligned}$$

Therefore, $\Pi(n_s, a_s; n_t, k_t)$ is supermodular in k_t and a_s. Moreover,

$$\frac{\partial^2 \Pi(n_s, a_s; n_t, k_t)}{\partial n_s \partial k_t} = \frac{1}{(n_s+1)^2}\frac{1}{2}\sqrt{\frac{a_s}{k_t}} > 0$$

so $\Pi(n_s, a_s; n_t, k_t)$ is also supermodular in k_t and n_s. It follows from Topkis's Theorem that the equilibrium matching $k_t = \rho(a_s, n_s)$ is increasing in both a_s and n_s.

Next consider the case where student (n_s, a_s) chooses to become a manager. Note that we can write

$$\omega(k) = \max_{k'} \sqrt{kk'} - \phi(k').$$

The necessary first-order and second-order conditions imply that

$$\phi''(k') \geq -\frac{1}{2k'}\phi'(k')$$

Let's define

$$\widehat{\Pi}(n_s, a_s; n_t, k_t) \equiv n_s\phi\left(\sqrt{a_s k_t}\right) + \frac{n_s}{n_s+1}\sqrt{a_s k_t} - \tau(k_t)$$

Then we have

$$\begin{aligned}\frac{\partial^2\widehat{\Pi}\left(n_s,a_s;n_t,k_t\right)}{\partial a_s\partial k_t} &= n_s\frac{1}{4}\phi''\left(\sqrt{a_sk_t}\right)+n_s\frac{1}{4\sqrt{a_sk_t}}\phi'\left(\sqrt{a_sk_t}\right)+\frac{1}{4\sqrt{a_sk_t}}\frac{n_s}{n_s+1}\\ &\geq -\frac{1}{4}\frac{1}{2\sqrt{a_sk_t}}\phi'\left(\sqrt{a_sk_t}\right)+\frac{1}{4\sqrt{a_sk_t}}\phi'\left(\sqrt{a_sk_t}\right)+\frac{1}{4\sqrt{a_sk_t}}\frac{n_s}{n_s+1}\\ &= \frac{1}{8\sqrt{a_sk_t}}\phi'\left(\sqrt{a_sk_t}\right)+\frac{1}{4\sqrt{a_sk_t}}\frac{n_s}{n_s+1}\\ &> 0\end{aligned}$$

and

$$\frac{\partial^2\widehat{\Pi}\left(n_s,a_s;n_t,k_t\right)}{\partial n_s\partial k_t}=\frac{1}{2}\sqrt{\frac{a_s}{k_t}}\phi'\left(\sqrt{a_sk_t}\right)+\frac{1}{\left(n_s+1\right)^2}\frac{1}{2}\sqrt{\frac{a_s}{k_t}}>0$$

Therefore, $\widehat{\Pi}\left(n_s,a_s;n_t,k_t\right)$ is supermodular in k_t and a_s, and supermodular in k_t and n_s. Thus, the equilibrium matching $k_t=\rho\left(a_s,n_s\right)$ is increasing in both a_s and n_s. ■

Proof of Proposition 4. Consider a type-(a_s,n_s) student whose equilibrium school choice is $k_t=\rho\left(n_s,a_s\right)$. If this student becomes a manager eventually, the optimal school choice k_t must satisfy

$$\frac{1}{2}n_s\phi'\left(\sqrt{a_sk_t}\right)\sqrt{\frac{a_s}{k_t}}+\frac{n_s}{n_s+1}\frac{1}{2}\sqrt{\frac{a_s}{k_t}}-\tau'\left(k_t\right)=0. \tag{31}$$

In contrast, if this student eventually becomes a worker, k_t must satisfy

$$\frac{1}{2}\omega'\left(\sqrt{a_sk_t}\right)\sqrt{\frac{a_s}{k_t}}+\frac{n_s}{n_s+1}\frac{1}{2}\sqrt{\frac{a_s}{k_t}}-\tau'\left(k_t\right)=0. \tag{32}$$

Now suppose $n_s=\widehat{n}(\sqrt{a_sk_t})$. If this student indeed chooses education level k_t, then k_t must solve both (31) and (32), which implies that

$$\widehat{n}(\sqrt{a_sk_t})\phi'\left(\sqrt{a_sk_t}\right)=\omega'\left(\sqrt{a_sk_t}\right). \tag{33}$$

Note that, by definition of $\widehat{n}\left(k\right)$, we have

$$\widehat{n}'\left(k\right)=\frac{\omega'\left(k\right)\phi\left(k\right)-\omega\left(k\right)\phi'\left(k\right)}{\left[\phi\left(k\right)\right]^2}=\frac{\omega'\left(k\right)-\widehat{n}\left(k\right)\phi'\left(k\right)}{\phi\left(k\right)}$$

So we have

$$\frac{\omega'\left(k\right)}{\phi'\left(k\right)}=\widehat{n}\left(k\right)+\frac{\phi\left(k\right)\widehat{n}'\left(k\right)}{\phi'\left(k\right)}$$

Therefore, condition (33) reduces to $\widehat{n}'\left(\sqrt{a_sk_t}\right)=0$, a contradiction. ■

References

[1] Abowd, J., F. Kramarz, and D. N. Margolis (1999): "High Wage Workers and High Wage Firms," *Econometrica*, 67(2), 251-333.

[2] Acemoglu, D. and D. Autor (2011): "Skills, Tasks and Technologies: Implications for Employment and Earnings," *Handbook of Labor Economics*, vol. 4B, Amsterdam, Elsevier, 1043-1172.

[3] Almlund, M., A. Duckworth, J. Heckman, and T. Kautz (2011): "Personality psychology and economics," *Handbook of the Economics of Education*, vol. 4, chapter 1, New York, Elsevier.

[4] Anderson, S. (2003): "Why Dowry Payments Declined with Modernization in Europe but Are Rising in India," *Journal of Political Economy* 111(2): 269-310.

[5] Bandiera, O., A. Prat, R. Sadun, and J. Wulf (2012): "Span of Control and Span of Activity," Working Paper, Harvard Business School.

[6] Becker, G. (1973): "A Theory of Marriage: Part I," *Journal of Political Economy*, 81, 813-846.

[7] Becker, G. (1974): "A Theory of Marriage: Part II," *Journal of Political Economy*, 82, S11-S26.

[8] Becker, G. (1985): "Human Capital, Effort, and the Sexual Division of Labor," *Journal of Labor Economics*, 3(1), 33-58.

[9] Becker, G. (1991): *A Treatise on the Family*, Cambridge, Mass., Harvard University Press.

[10] Brandt, L., A. Siow and C. Vogel (2010): "Large Shocks and Small Changes in the Marriage Market for Famine Born Cohorts in China,"Working Paper, University of Toronto.

[11] Borghans, L., A. L. Duckworth, J. J. Heckman, and Bas ter Weel (2008): "The Economics and Psychology of Personality Traits." *Journal of Human Resources* 43(4): 972-1059.

[12] Bruggink, T.H., and V. Gambhir (1996): "Statistical Models for College Admission and Enrollment: A Case Study for A Selective Liberal Arts College," *Research in Higher Education* 37(2): 221-240.

[13] Chiappori, P. A., R. J. McCann, and L. Nesheim (2010): "Hedonic Price Equilibria, Stable Matching, and Optimal Transport: Equivalence, Topology, and Uniqueness," *Economic Theory*, 42(2), 317-354.

[14] Chiappori, P. A., S. Oreffice, and C. Quintana-Domeque (2010): "Matching with A Handicap: The Case of Smoking in the Marriage Market," Working Paper, Columbia University.

[15] Chiappori, P. A., B. Salanie, and Y. Weiss (2010): "Partner Choice and the Marital College Premium," Working Paper, Columbia University.

[16] Choo, E., and A. Siow (2006): "Who Marries Whom and Why," *Journal of Political Economy*, 114(1), 175-201.

[17] Cleek, M. and T. Pearson (1985): "Perceived Causes of Divorce: An Analysis of Interrelationships," *Journal of Marriage and Family* 47 (1), 179-183.

[18] Cunha, F. and J. Heckman (2007): "The Technology of Skill Formation," *American Economic Review* 97(2), 31-47.

[19] Cunha, F., J. Heckman, and S. Schennach (2010): "Estimating the Technology of Cognitive and Non-cognitive Skill Formation," *Econometrica* 78(3), 883–931.

[20] DeYoung, C. (2006): "Higher-order Factors of the Big Five in a Multi-informant Sample," *Journal of Personality and Social Psychology* 91(6), 1138-1151.

[21] Eeckhout, J., and P. Kircher. (2011): "Assortative Matching with Large Firms: Span of Control over More versus Better Workers," Work in Progress, LSE.

[22] Epple, D., and R.E. Romano (2011): "Peer Effects in Education: A Survey of the Theory and Evidence," In *Handbook of Social Economics* 1 (11), Elsevier B.V., 1053-1163.

[23] Epple, D., R.E. Romano, and H. Sieg (2006): "Admission, Tuition, and Financial Aid Policies in the Market for Higher Education," *Econometrica* 74(4): 885-928.

[24] Gabaix, X., and A. Landier (2008): "Why Has CEO Pay Increased So Much?" *Quarterly Journal of Economics* 123(1): 49-100.

[25] Garicano, Luis (2000): "Hierarchies and the Organization of Knowledge in Production," *Journal of Political Economy* 108 (5): 874-904.

[26] Garicano, L., and E. Rossi-Hansberg (2004): "Inequality and the Organization of Knowledge." *American Economic Review* 94(2): 197-202.

[27] Garicano, L., and E. Rossi-Hansberg (2006): "Organization and Inequality in a Knowledge Economy," *Quarterly Journal of Economics* 121(6): 1383-1435.

[28] Gibbons, R., and M. Waldman (1999): "Careers in Organizations: Theory and Evidence," *Handbook of Labor Economics* 3: 2373–2437.

[29] Goleman, D. (1997): *Emotional Intelligence: Why It Can Matter More Than IQ.* New York: Bantam.

[30] Heckman, J.J., and B.E. Honore (1990): "The Empirical Content of the Roy Model," *Econometrica* 58(5): 1121-1149.

[31] Heckman, J.J. and G. Sedlacek (1985): "Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labour Market," *Journal of Political Economy*, 93 (6), 1077-1116.

[32] Heckman, J. J., J. Stixrud and S. Urzua (2006): "The Effects Of Cognitive and Noncognitive Abilities On Labor Market Outcomes and Social Behavior," *Journal of Labor Economics* 24(3), 411-482.

[33] Heckman, J. J., L. Malofeeva, R Pinto and P. Savelyev (2011). "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes," Presentation at the Asian Meeting of the Econometric Society, Seoul, Korea, August 12, 2011.

[34] Hirsch, B.T. (1978): "Earnings, Occupation, and Human Capital Investment," *Atlantic Economic Journal* 6(2): 31-40.

[35] John, O.P., L.P. Naumann, and C.J. Soto (2008): "Paradigm Shift to the Integrative Big Five Trait Taxonomy," *Handbook of Personality: Theory and Research*, 114-158.

[36] Kambourov, G., A. Siow and L. Turner (2012): "Social Skills, Marital Disolution, and Labor Separation," Work in Progress, University of Toronto.

[37] Keane, M.P., and K.I. Wolpin (1997): "The Career Decisions of Young Men," *Journal of Political Economy* 105(3): 473-522.

[38] Kremer, M. (1993): "The O-Ring Theory of Economic Development," *Quarterly Journal of Economics*, 551-575.

[39] Kristof, N. D. (2000): "Earning A's in People Skills at Andover," *New York Times*, June 10.

[40] Kuhn, P, and C. Weinberger (2005): "Leadership Skills and Wages," *Journal of Labor Economics* 23(3): 395-436.

[41] Legros, P., and A.F. Newman (2002): "Monotone Matching in Perfect and Imperfect Worlds." *Review of Economic Studies* 69(4): 925-942.

[42] Li H. and W. Suen (2001): "Technological Change and Labor Market Segregation," Working Paper, University of Toronto.

[43] Li, J. (2012): "A Theory of Wage Distribution and Wage Dynamics with Assignment and Pareto Learning," working Paper, Northwestern University.

[44] Lise, J., C. Meghir, and J.M. Robin (2011): "Matching, Sorting and Wages," Working Paper, Yale University.

[45] Lucas, R. E. (1978): "On the Size Distribution of Business Firms." *Bell Journal of Economics* 9(2): 508-523.

[46] Markon, K.E., R.F. Krueger, and D. Watson (2005): "Delineating the Structure of Normal and Abnormal Personality: An Integrative Hierarchical Approach," *Journal of Personality and Social Psychology* 88(1): 139-157.

[47] McCann R.J., X. Shi, A. Siow, and R. Wolthoff (2012): "Multisector Matching via Linear Programming," Work in Progress, University of Toronto.

[48] McCann R.J., and M. Trokhimtchouk (2010): "Optimal Partition of A Large Labor Force into Working Pairs," *Economic Theory* 42, 375-395.

[49] McCrae, R.R., and P.T. Costa Jr. (1999): "A Five-Factor Theory of Personality," *Handbook of Personality: Theory and Research* 2: 139-153.

[50] Neal, D., and S. Rosen (2000): "Theories of the Distribution of Earnings," *Handbook of Income Distribution* 1: 379-427.

[51] Ortin-Angel, P. and V. Salas-Fumas (2007): "Compensation Dispersion between and within Hierarchical Levels," *Journal of Economics & Management Strategy* 16(1): 53-79.

[52] Pollak, R. A. (2011): "Allocating Time: Individuals' Technologies, Household Technology, Perfect Substitutes, and Specialization," NBER Working Paper No. 17529.

[53] Posner, B.Z. (1981): "Comparing Recruiter, Student, and Faculty Perceptions of Important Applicant and Job Characteristics," *Personnel Psychology* 34(2): 329-339.

[54] Rajan, R.G., and J. Wulf. (2006): "The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies." *Review of Economics and Statistics* 88(4): 759-773.

[55] Rivera, L.A. (2011): "Ivies; Extracurriculars; and Exclusion: Elite Employers' Use of Educational Credentials," *Research in Social Stratification and Mobility* 29(1): 71-90.

[56] Rosen, S. (1978): "Substitution and Division of Labour," *Economica* 45 (179): 235-250.

[57] Rosen, S. (1981). "The Economics of Superstars." *American Economic Review* 71(5): 845–858.

[58] Rosen, S. (1982): "Authority, Control, and the Distribution of Earnings." *Bell Journal of Economics* 13(2): 311-323.

[59] Roy, A. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxord Economic Papers* (New Series), 3, 135-146.

[60] Sacerdote, B. (2011): "Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?" In *Handbook of Economics of Education* 1, 3(3).

[61] Sattinger, M. (1975): "Comparative Advantage and the Distributions of Earnings and Abilities," *Econometrica*: 455-468.

[62] Shimer, R., and L. Smith (2000): "Assortative Matching and Search," *Econometrica* 68(2): 343-369.

[63] Siow, A. (1998): "Differential Fecundity, Markets, and Gender Roles," *Journal of Political Economy* 106(2): 334-354.

[64] Siow, A. (2010): "Testing Becker's Theory of Positive Assortative Matching," Working Paper, University of Toronto.

[65] Smeets, V., and F. Warzynski (2008): "Too Many Theories, Too Few Facts? What the Data Tell Us about the Link between Span of Control, Compensation and Career Dynamics," *Labour Economics* 15(4): 687-703.

[66] Waldman, M. (1984): "Worker Allocation, Hierarchies and the Wage Distribution," *Review of Economic Studies* 51(1): 95-109.

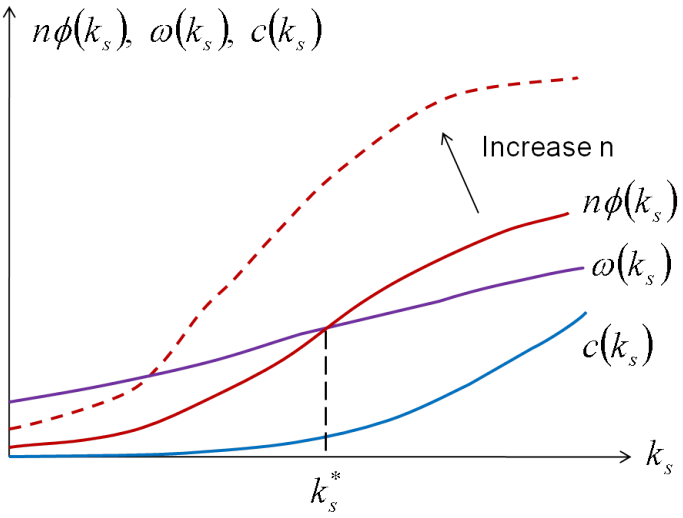
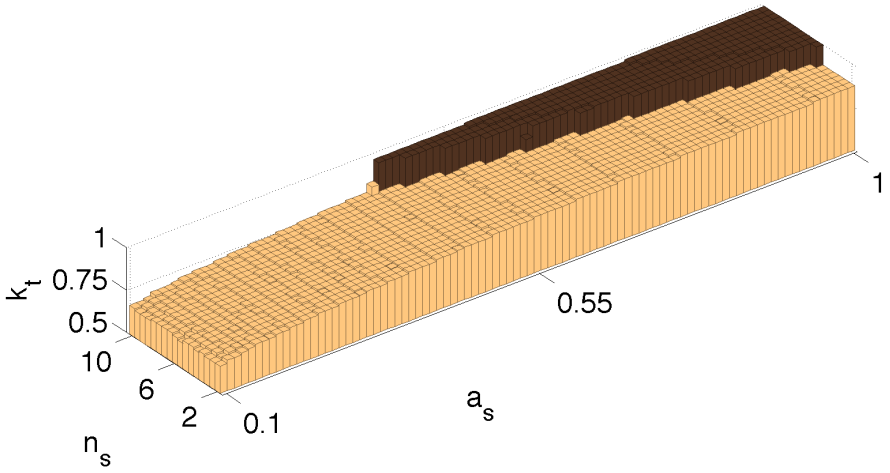
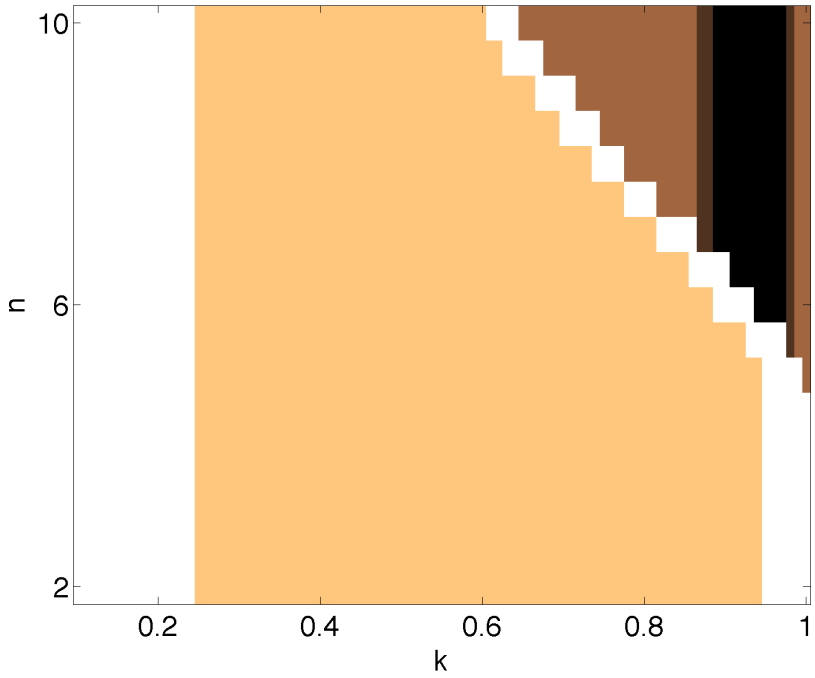


Figure 1: Polarization in the education choice



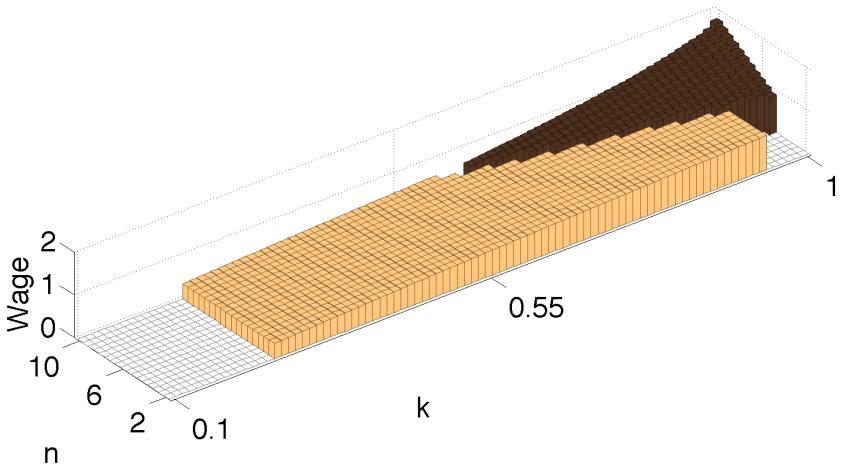
The cognitive skill k of the teacher that is chosen by a student with social skill n and initial cognitive skill a. Yellow students become workers later in life, and brown students become manager / teacher.

Figure 2: Education choice



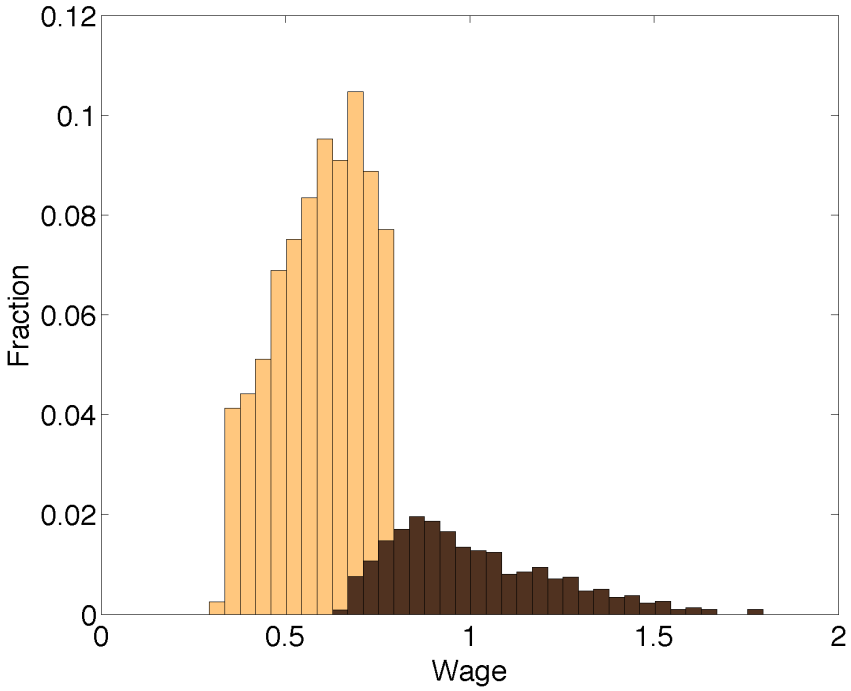
The occupation chosen by an adult with social skill n and cognitive skill k. Yellow adults are workers, light-brown adults are teachers, black adults are managers, and dark-brown adults can be managers or teachers. Note that the lack of smoothness is the result of discretization.

Figure 3: Occupation choice



The wage of an adult with social skill n and cognitive skill k. Yellow adults are workers, and brown adults are managers / teachers.

Figure 4: Labor market wages



Wage density. Yellow adults are workers, and brown adults are managers / teachers.

Figure 5: Labor market wage density