ELSEVIER

Journal of International Economics 71 (2007) 495 – 514

Journal of INTERNATIONAL ECONOMICS

www.elsevier.com/locate/econbase

Institutions and moral hazard in open economies☆

Jonathan Vogel

Department of Economics, 001 Fisher Hall, Princeton University, Princeton, NJ 08544, United States

Received 14 December 2004; received in revised form 10 February 2006; accepted 27 July 2006

Abstract

I investigate the interaction between international trade and national institutional development in an environment characterized by heterogeneous individuals choosing their education levels to maximize their utilities; and institutions alleviating moral hazard by allowing managers to better observe and verify the productive efforts of workers. Liberalized trade allows institutions to serve as independent sources of comparative advantage. In this setting, I examine the effect of trade liberalization on the distribution of income in institutionally developed and underdeveloped nations. Trade affects income via a direct effect on prices and an indirect effect on the incentives to invest in education.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Comparative advantage; Endogenous endowments; Inequality; Institutions; Moral hazard

JEL classification: B52; F10; F16

1. Introduction

Many of the consequences of international trade remain a mystery, among them the effects of trade on the growth and distribution of income.[1] Trade liberalization seems a necessary though insufficient cause of economic development. Trade allowed the Asian tigers to capitalize on their comparative advantages, helping grow their economies. Yet other countries, even those facing lower

☆ The author benefited from the advice of the anonymous referees, Robert Staiger, and especially Gene Grossman. The author is grateful to the National Science Foundation for financial support from grant no. SES 0211748. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the author and do not necessarily reflect the views of the National Science Foundation or any other organization.

E-mail address: jvogel@princeton.edu.

[1] Feenstra and Hanson (in press), for example, argues that international trade can explain much of the growth in the American wage gap between skilled and unskilled labor. Lawrence and Slaughter (1993) is an example of research that reaches the contrary conclusion.

0022-1996/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jinteco.2006.07.003

trade barriers, have not enjoyed the same success.[2] As Birdsall et al. (2005) note, "history and economic and political institutions have trumped other factors in determining economic success." If a nation's institutions and its access to foreign markets jointly determine its economic development, then we must understand the interaction between institutions and international trade. Such an understanding may also provide insights into the impact of trade on inequality within nations.

A growing literature documents the importance of institutional quality on economic growth and development. There is also a long history of theoretical and empirical investigation into the effects of international trade on growth and the distribution of income. However, relatively little work has been done on the effect of institutional quality on the pattern and consequences of international trade. There are at least two reasons to address this gap in the literature. First, not understanding the interaction between institutional quality and international trade may bias empirical studies of the effect of trade on incomes. I find, for example, that liberalizing trade directly influences agents' education decisions, and in this way may indirectly increase income inequality. This may help explain the lack of consensus about the role of international trade in widening America's wage gap. Second, not understanding the interaction between institutional quality and international trade may raise unrealistic expectations about the benefits of liberalized trade.

I consider a framework in which institutions alleviate moral hazard. There are two industries: X and Y. An agent works alone in the Y sector, but the X sector requires the joint effort of two workers. The two industries differ in that individual productive effort is observable and verifiable only in the Y sector. Neither individual effort nor output is observable or verifiable in the X sector, requiring managers in that sector to monitor employees to determine individual contributions. I assume that better trained managers are better able to monitor their employees. I also assume that managers in countries with more developed institutions are better able to monitor their employees: managerial training and national institutions mitigate moral hazard.

In order to capture the effect of the interaction between institutional quality and international trade, I consider a simple game in which agents choose their sector of employment, level of job training, firm, and level of effort. Agents differ according to "natural ability": agents with more ability obtain training while incurring a lower utility cost of education. This framework enables me to explore subtle interactions in a realistic but tractable setting.

Industries differ in the difficulty of judging individual contributions to firm output. Such judgments are straightforward in settings where units of output can be traced back to individual workers. Certain traditional forms of production–for example, agricultural harvesting, handloom weaving of textiles, and craft production of shoes–best exemplify the circumstances in which an employer can view both an employee's work and his product and draw a correlation between the two. Even industries in which more than one worker contributes to a given unit of output allow such straightforward judgments when each worker contributes to production in such a deliberately prescribed manner that specific units of output can be traced back to individual workers. This case obtains in assembly-line production. But the modern workplace is particularly characterized by another type of production, one in which multiple workers jointly produce each part of each unit of output. When teams share responsibility for output, employers often have no clear and systematic way to assign responsibility. Investment banking exemplifies team production.

That workers exert more productive effort when they are rewarded for their efforts is hardly controversial. Individual production industries, like agricultural harvesting, find it easy to provide

[2] Birdsall et al. (2005) provide a brief, but excellent comparison between the development success of Vietnam, which is still not a member of the WTO, and the relative failure of Nicaragua, a country with a similar recent history but better market access.

powerful incentives. A piece rate wage suffices. Farmers pay laborers for each piece of fruit they harvest. All things being equal, a worker choosing to exert less effort earns less than one exerting more. A laborer cannot easily fool the farmer. But an investment bank dependent on team production cannot provide such straightforward incentives. This is particularly true for junior staff members who are not responsible for bringing in or maintaining clients. Piece rate wages lose their effectiveness when employees work collaboratively on their analyses and presentations. The parts of the whole–especially in good work–cannot readily be attributed to individual workers, rendering it difficult to pay workers for their individual contributions. Management is thus forced to rely on imperfect performance measures if it wishes to judge the contributions of individual workers in team-production industries. Absent such performance measures, employers lack meaningful incentives and workers are unable to reap the fruits of their labor.

Management in team-production industries thus needs to establish performance metrics based on such criteria as attitude, effort, and quality of ideas to evaluate and reward employees. Management also needs to institute evaluation systems to measure how its employees carry out their work. Good management often institutes overlapping channels of communication to allow the reporting–and evaluation–of multiple perspectives on each worker's interactions with peers and supervisors.

But the firm–and whatever performance evaluation system its management embraces–forms only part of the picture when it comes to rewarding and incentivizing a high-performance culture. The broader institutional environment in which a team-production firm operates also influences the degree to which worker contributions can be attributed to specific individuals. This broader picture clearly involves nations and the nature and quality of their institutions. Countries differ, for example, in their accounting systems. Effective accounting systems–particularly those to which the state delegates some authority–regularize the reporting of data through which both the productivity of firms and the contributions of their various elements–down to the level of individual employees–can be assessed. Judicial systems also play important roles. The more competent a country's legal institutions, the more accurately courts can assign credit for specific contributions to production processes when disagreements enter the judicial system. The quality of the legal establishment influences incentives even when no disputes arise. Parties to contracts understand that more competent legal systems can verify a larger fraction of the observable facts. Employees thus exert effort, believing that employers will be legally obligated to reward their work.

Better trained managers and better national institutions thus improve the quality of performance measures. This assumption leads, in a very straightforward manner, to the first two results of the paper. First, better trained individuals enter the team production industry where they are rewarded for their ability to minimize moral hazard. Second, all else being equal, trade liberalization encourages countries with better developed national-level institutions to specialize in team production industries. Institutional-quality induced international specialization has important consequences for cross-country comparisons of income and education. The citizens of countries specializing in team production will choose to receive more education than equally able and advantaged citizens of countries with less developed institutions. This conclusion allows us to reconcile two distinct viewpoints within the development and growth literatures. Mankiw et al. (1992), for example, demonstrate starkly and undeniably positive effects of human capital accumulation on growth and income. On the other hand, Acemoglu et al. (2001, 2002) link reversals in economic fortune with reversals in institutional development. This paper establishes that human capital accumulation and institutional quality are inextricably linked. Individuals endogenously choose their levels of schooling understanding that institutional development determines the benefit to education.

The interaction between international trade and institutional quality can also help explain intra-national phenomena. My model predicts that trade tends to widen the distributions of education and income in countries with better developed institutions. It causes the opposite to occur in countries with under-developed institutions. This prediction is similar to typical Heckscher–Ohlin trade theory where the difference in real income between capitalists and laborers compresses in labor-abundant countries and widens in capital-abundant countries.

Compared to the literature on moral hazard, the literature on institutions and trade is relatively incomplete. While I consider the role of institutions in alleviating moral hazard in imperfect labor markets, in a related paper, Matsuyama (2005) considers the role of institutions in alleviating problems with imperfect credit markets. Like my paper, Costinot (2004) investigates the interaction between imperfect contract enforcement and international trade. Costinot reaches a conclusion similar to Matsuyama's and mine: institutions can act as independent sources of comparative advantage. The paper that most resembles this one is Grossman (2004): the perfect institution benchmark model from which my work diverges nearly replicates Grossman's benchmark model. Grossman abstracts from institutional quality and focuses on the pattern and consequences of trade when nations differ in their distributions of talent. As with the majority of the trade literature, both Costinot and Grossman exogenously fix the level (Costinot) or distribution (Grossman) of human capital/talent per worker. My paper deviates from the standard trade literature in assuming that factor endowments–particularly the distribution of human capital–are determined endogenously. I show that institutional quality is the common cause of comparative advantage and human capital accumulation (whereas Costinot finds that both institutions and human capital per worker are independent sources of comparative advantage). This insight leads to a set of conclusions relevant to both the international trade and development economics literatures: although increased human capital is certainly important, institutional reform lies at the heart of the development process. Moreover, trade liberalization impacts incentives to accumulate human capital, differentially affecting nations situated at different points within the development process. Such conclusions lie outside of the scope of the Costinot and Grossman frameworks, and all frameworks that treats endowments as exogenous.

The remainder of the paper is in six sections. I provide the setup of the model in Section 2. In it, I describe a two-sector economy in which training-augmented labor is the only input, I quantify agent utility, and I provide the game theoretic structure of the model. The Walrasian equilibrium of a world of perfect institutions, the case of the first best, is described in Section 3. This serves as a benchmark for what follows. In Section 4, I characterize the autarky equilibrium-occupational choice, job training decision, and utility of agents of different natural abilities. In Section 5, I consider the effects of institutional change within a small country open to trade. Section 6 deals with the effect of two large countries opening to trade with each other. Section 7 concludes.

2. Setup

I model a four-stage game in which an agent chooses her industry of employment in the first stage and her level of job training in the second stage. Production teams are formed in the third stage and effort levels are chosen in the final stage. I solve for a symmetric, subgame-perfect Nash equilibrium of the four-stage game. Given the temporal structure of the game, it should be understood that education always refers to industry-specific training. The ordering of events makes intuitive sense. After an agent chooses the industry in which she will seek employment, she determines how much industry-specific training to obtain. After choosing to become a computer

scientist, for example, an individual might elect to study at Carnegie Mellon. After completing her education, she selects the production team that best fits her training. The Carnegie Mellon graduate might get a job at Microsoft. Finally, once employed, she must decide how much and what types of effort to invest in her work.

Let

$$U = u(C_x, C_y) - \frac{1}{2t}(a^2 + d^2) - \frac{1}{2q}t^2 \tag{1}$$

denote the utility of an agent with skill level $q \in [q_{\min}, q_{\max}]$ where $0 < q_{\min} < q_{\max}$, obtaining observable and verifiable industry-specific training t, exerting efforts a and d, and consuming C_i units of good i. I assume that the sub-utility u is homothetic and that the marginal utility to consuming good i is infinite if $C_i = 0$ and $C_j > 0$ for $i \neq j$. Denoting by I the agent's income, the indirect utility function is

$$U = HI - \frac{1}{2t}(a^2 + d^2) - \frac{1}{2q}t^2 \tag{2}$$

where H is a function of p, the relative price of good Y.

Effort comes in two varieties: productive effort, a, and distortionary effort, d. Only productive effort has any impact on firm output. Although distortionary effort is completely unproductive, it is not useless from an employee's perspective. When institutions are imperfect, distortionary effort increases the worker's apparent productivity by increasing his performance measure. An employee who browses the World Wide Web at his desk does not increase his firm's output. But to his boss, he may appear to be working. The more training an individual receives, the easier it is for that individual to put forth effort. The results do not depend on the particular form in which the utility cost of effort enters the utility function. The results are qualitatively unchanged if the utility cost of effort is any increasing and convex function of a and d such that the two efforts do not enter the utility function as perfect substitutes.

There are two sectors in the economy: the individual production sector, Y, and the team production sector, X. In the Y sector, an agent's efforts are perfectly observable and verifiable and there are constant returns to productive effort. In particular, one unit of labor exerting productive effort a produces λa units of Y. In the X sector, agents produce in teams of two. The output of two workers exerting efforts $(\mathbf{a},\mathbf{d}) = (a_1,a_2) \times (d_1,d_2)$ is

$$x(a_1, a_2) = 2(a_1 a_2)^{\frac{1}{2}}. \tag{3}$$

As in Kremer's O-Ring theory of production (Kremer, 1993), the marginal productivity of one worker depends on the actions of the other worker. The assumption of complementaries in x captures the essence of team production. Production in the X sector is homogeneous of degree one so that, as in the Y sector, there are constant returns to productive effort. Output is symmetric in the productive efforts of agents one and two.

In the fourth stage, each individual inelastically supplies one unit of labor towards production. I assume that monitoring is costless but not necessarily perfect. After production occurs, firm owners must pay their employees according to the contracts into which they have entered. Firm owners are the residual claimants to firm profit.

Production teams are formed in the third stage via costless matching. I assume that each individual offers a menu of contracts to potential employees conditional on their education (which is observable and verifiable) while also announcing a set of contracts that she would accept

conditional on owner-education. The equilibrium is a set of contracts and pairs of individuals such that no individual can increase his utility by offering a contract to some other individual that would be accepted. I look for a symmetric equilibrium: if an agent with training t offers (or accepts) a certain menu of contracts, then all other agents in the team-production sector with equal training must offer (or accept) the same menu of contracts.

The labor force comprises a continuum of individuals indexed by their natural ability $q \in [q_{\min}, q_{\max}]$; the more talented an individual, the greater is her q and the easier it is for her to obtain education. In the second stage, agents choose their levels of job training. Skill level denotes the ease with which an agent obtains education. I explicitly assume that the cost of training is independent of institutional quality and is in terms of utility rather than in terms of foregone earnings. In reality, the relationship between the cost of training and institutional quality is ambiguous. On one hand, countries with better institutions provide easier access to education because schools are better equipped, teachers better prepared (and present), etc. According to this argument, the utility cost of education should be lower in countries with better institutions. On the other hand, the cost of education is the opportunity cost of foregone wages. In this case, education should be more costly in countries with higher wages (which may themselves be a product of better institutions). I take a neutral position so that neither of these opposing assumptions drive my results.

3. First best—perfect institutions

In this section, I examine a benchmark model in which worker effort levels, a and d, are observable and verifiable in both industries. As national institutions approach perfection, the results of the model with moral hazard replicate the results of this benchmark model. Hence, I call the benchmark model the case of perfect institutions. The perfect-institution model closely resembles Grossman's (2004) model.

I solve the model using backward induction, beginning in stage 4 in which individuals exert effort and firms produce output. An agent in the Y sector with t units of training and with natural ability q has utility $U^y(q,t) = Hp\lambda a - \frac{1}{2t}(a^2 + d^2) - \frac{1}{2q}t^2$. Optimization implies efforts $a = Hp\lambda t$ and $d = 0$, conditional on job training and natural ability. These effort levels lead to an output of $\lambda^2 pHt$, an income of $Ht(p\lambda)^2$, and a utility of

$$U^y(q,t) = \frac{1}{2}(Hp\lambda)^2 t - \frac{1}{2q}t^2. \tag{4}$$

Suppose that two agents with natural abilities q_1 and q_2 and job training levels t_1 and t_2 form a team in the X sector. They achieve the first-best outcome if they choose their actions $(\mathbf{a},\mathbf{d})$ to maximize the sum of their post-education utilities, Ω, where $\Omega = 2H(a_1 a_2)^{1/2} - \sum_{i=1}^{2} \frac{1}{2t_i}\left(a_i^2 + d_i^2\right)$. The first-best efforts are $d_i^{\mathrm{FB}} = 0$ and $a_i^{\mathrm{FB}} = H(t_i)^{3/4}(t_j)^{1/4}$ for $i \neq j$. If a firm's two workers choose their efforts accordingly, their firm produces $x = 2H(t_1 t_2)^{1/2}$. Because efforts are observable and verifiable, the first-best outcome is achievable: the firm owner pays agent $i \neq j$ a bonus of $b = (t_j/t_i)^{1/4}$ per unit of productive effort a_i for $i = 1, 2$. The first best is achievable if the firm owner is a member of the productive team or is not.[3]

Individuals in the X sector form teams in stage 3. Given the efforts chosen in stage 4, it is easily shown that Ω is a supermodular function of job training: $\partial^2\Omega/\partial t_1 \partial t_2 > 0$ for all $t_1, t_2 > 0$. The

[3] This is true under the assumption that an owner who is not part of the productive team can observe and make worker efforts verifiable when institutions are perfect.

supermodularity of Ω and the observability/verifiability of job training imply that both workers in a firm have the same level of job training.

In stage 2, individuals choose the amount of job training to receive. Agents in the Y sector maximize U^y over their choice of t_y. The equilibrium level of training as a function of natural ability is

$$t_y(q) = \frac{1}{2}(Hp\lambda)^2 q. \tag{5}$$

In the X sector agent i maximizes $\{(1/2)H^2(t_i t_j)^{1/2} - (1/2q_i)t_i^2\}$, which is her utility from half of the firm's income plus her dis-utility from taking actions a_i and d_i and from obtaining training level t_i, with respect to t_i. The optimal level of training as a function of natural ability, given that $t_i(q) = t_j(q) \equiv t_x(q)$, is

$$t_x^{\mathrm{FB}}(q) = (1/2)H^2 q. \tag{6}$$

In stage 1, individuals choose between entering the Y and X sectors. If an agent enters the Y sector in stage one, she anticipates utility $U^y(q) = \frac{q}{8}(Hp\lambda)^4$. If an agent enters the X sector in stage 1, she anticipates utility $U^x(q) = \frac{q}{8}H^4$. Both U^x and U^y are linear in q and both are equal to zero when $q=0$. Hence, either $U^y(q) > U^x(q)$, $U^y(q) < U^x(q)$, or $U^x(q) = U^y(q)$ for all $q \in [q_{\min}, q_{\max}]$. Because both X and Y are produced in autarky, $U^x(q) = U^y(q) \equiv U(q)$ for all $q \in [q_{\min}, q_{\max}]$. Otherwise, there would be no production of either Y or X. The fact that utility is constant across sectors fixes the relative price of good Y as a function of the technologies of production: $p = 1/\lambda$. This in turn implies that an individual receives the same level of job training in either sector: $t_y(q) = t_x(q) \equiv t(q)$.

If institutions are perfect, job training, $t(q)$, income, and utility, $U(q)$, are everywhere continuous and linearly increasing functions of natural ability. The allocation of talent across industries is indeterminate. Moreover, international trade implications are not particularly interesting: this is a normal Ricardian model with trade according to comparative advantage. If there are two countries with the same production technologies, the autarky relative price of Y is the same in both countries, $p = 1/\lambda$, independent of their talent distributions. For any distributions of talent, the pattern of trade between the two nations is completely indeterminate. An arbitrarily small trade cost would eliminate all international trade.

4. Second best—imperfect institutions

Consider the case in which neither firm output, x, nor worker efforts, a and d, are observable or verifiable in the team production sector. In this case, contracts cannot be based directly on x, a, or d.[4] Management could offer fixed-wage contracts, but fixed-wage contracts would offer employees no incentive to exert effort. Alternatively, management could offer contracts based on a verifiable performance measure $V(\theta, t_1, a_2, d_2)$ that, to some extent, reflects an employee's productive effort.

The extent to which unproductive effort increases a performance measure depends on the quality of management and the development of national institutions. In particular, I assume that

$$V(\theta, t_1, a_2, d_2) \equiv a_2 + \left(\frac{1-\theta}{t_1}\right) d_2, \tag{7}$$

[4] I maintain the assumption that an agent's training is observable and verifiable in order to abstract from the problem of adverse selection.

where $\theta \in [0,1]$ indexes national institutional quality and t_1 is the job training of the firm manager. The more developed are national institutions, the greater is θ and the less effect unproductive effort has on the performance measure. Similarly, the better trained is management, the better it is at monitoring its employee. Qualitatively, all results follow if $V = a_2 + \frac{f(\theta)}{t_1} d_2$ where $f \geq 0, f' < 0$ and $\lim_{\theta \to 1} f(\theta) = 0$.

I solve for a sub-game perfect Nash equilibrium of the four-stage game using backward induction. In stages 2 to 4, the actions of a worker in the Y sector are not directly affected by institutional quality; I focus exclusively on the X sector in those stages.

4.1. Stage 4—production

The first best is achievable regardless of whether a firm owner works in her firm if institutions are perfect. However, if national institutions are imperfect, all firm owners in the X sector will work in their own firms.[5] I use the subscript 1 to denote variables referring to the worker–owner, henceforth called the owner or employer, and the subscript 2 to denote the employee in the X sector. The owner's profit is $\Pi = x - bV - w$, where the owner pays her worker a bonus b per unit of the worker's performance measure, V, and also pays a fixed wage w to her worker. The worker's utility is $U_2 = H(bV + w) - \frac{1}{2t_2}\left(a_2^2 + d_2^2\right) - \frac{1}{2q_2} t_2^2$ under this payment structure. A worker with training t_2 who works for an owner with training t_1 exerts efforts

$$a_2 = Hbt_2 \text{ and } d_2 = Hbt_2\left(\frac{1-\theta}{t_1}\right). \tag{8}$$

Both productive and distortionary effort increase with H and b because the utility return to each unit of the performance measure, V, increases with these two variables. They both increase with the training of the worker, t_2, because a better trained worker loses less utility from expending productive and distortionary efforts. Finally, distortionary effort is decreasing in both national institutions, θ, and manager training, t_1, because better monitoring ameliorates the effect of distortionary effort on the performance measure.

The owner's utility is $U_1 = H\Pi - \frac{1}{2t_1}\left(a_1^2 + d_1^2\right) - \frac{1}{2q_1} t_1^2$. The owner chooses efforts $a_1 = Ht_1^{2/3}(bt_2)^{1/3}$ and $d_1 = 0$. Since the owner receives the residual profits of the firm, she will exert no distortionary effort.

4.2. Stage 3—production team choice

Individuals in the X sector form firms in stage 3. Before firms are formed, potential owners announce the payment system that they would use if they were to run a firm. A potential owner will choose a bonus to optimize total team utility because the fixed wage allows for any allocation of this utility. Let Ω be the total post-education utility derived from the production of an owner with training t_1 teamed with an employee with training t_2. The optimal bonus is

$$b = (t_1/t_2)^{1/4} A^{3/2}, \tag{9}$$

where

$$A \equiv \left(\frac{t_1^2}{t_1^2 + (1-\theta)^2}\right)^{1/2} \in [0, 1].$$

[5] I prove this in the Appendix.

The variable A gauges the quality of the performance measure of a manager with training t_1 in a country with institutional quality θ. A manager is better able to monitor her employee as A approaches 1. For a finite t_1, a manager monitors perfectly ($A=1$) only when national institutions are perfect ($\theta=1$). The bonus is increasing in the quality of the performance measure because the effect of the bonus on productive relative to distortionary effort is greater when national institutions are more developed and management is better trained.

By substituting the optimal bonus from Eq. (9) into total post-education firm utility Ω, it can be shown that Ω is strictly supermodular in the two arguments t_1 and t_2: $\frac{\partial^2\Omega(t_1,t_2)}{\partial t_1 \partial t_2}>0$.[6] The strict supermodularity of Ω implies that the owner and the employee have the same level of job training in equilibrium. If two agents with equal training form a team, each must be indifferent between being the owner and the employee in a symmetric equilibrium: $U_1(t_i,t_i)+\frac{1}{2q_1}t_i^2 = U_2(t_i,t_i)+\frac{1}{2q_2}t_i^2$. Post-education utilities are equal if and only if the fixed wage is set at zero, $w=0$. An agent with natural ability q anticipates utility

$$U^x(q) = \frac{1}{2}H^2tA - \frac{1}{2q}t^2, \tag{10}$$

if she enters the X sector.

4.3. Stage 2—job training

A type q agent in the X sector who chooses job training t has utility as given in Eq. (10). Such an agent's optimal job training, t_x^*, is implicitly defined by

$$\frac{1}{2}H^2A\frac{t_x^{*2}+2(1-\theta)^2}{t_x^{*2}+(1-\theta)^2} = \frac{1}{q}t_x^*, \tag{11}$$

where A is itself a function of t_x^*. There exists a solution to Eq. (11) such that $t_x^*(q)$ is positive for all $q>0$, more able agents obtain more education $\frac{\partial t_x^*}{\partial q}>0$,[7] and $t_x^*(q)$ approaches zero as q approaches zero.[8] This implicit solution represents a global maximum because $U_{tt}^x(t_x^*)<0$.

4.4. Stage 1—industry choice

In stage 1, type q agents will choose to join the X sector if $U^x(q)>U^y(q)$. At the optimal level of job training, an agent in the Y sector's utility is linearly increasing in q and is independent of national institutions (except through their effect on prices):

$$U^y(q) = \frac{1}{8}q\left(Hp\lambda\right)^4. \tag{12}$$

At $t_x^*(q)$ an individual in the X sector's utility is

$$U^x(q) = \frac{1}{4}H^2t_x^*A^3. \tag{13}$$

[6] The strict supermodularity of x is not a necessary condition for Ω to be strictly supermodular. An additively separable production function, which is only weakly supermodular, induces a strictly supermodular Ω.

[7] This result is obtained by totally differentiating the first-order condition for training.

[8] If optimal training did not converge to zero as skill approached zero then the right-hand-side of Eq. (11) would approach infinity while the left-hand-side would remain bounded from above.

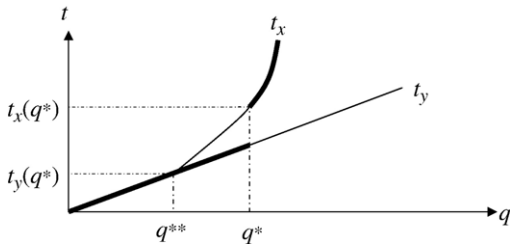

Fig. 1. Training as a function of ability.

Proposition 1. *In a closed economy with $\theta<1$, there exists a $q^*\in[q_{min},q_{max}]$, where $q^*_\theta<0$, such that agents enter the X sector if and only if $q>q^*$.*[9]

Call q^* the "utility cutpoint." Proposition 1 states that in a closed economy with imperfect institutions, talent allocation across sectors is determinate. More skilled workers opt into the X sector and less skilled workers opt into the Y sector. This differs from the benchmark model in which the talent allocation across sectors is indeterminate.

It should not be surprising that if institutions are imperfect, more talented individuals enter the X sector. More talented agents obtain more training, and with imperfect institutions, the return to training is greater in the X sector than in the Y sector. In both sectors, training facilitates the exertion of effort. However, only in the X sector does training also improve management's ability to alleviate moral hazard. It is this logic that leads to the first part of Proposition 2:

Proposition 2. *For any $\theta<1$*

1. *$t_x(q)>t_y(q)$ for all $q\geq q^*$ and $t_x(q)$ is convex in q for all $q\geq q^*$*
2. *$I^x(q)>I^y(q)$ for all $q\geq q^*$, where $I^z(q)$ denotes the income to an agent of skill q working in industry Z.*[10]

Part 1 of Proposition 2 states that if institutions are imperfect, any agent with natural ability greater than or equal to the utility cutpoint receives strictly more job training in the X sector than she would have obtained if she were to have entered the Y sector. In particular, there is a jump in job training at the utility cutpoint. Moreover, while training is a linear function of ability for all agents in the individual production sector, Y, it is a convex function of ability for all agents able enough to opt into the team production sector, X. I graph training as a function of natural ability for the case in which $\theta<1$ and $q^*\in(q_{\min},q_{\max})$ in Fig. 1. As in the appendix, q^{**} is defined as the greatest natural ability at which an agent would receive equal training if in the X or Y sector. The relevant portions of the graph are the t_y function for all $q<q^*$ and the t_x function for all $q>q^*$ (the bold sections).

It is also possible to compare incomes between the X and Y sectors. A worker in the Y sector taking action a has an income of $I^y(q)=p\lambda a(q)$, which in equilibrium equals

$$I^y(q) = (1/2)H^3q(p\lambda)^4. \tag{14}$$

An employee in the X sector earns an income of $I^x_2=bV$; in equilibrium $I^x_2(q)=Ht_xA$. An owner in the X sector earns an income of $I^x_1(q)=x-bV$; in equilibrium, $I^x_1(q)=Ht_xA$. Owners and

[9] See the Appendix for proof of Proposition 1.

[10] See the Appendix for the proof of both Parts 1 and 2 of Proposition 2.

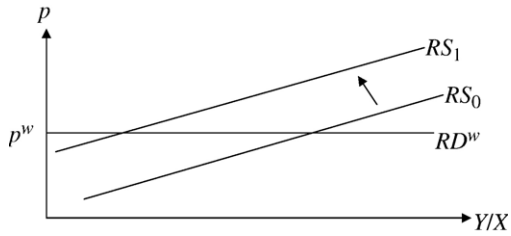

Fig. 2. An inward shift of the relative supply curve from RS_0 to RS_1.

employees have equal incomes in the X sector. Let $I^x(q)$ equal the income of an agent with natural ability q who is either an owner or an employee in the X sector:

$$I^x(q) = Ht_xA. \tag{15}$$

The second part of Proposition 2 states that if institutions are imperfect, any agent with natural ability greater than or equal to the utility cutpoint earns strictly more income in the X sector than she would have earned had she entered the Y sector. In particular, there is a jump in income at the utility cutpoint. This follows from the fact that an agent with ability q^* is indifferent between joining the X and Y sectors but receives strictly more job training if she enters X. In order for her utility to be equal in the two sectors while her utility cost of training is strictly higher in the X sector, her income must be strictly greater in the X sector than in the Y sector.

5. Small country

In this section I begin to investigate the interaction between trade and imperfect institutions, focusing on the distributional effects of institutional development in a small, price-taking country. When a small country opens to trade, it takes the world price, p^w, as given; i.e., its relative supply curve, RS, faces a flat world-relative-demand curve, RD^w, as in Fig. 2.

The effect of a change in institutional quality on a small open economy that does not completely specialize can be understood easily using a relative-supply–relative-demand graph, as in Fig. 2. If a country is not completely specialized, an improvement in national institutions has two effects on national production. First, it increases the fraction of the country's agents who work in the X sector because $q_\theta^* < 0$, as proven in Proposition 1. Second, it increases the efficiency of every firm in the X sector because $x_\theta > 0$. Both of these effects lead to an inward shift in the country's relative supply curve, in $(Y/X, p)$ space, from RS_0 to RS_1. I graph such an inward shift of the relative supply curve in Fig. 2.

Changes in institutional quality affect agents' educations, incomes, and utilities. Let $\Phi_k \equiv \frac{k(q_{\max})}{k(q_{\min})}$, where $k \in \{\text{income, job training, utility}\}$; Φ_k measures inequality in the distribution of k. For the purpose of exposition, suppose that k denotes income. Then $\Phi_{\text{income}} = \frac{I(q_{\max})}{I(q_{\min})}$ is the ratio of the income of the most able agent to that of the least able agent. When the difference between the incomes of the most and least skilled workers is small, Φ_{income} is close to 1; as the difference between the incomes grows, Φ_{income} increases. Unlike a Gini coefficient, Φ_{income} ignores all but the extremes of the income distribution.

Proposition 3. *In a small open economy for which $q^* \in (q_{min}, q_{max})$, a marginal increase in institutional quality, θ, generates a Pareto improvement. Although no agents are made worse off in absolute terms by the increase in θ, the measure of inequality, Φ, increases with a marginal improvement in θ.*

Proof. Choose θ such that $q^* \in (q_{\min}, q_{\max})$. q^* is a continuous, decreasing function of θ, implying that a marginal increase in θ leads to a marginal decrease in q^*. Let θ increase marginally. Then the new utility cutpoint, $\hat{q}^*$, satisfies $\hat{q}^* \in (q_{\min}, q^*)$. Any agent who originally chose the X sector–any agent with $q \geq q^*$–and any agent who switches into the X sector–any agent with $q \in (\hat{q}^*, q^*)$–experience increases in k. Hence, k increases for any agent with natural ability greater than the new utility cutpoint, $\hat{q}^*$. This implies that $k(q_{\max})$ increases. When prices are fixed, any agent who remains in the Y sector experiences no change in k. Thus, a marginal increase in θ generates a Pareto improvement and an increase in the measure of inequality. □

Proposition 3 states that for a small open economy that is not completely specialized, an improvement (deterioration) in institutions causes the ratio of the educations, incomes, and utilities of the most to the least skilled workers to increase (decrease). However, when national institutions improve, the increase in Φ_k is caused by an increase in k for the most skilled agents and not a decrease in k for the least skilled agents. This explains why institutional development generates a Pareto improvement.

Corollary 1. *In a small open economy for which $q^* > q_{max}$, a marginal increase in θ has no effect. In a small open economy for which $q^* < q_{min}$, a marginal increase in θ is Pareto improving and increases k for every agent in the economy.*

If a small open economy were completely specialized in the Y sector, a marginal improvement in institutional development would have no effect whatsoever on any agent's education, income, or utility.

6. Two large countries

Suppose that there are two countries, N (North) and S (South), that are identical in all respects except North has better, although still imperfect, institutions: $\theta^S < \theta^N < 1$. When these countries trade freely, the allocation of talent across sectors within a country and the pattern of trade across countries are determinate:

Proposition 4. *Suppose countries N and S open to free trade.*

1. *In country $J = N, S$ there exists a unique $q^*(\theta^J)$, where $q^*_{\theta,J}(\theta^J) < 0$, such that agents enter the X sector if and only if $q > q^*(\theta^J)$.*
2. *$q^*(\theta^N) \leq q_{max}$ while $q^*(\theta^S) \geq q_{min}$.*[11]
3. *Country N will export good X to country S in exchange for imports of good Y.*

Part 1 of Proposition 4 replicates Proposition 1 but for open economies. The intuition is the same as in Proposition 1: more talented agents work in the X sector because the return to education is greater there. Part 2 of Proposition 4 states that North must produce positive amounts of good X while South must produce positive amounts of good Y. Northerners are more efficient in team production than their equally able Southern peers. If Northerners optimally chose not to produce good X, the same would be true of Southerners. Because a positive amount of good X must be produced somewhere, the most skilled agents in North must work in the team production sector. Similar logic implies that the least skilled agents in South must work in the individual production sector.

[11] I prove Parts 1 and 2 of Proposition 4 in the Appendix.

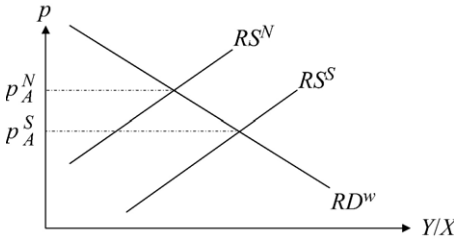

Fig. 3. Autarky equilibrium in North and South.

Part 3 of Proposition 4 goes further and states that national institutions act as an independent source of comparative advantage. When North and South are otherwise identical, superior national institutions give North a comparative advantage in team production. North exports the team production good, X, to South in exchange for South's exports of good Y. I include the proof of Part 3 of Proposition 4 below.

Proof. There are two countries–North, N, and South, S –with identical distributions of natural ability and identical technologies. National institutions are more developed in North than they are in South. More developed institutions in N imply that at any fixed price, a larger fraction of the workforce in N than in S will be employed in the X sector (recall from Part 1 of Proposition 4 that $q^*_{\theta^J}(\theta^J)<0$). Also, at any fixed price, a team of skill q in N produces more than a team of skill q in S (because $x_\theta>0$). These two facts imply that N will have a relative supply curve that is everywhere inside that of S in (Y/X, p) space, as in Fig. 3. The countries share the same relative demand curve because preferences are identical and homothetic. This implies that the autarky price in N, p_A^N, is strictly greater than the autarky price in S, p_A^S. I graph the relative demand facing both countries, RD^{w}, and each country's relative supply curves in Fig. 3. With free trade the world price will fall somewhere between the two autarky prices. N will export good X (and S will export good Y). □

Since the relative price of good Y falls in N, moving from autarky to free trade pushes N towards specializing in team production. Conversely, moving from autarky to free trade pushes S towards specializing in individual production. These changes in price and movements towards specialization that accompany trade liberalization have powerful consequences on the distributions of job training, income, and utility in both countries.

6.1. Job training, income, and utility distributions

In what follows I focus on the effects that trade liberalization has on the intra- and international distributions of job training when neither country completely specializes. Similar arguments lead to qualitatively similar conclusions when considering income or utility instead of job training. International trade influences job training decisions through its effect on relative prices. Recall that when the countries move from autarky to free trade, the relative price of the individual production good, Y, decreases in North and increases in South. Job training in the Y sector increases with the relative price of good Y while job training in the X sector decreases with the relative price of good Y.

Suppose that neither N nor S is completely specialized when trade is free. After liberalizing trade, the most (least) talented agents in North will increase (decrease) their levels of job training. Trade liberalization causes Φ_{educ} to increase in North. Consider a highly talented Northern individual who

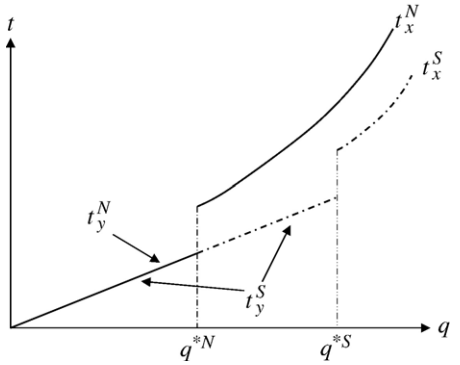

Fig. 4. Free trade training in N and S (dashed curved) in Y and X.

was in the team production industry in autarky. Because trade decreases the cutpoint in North, this agent remains in the X sector after trade opens. Because this agent does not switch industries when trade opens, it is easy to see that he chooses to receive more education when trade is free. Trade liberalization increases the relative price of his output. The increased price of his output increases the marginal benefit of his training and strengthens his incentives to invest in education. Next, consider a Northern agent with little natural ability who works in the individual production industry under free trade. Because the Northern cutpoint decreases with trade liberalization, this individual must have been in the individual production industry in autarky. Because this agent does not switch industries when trade opens, it is easy to see that he chooses to receive less education when trade is free. Trade liberalization reduces the relative price of his output. The decreased price of his output reduces the marginal benefit of his training and weakens his incentives to invest in education.

Recall that if institutions are perfect, trade liberalization has no effect on job training. However, if institutions are imperfect, trade liberalization fattens the tails of the education-distribution function in North. This implies that trade liberalization increases inequality in North, at least according to the Φ measure of inequality. On the other hand, when South liberalizes trade, the price of good Y increases. The most (least) talented Southerners will decrease (increase) their levels of job training. In South, trade liberalization compresses the distribution of education, yielding a more equitable distribution of education: Φ_{educ} decreases in South with the liberalization of trade. Trade liberalization increases education, income, and utility inequality in developed countries and decreases inequality in developing nations.

Under free trade, there exist Northerners who receive more job training than their equally skilled Southern peers while there do not exist any Southerners who receive more education than their equally skilled Northern peers. Equally able agents who work in the individual production industry receive the same amount of training independent of their country of origin: type $q < q_{\text{FT}}^{*N}$ have the same training level in S and N as in Fig. 4. With free trade, North's institutional advantage does not confer its benefits to those Northerners who work in individual production any more than it does to Southerners. As long as institutional differences are not too large, free trade allows low skill Southerners to catch up with low skilled Northerners. On the other hand, any Northerner in the team production sector receives more training than his equally able Southern counterpart: type $q > q_{\text{FT}}^{*N}$ agents have more training in N than in S as in Fig. 4.

Even when trade is free, technology is identical, and job training is equally obtainable across countries, a country with better moral-hazard-reducing institutions will have a better trained workforce and higher real GDP per capita.

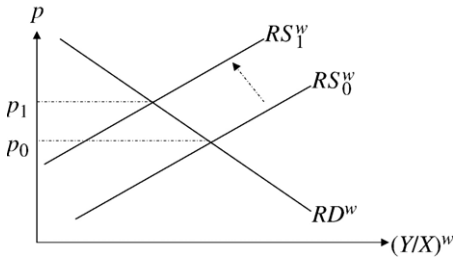

Fig. 5. An increase in θ^S causes an inward shift in RS^w.

Although there are benefits to trade liberalization for both countries (in the aggregate), free trade does not bring about complete convergence–in education, in income, or in utility. If underdeveloped countries are to become developed, they must improve the quality of their institutions. Interestingly, I have shown that improvements in institutional quality bring about increased inequality (at least in the small country case). According to this model, inequality is a natural consequence of the development process. However, this apparently harsh conclusion of the model is mitigated by the fact that improvement at the top of the distribution rather than deterioration at the bottom causes the increase in inequality.

6.2. International spillovers from institutional change

If trade is free, improvements in Southern institutions affect Northerners. If institutions improve in South, then the relative price of the individual production good increases in North. Similarly, when South liberalizes trade with North, the relative price of the individual production good increases in South. Because trade's effects work through its influence on prices, if trade is free and institutions improve in South, the effect on North is qualitatively the same as the effect on South of liberalizing trade with North.

If institutions improve in South, then at a fixed price production of good X increases in South (both because the utility cutpoint decreases and because each team in the X sector becomes more productive) and production of the individual production good decreases in South (both because the utility cutpoint decreases and because it is the most productive agents in the Y sector who move to the X sector). In Fig. 5, these changes lead to an inward shift (in $(Y/X, p)$ space) of the world's relative supply curve from RS_0^w to RS_1^w, causing the equilibrium relative price of good Y to increase from p_0 to p_1.

Because a Southern improvement in institutional quality increases the relative price of the individual production good, the most (least) talented Northerners would experience a decrease (increase) in their educations, incomes, and utilities: Φ^N decreases as θ^S increases. This result follows from the fact that the education, income, and utility of an agent in the team production sector is decreasing in the relative price of the individual production good. This is a standard conclusion in trade theory in which owners of the input used intensively in the sector in which the country has a comparative advantage prefer not to have their comparative advantage eroded.

7. Conclusion

In an economy in which managerial training mitigates moral hazard, job training yields greater rewards in team production sectors. More able agents, those for whom acquiring education is less

costly, self-select into the team production industry while less-able agents self-select into the individual production industry. When not only agents but also countries differ in their abilities to ameliorate moral hazard, a parallel international self-selection occurs. Countries with developed institutions specialize in the team-production industry while countries with inferior institutions specialize in the individual production industry.

The interaction between international trade and institutional quality has important consequences for the distributions of education and income. As a country with developed institutions liberalizes trade with countries that have weaker institutions, the developed country can anticipate a growing polarization in its distribution of education. The most talented agents will obtain comparatively more education while the least talented will obtain comparatively less. Trade liberalization also increases income inequality in developed countries, directly by increasing the relative price of the team production goods and indirectly through the effect of trade liberalization on education. In developing countries, the direct and indirect effects of trade on income inequality are reversed: both tend to decrease income inequality.

Agents in institutionally developed nations tend to obtain more education and earn higher incomes than their equally skilled counterparts in nations with less developed institutions. While training levels are lower in underdeveloped nations, this model suggests that the correct policy prescription is not necessarily to increase education. The return to education is lower in South than in North because South's institutions are inferior. Unless institutions improve in South, it is optimal for Southerners to receive relatively less education than Northerners. Institutional development appears to be the key to economic development, even though increased inequality is a natural consequence of this development.

Endogenizing institutional development presents an interesting extension to this model. Of particular interest is the interaction between international trade and the incentive to develop institutions. Corollary 1 provides insight into an extreme example. If a country has sufficiently underdeveloped institutions, then it may specialize in the individual production industry if it liberalizes international trade. In this case, there is no marginal benefit to institutional development; with no one working in the team production industry, no one benefits from marginal improvements in institutions. However, if this nation were in autarky, the marginal benefit to institutional development would be positive because the country would produce positive quantities of the team production good. Although this is a highly specialized example, it highlights the fact that trade, by altering the mass of agents who benefit from institutional development, influences the incentive to invest in institutions.

Appendix A. Owner-worked firms

In this section, I show that when institutions are imperfect, if a firm owner does not work in her own firm her profits are negative. If neither of a firm's employees is the owner, I need to make an assumption about the performance measure according to which the employees are paid. Recall that in an owner-managed firm, the owner/manager is not paid a salary or a bonus. Instead, she is the residual claimant to profit. Such an arrangement is not feasible if the manager is not also the owner because profit is neither observable nor verifiable by assumption. I assume that when the owner does not work in her firm, worker 1's performance measure is a function of worker 2's training and vice versa. In effect, I assume that both workers manage each other.

Suppose there is a firm in which the owner does not herself work. The owner's profit is greatest when she employs two workers with identical levels of training t. When neither employee is the firm

owner, both employees choose their efforts a and d according to Eq. (8): $a=Hbt$ and $d=Hb(1-\theta)$. The optimal bonus, chosen to maximize residual profit, is $b=\frac{1}{2}A^2$. At this bonus, a worker's post-education utility is $U=\frac{1}{8}H^2tA^2+Hw$. The owner will offer the lowest fixed wage, w, at which the contract is incentive compatible.

The owner offers the fixed wage that equates her employees' post-education utilities when working for her with their post-education utilities when working in (or owning) an owner-managed firm. This fixed wage equals $w=\frac{1}{8}HtA(4-A)$. Thus, the owner's costs are $\frac{7}{8}HtA^2+\frac{1}{2}HtA$; her revenues are $x=2a=HtA^2$; and her residual profit is $\frac{1}{8}HtA(A-4)$. This residual profit is strictly negative. No agent will choose to own a firm that she does not herself manage.

Appendix B. Proof of Proposition 1

I prove Proposition 1 in three steps.

First I prove that if there exists a q^* such that $U^x(q^*)=U^y(q^*)$, then this q^* is unique and for all $q>q^*$ agents enter the X sector while for all $q<q^*$ agents enter the Y sector. Suppose that $\theta<1$. Then $U^x\geq U^y \Leftrightarrow 2H^2t_xA^3\geq q(Hp\lambda)^4 \Leftrightarrow$

$$A^4\left(\frac{t_x^2+2(1-\theta)^2}{t_x^2+(1-\theta)^2}\right)\geq(p\lambda)^2. \tag{16}$$

The first equivalence simply substitutes for the utilities using Eqs. (12) and (13). The second follows from the equilibrium relation between q, t_x, and A defined in Eq. (11). The left-hand side of Eq. (16) is strictly increasing in q while the right-hand side is constant. Hence, if there exists a q^* at which $U^x(q^*)=U^y(q^*)$, then this q^* is unique. Moreover $U^x(q)>U^y(q)$ if and only if $q>q^*$.

Second, I show that $q^*(\theta)$ exists and $q^*(\theta)\in[q_{\min},q_{\max}]$. To obtain a contradiction, suppose that $q^*(\theta)<q_{\min}$. This assumption implies that $U^x(q_{\min})>U^y(q_{\min})$, which implies that $U^x(q)>U^y(q)$ for all $q\in[q_{\min}, q_{\max}]$. Hence, no one enters the Y sector. If no one enters the Y sector, no good Y is produced and p goes to infinity. If p approaches infinity, $p\lambda$ approaches infinity for any fixed $\lambda>0$. This contradicts our assumption that $q^*(\theta)<q_{\min}$. Showing that $q^*(\theta)\leq q_{\max}$ follows similar logic. Hence, for any $\theta<1$, there exists a $q^*(\theta)\in[q_{\min},q_{\max}]$ such that $U^x(q^*(\theta))=U^y(q^*(\theta))$.

Finally, I prove that $q^*_\theta<0$. It is straightforward to show that the utility cutpoint is decreasing in the quality of national institutions. Since U^x is increasing in θ for all q while U^y is independent of θ, the better are a country's national institutions, the larger the measure of agents employed in the X sector; i.e., $q^*_\theta<0$. □

Appendix C. Proof of Proposition 2

To prove Part 1 of Proposition 2, I require two lemmas. In the first lemma, I show that there exists a $q(\theta)$ such that $t_x(q)$ is convex if and only if $q>q(\theta)$. In the second lemma, I define q^{**} to be the greatest natural ability at which an agent would get the same amount of education if in the X or Y sector: $t_x(q^{**})=t_y(q^{**})$. I show that q^{**} exists and that $q^{**}<q^*$ for all $\theta<1$. I then combine Lemmas 1 and 2 to prove Part 1 of Proposition 2.

Lemma 1. *For any $\theta<1$ there exists a $q(\theta)$, where $q'(\theta)<0$ and $\lim_{\theta\to 1}q(\theta)=0$, such that t_x is strictly convex in q if and only if $q>q(\theta)$.*

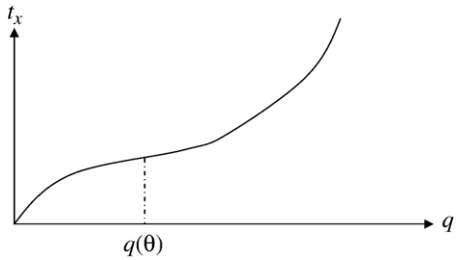

Fig. 6. Training in the X sector as a function of ability.

Proof. $t'' = \frac{-t}{(qU_{tt})^3}\beta$, where $\beta = \left[-2U_{tt}^2 - \frac{2}{q}U_{tt} + \frac{t}{q}U_{ttt}\right]$. Therefore, training in the X sector is convex in q if and only if $\beta > 0$ because $\frac{-t}{(qU_{tt})^3} > 0$. It remains to find the condition under which $\beta > 0$. Substituting U_{tt} and U_{ttt} into β and using Eq. (11) to substitute out all q terms yields

$$\beta = \left[\frac{H^4A^2(1-\theta)^2}{4\left(t^2 + (1-\theta)^2\right)^4}\right]\left(t^4 - 10t^2(1-\theta)^2 - 8(1-\theta)^4\right).$$

Because the term in square brackets is positive, $\beta > 0 \Leftrightarrow t^4 - 10t^2(1-\theta)^2 - 8(1-\theta)^4 > 0$. Hence,

$$\beta > 0 \Leftrightarrow t_x^2 > t_0^2 \equiv \frac{1}{2}(1-\theta)^2\left(10 + \sqrt{132}\right). \tag{17}$$

The fact that t_x* is strictly increasing in q implies if there exists a $q(\theta)$ at which $t_x^2 > t_0^2$, then job training is strictly convex in q for all $q > q(\theta)$.

It remains to prove that there exists a $q(\theta)$ at which $t_x^2 > t_0^2$. I prove the existence of such a $q(\theta)$ by substituting $t_x^2 = t_0^2$ into the job-training first-order condition for utility maximization. If there exists a $q(\theta)$ at which $U_t(t_x^2 = t_0^2) > 0$ then I know that at that $q(\theta)$ the optimal level of job training satisfies $t_x^*(q(\theta))^2 > t_0^2$, since utility is concave in job training. $U_t > 0$ at $t_x^2 = t_0^2$ if and only if $q > q(\theta) \equiv (1-\theta)K$, where K is defined by the expression: $K \equiv \frac{2}{H^2}\left(\frac{(6+\sqrt{33})^{3/2}}{7+\sqrt{33}}\right) > 0$.

Finally, $q'(\theta) < 0$ and $\lim_{\theta \to 1} q(\theta)$.[12] For all $\theta < 1$ the graph of optimal job training in the X sector as a function of natural ability is shown in Fig. 6.

I have shown that for all $\theta < 1$ there exists a $q(\theta)$, with $\lim_{\theta \to 1} q(\theta) = 0$ and $q'(\theta) < 0$, such that t_x^* is strictly convex in q for all $q > q(\theta)$ and t_x is strictly concave in q for all $q < q(\theta)$. □

Lemma 2. *Define q^{**} to be the greatest natural ability at which an agent would get the same amount of training if in the X or Y sector: $t_x(q^{**}) = t_y(q^{**})$ and $t_x(q) \neq t_y(q)$ for all $q > q^{**}$. Then $q^{**} > 0$ exists and $q^{**} < q^*$ for all $\theta < 1$. Moreover, $t_x(q) < t_y(q)$ for sufficiently low q.*

Proof. First, I prove that $q^{**} < q^*$, if q^{**} exists. Suppose that q^{**} exists. To obtain a contradiction, suppose that $q^{**} \geq q^*$ for some $\theta < 1$. By Proposition 1, $q^{**} \geq q^*$ if and only if $U^x(q^{**}) \geq U^y(q^{**})$, where U^y and U^x are defined in Eqs. (4) and (10). Substituting $t_x(q^{**}) = t_y(q^{**})$, which is true

[12] If $V = a_2 + \frac{f(\theta)}{t_1}d_2$ then this all holds except now $q(\theta) \equiv f(\theta)\,K$. Of course, in this case it is still true that $q' < 0$ and $\lim_{\theta \to 1} q(\theta) = 0$.

by the definition of q^{**}, into the equation for U^x, where $t_y(q)$ is defined in Eq. (5), implies that $U^x(q^{**}) \geq U^y(q^{**}) \Leftrightarrow A^3(q^{**}) \geq (p\lambda)^2$. At $t(q^{**})$ the marginal cost to education is equal in the X and Y sectors for a type q^{**} agent (since the total cost is $\left(\frac{1}{2}q^{**}\right)t(q^{**})^2$ in both industries). The fact that education is optimally chosen implies that the marginal benefits must also be equal; i.e., $(p\lambda)^2 = A(q^{**})\left(\frac{t(q^{**})^2+2(1-\theta)^2}{t(q^{**})^2+(1-\theta)^2}\right)$.Hence,

$$U^x(q^{**}) \geq U^y(q^{**}) \Leftrightarrow A^3(q^{**}) \geq (p\lambda)^2 \Leftrightarrow A^3(q^{**}) \geq A(q^{**}) \times \left(\frac{t(q^{**})^2+2(1-\theta)^2}{t(q^{**})^2+(1-\theta)^2}\right) \Leftrightarrow \frac{t(q^{**})^2}{t(q^{**})^2+(1-\theta)^2} \geq \frac{t(q^{**})^2+2(1-\theta)^2}{t(q^{**})^2+(1-\theta)^2}$$

The second equivalence follows from the fact that if education is optimally chosen to be equal in both industries then the marginal benefits must be equal in both industries. The final equivalence follows from the definition of $A(q^{**})$. The final relationship is clearly violated for all $\theta<1$. Hence, $U^x(q^{**})<U^y(q^{**})$. By Proposition 1 this implies that if q^{**} exists, then $q^{**}<q^*$.

To complete the proof of Lemma (2), it must be shown that $q^{**}>0$ exists. To obtain a contradiction, suppose q^{**} does not exist. If q^{**} does not exist, then either $t^x(q)>t^y(q)$ or $t^x(q)<t^y(q)$ for all $q>0$. First suppose that $t^x(q)>t^y(q)$ for all q. Substitute in $t^y(q)$ from Eq. (5) for $t^x(q)$ on the right-hand side of Eq. (11). Then $t^x(q)>t^y(q)$ for all q implies that

$$A\frac{\left(t_x^*(q)\right)^2+2(1-\theta)^2}{\left(t_x^*(q)\right)^2+(1-\theta)^2}>(p\lambda)^2 \text{ for all } q>0. \tag{18}$$

The right-hand side of Eq. (18) is strictly greater than zero. The left-hand side of Eq. (18) is continuous in q and approaches zero as q approaches zero. Hence, there exists a range of q above $q=0$ in which Eq. (18) is necessarily violated. This implies that $t^x(q)<t^y(q)$ for sufficiently low q. Second, suppose that $t^x(q)<t^y(q)$. Again, substitute in $t^y(q)$ from Eq. (5) for $t^x(q)$ on the right-hand side of Eq. (11) and assume that $t^x(q)<t^y(q)$ for all q. Then the inequality in Eq. (18) is reversed, which implies that $U^x(q)<U^y(q)$ for all q. This yields another contradiction. Therefore, q^{**} exists.

I use Lemmas 1 and 2 to prove Part 1 of Proposition 2.

First I prove that $t_x(q)$ is convex in q for all $q \geq q^{**}$. I proved in Lemma 2 that $t_x(q)<t_y(q)$ for all q in some range above $q=0$. In Lemma 1 I proved that $t_x(q)$ must be concave in some region above $q=0$. Moreover, both t_x and t_y approach zero as q approaches zero. Hence, in the region over which $t_x(q)$ is concave, $t_x(q)$ never intersects $t_y(q)$ at any $q>0$. In Lemma 2 I proved that there exists a q^{**} such that $t_x(q^{**})=t_y(q^{**})$. Thus, t_x must be convex in q at q^{**}.

Here I prove that $t_x(q)>t_y(q)$ for all $q \geq q^*$. To do so, I show that $t_x(q)>t_y(q)$ for all $q>q^{**}$. By definition, $t_x(q^{**})=t_y(q^{**})$. The fact that q^{**} is the greatest q at which $t_x(q)=t_y(q)$ implies that either $t_x(q)>t_y(q)$ or $t_x(q)<t_y(q)$ for all $q>q^{**}$. Because $t_y(q)$ is linear in q for all q while $t_x(q)$ is convex in q for all q greater than q^{**}, I find that $t_x(q)>t_y(q)$ for all $q>q^{**}$. To complete the proof I combine the fact that $t_x(q)>t_y(q)$ for all $q>q^{**}$ with the fact that $q^*>q^{**}$ from Lemma 2 to prove that $t_x(q)>t_y(q)$ for all $q \geq q^*$.

Finally, I prove Part 2 of Proposition 2.

Suppose that $\theta<1$. Then from the definitions of $I^y(q)$ and $I^x(q)$ in Eqs. (14) and (15),

$$I^x(q)>I^y(q) \Leftrightarrow t_x A>(1/2)H^2 q(p\lambda)^4. \tag{19}$$

Setting $U^x(q^*)=U^y(q^*)$ implies that $q^* = \frac{2t_x(q^*)A^3(q^*)}{H^2(p\lambda)^4}$. Substituting this into Eq. (19) yields $I^x(q^*)>I^y(q*) \Leftrightarrow 1>A^2(q^*)$, which is satisfied for all $q^* \leq q_{\max}$ when $\theta<1$. Hence, $I^x(q^*)>I^y(q^*)$ for all $\theta<1$.

It can be shown that there exists a unique q, denoted q^{***}, at which $I^x(q^{***})=U^y(q^{***})$ and that $I^x(q)>I^y(q)$ for all $q>q^{***}$, by using a similar argument to the one that showed that for all $\theta<1$ there exists a unique q, q^*, at which $U^x(q^*)=U^y(q^*)$ and that $U^x(q)>U^y(q)$ for all q greater than q^*. The fact that $I^x(q^*)>I^y(q^*)$ implies that $q^*>q^{***}$ and, thus, that there is an income jump at q^* for all $\theta<1$. □

Appendix D. Proof of Proposition 4

First I prove Part 1 of Proposition 4.

For notational simplicity, I drop the J superscript. The proof that if there exists a q^* such that $U^x(q^*)=U^y(q^*)$ then this q^* is unique, and that agents enter the X sector if and only if $q>q^*(\theta^J)$ is exactly the same as the proof of Proposition 1.

I show that such a q^* must exist. From the proof of Proposition 1: $U^x \geq U^y \Leftrightarrow$

$$A^4\left(\frac{t_x^2+2(1-\theta)^2}{t_x^2+(1-\theta)^2}\right)\geq(p\lambda)^2. \tag{20}$$

As $q\to 0$ the left-hand side of Eq. (20) approaches zero because $\lim_{q\to 0} t_x=0$ and $\lim_{t_x\to 0}A=0$. As $q\to\infty$ the left-hand side of Eq. (16) approaches 1 since $\lim_{q\to\infty}t_x=\infty$ (as proven in Lemma 2) and $\lim_{t_x\to\infty}A=1$. To obtain a contradiction suppose that q^* does not exist. Then $(p\lambda)^2>1$. However, if $(p\lambda)^2>1$ then no one enters the X sector in either country N or country S. If no one enters the X sector from either country, then the relative price of good Y goes to zero. If p approaches zero, then $(p\lambda)^2<1$, contradicting the assumption that q^* does not exist. The proof that $q^*_{\theta^J}(\theta^J)<0$ is exactly as in the proof of Proposition 1.

Next I prove Part 2 of Proposition 4.

I only prove the fact that $q^*(\theta^N)\leq q_{\max}$. The proof that $q^*(\theta^S)\geq q_{\min}$ follows a similar argument. To obtain a contradiction, suppose that $q^*(\theta^N)>q_{\max}$. $q^*(\theta^N)<q^*(\theta^S)$ because $q^*_\theta<0$ and $\theta^N>\theta^S$. Thus, $q^*(\theta^N)>q_{\max}$ implies that $q^*(\theta^S)>q_{\max}$. If $q^*(\theta^N)>q_{\max}$ and $q^*(\theta^S)>q_{\max}$, then the relative price of good Y converges to zero, implying that $q^*(\theta^N)<q_{\max}$ and $q^*(\theta^S)<q_{\max}$. This contradicts our assumption. Therefore, $q^*(\theta^N)>q_{\max}$.

References

Acemoglu, Daron, Johnson, Simon, Robinson, James, 2001. Colonial origins of comparative development: an empirical investigation. American Economic Review 91, 1369–1401.

Acemoglu, Daron, Johnson, Simon, Robinson, James, 2002. Reversal of fortune: geography and institutions in the making of the modern world income distribution. Quarterly Journal of Economics 117, 1231–1294.

Birdsall, Nancy, Rodrik, Dani, Subramanian, Arvind, 2005. How to help poor countries grow. Foreign Affairs 84.

Costinot, Arnaud, 2004. Imperfect contract enforcement, firm size, and the pattern of trade. mimeo Princeton University.

Feenstra, Robert C., Hanson, Gordon H., in press. Global production sharing and rising inequality: a survey of trade and wages. In: Kwan Choi, James Harrigan, eds., Handbook of International Trade.

Grossman, Gene M., 2004. The distribution of talent and the pattern and consequences of international trade. Journal of Political Economy 112, 209–239.

Kremer, Michael, 1993. The O-Ring theory of economic development. Quarterly Journal of Economics 108, 551–575.

Lawrence, Robert Z., Slaughter, Matthew J., 1993. International trade and American wages in the 1980s: giant sucking sound or small hiccup? Brookings papers on economic activity. Microeconomics 161–210.

Mankiw, N. Gregory, Romer, David, Weil, David N., 1992. A contribution to the empirics of economic growth. Quarterly Journal of Economics 107, 407–437.

Matsuyama, Kiminori, 2005. Credit market imperfections and patterns of international trade and capital flows. Journal of the European Economic Association 3, 714–723.