Investment, Nontransferabilities, and Matching Policies

Thomas Gall, Patrick Legros, Andrew Newman

Very Preliminary
This Version: June 24, 2008

Abstract

Mobility depends essentially on investment, which often occurs in environments in which individuals match (school) or will match after investing (the labor market). Where partners can transfer surplus to each other only imperfectly (NTU), the pattern of matching will typically be ine¢cient, providing a possible justi…cation for "associational redistribution": a social planner who could enforce a matching outcome that di¤ers from the market outcome may raise aggregate social surplus. We show that this static ine¢ciency due to NTU can be exacerbated in a dynamic environment in which individuals' productive types are determined by investments made before they match. In contrast to TU models, which always have an e¢cient equilibrium, there will typically be investment distortions, with high types over-investing and low types under-investing; this result occurs despite symmetric information about types. Moreover, if investment itself takes place in a matching environment (e.g. schools), the e¤ects can be further magni…ed. We study the timing of associational redistribution policies, assessing the di¤erential e¤ects of early-stage and later-stage policies as well as complementarities between them. We discuss implications of our results for inequality, intergenerational mobility, underclass phenomena, and possible policy remedies such as school integration and a¢rmative action.

1 Introduction

The allocation on the labor market determines incentives for education acquisition, both in terms of worker assignment and market wages. Distortions of the labor market allocation, for instance due to moral hazard problems both within …rms and between borrowers and lenders, thus have repercussions for individuals' decisions on education investment. On the other hand, qualities of the education investment environment may in turn a¤ect the labor market allocation through the distribution of educational attainment present on the labor market. That is, education acquisition and labor market are intricately linked, so that e¤ects of market frictions and potential remedies on one cannot be analyzed without considering the other.

This paper studies a model in which individuals match into …rms on the labor market based on their productive types that are determined by investments made before they match. Within …rms agents jointly produce output according to a production function that has increasing di¤erences in agents' investments. While individuals' types are publicly observable, …rm members face a moral hazard problem that induces dependence of surplus size and surplus distribution between …rm member, that is utility is non-transferable (NTU). This market friction implies segregation in investment is always the labor market outcome.[1] Since higher investment by …rm members increases potential surplus and thus reduces incentives to shirk, the cost of the second best contract decreases in …rm members' investment levels.

The non-transferability has a twofold e¤ect on investment incentives: an exclusion e¤ect that leads agents with high investment cost to under-invest and leave the industrial sector, and over-investment by agents with low investment costs due the decreasing …xed cost of moral hazard in investment. Investment choice inequality exceeds initial inequality in investment cost types between agents who match into industrial …rms and those who stay solitary. Among agents in the industrial sector investment inequality falls short of initial inequality. The between-sector-inequality in investment choices is exacerbated as the scope for moral hazard grows interpretable as higher complexity of the industrial task. If investment cost and parental investment choice correlate this enables persistent investment inequality in a growth model, possibly o¤ering a dynamic reason for integrative policies.

Over- and under-investment can also occur in a second best sense as pa-

[1] Some evidence for this is provided by Dunne et al. (2004) who …nd that wage dispersion between plants exceeds the one within plants for US manufacturing …rms. Iranzo et al. (2008) on the other hand …nd the contrary for a large data set of Italian manufacturing …rms pointing to mixing into …rms on the labor market. In contrast to the US wages in Italy are generally determined by centralized bargaining in each industry sector.

rameter constellations exist such that aggregate social surplus increases if a social planner could enforce a matching outcome that entails some mixing. This provides a possible justi…cation for "associational redistribution" on the labor market. Such associational redistribution has to be accompanied by adequate tax-subsidy schemes to ensure appropriate investment incentives, however, making it an informationally demanding policy.[2] In contrast to an incomplete information model of statistical discrimination based on self-ful…lling prophecies (as for instance Coate and Loury, 1993, Mailath et al., 2000), associational redistribution on its own does not generally suf…ce to restore e¢ ciency in the present setting since undereducated agents indeed face high education cost. Likewise, pure tax-subsidy schemes that condition on investment or …rm output cannot achieve an increase in aggregate social surplus since under segregation the labor market induces e¢ cient investments conditional on an individual's match.

If investment itself takes place in a matching environment (e.g. schools) policies may also involve associational redistribution at the investment stage.[3] Assuming there exist peer e¤ects inducing a positive linear cost externality we …nd that a move from schools that are segregated with respect to student investment to an integrated schooling system may raise aggregate social surplus even if segregation on the labor market is e¢ cient.[4] That is, there exists a justi…cation for school integration on the grounds of raising aggregate social surplus even when both production in …rms and peer e¤ects at the school stage exhibit increasing di¤erences.[5] A similar point arises when considering the problem of resource allocation among students: we …nd that over-proportional spending on weaker students tends to raise aggregate so-

[2] A policy re‡ecting associational redistribution on the labor market is a¢ rmative action, i.e. positive discrimination based on observable characteristics correlated with education cost. Empirically, the relationship between a¢ rmative action on labor market and output e¢ ciency remains inconclusive (see e.g. the survey by Holzer and Neumark, 2000).

[3] Examples are busing, school desegregation, or relocation such as the moving to opportunity program – neither of which seems to have had much of an impact (see Rivkin, 2000, Katz et al., 2001, Sanbonmatsu et al., 2006)

[4] Note that also teachers may segregate, that is schools in districts populated by poor and minority households have systematically lower teacher quality as reported by Hanushek et al. (2004), Boyd et al. (2005) for the US and Bonesrønning et al. (2005) for Norway. Hanushek et al. (2005) provide some evidence that teacher quality indeed a¤ects students' academic achievement.

[5] This seems to be particularly interesting given the large variation between countries documented in the PISA study (OECD, 2007) in the correlation between children's educational outcomes and their parents' socioeconomic status. This is consistent with variation of early-stage policies across emphasizing or de-emphasizing segregation at school.

cial surplus.[6]

Methodologically we employ a one-sided non-transferable utility matching market with a continuum of agents a la Legros and Newman (2007) to model the labor market. Before entering the labor market agents choose their investment based on individual investment cost type and anticipating the labor market outcome. On the labor market matching is either into coalitions of size one (a basic task giving output linear in investment) or size two (industrial …rms where production has investment complementarities).

This work is connected to a number of theoretical contributions analyzing investments prior to assignment markets. Felli and Roberts (2002), Cole et al. (2001) …nd near e¢ cient or e¢ cient investments always to be in the equilibrium set when utility is perfectly transferable. Peters and Siow (2002) consider investment prior to two-sided matching markets with completely non-transferable utility, …nding the resulting allocation to be Pareto-e¢ cient. In a similar setting Peters (2006) …nds over-investment when the matching pattern is …xed exogenously to positive assortative. Bidner (2007) and Hopkins (2005) also …nd investment distortions in a matching-cum-investment setting. In contrast to the present framework their …ndings rely on asymmetric information, where over-investment is generated by wasteful signalling. Gall et al. (2006) …nd in a setup similar to the present model, but where production has decreasing di¤erences, that optimal timing of education acquisition, that is before the market or within …rms, depends on the degree of market imperfection. More elaborate models on sorting and human capital acquisition are Bénabou (1996) and Epple and Romano (1998) who analyze school and neighborhood choice in the presence of peer e¤ects.

This paper is organized as follows. Section 2 presents the framework and section 3 derives the labor market equilibrium and investment choices when there is moral hazard. In Section 4 we examine suitability of a number of policy instruments with respect to aggregate surplus while section 5 concludes. The more involved proofs can be found in the appendix.

[6] Resource allocation can be interpreted as the choice between centralization and decentralization of school …nance Hoxby (1996). Card and Payne (2002) …nd that after the ongoing decentralized system of school …nance was declared unconstitutional in many states of the US the subsequent more equal spending on schools was re‡ected in a more equal distribution of SAT scores.

2 Model

2.1 Agents

The economy is populated by a continuum of agents I endowed with unit Lebesgue measure. An agent $i \in I$ is characterized by education cost type $\mu_i \in [\underline{\mu}; \overline{\mu}]$. Assume for now that education cost types are given exogenously and their distribution $F(\mu)$ is differentiable and has full support. Agents have the opportunity to invest in education e_i at a cost

$$\frac{e_i^2}{8\mu_i};$$

that depends negatively on own education cost type. Therefore higher types are associated with lower cost. Based on education choice agents enter the labor market and match into firms. An agent's utility is linear in monetary income from wages.

2.2 Production

A pair of agents $(i; j)$ with educational attainments e_i and e_j may form an industrial firm and jointly produce output. Production in firm $(i; j)$ may either succeed yielding output y or fail, in which case output is 0. Let

$$y = \sqrt{e_i e_j}.$$

The probability of success depends on agents' efforts within a firm. An agent may either exert effort or shirk. If both agents exert effort the success probability is 1, otherwise it is 0. Exerting effort causes a disutility of 1. Agent i's utility from choosing e_i, exerting effort in a firm and obtaining wage w_i with certainty is

$$u(e_i) = w_i - 1 - \frac{e_i^2}{8\mu_i}.$$

Effort is not observable. There exists, however, a monitoring technology that detects, with probability q, whether a worker shirks (q is the same for both workers and detection of one is independent of the other's effort). The monitoring device comes at a monetary cost $c^M(Á; q)$ that depends on the quality of monitoring $q \in [0; 1]$ and on a parameter $\phi \geq 0$. Suppose

$$c^M(\phi; q) = \phi q.$$

Given a choice of q and a wage w an agent has an incentive to exert effort if

$$w - 1 \geq (1 - q)w \Leftrightarrow w \geq \frac{1}{q}, \tag{1}$$

since education choice is sunk. The sum of wages in firm $(i; j)$ have to equal output so that

$$w_i + w_j = \sqrt{e_i e_j} - \phi q. \tag{2}$$

If $\phi > 0$ condition (1) must hold with equality, since otherwise there is scope for Pareto improving reorganization within the firm. Therefore q is given by

$$q = \frac{1}{\min\{w_i; w_j\}}. \tag{3}$$

This means the monitoring cost minimizing and output maximizing wage profile splits the surplus, that is $w_i = w_j$. Denote the optimal monitoring intensity when output is shared equally by $q^*(y)$. Using (2) it is given by

$$q^*(y) = \frac{y - \sqrt{y^2 - 8\phi}}{2\phi}. \tag{4}$$

The industrial production technology is interpreted best as relying on coordination efforts that are hard to measure. Human capital input determines the scope of the project, and thus high education translates into high-powered incentives requiring only a low intensity of monitoring.

Additionally, agents have access to a basic technology that does not require teamwork. Equally, the basic task may allow costless monitoring of effort or output because it requires less sophisticated organization of production. Output generated by an agent i employing the basic technology is

$$y^B(e_i) = \frac{e_i}{4}.$$

For simplicity assume effort cost to be zero in the basic task.

2.3 Timing

The timing of events in the economy is as follows.

(1) Nature chooses education cost types.

(2) Agents choose education levels.

(3) Labor market opens, …rms form and choose wages and monitoring intensity.

(4) Agents choose e¤ort levels within …rms.

(5) Monitoring result is realized, wages are paid.

(6) Output is realized.

At the labor market stage there is perfect information on agents' education choices and team composition. Wage contracts can also condition on the outcome of monitoring and monitoring intensity. That is, competition is for education levels in labor contracts. Agents enter the labor market without cash holdings, all payments within …rms have to be funded by the project's output.[7] When acquiring education agents rationally anticipate the labor market outcome.

2.4 Equilibrium Concept

De…ne now formally the equilibrium concept for the model economy. Let $\Phi(I)$ denote the set of all partitions P of the agent apace I into coalitions of size $|P_i| \leq 2$ for all $P_i \in P$. Now it is convenient to introduce the notion of a labor contract. A labor contract $(w_i; w_j)(e_i; e_j)$ speci…es success wages of agents i and j in a …rm $(i; j)$ with education levels e_i and e_j. The set of feasible labor contracts $w^I(e_i; e_j)$ speci…es all wages $(w_i^I; w_j^I)$ in …rm $(i; j)$ such that both incentive compatibility (1) and feasibility (2) hold. Likewise $w^B(e_i) = e_i/4$ de…nes wage of an agent with e_i in a basic …rm.

De…nition 1 *An allocation of education investments* $e^* = (e_i^*)_{i \in I}$, *a measure consistent partition* $P^* \in \Phi(I)$, *and wages* $w^* = (w_i^*)_{i \in I}$ *is called an equilibrium if*

[7] This assumption ensures that segregation is the labor market outcome. Dropping it would not change the bene…cial role of associational redistribution at the schooling stage.

(i) *for all* $i \in I$ *either* $w_i^* = w^B(e_i)$ *if* $\{i\} = P_i^*$ *or* $(w_i^*, w_j^*) \in w^I(e_i^*, e_j^*)$ *if* $\{i, j\} = P_i^*$.

(ii) $w_i^* \geq w^B(e_i^*)$ *for all* $i \in I$ *and there does not exist* $w' \in w^I(e_i^*, e_j^*) \neq (w_i^*, w_j^*)$ *such that* $u_i(e_i^*, w_i') > u_i(e_i^*, w_i^*)$ *and* $u_i(e_j^*, w_j') > u_i(e_j^*, w_j^*)$.

(iii) *for all* $i \in I$ *there are no* $w' \in w^I(e_i', e_j^*)$, $e_i' \neq e_i^*$, $j \neq i \in I$, *such that* $u_i(e_i', w_i') > u_i(e_i^*, w_i^*)$ *and* $u_j(e_j^*, w_j') > u_j(e_j^*, w_j^*)$ *and there is no* $e_i' \neq e_i^*$ *such that* $u_i(e_i', w_i^B(e_i')) > u_i(e_i^*, w_i^*)$.

(iv) *for all* $i \in I$ *there is no* $w' \in w^I(e_i', e_j')$, $e_i' \neq e_i^*, e_j' \neq e_j^*$, *such that* $u_i(e_i', w_i') > u_i(e_i^*, w_i^*)$ *and* $u_i(e_j', w_j') > u_i(e_j^*, w_j^*)$.

This postulates that in equilibrium payoffs are feasible, the labor market outcome is a stable match with respect to the chosen education investments, that education investments are individually optimal given the labor market outcome, and finally that a pairwise deviation in education investments is not profitable for both agents. The latter ensures that the equilibrium investments are unique, since it is well known that in models with ex-ante investments, coordination failure may lead to multiple equilibria in which only individual deviations are allowed (e.g. Cole et al., 2001, Mailath et al., 2004, Nosaka, 2007).[8]

2.5 Benchmark $\phi = 0$

In a frictionless world monitoring is costless, that is $\phi = 0$. In this case $q = 1$ and joint surplus in a firm (i, j) is given by

$$\pi(e_i, e_j) = \sqrt{e_i e_j} - 2 - \frac{e_i^2}{8\theta_i} - \frac{e_j^2}{8\theta_j}.$$

Optimal education choices in a firm (i, j) satisfy $e_i = 2\theta_i^{3/4}\theta_j^{1/4}$ and $e_j = 2\theta_j^{3/4}\theta_i^{1/4}$. That is, joint surplus depends on education cost types θ_i and θ_j:

$$\pi(\theta_i, \theta_j) = \sqrt{\theta_i \theta_j} - 2.$$

[8] Alternatively, one can interpret our analysis as focusing only on the "best" equilibria. As we shall show, there is still a potential beneficial role for policy, unlike in the corresponding transferable utility case.

Since $\sqrt{\mu_i\mu_j}$ has increasing differences in its arguments segregation maximizes aggregate surplus. The industrial technology yields higher surplus than the basic task if $\mu_i - 2 > \mu_i/4$. Hence, in the benchmark allocation agents with $\mu_i \leq 8/3 =: \hat{\mu}^0$ use the basic technology, and agents with $\mu_i > \hat{\mu}^0$ match into homogenous industrial firms.[9] Optimal education choice is

$$e^0(\mu_i) = \begin{cases} \mu_i & \text{if } \mu_i \leq \hat{\mu}^0 \\ 2\mu_i & \text{if } \mu_i > \hat{\mu}^0 \end{cases}$$

The decentralized labor market reaches the benchmark allocation since when $\phi = 0$ utility can be transferred perfectly. Indeed, once education choices are sunk, the labor market is segregated because the total surplus is supermodular in educational attainments. Anticipating segregation, an agent with educatin cost μ maximizes $e/2 - 1 - e^2/(8\mu)$ and therefore chooses $e(\mu) = 2\mu$ obtaining a utility of $\mu/2 - 1$ in firms. Alternatively, the agent could invest in order to choose the basic task and in this case invest e to maximize $e/4 - e^2/(8\mu)$, or $e = \mu$ giving utility of $\mu/8$. An agent's optimal decision in the decentralized economy is therefore first-best.

3 The Economy with Moral Hazard

Suppose for the rest of this paper that the economy suffers from frictions, that is $\phi > 0$. To solve for the equilibrium allocation we proceed backwards in time and start with the labor market matching based educational attainment.

3.1 Labor Market Matching

On the labor market agents match into firms based on their education choices e_i, or are employed in the basic task. Since higher education implies less need for costly monitoring one might expect possible gains from trade between high and low education types that lead to heterogeneous firms. The following proposition asserts that this is not the case.

[9] Note that the cutoff $\hat{\mu}$ depends on the matching pattern. Increasing differences implies, however, that any matching pattern other than segregation in education cost type can be reorganized into segregation yielding a strict increase in aggregate surplus.

Proposition 1 *Agents in the industrial sector segregate in educational attainment for all $\phi \geq 0$.*

Proof: Suppose agents i and j with educational attainments $e_i \neq e_j$ match in equilibrium and share the output according to wages w_i and w_j. Let $e_i > e_j$ without loss of generality. Wages w_i and w_j must be incentive compatible and individually rational. Denote the cost minimizing detection probability at wages w_i and w_j by $q(w_i, w_j)$ given by equation (3).

Both i and j must find their wages profitable compared to their segregation payoffs, that is

$$w_i \geq \frac{\sqrt{e_i e_i} - \phi q^*(e_i, e_i)}{2} \text{ and } w_j \geq \frac{\sqrt{e_j e_j} - \phi q^*(e_j, e_j)}{2}. \tag{5}$$

Since the lower wage determines q we have that $q^*(e_i, e_i) < q(w_i, w_j) \leq q^*(e_j, e_j)$. Note that $\sqrt{e_j e_j} - \phi q(w_i, w_j) \leq \frac{2}{q(w_i, w_j)}$ since $q^*(e_j, e_j)$ is defined as the minimal incentive compatible detection probability in a (e_j, e_j) firm. Incentive compatibility in the (i, j) firm requires $w_j \geq 1/q(w_i, w_j)$, so that

$$w_j \geq \frac{\sqrt{e_j e_j} - \phi q(w_i, w_j)}{2}. \tag{6}$$

Aggregate payoffs in a (i, j) firm must be feasible, $w_i + w_j = \sqrt{e_i e_j} - \phi q(w_i, w_j)$. Using the first part of (5) we have that

$$\sqrt{e_i e_j} - \phi q(w_i, w_j)) - w_j \geq \frac{\sqrt{e_i e_i} - \phi q^*(e_i, e_i)}{2}.$$

Combined with (6) this yields

$$0 > \sqrt{e_i e_j} - \frac{\sqrt{e_i e_i} + \sqrt{e_j e_j}}{2} \geq \frac{\phi q(w_i, w_j) - \phi q^*(e_i, e_i)}{2} > 0,$$

a contradiction. Hence, matches of agents that are heterogeneous in educational attainments cannot occur in equilibrium. $\square$

Intuitively, although gains from trade may exist they cannot be realized due to the non-transferability. Agents choosing low education would have to compensate agents with high education. This requires low profit shares for the low education agents in mixed firms, which offsets the efficiency gains from possibly lower monitoring intensity. In some cases, incentive compatibility of the low type cannot be satisfied while paying the high type his opportunity cost.

3.2 Education Choice

When choosing educational attainment agents correctly anticipate the labor market outcome. Since there is segregation for industrial firms output is shared equally and monitoring intensity is $q^*(y)$ with $y = e$ where e is the education level common to both agents. In a rational expectation equilibrium an agent i's expected payoff from education choice e is

$$u_i(e) = \frac{e + \sqrt{e^2 - 8\chi}}{4} - 1 - \frac{e^2}{8\theta_i};$$

if i expects to be matched into the industrial sector. If the agent anticipates participating in the basic activity, his expected utility is instead $\theta_i/8$. The necessary first order condition is

$$\frac{1}{e} + \frac{1}{\sqrt{e^2 - 8\chi}} = \frac{1}{\theta_i}. \tag{7}$$

Checking the second derivative reveals that the necessary condition is sufficient as well and (7) implicitly defines the optimal education choice $e(\theta_i)$ in an industrial firm. Note that

$$\frac{1}{e} + \frac{1}{\sqrt{e^2 - 8\chi}} > \frac{2}{e}; \tag{8}$$

and therefore $e(\theta_i) > 2\theta_i = e^0(\theta_i)$.

Lemma 1 *There exists a unique $\hat{\theta}^{\chi}$ such that agents with $\theta_i \leq \hat{\theta}^{\chi}$ are employed in the basic task, and agents with $\theta_i > \hat{\theta}^{\chi}$ match into industrial firms.*

This means when $\chi > 0$ education choice is given by

$$e^{\chi}(\theta_i) = \begin{cases} \theta_i & \text{if } \theta_i \leq \hat{\theta}^{\chi} \\ e(\theta_i) & \text{if } \theta_i > \hat{\theta}^{\chi}; : e(\theta_i) \text{ defined by (7).} \end{cases} \tag{9}$$

The following proposition compares the laissez faire market allocation with the benchmark allocation.

Proposition 2 *Properties of the market equilibrium allocation:*

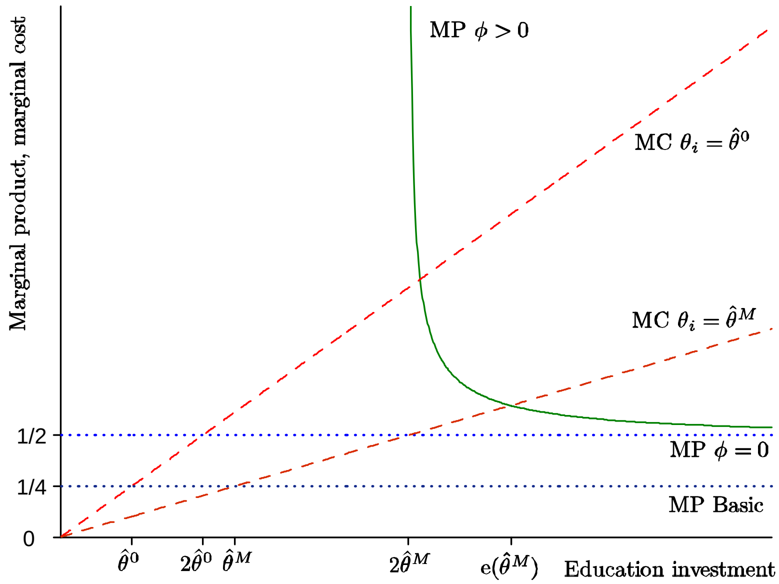

Figure 1: Investment Incentives

(i) *Participation in the industrial sector is lower in the market than in the benchmark allocation,* $\hat{µ}^0 < \hat{µ}^Á$*.*

(ii) *Agents with* $µ_i$ 2 $[\hat{µ}^0; \hat{µ}^Á]$ *under-invest in education compared to the benchmark allocation,* $e^Á(µ_i) < e^0(µ_i)$*, while agents with* $µ_i > \hat{µ}^Á$ *over-invest,* $e^Á(µ_i) > e^0(µ_i)$*.*

(iii) *Absolute over-investment* $e^Á(µ_i)$ ¡ $e^0(µ_i)$ *decreases in* $µ_i$ *for* $µ_i > \hat{µ}^Á$*.*

Figure 1 captures the intuition of our results. It depicts individual marginal cost and marginal product of education. The dashed lines represent marginal cost $\frac{e_i}{4µ_i}$ of education cost types $\hat{µ}^0$ and $\hat{µ}^Á$. The dotted lines show marginal product of education in the industrial sector when $Á = 0$ (upper line) and in the basic sector (lower line). The solid line gives the marginal product of education on the labor market when $Á > 0$. At $e =$ p$\overline{8Á}$ it tends to in…nity and as e grows out of bounds the marginal product approaches its benchmark value 1=2.

The market friction induced by the moral hazard problem in …rms a¤ects the marginal product of education in industrial production. For very low education choices monitoring cost makes industrial production unpro…table thus decreasing incentives to acquire education. For intermediate education choices acquiring more education lowers monitoring cost thus boosting education investments. As output gets large monitoring becomes less important and education incentives given by the marginal product of education approach the …rst best benchmark.

3.3 Positive Analysis

First we are interested in whether the labor market equilibrium ampli…es type inequality. That is, is the distribution of education choices more unequal than the education cost type distribution? This may have a large range of implications for a dynamic economy where this generation's educational attainments is at least instrumental in determining their o¤spring's cost of education (see e.g. the survey by Heckman, 2008, who argues that evidence points to positive correlation of parental attainment and their o¤spring's educational cost type). Indeed we …nd an ampli…cation e¤ect of the labor market.

Proposition 3 (Inequality) *There is an inter-sectoral education gap as agents with* $\mu_i \in [\underline{\mu}; \hat{\mu}^{Á}]$ *choose* $e^{Á}(\mu_i) \leq \hat{\mu}^{Á}$ *and agents with* $\mu_j \in [\hat{\mu}; \overline{\mu}]$ *choose* $e^{Á}(\mu_j) > 2\hat{\mu}^{Á} > e^{Á}(\mu_i)$.
Within the industrial sector inequality is decreasing: $\forall \mu_j; \mu_h \in [\hat{\mu}; \overline{\mu}]$

$$\mu_h > \mu_j \Rightarrow \frac{\mu_h}{\mu_j} > \frac{e^{Á}(\mu_h)}{e^{Á}(\mu_j)}.$$

Proof: The …rst part is immediate from (9). To show the second part note that for $\mu_j \geq \hat{\mu}^{Á}$ by the …rst order condition (7)

$$\frac{e^{Á}(\mu_j)}{\mu_j} = 1 + \frac{e^{Á}(\mu_j)}{\sqrt{e^{Á}(\mu_j)^2 - 8Á}} > 1 + \frac{e^{Á}(\mu_h)}{\sqrt{e^{Á}(\mu_h)^2 - 8Á}} = \frac{e^{Á}(\mu_h)}{\mu_h}.$$

since $\mu_h > \mu_j$. ¤

That is, inequality across sectors is ampli…ed, but within the industrial sector inequality decreases as agents with lower education cost type catch up. Trivially, within the basic sector inequality remains constant as $e^{Á}(\mu_i) = \mu_i$ for all $\mu_i < \hat{\mu}^{M}$. This may point to a dynamic reason for associational redistribution using negative matching at the schooling stage.

Next we characterize types of industry we would expect to be most heavily affected by the moral hazard problem.

Proposition 4 (Technology) *Suppose monitoring becomes more costly, that is ϕ increases to $\phi' > \phi$. Then $e^{\phi}(\theta) < e^{\phi'}(\theta)$ for $\theta \geq \hat{\theta}^{\phi'}$ and $e^{\phi}(\theta) \geq e^{\phi'}(\theta)$ for $\theta < \hat{\theta}^{\phi'}$, that is given ϕ' investments increase for agents who match into industrial firms and decrease for agents choosing the basic task. Participation in the industrial sector decreases, $\hat{\theta}^{\phi} < \hat{\theta}^{\phi'}$.*

Proof: Note first that $e^{\phi}(\theta_i) = \theta_i$ using the basic technology for $\theta_i \leq \hat{\theta}^{\phi}$ independently of ϕ. Consider now an agent i with education cost $\theta_i > \hat{\theta}^{\phi}(\phi)$, where we make explicit in this proof the dependence of the education level on ϕ. Since the function $(e + \sqrt{e^2 - 8\phi})/2 - 1 - e^2/(8\theta)$ is supermodular in $e; \phi$, Topkis' theorem implies that in an equilibrium allocation, the educational choice $e^{\phi}(\theta_i; \phi)$ is increasing in ϕ. The cutoff value $\hat{\theta}^{\phi}(\phi)$ is defined by

$$\frac{e^{\phi}(\theta;\phi) + \sqrt{e^{\phi}(\theta;\phi)^2 - 8\phi}}{4} - \frac{e^{\phi}(\theta;\phi)^2}{8\theta} = 1 + \frac{\theta}{8};$$

By the envelope theorem, the left hand side is decreasing in ϕ. It follows that $\hat{\theta}^{\phi}(\phi)$ is increasing in ϕ. This proves the statement since $e^{\phi}(\theta_i)(\phi) = \theta_i < e(\theta_i)$ for $\theta_i < \hat{\theta}^{\phi}$ and thus an increase in ϕ to ϕ' implies a strict decrease in $e^{\phi}(\theta_i)$ for all $i \in I$ with $\hat{\theta}^{\phi} < \theta_i < \hat{\theta}^{\phi'}$. ¤

As the cost of monitoring increases, avoiding monitoring by increasing education becomes more profitable and thus returns to education in industrial firms increase. Since the cost of monitoring increases nevertheless, previously indifferent individuals must now prefer the basic task and therefore participation in the industrial sector decreases. That is, technologies that require inputs that are hard to measure (e.g. human capital intensive technologies that mainly require coordination effort) will suffer most from the contracting friction. If ϕ can be affected by policy such as contracting law, for instance, this may provide an effective instrument to improve welfare.

3.4 Efficiency

Let us now determine the allocation that maximizes aggregate output when $\phi > 0$. Note first that within firms the cost of monitoring is minimized if surplus is split equally and $q = q^*(y)$. Since education investments are endogenous, start by deriving efficient investments conditional on abilities μ_i and μ_j of matched agents i and j who split the surplus. The optimization problem in the firm is

$$\max_{e_i,e_j} \frac{\sqrt{e_ie_j} + \sqrt{e_ie_j - 8\chi}}{2} - 2 - \frac{e_i^2}{8\mu_i} - \frac{e_j^2}{8\mu_j}. \tag{10}$$

First order conditions yield

$$e_j^2 = \frac{\mu_j}{\mu_i} e_i^2. \tag{11}$$

Defining the cost ratio as $\kappa := \sqrt{\mu_j/\mu_i}$, we have that

$$1 + \frac{\sqrt{\kappa e_i^2}}{\sqrt{\kappa e_i^2 - 8\chi}} = \frac{e_i}{\mu_i \sqrt{\kappa}}. \tag{12}$$

Denote the solution to this equation by $e^*(\mu_i; \mu_j)$.

Proposition 5 *(i) The labor market induces efficient investments conditional on the match obtained on the labor market, $e^{\chi}(\mu_i) = e^*(\mu_i; \mu_i)$.*

(ii) There exists μ^C such that if $\mu_i \leq \mu^C$ negative assortative matching of agents with $\mu_i > \mu^M$ and education $e^*(\mu_i)$, into firms maximizes output.

(iii) Let $\mu_i < \mu^C$ for all $i \in I$ and define μ^{med} by $F(\mu^{med}) = \frac{1}{2}(1 + F(\hat{\mu}^{\chi}))$. Then agents with $\mu_i > \mu^{med}$ over-invest and agents with $\hat{\mu}^{\chi} < \mu_i < \mu^{med}$ under-invest on the labor market.

Proposition 5 states that the labor market is doing quite well at setting incentives conditional on segregation into homogenous firms. This matching pattern might not maximize aggregate output, however. If the education cost type space is sufficiently compressed monitoring cost savings in heterogeneous firms are sufficiently large and negative assortative matching in the industrial sector maximizes aggregate output.

4 Policy

This section examines the effects of potential labor market policies that aim to improve on the laissez faire outcome.

4.1 Tax-subsidy Schemes on the Labor Market

First we note that relying on tax schemes alone cannot improve output, since the labor market induces efficient investments conditional on the matching. Define a tax-subsidy scheme by monetary transfers to agents or firms that may condition on educational achievement $t^e(e_i)$ or on the output in firm $t^y(y)$. Since laissez faire education choices were efficient conditional on the match the tax-subsidy scheme may be expected to achieve relatively little in this setting. This intuition is confirmed by the following proposition.

Proposition 6 *Tax-subsidy schemes that condition only on output or education cannot increase aggregate surplus.*

To prove this proposition we use a simple revealed preference argument. Absent externalities across agent types all that any tax-subsidy scheme can achieve is to induce education choices that have not been chosen by the agents under laissez faire. It cannot, however, change the matching based on agents' types μ_i. Since education choices on the market were constrained efficient based on types μ_i the conclusion follows.[10]

4.2 Education Cost Formation

A redistribution of education cost types while pursuing a laissez faire policy on the labor market could, however, increase aggregate output. The reason for this is that the marginal effect of education cost on surplus depends on education cost type as stated in the following lemma.

Lemma 2 *In a labor market equilibrium it holds for all $\mu_i > \hat{\mu}^{\phi}$ that*

$$\frac{\partial u_i(\mu_i)}{\partial \mu_i} = \frac{e^{\phi}(\mu_i)^2}{8\mu_i^2} > \frac{1}{2} \text{ and } \frac{\partial^2 u_i(\mu_i)}{\partial \mu_i^2} < 0.$$

Proof: The first part follows from the envelope theorem and the fact that $e^{\phi}(\mu_i) > 2\mu_i$ by Proposition 2 for $\mu_i > \hat{\mu}^{\phi}$. Let $e(\mu_i)$ denote the solution to individual i's first order condition (7) in the industrial sector. The second

[10] This reasoning depends on the absence of matching externalities across types. In the present setting this is ensured since there is always segregation in the labor market equilibrium. In the appendix we provide an example showing that if a tax scheme may alter the equilibrium matching pattern it also may increase surplus.

derivative of individual i's utility is given by

$$\frac{\partial^2 u_i(\theta_i)}{\partial \theta_i^2} = \frac{e(\theta_i)}{2\theta_i^2}\left(\frac{\partial e(\theta_i)}{\partial \theta_i} - \frac{e(\theta_i)}{\theta_i}\right).$$

By the implicit function theorem

$$\frac{\partial e(\theta_i)}{\partial \theta_i} = \frac{e(\theta_i)^2}{\theta_i^2}\left(1 + \frac{e(\theta_i)^3}{(e(\theta_i)^2 - 8\chi)^{3/2}}\right)^{-1}.$$

By the first order condition (7)

$$\frac{e(\theta_i)}{\theta_i} = 1 + \frac{e(\theta_i)}{\sqrt{e(\theta_i)^2 - 8\chi}}.$$

Since indeed

$$1 + \frac{e(\theta_i)}{\sqrt{e(\theta_i)^2 - 8\chi}} < 1 + \frac{e(\theta_i)^3}{(e(\theta_i)^2 - 8\chi)^{3/2}},$$

as $e(\theta_i) > \sqrt{e(\theta_i)^2 - 8\chi}$ the second statement in the lemma follows. ∎

The essence of Lemma 2 is depicted in Figure 2. Individual surplus in the basice sector is linear in θ while individual surplus in a an industrial firm, which equals average surplus in a firm due to segregation, is concave in θ. Note that for $\theta_i < \hat{\theta}^{\chi}$ the basic technology and otherwise for $\theta_i > \hat{\theta}^{M}$. Surplus from the optimal education and technology choice $u(e^{\chi}(\theta_i))$ is given by the upper contour of the two functions.

Note that marginal surplus has a discontinuity at $\hat{\theta}^{\chi}$ and the optimization problem of the social planner may not be well-defined. Therefore, to develop our main argument, consider first a change in education cost type θ by a discrete amount Δ. The associated change in aggregate surplus is

$$u_i(e^{\chi}(\theta_i + \Delta)) - u_i(e^{\chi}(\theta_i + \Delta)),$$

which is if $\theta_i < \hat{\theta}^{\chi} < \theta_i + \Delta$

$$\frac{\hat{\theta}^{\chi} - \theta_i}{8} + u_i(e^{\chi}(\hat{\theta}^{\chi} + \Delta')) - u_i(e^{\chi}(\hat{\theta}^{\chi}));$$

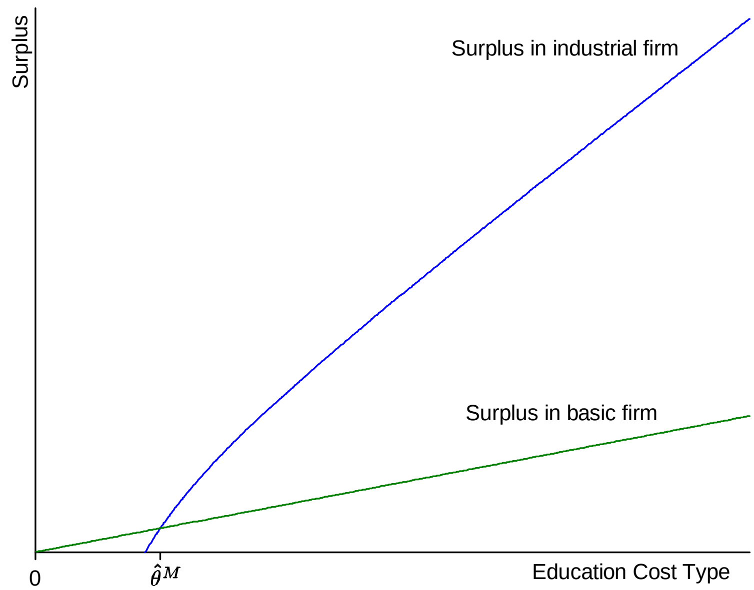

Figure 2: Individual surplus depending on own education cost θ_i.

with $\Delta' = \theta_i + \Delta - \hat{\theta}^{\chi}$. Suppose that $\bar{\theta} > \hat{\theta}^{\chi} > \underline{\theta}$. Note that Lemma 2 implies there exists $\theta_i < \hat{\theta}^{\chi}$ and $\Delta > \hat{\theta}^{\chi} - \theta_i$ such that

$$u(e^{\chi}(\theta_i + \Delta)) + u(e^{\chi}(\bar{\theta} - \Delta)) > u(e^{\chi}(\theta_i)) + u(e^{\chi}(\bar{\theta}));$$

That is, aggregate surplus increases if education cost types can be redistributed from the highest type to a type sufficiently low as to prefer working in the basic sector. In the following we propose two examples of education formation that entail the possibility that a social planner may influence the education cost type distribution. Suppose for this that agents are characterized by genetic types $\alpha_i \in [\underline{\alpha}, \bar{\alpha}]$ and that agent i's education cost $\theta_i = f(\alpha_i, x)$ is a function of own type α_i and a policy variable x. Suppose α is distributed according to a differentiable distribution function $G(\alpha)$ that has full support. This is consistent with the assumption on $F(\theta)$.

4.2.1 Resource Allocation

Suppose first that education cost type θ_i is determined by

$$f(\alpha_i, x_i) = \alpha_i + \sqrt{x_i}.$$

x_i is best interpreted as the spending of resources on agent i which has a decreasing marginal effect on i's education cost type, for instance spending per school. Suppose that spending resources is distortion-free in the sense that it is associated with constant marginal social cost γ. Since agents segregate in education achievement on the labor market there are matching externalities and it suffices to consider the point-wise optimization problem.

$$\max_{x_i} u(e^{\phi}(\alpha_i + \sqrt{x_i})) - \gamma x_i. \tag{13}$$

Solving the optimization problem for an optimal policy x_i^* it turns out that the optimal spending of resources should favor agents that have intermediate to low types α_i.

Proposition 7 *There exists $\bar{\alpha} < \hat{\theta}^{\phi}$ such that (i) x_i^* is strictly decreasing in α_i for all $\alpha_i > \bar{\alpha}$ and (ii) $\alpha_i + \sqrt{x_i^*} = \hat{\theta}^{\phi} + \sqrt{\hat{x}}$, $\sqrt{\hat{x}} > 0$, for all $\alpha_i \in [\bar{\alpha}, \hat{\theta}^{\phi}]$.*

Proof: Due to the discontinuity at $\hat{\theta}^{\phi}$ we need to distinguish between several cases. Suppose first $\alpha_i \geq \hat{\theta}^{\phi}$. Then x_i^* defined by the first order condition

$$\frac{\partial u(e^{\phi}(\alpha_i + \sqrt{x_i^*}))}{\partial \theta_i} = 2\gamma\sqrt{x_i^*},$$

solves problem (13) for α_i. Define $\hat{x}$ as the solution x_i^* associated to $\alpha_i = \hat{\theta}^{\phi}$.

Turn now to $\alpha_i < \hat{\theta}^{\phi}$, with $e^{\phi}(\theta_i) = \theta_i$ and $u(\theta_i) = \frac{\theta_i}{8}$. Suppose first that $\sqrt{x_i} \leq \hat{\theta}^{\phi} - \alpha_i$ and denote $\Delta = \alpha_i + \sqrt{x_i} - \hat{\theta}^{\phi}$. Then optimization reduces to choosing Δ. Using the above result optimally $\Delta = \sqrt{\hat{x}}$ and $\sqrt{x_i^*} = \hat{\theta}^{\phi} + \sqrt{\hat{x}} - \alpha_i$. If $\sqrt{x_i^*} < \hat{\theta}^{\phi} - \alpha_i$ on the other hand, optimal investment x_i^* must satisfy the first order condition

$$\sqrt{x_i^*} = \frac{1}{16\gamma}.$$

The regime $\sqrt{x_i^*} + \alpha_i \geq \hat{\theta}^{\phi}$ obtains if

$$u(e^{\phi}(\hat{\theta}^{\phi} + \sqrt{\hat{x}})) - \gamma(\hat{\theta}^{\phi} + \sqrt{\hat{x}} - \alpha_i)^2 - \frac{\alpha_i}{8} \geq \frac{2}{16^2\gamma} - \gamma\frac{1}{16^2\gamma^2}.$$

Since $\hat{\theta}^{\phi} + \sqrt{\hat{x}} - \alpha_i > \frac{1}{16\gamma}$ the left hand side of the above inequality strictly increases in α_i. This implies immediately the existence of $\bar{\alpha}$ such that optimal resource spending is $x_i^* = (\hat{\theta}^* + \sqrt{\bar{\alpha}} - \alpha_i)^2$ for $\bar{\alpha} \le \alpha_i \le \hat{\theta}^*$ and $x_i^* = (\frac{1}{16\gamma})^2$ for $\alpha_i < \bar{\alpha}$. ∎

Note that depending on social cost γ and the distribution of α there may be types $\underline{\alpha} \le \alpha_i < \bar{\alpha}$ such that optimal resource spending is lowest for these types. This means acquisition of sufficient education for the industrial sector is prohibitively costly.

4.2.2 School Choice

Education cost θ_i may also be affected by an agent's environment when acquring education, for instance by peer effects at school. Suppose that education is acquired in schools of size 2. Education cost types are a function of own type and the match's type, $\theta_i = f(\alpha_i, \alpha_j)$ if i and j match into the same school. Let $f(\alpha_i, \alpha_j)$ increase in both its arguments. Assume that agents do not have access to funds and cannot contract on future income at the schooling stage. That is, the matching market for schools has strictly non-transferable utility which implies segregation in α since an agent's equilibrium payoff increases in $f(\alpha_i, \alpha_j)$ which in turn increases in α_j.

To provide a benchmark assume that the peer effect is linear:

$$f(\alpha_i, \alpha_j) = \lambda\alpha_i + (1 - \lambda)\alpha_j,$$

where $\lambda \in (0, 1)$ is a positive constant, so that $\theta_i = \alpha_i$ under laissez faire. A linear peer effect implies a convenient independence of efficient education cost formation and the matching pattern in schools. Consider now the impact of associational redistribution, that is implementing a matching function $\mu(\alpha)$. We have the following result.

Proposition 8 *Let $\bar{\alpha} > \hat{\theta}^* > \underline{\alpha}$. Then there exists a matching function $\mu(\alpha)$ such that $\mu(\bar{\alpha}) = \alpha_0$ with $\alpha_0 < \hat{\theta}^*$ and $\mu(\alpha)$ is decreasing in its argument for all $\alpha \ge \alpha_0$, so that*

$$\int_{i \in I} u(e^*(f(\alpha_i, \mu(\alpha_i))))di > \int_{i \in I} u(e^*(f(\alpha_i, \alpha_i)))di.$$

Proof: Let $\alpha_i, \alpha_j \in [\hat{\theta}^*, \bar{\alpha}]$ and suppose that $\alpha_i \neq \alpha_j$. Then $f(\theta_i, \theta_i) + f(\theta_j, \theta_j) = 2f(\theta_i, \theta_j)$ since $f(.)$ is linear. Moreover, using Lemma 2 this implies that

$$2u(e^{\varphi}(f(\alpha_i,\alpha_j))) > u(e^{\varphi}(f(\alpha_i,\alpha_i))) + u(e^{\varphi}(f(\alpha_j,\alpha_j))). \qquad (14)$$

Let $\varepsilon > 0$ and choose $\alpha_i = \hat{\theta}^{\varphi} - \varepsilon$. Then we can define δ implicitly by

$$u(e^{\varphi}(\hat{\theta}^{\varphi} - \varepsilon)) + u(e^{\varphi}(\bar{\alpha} - \delta)) = 2u\left(e^{\varphi}\left((1-a)(\bar{\alpha} - \delta) + a\hat{\theta}^{\varphi} - a\varepsilon\right)\right).$$

δ is indeed well-defined and exists since a single crossing property holds in that above expression strictly increases in δ by the implicit function theorem and Lemma 2. Moreover, $\delta > 0$ for $\varepsilon > 0$ sufficiently close to 0 by (14). Moreover, invoking the implicit function theorem again and using Lemma 2 yields

$$\frac{\partial \delta}{\partial \varepsilon} < 0. \qquad (15)$$

Combining these facts means that there always exists $\varepsilon > 0$ and $\delta(\varepsilon) > 0$ such that letting any $\alpha_i \in [\hat{\theta}^{\varphi} - \varepsilon, \hat{\theta}^{\varphi}]$ match with any $\alpha_j \in [\bar{\alpha} - \delta(\varepsilon), \bar{\alpha}]$ improves aggregate surplus:

$$u(e^{\varphi}(a\alpha_i + (1-a)\alpha_j)) > au(e^{\varphi}(\alpha_i)) + (1-a)u(e^{\varphi}(\alpha_j)).$$

This must also be true for all $\alpha_h \in [\hat{\theta}^{\varphi}, \bar{\theta} - \delta(\varepsilon)]$ as argued above. To construct a negative assortative matching that increases point-wise surplus we need that

$$G(\hat{\theta}^{\varphi}) - G(\hat{\theta}^{\varphi} - \varepsilon) = G(\bar{\alpha}) - G(\bar{\alpha} - \delta(\varepsilon)).$$

(15) and full support of $G(\alpha)$ ensure that there exists a tuple $(\varepsilon, \delta(\varepsilon))$, $\varepsilon, \delta > 0$ to satisfy above equality. Let $\alpha_0 = \hat{\theta}^{\varphi} - \varepsilon$. Since a negative assortative matching of all agents in $[\alpha_0, \bar{\alpha}]$ achieves a point-wise increase of surplus, aggregate surplus must increase as well since the labor segregates in education e_i. Since $f(.)$ is linear any matching of agents $\alpha \in [\underline{\alpha}, \alpha_0]$ achieves the same surplus. This proves the statement. $\square$

Intuitively, we show that for each ε neighborhood to the left of $\alpha = \hat{\theta}^{\varphi}$ there exists a δ neighborhood to the left of $\bar{\alpha}$ so that any matching between these two neighborhoods increases surplus compared to school segregation. Since δ decreases in ε we choose ε small enough such that negative assortative matching between all agents with $\alpha_i > \alpha_0 = \hat{\theta}^{\varphi} - \varepsilon$ strictly increases aggregate surplus. This means, of course, that switching from school segregation to school integration induces some agents to switch from the basic sector to the industrial sector. The following corollary provides a condition such that aggregate surplus increases if the entire economy integrates at school level and ends up in the industrial sector.

Corollary 1 *Let $\underline{\theta} < \hat{\theta}^M$. Negative assortative matching at the schooling stage for the entire economy increases aggregate surplus if $\underline{\theta}$ is sufficiently close to $\hat{\theta}^\phi$ or the measure of agents in the basic sector, $G(\hat{\theta}^\phi)$, is sufficiently small.*

The first condition is straightforward and the second one follows from the fact that heterogenous matching among agents with $\alpha_i > \hat{\theta}^\phi$ strictly increases surplus.

4.3 Associational Redistribution on the Labor Market

Finally, turn back to the labor market and treat the matching pattern as a policy variable. We propose a policy that reaches a second best allocation when segregation is not the output-maximizing matching pattern. Let $m(e_i)$ denote a matching function that assigns an agent with educational attainment e_i to an agent with $m(e_i)$. If $\overline{\theta}$ is sufficiently small, negative assortative maximizes output by Proposition 5. By part (i) of Proposition 5 we also know that laissez faire matching provides constrained efficient investment incentives through competition and wages. Any policy that affects the matching pattern needs to ensure that investment incentives remain adequate. Consider therefore transfers $t(e_i)$ that condition on individual educational attainment e_i.

Given a matching function $m(e_i)$ and a transfer schedule $t(e_i)$, an agent i's optimization problem in the industrial sector becomes

$$\max_{e_i} \frac{\sqrt{e_i m(e_i)} + \sqrt{e_i \mu(e_i) - 8\phi}}{4} - 1 - \frac{e_i^2}{8\theta_i} + t(e_i).$$

Assuming $t(e_i)$ is differentiable the first order condition is

$$\frac{\sqrt{m(e_i)}}{8\sqrt{e_i}} + \frac{\sqrt{e_i}}{8\sqrt{m(e_i)}}\frac{\partial m(e_i)}{\partial e_i} + \frac{m(e_i) + e_i\frac{\partial m(e_i)}{\partial e_i}}{8\sqrt{e_i m(e_i) - 8\phi}} + \frac{\partial t(e_i)}{\partial e_i} = \frac{e_i}{4}. \qquad (16)$$

Since $F(\theta)$ has full support and is differentiable $m(e_i)$ is differentiable. Denote the constrained efficient investment that is given by the solution to problem (10) by $e_i(e_j)$. It is defined implicitly by the first order condition:

$$\frac{\sqrt{e_j}}{\sqrt{e_i(e_j)}} + \frac{e_j}{\sqrt{e_i(e_j)e_j - 8\phi}} = \frac{e_i(e_j)}{\theta_i}. \tag{17}$$

Note that $e_i(e_j)$ may be locally decreasing in θ_i if $\theta_i < \theta^c$, that is if negative assortative matching is indeed efficient. Suppose this is not the case, for instance because $F(\theta)$ is symmetric on $[\hat{\theta}^{\phi}, \bar{\theta}]$ implying $\frac{\partial m(e_i)}{\partial e_i} = -1$. Then a transfer scheme can be computed that induces agents to choose education investments that satisfy (17) given a matching function $m(e_i)$.

Proposition 9 *Let $\bar{\theta}$ be sufficiently small and $F(\theta)$ symmetric on $[\hat{\theta}^{\phi}, \bar{\theta}]$. The efficient allocation can be implemented by a matching function $m(e_i)$ with $\frac{\partial m(e_i)}{\partial e_i} = -1$ for $e_i \geq e^*(\hat{\theta}^M, \overline{\theta^M})$ and $m(e_i) = e_i$ otherwise, and a monetary payment $t(e_i)$ such that*

$$\frac{\partial t(e_i)}{\partial e_i} = \frac{m(e_i)}{8\sqrt{e_i m(e_i)}} - \frac{e_i \frac{\partial m(e_i)}{\partial e_i}}{\sqrt{e_i m(e_i) - 8\phi}}\, y(e_i; e_j);$$

for $e_i \geq e^*(\hat{\theta}^M, \overline{\theta^M})$.

Proof: It is straightforward that given the assumption on $\frac{\partial t(e_i)}{\partial e_i}$ conditions (16) and (17) coincide. Since $\frac{\partial m(e_i)}{\partial e_i} = -1$ the induced choice $e_i(e_j)$ is strictly increasing in θ_i. Symmetry of $F(\theta)$ implies then that $m(e_i)$ induces negative assortative matching for all $e_i \geq e^*(\hat{\theta}^M, \overline{\theta^M})$ which is efficient when $\bar{\theta}$ is small enough. Monotonicity of $e_i(e_j)$ in θ_i implies that agents with $\theta_i < \hat{\theta}^{\phi}$ choose $e_i = \theta_i$ and work in the basic task. Note that decreasing differences of $u(e^*(\theta_i, \theta_j))$ only apply to $\theta_i, \theta_j \geq \hat{\theta}^{\phi}$ since $\theta_h < \hat{\theta}^{\phi}$ obtain constant payoff $u(\theta_h) = \theta_h/8 > u(e^*(\theta_h, \theta_h))$. $\square$

This means affirmative action on the labor market not only requires a lot of information, but also provision of incentives for constrained efficient education choice incurs a positive cost.

4.4 Interaction of early and later stage policies

Effectiveness of associational redistribution at school stage might depend on the policy choice at the labor market stage and vice versa. Since general treatment of possible interactions proves to be a complex problem we present here a simple example where policies may be complements or substitutes. Assume the distribution of types α is discrete, that is measure $1/2$ of the population has type α_0 and the remainder $\alpha_1 > \alpha_0$.

Let us first assume that associational redistribution are free of distortions both at the school and the labor market stage. That is, let redistribution at the labor market stage follow a version of the scheme described in Proposition 9 and suppose that peer effects are linear. Then associational redistribution at school is independent of labor market policy in the sense that it is beneficial both when the is redistribution on the labor market and when there is not.

The case without labor market policy is described. Suppose now we have negative assortative matching on the labor market that is at least efficient constrained on matches, that is investments obey conditions (11) and (12). Consider now moving from segregation at school to integration causing θ_0 to become $\theta_0' = \theta_0 + \varepsilon$ and $\theta_1' = \theta_1 - \varepsilon$. Total utility in a firm with types θ_0 and θ_1 is given by

$$\begin{aligned} v(\theta_i, \theta_j) &= \frac{\sqrt{e_0(\theta_0, \theta_1) e_1(\theta_0, \theta_1)} + \sqrt{e_0(\theta_0, \theta_1) e_1(\theta_0, \theta_1) - 8\phi}}{2} - 2 \\ &\quad - \frac{e_0(\theta_0, \theta_1)^2}{8\theta_0} - \frac{e_1(\theta_0, \theta_1)^2}{8\theta_1}. \end{aligned}$$

where $e_1(.)$ and $e_0(.)$ are implied by (11) and (12). Taking the total differential yields

$$\frac{\partial v(\theta_0, \theta_1)}{\partial \theta_0} \cdot \theta_0 + \frac{\partial v(\theta_0, \theta_1)}{\partial \theta_1} \cdot \theta_1 = \left(\frac{e_0(\theta_0, \theta_1)^2}{8\theta_0^2} - \frac{e_1(\theta_0, \theta_1)^2}{8\theta_1^2} \right) \varepsilon,$$

where we used the envelope theorem. From the optimality of investments conditional on the match (11) we know that $e_0(\theta_0, \theta_1) = e_1(\theta_0, \theta_1)\sqrt{\theta_1/\theta_0}$ and therefore

$$\left(\frac{e_0(\theta_0, \theta_1)^2}{8\theta_0^2} - \frac{e_1(\theta_0, \theta_1)^2}{8\theta_1^2} \right) \varepsilon > 0 \iff \varepsilon > 0.$$

That is, for any tuple (θ_0, θ_1) a marginal increase of θ_0 and an equal decrease of θ_1 increase aggregate utility of agents in firms (i, j) with $\theta_i = \theta_0$ and $\theta_j = \theta_1$.

Note that the reverse, namely that effectiveness of associational redistribution on the labor market is independent of policy at the school stage, does not hold in general. The reason is that the optimal matching pattern, for instance negative assortive or segregation, depends on the distribution of education cost μ_i. This may be altered by a move to school integration thus altering the distribution in second order statistical sense. In the Appendix we provide a numerical example of an economy where associational redistribution on the labor market is only increasing output if there is no associational redistribution at school. Moreover, the numerical example allows us to deduce the magnitude of welfare gains. It appears that effects of policy at the school stage are higher by an order of magnitude than those of labor market policy.

5 Conclusion

We have presented a model of a labor market that suffers from informational frictions, that is moral hazard in production within firm. This market friction introduces a setup cost that decreases in project size. It prevents low types from entering the industrial labor market. The labor market, however, is doing quite well at pricing educational investments conditional on an individual's match. Therefore transfer schemes that condition on profit or education cannot improve welfare. Sometimes welfare might be improved by combining a transfer scheme with associational redistribution, that is a policy that sets the matching pattern or affirmative action, on the labor market. Informational requirements to implement such policies are quite demanding, since the matching policy has to be accompanied by the correct transfer scheme to fine-tune incentives.

A policy that is less hard to implement may involve a redistribution of education cost types. This can increase aggregate surplus since within the industrial sector surplus is concave in education cost. A redistribution of education cost types may be achieved for instance by over-proportionally spending resources on the education of low education cost types or, in case there are peer effects at school, by associational redistribution of individuals across schools to increase heterogeneity at any given school. That is, affirmative action seems to be effective at an early stage in the form of school integration if at all contrary to Fryer and Loury (2005).

Indeed the PISA studies (OECD, 2001, 2004, 2007) find that the dependence of children's educational outcomes on their parents' socioeconomic status strongly varies between countries. Moreover, the strength of this relationship seems to be heavily influenced by variation of educational outcomes between schools. This suggests that the extent of school segregation conditional on parental socioeconomic status, which is a policy variable, has a significant effect. Germany and Belgium, for instance, are among the countries with the highest correlation between educational outcomes and parental socioeconomic status which is largely driven by variation between schools. This means there is ample room for policy interventions in these countries in order to decrease the dependence between children's and parents' education. If parental socioeconomic status can be interpreted as a proxy for individual education cost type α_i, a move towards school integration in these countries is associated to sizable gains in output.

A Appendix

Example: A tax scheme increases aggregate output

Let there be two possible education cost types, $\theta_i \in \{L; H\}$, and four education levels $e_i \in \{e_0; e_1; e_2; e_3\}$. Denote the measure of L agents by $\mu \in (0; 1/2]$ and the measure of H agents by $1 - \mu$.[11] An L agent may invest in e_0 at zero cost, in e_1 at cost $1 + \varepsilon/4$, investing in e_2 or e_3 is not feasible (say because of infinite cost). An H agent may invest in e_0, e_1, or e_2 at zero cost, investing $e_i = e_3$ comes at cost 1. Agents may form pairs to jointly produce output at zero production cost. Resulting joint payoffs are given in Table 1.

	e_0	e_1	e_2	e_3
e_0	0	1	2	3
e_1	1	2	4	$4+\varepsilon$
e_2	2	4	$4+\varepsilon$	5
e_3	3	$4+\varepsilon$	5	$6+\delta$

Table 1: Joint payoffs depending on education choice

Assume that $0 < \varepsilon < \delta < 1$ and $2\varepsilon > \delta$. When remaining solitary an agent's payoff is 0 independently of education choice. Suppose a very simple

[11] Scarcity of L agents facilitates exposition, the necessary condition is $\mu \leq \frac{8+4\varepsilon/2}{8+5\varepsilon/2}$.

form of non-transferability in that output is shared equally and agents are liquidity constrained. That is, starting at equal sharing agents may only transfer up to b units of surplus within …rms. Let ²=2 < b < 1.

In a market allocation there will be segregation on the matching market due to the non-transferability. To see this note that an agent of type e_i can always obtain a minimum payo¤ by matching with another e_i agent and splitting the surplus. These segregation payo¤s are given by 0 for e_0, 1 for e_1, 2 + ²=2 for e_2, and 3 + ±=2 for e_3. Since an agent can transfer at most b < 1 within a …rm, heterogeneous …rms cannot emerge on the market.

Since 1 + ²=4 > 1 L agents cannot choose e_i 6= e_0 in equilibrium. Similarly, since ± > ² and therefore 1 + ±=2 > 1 + ²=2 type H agents …nd it optimal to choose e_i = e_3. That is, the competitive market allocation has measure ¹=2 of (e_0; e_0) …rms and measure (1 ¡ ¹)=2 of (e_3; e_3) …rms. Total surplus is (1 ¡ ¹)(3 + ±=2 ¡ 1).

Now suppose there is a tax $t(e)$ conditional on education choice. Let $t(e_0)$ = 0, $t(e_1)$ = ¡²=2, $t(e_2)$ = 0, and $t(e_3)$ = ²=2. Note that the tax is subject to the non-transferability in that j$t(e)$j < b for all education choices. Consider …rst H types. Since 2 + ²=2 > 3 + ±=2 ¡ ²=2 ¡ 1 when facing the tax H types choose e_2. Since 2 + ²=2 > 2 ¡ ²=2 + ²=2 the subsidy does not induce them to choose e_1.

As 1 + ²=2 ¡ (1 + ²=4) > 0 the subsidy ensures that L types now choose e_1. Since 2 + ²=2 < 2 + b and 2 ¡ b > 1 there exists a distribution of surplus in (e_1; e_2) …rms such that there cannot be positive measure of both (e_1; e_1) and (e_2; e_2) …rms in the economy since at least agents of one education level would be better o¤ in the heterogeneous …rm. Hence, the market allocation under the tax scheme is given by all L agents choosing e_1 and all H agents choosing e_2. There is measure ¹ of (e_1; e_2) …rms while the remainder of …rms is homogenous. Aggregate surplus in the economy is now ¹(4 ¡ 1 ¡ $\frac{3}{4}$²) + (1=2 ¡ ¹)(4 + $\frac{1}{2}$²) > (1 ¡ ¹)(3 + ±=2 ¡ 1) since ¹ · 1=2.

There are several points worth mentioning here. First, although the tax schedule appeared to raise positive net revenue given the laissez faire outcome, in the market allocation net revenue from the tax/subsidy scheme is strictly negative. Second, private returns to education highly depend on the market outcome. Third, this generates a matching externality of education investment which implies there is scope in government intervention in form of an education subsidy in this example. That is, labor market policy has to consider effects on investment incentives and the ensuing changes in the composition of the labor force which may easily dominate direct labor market effects.

Example: A case where effectiveness of labor market policy depends on school integration

Assume the model's parameters as follows: $\theta_0 = 12.6$ and $\theta_1 = 14.8$ and both have measure $1/2$, and $\phi = 99$. At these parameters all agents go to the industrial sector even under segregation, and $\hat{\theta}^C = 13.6735$, that is the sufficient condition in Proposition 5 does not hold. Suppose integration at the schooling stage leads to modified education cost types $\theta_0 + \varepsilon$ and $\theta_1 - \varepsilon$. Table 2 gives the sum of agents' utilities in firm (θ, θ') depending on ε. The

ε	(θ_0, θ_0)	(θ_1, θ_1)	(θ_0, θ_1)	Difference Δ
0	1.6122	5.3887	3.5006	$-3.1349e-4$
0.5	2.5266	4.5812	3.5539	$+6.1877e-5$
1	3.4047	3.7468	3.5757	$+3.4804e-6$

Table 2: Aggregate payoffs depending on policy

last column gives the difference of aggregate payoffs between segregation and integration on the labor market, $\Delta = v(\theta_0, \theta_0) + v(\theta_1, \theta_1) - 2v(\theta_0, \theta_1)$. The gains in output by school integration are sizable, for $\varepsilon = 0.5$ aggregate output increases by 1.53 % and for $\varepsilon = 1$ by 2.15 % compared to the laissez faire labor market allocation.

B Proofs

B.1 Proof of Lemma 1

An industrial firm is preferred to the basic task if

$$\frac{e(\theta_i) + \sqrt{e(\theta_i)^2 - 8\phi}}{4} - 1 - \frac{e(\theta_i)^2}{8\theta_i} \geq \frac{\theta_i}{8}. \qquad (18)$$

Using the envelope theorem the slope of the LHS of the above condition is

$$\frac{e(\theta_i)^2}{8\theta_i^2} > \frac{1}{2},$$

the slope of the RHS of (18) is $1/8 < 1/2$. Hence, there exists at most one $\theta_i \in [\underline{\theta}, \overline{\theta}]$ such that (18) holds with equality. That is, define $\hat{\theta}^\psi$ implicitly by (18) holding with equality and by $\max\{\sqrt{2\psi}, \underline{\theta}\}$ in case (18) is positive for $\theta_i = \max\{\sqrt{2\psi}, \underline{\theta}\}$.

B.2 Proof of Proposition 2

Let us show first that indeed $\hat{\theta}^0 < \hat{\theta}^\psi$. Recall that $\hat{\theta}^0 = 8/3$. Using (7) an interior cutoff $\hat{\theta}^*$ is defined by

$$\frac{e(\theta_i)}{4\theta_i} - 1 - \frac{e(\theta_i)^2}{8\theta_i} - \frac{\theta_i}{8} = 0. \tag{19}$$

At $\theta_i = \hat{\theta}^0 = 8/3$ we have

$$2e - e^2 - \frac{256}{9} < 0 \ \forall e \in \mathbb{R}.$$

That is, for $\theta_i = \hat{\theta}^0$ there does not exist e such that the industrial technology is preferred to the basic task. Since (19) increases in θ we have that indeed $\hat{\theta}^0 < \hat{\theta}^*$. As for (ii), over- and under-investment follow trivially from (8) and (9). To prove (iii) we need to show that the derivative of the difference $e^\psi(\theta_i) - e^*(\theta_i)$ with respect to θ_i is negative for $\theta_i \geq \hat{\theta}^\psi$:

$$\frac{\partial e^\psi(\theta_i)}{\partial \theta_i} < 2.$$

Using the implicit function theorem on the first order condition (7) this is

$$\frac{1}{2\theta_i^2} < \frac{1}{e^\psi(\theta_i)^2} + \frac{e^\psi(\theta_i)}{(e^\psi(\theta_i)^2 - 8\psi)^{3/2}}.$$

Using the identity (7) this becomes

$$\frac{1}{2}\frac{\left(e^\psi(\theta_i) + \sqrt{e^\psi(\theta_i)^2 - 8\psi}\right)^2}{e^\psi(\theta_i)^2(e^\psi(\theta_i)^2 - 8\psi)} < \frac{1}{e^\psi(\theta_i)^2} + \frac{e^\psi(\theta_i)}{(e^\psi(\theta_i)^2 - 8\psi)^{3/2}}.$$

That is

$$e^\psi(\theta_i)\sqrt{e^\psi(\theta_i)^2 - 8\psi} < e^\psi(\theta_i)^2 - 4\psi + \frac{e^\psi(\theta_i)^3}{\sqrt{e^\psi(\theta_i)^2 - 8\psi}}.$$

This condition holds if $4\psi < e^{\psi}(\theta_i)^2$, which is the case for $\theta_i > \hat{\theta}^{\psi}$.

B.3 Proof of Proposition 5

(i) is obvious since (12) coincides with (7) when $\theta_i = \theta_j$.

Turn now to (ii) and start by applying the envelope theorem to the solution of problem (10). The derivative of a match's value function

$$v(\theta_i; \theta_j) = \max_{e_i; e_j} \frac{\sqrt{e_i e_j} + \sqrt{e_i e_j - 8\psi}}{2} - 2 - \frac{e_i^2}{8\theta_i} - \frac{e_j^2}{8\theta_j}.$$

with respect to θ_i is given by

$$\frac{\partial v(\theta_i; \theta_j)}{\partial \theta_i} = \frac{e_i^2}{8\theta_i^2}.$$

The cross partial derivative of the value function is then

$$\frac{\partial^2 v(\theta_i; \theta_j)}{\partial \theta_i \partial \theta_j} = \frac{e_i}{4\theta_i^2}\frac{\partial e_i}{\partial \theta_j}.$$

Applying the implicit function theorem to (12) we find that

$$\frac{\partial e_i}{\partial \theta_j} < 0 \iff -\sqrt{\frac{\kappa e_i^2}{\kappa e_i^2 - 8\psi}}\,\frac{16 - \kappa e_i^2}{\kappa e_i^2 - 8\psi} > 1.$$

Solving this expression for κe_i^2 yields

$$8\psi < \kappa e_i^2 < 4(\sqrt{5} + 1)\psi.$$

Define

$$\theta^c = \frac{4\sqrt{\psi}}{\sqrt{\sqrt{5} + 1} + \sqrt{\sqrt{5} - 1}}.$$

This means $e^{\psi}(\theta_i)^2 < 4(\sqrt{5} + 1)\psi$ for all $\theta_i < \theta^c$. With (i) this implies $e^*(\theta_i; \theta_i)^2 < 4(\sqrt{5} + 1)\psi$ for all $\theta_i < \theta^c$ and since using the implicit function theorem we know that $e^*(.)$ increases in both arguments it follows that $e^*(\theta_i; \theta_j)^2 < 4(\sqrt{5} + 1)\psi$ for all $\theta_i; \theta_j < \theta^c$. Hence, if $\theta_i < \theta^c$, $\frac{\partial^2 v(\theta_i;\theta_j)}{\partial \theta_i \partial \theta_j}$ and v has decreasing differences in its arguments. Statement (ii) follows.

To show (iii) apply the implicit function theorem to (12) to verify that the derivative of e_i with respect to κ exists and is negative. Under the assumption efficient matching must be negative assortive as shown above. Since $\kappa = 1$ in segregation and $\kappa < 1$ for $\mu^M < \mu_i < \mu^{med}$) and $\kappa > 1$ otherwise the statement follows.

B.4 Proof of Proposition 6

Start with $t^y(y)$. Recall that output y in a firm $(e_j; e_i)$ is given by $y(e_i; e_j) = \frac{\sqrt{e_i e_j} + \sqrt{e_i e_j - 8\phi}}{2}$. By Proposition 1 the industrial sector segregates, only firms $(e_i; e_i)$ emerge and agents split the surplus in firms. Consider a tax on profits of $t^y(y)$. Agent i's optimization problem in the industrial sector is

$$\max_{e_i} \frac{t^y(y(e_i))}{2} - 1 - \frac{e_i^2}{8\mu_i}.$$

Assume that the maximization problem is well defined (otherwise the statement follows trivially). Let $e^t(\mu_i)$ denote the solution of this problem. Since segregation in education e_i obtains on the labor market independently of t^y, it suffices to calculate the change in aggregate surplus of each agent type μ_i separately. Aggregate surplus of two μ_i agents including the tax-subsidy revenue is

$$\begin{aligned} S(t^y) &= t^y(y(e^t(\mu_i))) - \frac{e^t(\mu_i)^2}{4\mu_i} - 2 + y(e^t(\mu_i)) - t^y(y(e^t(\mu_i))) \\ &= 2\left(\frac{y(e^t(\mu_i))}{2} - 1 - \frac{e^t(\mu_i)^2}{4\mu_i}\right). \end{aligned}$$

But by Proposition 5 market equilibrium investments $e^\phi(\mu_i)$ are conditionally efficient and therefore

$$S(t^y) < 2\left(\frac{y(e^\phi(\mu_i))}{2} - 1 - \frac{e^\phi(\mu_i)^2}{4\mu_i}\right);$$

for all $t^y(y) \neq id$. The case of a tax-subsidy scheme conditioning on education is treated in a similar way. Given an education tax-subsidy scheme t^e, agent i's optimization problem for the industrial sector is

$$\max_{e_i} \frac{y(e_i)}{2} - 1 - \frac{e_i^2}{8\mu_i} + t(e_i).$$

Let $e^t(\mu_i)$ denote the solution to this problem. Aggregate surplus of the two agents in the firm is then

$$S(t^e) = 2\left(\frac{y(e^t(\mu_i))}{2} - 1 - \frac{e_i^2}{8\mu_i} + t(e_i)\right) - 2t(e_i).$$

Hence,

$$S(t^e) < 2\left(\frac{y(e^{*}(\mu_i))}{2} - 1 - \frac{e^{*}(\mu_i)^2}{4\mu_i}\right);$$

for all $t^e(e_i) \neq id$ again.

References

Bénabou, R.: 1996, 'Equity and Efficiency in Human Capital Investment'. *Review of Economic Studies* 63, 237–264.

Bidner, C.: 2007, 'A Spillover-based Theory of Credentialism'. *mimeo University of British Columbia*.

Bonesrønning, H., T. Falch, and B. Strøm: 2005, 'Teacher Sorting, Teacher Quality, and Student Composition'. *European Economic Review* 49, 457–483.

Boyd, D., H. Lankford, S. Loeb, and J. Wyckoff: 2005, 'Explaining the Short Careers of High-Achieving Teachers in Schools with Low-Performing Students'. *American Economic Review* 95(2), 166–171.

Card, D. and A. A. Payne: 2002, 'School finance reform, the distribution of school spending, and the distribution of student test scores'. *Journal of Public Economics* 83(1), 49–82.

Coate, S. and G. C. Loury: 1993, 'Will Affirmative-Action Policies Eliminate Negative Stereotypes?'. *American Economic Review* 5, 1220–1240.

Cole, H. L., G. J. Mailath, and A. Postlewaite: 2001, 'Efficient Non-contractible Investments in Large Economies'. *Journal of Economic Theory* 101, 333–373.

Dunne, T., L. Foster, J. Haltiwanger, and K. R. Troske: 2004, 'Wage and Productivity Dispersion in United States Manufacturing: The Role of Computer Investment'. *Journal of Labor Economics* 22(2), 397–429.

Epple, D. and R. E. Romano: 1998, 'Competition Between Private and Public Schools, Vouchers, and Peer-group E¤ects'. *American Economic Review* 88(1), 33–62.

Felli, L. and K. Roberts: 2002, 'Does Competition Solve the Hold-up Problem?'. *CEPR Discussion Paper Series* 3535.

Fryer, R. G. and G. C. Loury: 2005, 'A¢ rmative Action and Its Mythology'. *Journal of Economic Perspectives* 19(3), 147–162.

Gall, T., P. Legros, and A. F. Newman: 2006, 'The Timing of Education'. *Journal of the European Economic Association* 4(2-3), 427–435.

Hanushek, E. A., J. F. Kain, and S. G. Rivkin: 2004, 'Why Public Schools Lose Teachers'. *Journal of Human Resources* 39(2), 326–354.

Hanushek, E. A., J. F. Kain, and S. G. Rivkin: 2005, 'Teachers, Schools, and Academic Achievement'. *Econometrica* 73(2), 417–458.

Heckman, J. J.: 2008, 'Schools, Skills, and Synapses'. *NBER Working Paper Series* 14064.

Holzer, H. and D. Neumark: 2000, 'Assessing A¢ rmative Action'. *Journal of Economic Literature* 38(3), 483–568.

Hopkins, E.: 2005, 'Job Market Signalling of Relative Position, or Becker Married to Spence'. *mimeo University of Edinburgh*.

Hoxby, C.: 1996, 'Are E¢ ciency and Equity in School Finance Substitutes or Complements?'. *Journal of Economic Perspectives* 10(4), 51–72.

Iranzo, S., F. Schivardi, and E. Tosetti: 2008, 'Skill Dispersion and Firm Productivity: An Analysis with Employer-Employee Matched Data'. *Journal of Labor Economics* 26(2), 247–285.

Katz, L. F., J. R. Kling, and J. B. Liebman: 2001, 'Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment'. *Quarterly Journal of Economics* 166(2), 607–654.

Legros, P. and A. F. Newman: 2007, 'Beauty Is a Beast, Frog Is a Prince: Assortative Matching with Nontransferabilities'. *Econometrica* 75, 1073–1102.

Mailath, G. J., A. Postlewaite, and L. Samuelson: 2004, 'Sunk Investments Lead to Unpredictable Prices'. *American Economic Review* 94(4), 896–918.

Mailath, G. J., L. Samuelson, and A. Shaked: 2000, 'Endogenous Inequality in Integrated Labor Markets with Two-Sided Search'. *American Economic Review* 90(1), 46–72.

Nosaka, H.: 2007, 'Specialization and Competition in Marriage Models'. *Journal of Economic Behavior & Organization* 63(1), 104–119.

OECD: 2001, *Knowledge and Skills for Life: First Results from PISA 2000*. OECD, Paris.

OECD: 2004, *Learning for Tomorrow's World: First Results from PISA 2003*. OECD, Paris.

OECD: 2007, *PISA 2006 Science Competencies for Tomorrow's World*. OECD, Paris.

Peters, M.: 2006, 'The Pre-marital Investment Game'. *mimeo University of British Columbia*.

Peters, M. and A. Siow: 2002, 'Competing Pre-marital Investments'. *Journal of Political Economy* 110, 592–608.

Rivkin, S. G.: 2000, 'School Desegregation, Academic Attainment, and Earnings'. *Journal of Human Resources* 35(2), 333–346.

Sanbonmatsu, L., J. R. Kling, G. J. Duncan, and J. Brooks-Gunn: 2006, 'Neighborhoods and Academic Achievement: Results from the Moving to Opportunity Experiment'. *NBER Working Paper Series* 11909.