International Trade, Factor Mobility and the Persistence of Cultural-Institutional Diversity[§]

Marianna Belloc[*] and Samuel Bowles[**]

This version: 16 January, 2010

Abstract

Cultural and institutional differences among nations may result in differences in the ratios of marginal costs of goods in autarchy and thus be the basis of specialization and comparative advantage, as long as these differences are not eliminated by trade. We provide an evolutionary model of endogenous preferences and institutions under autarchy, trade and factor mobility in which multiple asymptotically stable cultural-institutional conventions may exist, among which transitions may occur as a result of decentralized and un-coordinated actions of employers or employees. We show that: *i*) specialization and trade may arise and enhance welfare even when the countries are identical other than their cultural-institutional equilibria; *ii*) trade liberalization does not lead to convergence, it reinforces the cultural-institutional differences upon which comparative advantage is based and may thus impede even Pareto-improving cultural-institutional transitions; and *iii*) by contrast, greater mobility of factors of production favors decentralized transitions to a superior cultural-institutional convention by reducing the minimum number of cultural or institutional innovators necessary to induce a transition.

JEL CODES: C73, D23, F15, F16

KEYWORDS: institutions, partnership, evolutionary game theory, culture, trade integration, factor mobility, globalization

---

[§] We would like to thank Pranab K. Bardhan, Ennio Bilancini, Eric Brousseau, Arnaud Costinot, Simone D'Alessandro, Michele Di Maio, Francesco Drago, Maurizio Franzini, Roberto Galbiati, Pierre Garrouste, Peter Grajzl, Sung Ha Hwang, Eric Nilsson, Ugo Pagano, Emmanuel Raynaud, Robert Rowthorn, Massimiliano Tancioni, Anne van Aaken, Pietro Vertova, Peyton Young, and Giulio Zanella for helpful comments. Thanks also to the Behavioral Science Program of the Santa Fe Institute and the U.S. National Science Foundation for support of this project. The mathematical appendix for this paper will be posted on the authors' website.

[*] Sapienza University of Rome (marianna.belloc@uniroma1.it).

[**] Santa Fe Institute and University of Siena (samuel.bowles@gmail.com).

1. Introduction

Among history's great puzzles are the many instances of centuries-long persistence of institutional and cultural differences between populations, often enduring long after their initial causes have disappeared. Institutions and elite cultures that owed their origin to the 16th century exploitation of slaves and coerced Native American labor in the mines and fields persisted long after sugar and gold had lost their central role in the Latin American economies (Sokoloff and Engerman, 2000). Current levels of distrust in distinct African populations reflect the enduring effects of variations in exposure to the slave trade that ended two centuries ago (Nunn and Wantchekon, 2008). Differing levels of cooperation and civic values among Italian urban areas appear to be the legacy of autonomous city-state institutions or their absence half a millennium earlier (Guiso, Sapienza, and Zingales, 2009). The effects of the differing tax and land tenure systems imposed by the British Raj in the 18th and 19th century persisted in post-Independence India (Banerjee and Iyer, 2005).

In epochs and social orders marked by limited contact and restricted competition among geographically separated areas, persistent cultural and institutional differences are hardly surprising. But this is not the case in a globally integrated world economy. In this paper we explain how the decentralized updating of both preferences and contractual choices can support durable cultural and institutional differences that may provide a basis for specialization, comparative advantage, and hence trade, which in turn stabilizes the cultural and institutional differences. Our model hinges on the codetermination of institutions, cultures, and economic specialization, a nexus long-studied by economists with a historical bent (Gerschenkron, 1944; Kindleberger, 1962; Sokoloff and Engerman, 2000), but only recently modeled by economic theorists (Costinot, 2009; Bardhan, Mookherjee, and Tsumagari, 2009; Levchenko, 2007; Olivier, Thoenig, and Verdier, 2008).

We develop a two-good/two-country model with endogenous preferences and institutions in which employee-employer relations are shaped by social norms and governed by either joint residual claimancy under share contracts (partnerships) or forcing contracts. Goods differ in the extent to which their production depends on what we define below as qualitative labor, namely that which is not verifiable and hence cannot be cost effectively ordered by explicit contracts over labor input. Where non-verifiable aspects of labor are important to production, social norms such as reciprocity or a positive work ethic may be required for high levels of productivity. We refer to differences across economies in the kinds of contracts that are offered as institutional differences, while between-economy variations in preferences (including social norms) are termed cultural differences.

The main novelty of our approach (one shared with Greif, 1993, 1994, 2002, Galor and Moav, 2002, and Doepke and Zilibotti, 2008) is that, rather than treating institutions and preferences as exogenous or determined by a national-level constitutional bargain, we model the interacting dynamics of both as the result of decentralized non-cooperative interactions among economic agents. Like Guiso, Sapienza and Zingales (2009), Tabellini (2008) and Spolaore and Wacziarg (2008), we study the economic importance of cultural differences and model cultural evolution (Bowles, 1998; Bisin and Verdier, 2001; Fershtman and Bar-Gill, 2005).

In our model, the optimal form of contract to offer depends on the preferences which prevail in a given country. Partnership contracts, for example, are more profitable where social preferences like the work ethic or reciprocity are common. The distribution of preferences in turn is based on a cultural updating process in which the payoffs to different preferences (and the behaviors they support) depend on the distribution of contracts in the economy. It is this mutual dependence of preferences and contracts and the differences among goods in the extent that monitoring and hence forcing contracts are cost-effective that supports the multiplicity of equilibria in our model and the resulting country differences in comparative advantage. The strategic complementarity of preferences and contracts that supports multiple equilibria in our model thus plays a role analogous to technology-based economies of scale in Paul Krugman's (1987) model of intra-industry trade among countries with identical factor endowments and technologies.

Transitions may occur among these cultural-institutional conventions when behavioral or contractual innovators deviate from the status quo convention due to individual experimentation and other forms of idiosyncratic play. We derive three key results.

First, for historical reasons two otherwise identical countries may experience different cultural-institutional conventions, and these cross-country differences in the institutional and cultural environment, like differences in technologies in the Ricardian approach or factor endowments in the standard Heckscher-Ohlin model, are an independent source of comparative advantage. In the absence of idiosyncratic play, a nation's cultural-institutional convention may persist indefinitely even when a Pareto-superior convention exists and when the status quo convention confers absolute disadvantage with respect to other countries in all goods. The source of persistent inefficiency in this model is the coordination failure arising from the decentralized nature of preference formation and contractual choice.

Second, economic integration reinforces rather than destabilizes institutional and cultural diversity and may impede transitions, even to Pareto-improving conventions. Our second result contradicts the view, popular among critics of trade liberalization since John Maynard Keynes (1933), that trade will lead to institutional and cultural convergence. This is especially thought to be true when one nation's cultural-institutional equilibrium confers absolute advantage in both products. But since trade allows countries to specialize in the goods that are relatively more advantaged given their institutions and preferences, it increases the joint surplus in the cultural-institutional status quo, thereby also raising the cost of local deviations from the prevalent preferences and contracts and thus making cultural-institutional transitions less likely. It may also increase the number of behavioral or contractual innovators required to induce a transition to the superior convention.

Third, in contrast to trade, factor market integration facilitates convergence to superior culture and institutions. The reason is that factor mobility lowers the expected costs of deviating from one's nation's status quo and reduces the minimum number of innovators necessary to induce Pareto-improving cultural-institutional transitions. Factor market integration thus reduces both the size and (loosely speaking) the depth of the basin of attraction of the inferior equilibrium. We begin with the basic assumptions of our model and the empirical evidence motivating them (Section 2). We then develop a model of endogenous preferences and contractual choice and embed it in a standard 2x2 model of international exchange, illustrating cultural-institutional comparative advantage (Section 3). In section 4 we introduce the model's dynamics and show that multiple asymptotically stable cultural-institutional equilibria may exist. We then explore the persistence of cultural and institutional differences following trade integration (Section 5), and factor mobility (Section 6). Section 7 discusses related literature and concludes.

2. Goods, preferences and contracts

An economy is populated by employers and employees. Employers hire employees to produce one of the two goods, the employment relationship being a random employee-employer pair for a single interaction in which the employer offers a contract under which the employee works. Labor is perfectly mobile across industries but (initially) immobile across countries. Our model is based on four distinctive assumptions that we believe are of board empirical relevance.

First, there are two aspects of labor. Qua*N*titative labor (denoted by the subscript *N*) includes time at work, compliance with directions, simple effort readily measured either by input or output, and other aspects of work that are readily observable, either directly or that may be inferred from the associated outputs. By contrast, qua*L*itative work (denoted by the subscript *L*) consists of care, creativity, problem solving and other non-routine aspects of work that are not verifiable, and hence not cost-effectively subject to explicit contracts conditional on individual performance. Production of all goods requires quantitative labor and is also enhanced by qualitative labor (though, as we will see, in differing degree). Each employee may provide either quantitative labor alone or both quantitative and qualitative labor.

Second, there are two goods. One is intensive in quantitative labor and termed transparent (*t*-good) because the labor activities that are readily observed are relatively more important in its production. The production of the opaque good (*o*-good), by contrast, depends more intensively on qualitative aspects of work. Examples of the latter are knowledge-intensive goods (and services), personal services, and quality-sensitive agricultural products such as tobacco, many vegetables and fruits and wine. Transparent goods include standardized manufactured goods (exemplified by any good the production of which is cost effectively compensated by piece rates), most grains and sugar. Hence, denoting by Q_L^i the quantity of good i ($i = t,o$) obtained using one unit of both qualitative and quantitative labor, and by Q_N^i the output obtained with a single unit of quantitative labor only, we have:

$$\frac{Q_L^o}{Q_N^o} > \frac{Q_L^t}{Q_N^t}, \tag{1}$$

that is, the increase in production obtained employing quantitative and qualitative labor rather than quantitative labor alone is relatively greater in the opaque than in the transparent sector.

Our third assumption is that some employees have preferences over the form of the contract under which they work *per se*, that is, in addition to the material payoffs. For some individuals, close supervision and threats of sanctions for non-compliance signal distrust or otherwise offend reciprocal or other social preferences essential to mutually beneficial exchange. This is found in a large number of natural environments (Bewley, 1999) and experimental studies (Fehr, Klein and Schmidt, 2007; Falk and Kosfeld, 2006; surveyed in Bowles, 2008, and Bowles and Polania, 2009). We simplify by assuming just two kinds of employees. We term Reciprocators (denoted by the superscript *R*) those who care about the form of the contract *per se*: in a dyadic interaction their utility is increasing in their own payoffs and may be either increasing or decreasing in the payoffs of the other depending on the individual's belief about the type of the other, in the spirit of Rabin (1993), Levine (1998) and Fehr and Falk (2002). Individual *i*'s utility depends on his own material payoff including the disutility of labor (π_i) and the payoff of the other individual (π_j):

$$U_i = \pi_i + \alpha_i \gamma_{ij} \pi_j, \tag{2}$$

where α_i (>0 for Reciprocator and =0 otherwise) is the strength of *i*'s reciprocity preferences and γ_{ij} (= −1, 1) is *i*'s belief about *j*'s type, the latter depending on the form of contract that *j* offers *i*. In the model below a partnership (denoted hereafter by the superscript *P*) in which the employer and the employee are joint residual claimants on the firm's output and the employee is free to choose any type of work signals the good will and trust of the employer, leading to γ_{ij} = 1; while a Forcing contract (superscript *F*), under which the employer's close surveillance and the threat of termination is designed to implement a quantitative labor, signals distrust with γ_{ij} = −1 as a result.

Other individuals, who we will term Homo economicus (superscript *E*), care only about their own material payoffs (α_i=0) so that $U_i = \pi_i$. We refer to preferences of this kind as self-regarding. As we will see, from this it follows that social preferences such as a strong work ethic, truth telling and intrinsic motivation may be essential to the production of opaque goods, because Forcing contracts and other kinds of explicit incentives appealing to conventional self-regarding motives cannot elicit qualitative labor in the production of opaque goods due to the lack of verifiability of an essential input.

The final assumption is that while both cultures and institutions are endogenous, neither is the result of instantaneous individual maximization or collective choice. Rather they are durable characteristics of individuals and organizations that evolve in a decentralized environment under the influence of long-run society-wide payoff differences. Institutions and preferences are acquired and abandoned by a trial and error process often taken place at critical times, the birth of a firm, for example for contractual forms, or early childhood or adolescence for preference formation. Because childhood, socialization and the other processes by which preferences are acquired take place under the influence of religious values, schooling and other effects operating at the national level, we represent this process of cultural evolution by a society-wide dynamic operating prior to economic matching for production. Thus individuals do not condition their preferences on the kind of contract they are offered in any period; rather they best respond to the distribution of contracts in the previous period. Similarly firms do not condition their contractual offers on the type of the employees with whom they are paired in a given period; rather they best respond to the distribution of employee types (Reciprocator, Homo economicus) in the previous period.

The underlying idea motivating our model is that goods differ in the kinds of contracts that are appropriate for their production and that strategic complementarities between contracts and the nature of social norms may result in a multiplicity of cultural-institutional-specialization equilibria. The implied correspondence between preferences, contracts and specialization is widely observed. In Thailand the wholesale rice market approximates a standard economic textbook impersonal exchange among parties whose identity is effectively irrelevant to the transaction (Siamwalla, 1978). The raw rubber market, by contrast, is highly personal and is based on long-standing relationships of trust. The difference is explained by the fact that the quality of rice is readily assayed by the buyer, while the quality of raw rubber is impossible to determine when it is purchased. In the absence of trust among Thai buyers and sellers, trade in raw rubber would be more expensive. Raw rubber is an opaque good, rice is transparent.

Economic historians have used similar distinctions. Eric Nilsson (1994) studied the effects on comparative advantage and specialization resulting from the emancipation of slaves at the time of the U.S. Civil War. Cotton, according to Nilsson, was a “slave commodity” for which kinds of labor beyond that which could be coerced from the worker were of little importance. For other commodities – manufactures and tobacco in Nilsson’s empirical study – variations in the labor quality were more important, and impossible to secure by coercion. Nilsson exploited the natural experiment provided by the end of slavery to study the effect of an exogenous institutional shock on production specialization in 169 counties in the Confederacy. He found that the end of slavery brought about a significant shift away from the “slave commodity” (cotton) and towards manufactures and tobacco. Stefano Fenoaltea’s (1984) study of slave and non-slave production makes a similar distinction between “care intensive” and “effort intensive” productive activities the former being opaque in our terminology and the latter transparent. A similar distinction between sugar and tobacco was made in the much earlier study of Cuba by Fernando Ortiz (1963) who contrasted the coerced labor and hierarchical and authoritarian culture of the sugar plantation regions with the self-motivated labor and liberal culture of the tobacco family-farming areas.

Norms and preferences influencing economic behavior differ significantly among societies (Inglehart, 1977; Henrich, Boyd, Bowles et al., 2005). In particular, there is some evidence that reciprocal social preferences are more prevalent in the higher income countries. Among subjects in 15 countries, the level of cooperation sustained in a public goods experiment in which the altruistic punishment of free riders was possible was much higher in wealthier nations (Herrmann, Thoni, and Gaechter, 2008). For these reasons we represent an economy whose cultural-institutional equilibrium is characterized by partnerships and extensive social preferences such as trust and the positive work ethic as having a “good” cultural-institutional environment and, as a result, enjoying absolute advantage with respect to other countries in which forcing contracts and high levels of monitoring may elicit quantitative (but not qualitative) labor services from entirely self-regarding economic agents. This view is consistent with the observation that opaque goods make up a substantial fraction of the output of the more advanced economies (production and distribution of information-intensive goods and many services ranging from health care to entertainment and other recreational services), whereas poorer nations produce large shares of agricultural and manufactured goods that are closer to the transparent pole of the opaque-transparent continuum.

3. Cultural-institutional equilibrium under autarchy

Employers maximize profits, while employees maximize utility. Agents consume a composite bundle (indicated by the superscript c) of the two goods produced. For simplicity, we assume that the composite good is made up of one unit of the transparent and one unit of the opaque good, and prices have no effect on consumption proportions. Denoting by p^t and p^o the prices of the t-good and the price of the o-good, we define $\rho^o = p^o/(p^t+p^o)$ and $\rho^t = p^t/(p^t+p^o)$ respectively the value of the opaque good in terms of the composite good (how many units of the c-good one can purchase with one unit of the o-good) and the value of the transparent good in terms of the composite good (how many units of the c-good one can purchase with one unit of the t-good). Markets are competitive in the sense that employers take the price of the good as exogenously given.

The (risk neutral) utility function of employees is additive in consumption of the composite good, the subjective utility associated with the contract (for the reciprocal agents) and the disutility associated with the type of labor provided in production. Supplying quantitative labor incurs a cost η (>0), while supplying both quantitative and qualitative labor costs $\delta > \eta$.

The employer may offer the employee a Forcing (F) contract or a Partnership (P). The key difference between the two is that in the former the motivation to work is provided by the fear of being fired (as in many secondary labor market jobs), while in the second the primary motivation is gain sharing based on joint residual claimancy (as in many legal practices, financial consulting, and software design). We assume that the wage rate ($w > 0$) in the Forcing contract is the result of an economy wide bargaining process that results in a wage satisfying the participation constraint of all workers. Under the Forcing contract the employee is offered a fixed compensation, is closely monitored at a cost μ (>0) to the employer, and required to provide (at least) quantitative labor as a condition of being paid. Under the Partnership the "employee" is offered half of the revenue of the Partnership and selects any type of labor without supervision.

We have already introduced the two types of employees: the reciprocal type and the conventional Homo economicus. In the F-contract quantitative work is sufficient to secure compliance and therefore the payment of the wage so the E-type employees offer quantitative labor, incurring the associated disutility η. If offered a P-contract, the E-worker also provides quantitative labor only since the workers share of increased output associated with qualitative labor is less than the greater disutility required (i.e. $\rho Q_N^i / 2 - \eta > \rho Q_L^i / 2 - \delta$, with $i=t,o$). By contrast, reciprocal employees have preference on the contract that is offered by the employer *per se*. Under a Forcing contract the R-worker values the payoff of the employer negatively ($\gamma_{ij} = -1$), and so like the E-worker provides quantitative labor only (also at a cost η). Under the Partnership, however, the R-worker's positive valuation of the payoff to the partner ($\gamma_{ij} = 1$) is sufficient to offset the greater disutility of labor, and so the reciprocal type employee provides, in addition, qualitative aspects of work contributing to production (at a greater cost δ).

Table 1 reports the matrix of payoffs measured in number of units of the composite good commanded. Since in autarchic equilibrium both goods are produced in fixed proportions at the national economy level we may assume that all firms produce a joint product in the proportions given by the composite consumption good.

To exclude cases where cultural-institutional differences could not occur in equilibrium we assume that $\rho Q_N^i > 2(w+\mu)$ and $(1+\alpha)\rho Q_L^i/2-\delta > \rho Q_N^i/2-\eta$, i denoting the good. From these assumptions we know that $\{P,R\}$, that is, the Partnership contract matched with the reciprocal employee, is the joint surplus maximizing outcome. But that does not guarantee that $\{P,R\}$ will be observed in practice in a dynamic setting because the "inferior" convention $\{F,E\}$ is also asymptotically stable. We term $\{P,R\}$ and $\{F,E\}$ cultural-institutional equilibria (a third Nash equilibrium in mixed strategies exists and is unstable as we will see in Section 4).

	Employee/Preferences	
Employer/Contract	Reciprocator	Homo economicus
Partnership	$\boldsymbol{\rho Q_L^i/2,\ (1+\alpha)\rho Q_L^i/2-\delta}$	$\rho Q_N^i/2,\ \rho Q_N^i/2-\eta$
Forcing contract	$\rho Q_N^i - w-\mu,\ w-\eta-\alpha(\rho Q_N^i - w-\mu)$	$\boldsymbol{\rho Q_N^i - w-\mu,\ w-\eta}$

***Table 1*: Matrix of payoffs.** (NOTE: Payoffs in bold type indicate pure stable Nash equilibria)

Assume now that the world economy comprises two countries, 1 and 2, identical in all relevant respects (same relative labor endowment, same technology, same demand function, same economy wide wage bargaining process), except for their recent histories, which have given them different cultural and institutional conventions. Let us suppose that country 1 is near the $\{P,R\}$ equilibrium so that virtually all pairs are reciprocal types working under Partnership contracts, while country 2 is near the $\{F,E\}$ equilibrium. In Figure 1 we represent the production possibility frontiers of the two countries, the slope of the dashed lines indicating the international terms of trade. Because $Q_L^o > Q_N^o$ and $Q_L^t > Q_N^t$, the $\{P,R\}$ country enjoys an absolute advantage in the production of both goods. However, the two countries enjoy comparative advantage in the production of different goods. Country 1, where the established cultural-institutional equilibrium is able to elicit qualitative labor in all the employment relations, is superior in the production of both commodities, but has a greater advantage in the production of the *o*-good where qualitative aspects of work are relatively more important. By contrast, country 2 has a culture and institutions for which employees are willing to provide quantitative labor only; this country, as a consequence, has comparative advantage in the production of the *t*-good that is relatively less intensive in non-verifiable labor services. In autarchic equilibrium there will be only one relative price in each country such that both goods are produced: this price ratio must be equal to the marginal rate of transformation (*MRT*) in the two countries, namely: $p^o_{1A} / p^t_{1A} = Q^t_L / Q^o_L = MRT_1$ and $p^o_{2A} / p^t_{2A} = Q^t_N / Q^o_N = MRT_2$, where p^o_{1A} / p^t_{1A} and p^o_{2A} / p^t_{2A} are the autarchic relative prices in the two countries. Therefore, given (1), we have:

$$\frac{p^o_{1A}}{p^t_{1A}} = \frac{Q^t_L}{Q^o_L} < \frac{Q^t_N}{Q^o_N} = \frac{p^o_{2A}}{p^t_{2A}}. \tag{3}$$

Providing that the international terms of trade, p^o_T / p^t_T (the subscript "*T*" refers to trade), falls strictly between the autarchic relative prices of the two countries, specialization and trade will take place. Given the linearity of the two production possibility frontiers, country 1 will specialize entirely in the production of (and will export) the opaque good, while country 2 will specialize in the production of (and will export) the transparent good.

Unless the two economies happen to be of the "right" size, given the fixed proportions in the composite consumption good there will either be excess supply of one of the two goods under complete specialization following trade integration. To retain the valuable simplifications due to both complete specialization and fixed proportions in consumption we could (artificially, but harmlessly) assume that under trade integration the "smaller" nation specializes and that firms in the other country produce a joint product of the two goods in the proportions necessary to satisfy global demands for the two goods. We opt for the simpler assumption that the countries are of a size to equilibrate world commodity markets, thereby avoiding notation clutter associated with joint production in one country.

Compared to autarchy, trade benefits both classes of individuals in country 1 and employers in country 2. The resulting gains from trade are illustrated below. When cross-country barriers to trade are removed and in absence of transportation costs, the relative price of the opaque (transparent) good increases in country 1 (country 2), whereas the relative price of the transparent (opaque) good decreases. It follows that $\rho_T^o > \rho_{1A}^o$ and $\rho_T^t > \rho_{2A}^t$, where (recall that) $\rho^o = p^o/(p^o + p^t)$ and $\rho^t = p^t/(p^o + p^t)$: in both countries the good in which the country specializes becomes relatively more valuable in terms of the *c*-good (with one unit of the *o*-good (*t*-good) in country 1 (country 2) one can purchase a greater number of units of the *c*-good under trade than in autarchy). Thus, as expected, $\rho_T^o Q_L^o > \rho_{1A}^o Q_L^o$ and $\rho_T^t Q_N^t > \rho_{2A}^t Q_N^t$: the *c*-good value of output in the two countries increases. All the other terms (δ, η, w, μ and γ) in the payoff matrix (Table 1) are measured in units of the composite goods and so remain unaltered.

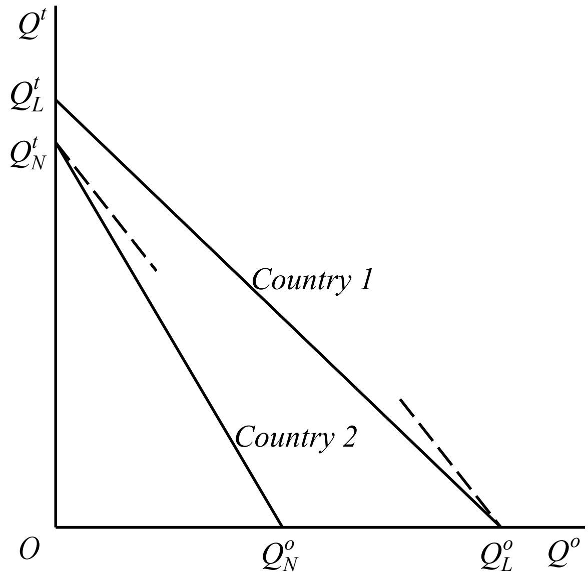

***Figure 1*: Production possibility frontiers in the two countries.**
(NOTE: Each country has a normalized labor endowment of 1)

4. Dynamics

We now study the asymptotic stability properties of the two conventions to provide a framework for understanding the impediments to transitions from one convention to the other. Writing the fraction of the employees who were Reciprocators in the previous period as ω and using the payoffs in Table 1, the expected payoffs to employers offering the *P*- and *F*-contracts are:

$$v_P = \omega \frac{\rho^i Q_L^i}{2} + (1-\omega)\frac{\rho^i Q_N^i}{2},$$
$$v_F = \omega[\rho^i Q_N^i - (w+\mu)] + (1-\omega)(\rho^i Q_N^i - w - \mu) = \rho^i Q_N^i - w - \mu. \tag{4}$$

Similarly, writing the fraction of the employers offering Partnership contracts in the previous period as ϕ, the expected payoffs to the R- and E-employees are respectively:

$$v_R = \phi\left[(1+\alpha)\frac{\rho^i Q_L^i}{2} - \delta\right] + (1-\phi)[w - \eta - \alpha(\rho^i Q_N^i - w - \mu)],$$
$$v_E = \phi\left(\frac{\rho^i Q_N^i}{2} - \eta\right) + (1-\phi)(w-\eta). \tag{5}$$

These expected payoff functions are illustrated in Figure 2.

To model the mutual dependence of social preferences and contracts, we represent the updating of new preferences as a now standard cultural evolution process in which individuals periodically update their behavioral norms and institutions (perhaps only during adolescence, or upon the founding of a new firm) after having taken into account information about the frequency distribution of various behaviors in the population, the payoffs associated with various behaviors in recent periods, or other facts (Bowles, 2004; Bisin and Verdier, 2001). Suppose that both employers and employees periodically update the contracts they offer and their preferences (respectively) by best responding to the distribution of play in the other class in the previous period.

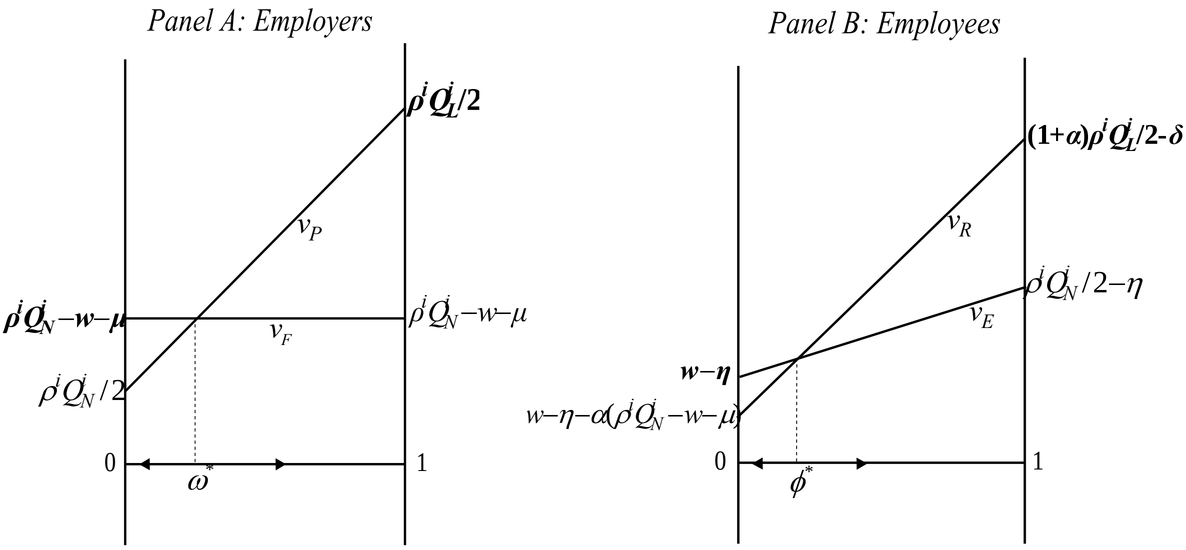

***Figure 2*: Expected payoffs under autarchy to *P*- and *F*-employers (*panel A*) and to *R*- and *E*-employees (*panel B*).** (NOTE: ϕ is the fraction of the employers offering Partnerships and ω the fraction of the employees being Reciprocators in the previous period. The vertical intercepts are from Table 1; payoffs in bold type refer to the stable pure Nash)

The updating process works as follows. At the beginning of each period, individuals are exposed to a cultural model randomly selected from their class: for instance, an employee, named A, has the opportunity to observe the behavior of another employee, named B, and to know her payoff. If the employee B is the same type as the employee A, A does not update. But if B is a different type, A compares the two payoffs and if B has the greater payoff, A switches to B's type with a probability equal to β (>0) times the payoff difference, retaining her own type otherwise. It is easily shown (Bowles, 2004) that this process gives the replicator dynamic equations:

$$\begin{aligned} \frac{d\phi}{d\tau} &= \phi(1-\phi)\beta[v_P(\omega) - v_F(\omega)], \\ \frac{d\omega}{d\tau} &= \omega(1-\omega)\beta[v_R(\phi) - v_E(\phi)], \end{aligned} \tag{6}$$

where τ denotes time. Notice that, reflecting our fourth assumption, people do not condition their updating on an already known kind of employment contract in which they will engage. We are now interested in the stationary states, such that dϕ/dτ=0 and dω/dτ=0. It is easy to see that:

$$\frac{\mathrm{d}\phi}{\mathrm{d}\tau}=0 \text{ for } \phi=0,\ \phi=1 \text{ and } \omega^*=\frac{\dfrac{\rho^i Q_N^i}{2}-(w+\mu)}{\dfrac{\rho^i Q_L^i}{2}-\dfrac{\rho^i Q_N^i}{2}},$$
$$\frac{\mathrm{d}\omega}{\mathrm{d}\tau}=0 \text{ for } \omega=0,\ \omega=1 \text{ and } \phi^*=\frac{\alpha[\rho^i Q_N^i-(w+\mu)]}{\left[(1+\alpha)\dfrac{\rho^i Q_L^i}{2}-\delta\right]-\left(\dfrac{\rho^i Q_N^i}{2}-\eta\right)+\alpha[\rho^i Q_N^i-(w+\mu)]}. \tag{7}$$

The resulting dynamical system is illustrated in Figure 3 where the arrows indicate the out-of-equilibrium adjustment given by the replicator equations. The states where $\mathrm{d}\phi/\mathrm{d}\tau=0$ and $\mathrm{d}\omega/\mathrm{d}\tau=0$ are cultural-institutional equilibria. The state (ϕ^*,ω^*) is stationary, but it is a saddle: small movements away from ϕ^* or ω^* are not self-correcting. (Two additional unstable stationary states, namely (ϕ = 1, ω = 0) and (ϕ = 0, ω = 1) are of no interest.) The asymptotically stable states are (1,1) (corresponding to {*P*,*R*} in Table 1) and (0,0) (corresponding to {*F*,*E*} in Table 1).

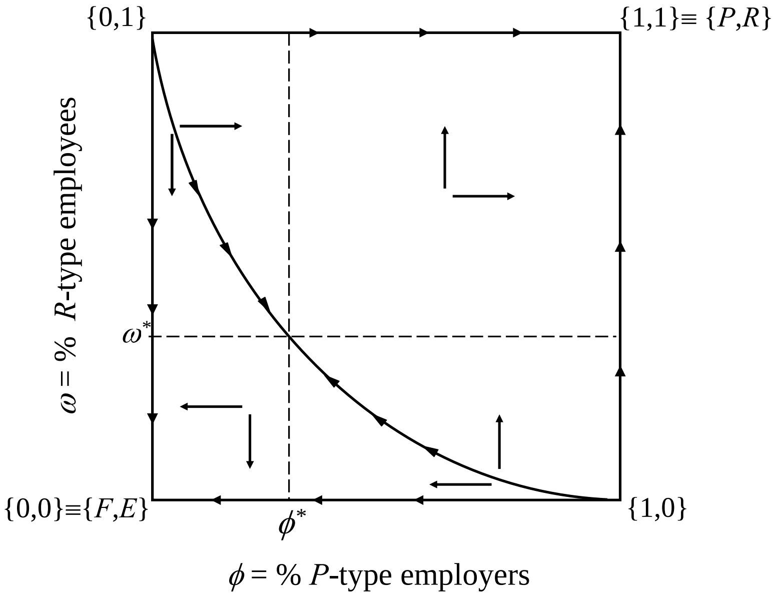

***Figure 3*: Co-evolution of preferences and institutions, and persistence of two cultural-institutional equilibria in a given country.**

In this deterministic setting, the initial state determines which of these two asymptotically stable states occurs. Of course institutions (and, in some cases, even cultural preferences) may be altered by a joint decision of a hypothetical representatives of one or both classes (Acemoglu and Robinson, 2006). But non-cooperative (that is decentralized, bottom-up) transitions are also possible. To study such a process we assume that occasional idiosyncratic (non-best response) updating of both preferences and contractual offers occurs (Kandori, Mailath, and Rob, 1993; Young, 1993, 1998). Suppose that with probability $1-\varepsilon$ myopic best response updating occurs as described above, but with a small probability ε the employee chooses randomly from the two behavioral traits and the employer likewise randomizes her contractual offer. The behavioral or contractual innovations represented by idiosyncratic play may be due to deliberate experimentation, error, or any other reason for non-best response play. We assume throughout that the rate of idiosyncratic play is sufficiently small that the equilibrium conventions described above are persistent, defined as having an expected duration of more than one period (i.e. $\varepsilon <$ critical number that would induce a transition to the other convention), that is for the $\{P,R\}$ equilibrium $\varepsilon < 1-\omega^*$ and $\varepsilon < 1-\phi^*$, whereas for the $\{F,E\}$ equilibrium $\varepsilon < \omega^*$ and $\varepsilon < \phi^*$. Jointly these persistence conditions imply $\varepsilon < 1/2$.

The resulting perturbed Markov process is ergodic (see Young, 1998), so over the long run both $\{P,R\}$ and $\{F,E\}$ will occur, with infrequent transitions between the basins of attraction of these two equilibria. In the absence of system-level exogenous shocks, for even moderately large populations and plausible rates of idiosyncratic play cultural-institutional equilibria will persist over very long periods and the system will spend more time at the convention with the larger basin of attraction. Thus the $\{P,R\}$ equilibrium will be more persistent if $\phi^*\omega^* < (1-\phi^*)(1-\omega^*)$ that is, if $\{P,R\}$ is the risk-dominant equilibrium, and conversely for the $\{F,E\}$ equilibrium.

5. Trade integration and the persistence of inefficient equilibria

Differences in the preferences and institutions prevailing in each country are a source of comparative advantage, and opening up to trade enables the two otherwise identical countries to enjoy welfare gains. But how does trade exposure affect the cultural and institutional environment in a given country? Will the two countries' different cultural-institutional equilibria persist after the two countries open up to international exchange? Does economic integration make cultural and institutional convergence more likely? These two questions may be translated as follows: will integration eliminate one of or both the critical values, ϕ^* and ω^*? If the answer is no, so that both asymptotically stable equilibria persist following integration, will trade decrease the costs of deviating from the status quo contract and preference, thereby facilitating a convergence to the other cultural-institutional equilibrium?

Figure 4 shows how the expected payoffs for each group of individuals change as a result of trade (expected payoff lines after trade drawn in dashed type). Payoffs received by the individuals in equilibrium are emphasized in bold fonts in the relevant panel.

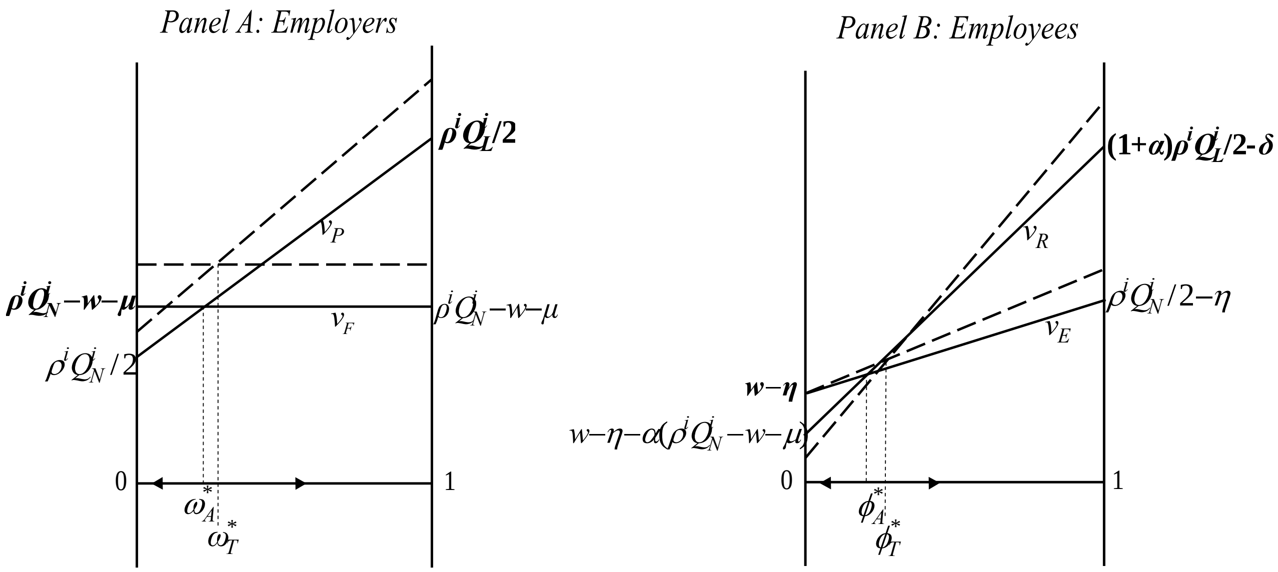

***Fig. 4*: Payoff changes to *P*- and *F*-employers (*panel A*) and *R*- and *E*-employees (*panel B*) after trade openness.** (NOTE: ϕ is the fraction of the employers offering Partnerships and ω the fraction of the employees being Reciprocators in the previous period. Dashed lines represent expected payoff lines after trade integration)

Trade increases the amount of the composite good that may be purchased with one unit of the good in which each country specializes, i.e. increases ρ^i (where i=o in country 1 and i=t in country 2). In the pictures, these effects are easily seen by reference to the resulting movements of the intercepts in each vertical axis. The corresponding positive changes in their difference (see appendix A.1.1) ensure that: *i*) after trade, the critical values of ϕ and ω remain within the unit interval in both countries, implying that trade integration does not destroy the cultural-institutional differences upon which specialization is based; *ii*) trade increases the cost of deviating from the status quo cultural-institutional convention for both groups in both countries, implying that non-coordinated convergence from one equilibrium to the other is less likely under trade integration rather than under autarchy (see appendix A.1.1). The reason is that deviating from the convention almost always entails a mismatch, the result being forgoing some of or all the surplus, the value of which is higher after trade integration.

Though we do not pursue this extension here, in a more complete model with state dependent rates of idiosyncratic play (Bergin and Lipman, 1996) the increased cost of innovating plausibly would reduce the rate of innovation, thereby prolonging the expected duration of each of the conventions. The fact that the cost of deviating increases may not only discourage the experimentation and error on which idiosyncratic play is based; it will also increase the selection pressures operating against individuals and firms that have innovated as long as these innovators constitute less than the minimum number needed to induce a transition. This can be seen from equations (6), along with the fact that trade increases both $[v_P(\omega) - v_F(\omega)]$ and $[v_R(\phi) - v_E(\phi)]$ when $\omega = 1 = \phi$ and increases both $[v_F(\omega) - v_P(\omega)]$ and $[v_E(\phi) - v_R(\phi)]$ when $\omega = 0 = \phi$. Thus trade will not induce a non-cooperative transition from the $\{F,E\}$ to the $\{P,R\}$ equilibrium despite the fact that the $\{P,R\}$ institutions and culture confer absolute advantage in both goods.

In addition to increasing the incentive not to innovate and the selection pressures operating against those who do, trade may even increase the number of innovators necessary to induce a transition from the $\{F,E\}$ to the $\{P,R\}$ equilibrium. To see this we study the effect of trade (that is, the increase in ρ^i with $i=t,o$) on ϕ^* and ω^*. In the case of ω^* the result is unambiguous: trade increases the critical fraction of reciprocal workers necessary to induce the *F*-type employers to best respond by adopting *P*-contracts (see appendix A.1.2):

$$\frac{\mathrm{d}\omega^*}{\mathrm{d}\rho^i} = \frac{\left(\frac{Q_L^i}{2} - \frac{Q_N^i}{2}\right)(w+\mu)}{\left(\frac{\rho^i Q_L^i}{2} - \frac{\rho^i Q_N^i}{2}\right)^2} > 0. \tag{8}$$

The reason can be seen by noting that the critical values ϕ^* and ω^* are simply given by the cost (for respectively employees and employers) of deviating from the $\{F,E\}$ equilibrium divided by the sum of this cost and the cost of deviating from the $\{P,R\}$ equilibrium. While the costs of deviating from both equilibria increase for the employers, trade increases the cost of deviating from the $\{F,E\}$ equilibrium of country 2 proportionally more.

The effect of trade on ϕ^* cannot be signed in general, but (under plausible conditions) it too may increase following integration. We have (see appendix A.1.2)

$$\frac{\mathrm{d}\phi^*}{\mathrm{d}\rho^i}=\frac{Q_N^i(-\delta+\eta)+\left[(1+\alpha)\dfrac{Q_L^i}{2}-\dfrac{Q_N^i}{2}\right](w+\mu)}{\left\{\left[(1+\alpha)\dfrac{\rho^i Q_L^i}{2}-\delta\right]-\left(\dfrac{\rho^i Q_N^i}{2}-\eta\right)+\alpha(\rho^i Q_N^i-w-\mu)\right\}^2}>0 \tag{9}$$

if and only if $[(1+\alpha)Q_L^i/2-Q_N^i/2](w+\mu)-Q_N^i(\delta-\eta)>0$. This will be the case if (given the strength of reciprocal preferences α) the extent to which the Reciprocator's evaluation of his output share under the Partnership contract exceeds that in the Forcing contract is large relative to the disutility of providing qualitative labor in addition to quantitative labor, and the output obtained employing quantitative labor only relative to the cost of hiring labor in Forcing contracting is small.

Thus removing impediments to international exchange need not destabilize and, indeed, may even fortify the preexisting cultural and institutional differences on which specialization and trade are based even if there exists an alternative cultural-institutional equilibrium that confers absolute advantage and to which a transition would be Pareto-improving. Trade impedes cultural-institutional convergence because it raises the costs of deliberate or accidental experimentation with uncommon preferences and contracts. Under plausible conditions it also increases the number of cultural or institutional innovators necessary to induce a decentralized transition from the high productivity equilibrium.

A best-response-induced transition to the superior culture and institutions, however, can be induced by a tariff. It is readily shown that there exists a one-time tariff protecting the opaque good in country 2 such that a cultural-institutional transition will occur, country 2 adopting the $\{P,R\}$ cultural-institutional nexus. Assuming that the international price ratio is not affected by the tariff, θ_ω^* and θ_ϕ^* are the ad-valorem tariff rates on the opaque (imported) good which will implement an (after tax) domestic price ratio in country 2 such that, respectively, $\omega_2^*=0$ and $\phi_2^*=0$. The transition-inducing tariff is given by $\theta^*=\min[\theta_\omega^*,\theta_\phi^*]$. Using equations (7) it can be shown (see appendix A.1.3) that:

$$(1+\theta_\omega^*)=\left(\frac{Q_N^t}{2(w+\mu)}-1\right)\frac{p_T^t}{p_T^o} \text{ and } (1+\theta_\phi^*)=\left[\frac{\alpha Q_N^t}{(w+\mu)}-1\right]\frac{p_T^t}{p_T^o}. \tag{10}$$

It is readily seen that $\theta^*_\omega < \theta^*_\phi$ if and only if α>1/2, in which case $\theta^* = \theta^*_\omega$.

The logic of the transition-inducing tariff is exactly the opposite of the mechanism underlying the fact that trade liberalization is transition-impeding. The tariff makes the transparent good less valuable in terms of the units of the composite good it can command and hence reduces the joint surplus available to the employer and the employee. So rather than increasing the cost of deviation from the $\{F,E\}$ convention as in the case of trade liberalization, the tariff reduces the cost of deviation. The level that entirely eliminates the cost of deviation for either of the two classes is the transition-inducing tariff, θ^*. Any tariff greater than θ^*_ω makes the Partnership a strict best response for the employers. Similarly, a tariff greater than θ^*_ϕ ($> \theta^*_\omega$ if and only if α>1/2) reduces profits under the Forcing contract to zero and thus makes employees indifferent to being Reciprocator or Homo economicus (if the employer is making zero profits the reciprocal employee is not offended by a Forcing contract).

6. Factor market integration and transitions to efficient equilibria

As Samuelson's factor price equalization theorem showed (Samuelson, 1949), the effects of the removal or reduction of the economic importance of national boundaries may be independent of whether integration is accomplished through the elimination of barriers to trade in commodities or through the mobility of factors of production. Where comparative advantage is based on country differences in culture and institutions, however, this is not the case.

In contrast to trade integration, factor market integration facilitates a Pareto-improving cultural-institutional transition in country 2. It does this by having the opposite of the two effects of trade integration: in the neighborhood of the equilibrium, it reduces the costs of the idiosyncratic play that induces transitions, and it reduces the number of innovators required to induce a transition. Under factor market integration, cultural and institutional innovators may enjoy an advantageous match not only with rare innovators from their own economy but also with the prevalent type of agent from the other country. Thus factor market integration provides a kind of innovation insurance, in contrast to commodity market integration which makes possible gains from trade that heighten the opportunity costs of the frequent mismatches that innovators may expect when paired with agent from their own country.

As we are interested in convergence to superior cultural-institutional conventions, we model the effect of factor market integration on the stability of country 2's inferior $\{F,E\}$ conventions. Suppose that some matches are made entirely with one's own nationals while others are made randomly in the global population. As pictured in Figure 5, there are now three factor markets, two of them national-specific and the third, a common pool without country identification. The common pool is populated by agents drawn at random from the two country-specific pools and hence has the same distribution of types as the meta-population (both countries combined). For both employers and employees in both countries let n be the fraction of matches made with individuals from one's own nation, the complement, $1-n$, being matches in the common pool. In the autarchic factor markets we have thus far assumed $n = 1$. But, if $n < 1$, one's expected match is n times the fraction of agents in one's own country plus $1-n$ times the distribution of types in the common pool. It is readily confirmed that if the countries are in the neighborhood of the $\{F,E\}$ and $\{P,R\}$ equilibria respectively, the expected difference, conditional on being resident in country 1 or country 2, in the likelihood of an employer meeting a Reciprocator and a Homo economicus or an employee meeting a Partnership or a Forcing contract is approximately $n(1-2\varepsilon)$ (which must be positive by the persistence conditions given in Section 4). Thus n is a measure of the degree of national specificity of the factor markets and $1-n$ is the degree of factor market integration.

One may image the two countries as two "villages" within which all production takes place under autarchy. But with factor market integration some (a random draw from each of the two villages) go to the cosmopolitan "city" where they make random matches with members of the other class. In this model n is not chosen by the individual agents; it is a characteristic of the two countries' cultures, language differences, geographical distance, immigration policies and other influences on factor movement that are exogenous from the standpoint of the individual employer or employee. To avoid considerable notational clutter for no additional insight we assume that n does not vary among countries. When factors of production are matched in the pool we assume that the product produced is determined by the nationality of the employer, reflecting the fact that the physical assets of the employer are product-specific while the skills of the worker are less so (notice however that this assumption may easily be relaxed without altering the conclusions in any relevant way). In the case of autarchy, the prices at which the output is sold are also determined by the nationality of the employer. Thus, for example, when an employee from country 2 is matched with an employer from country 1, the pair will produce the opaque good to be sold either at the prevailing international prices (in the case of trade integration) or at the autarchic prices of country 1 (in the absence of trade integration).

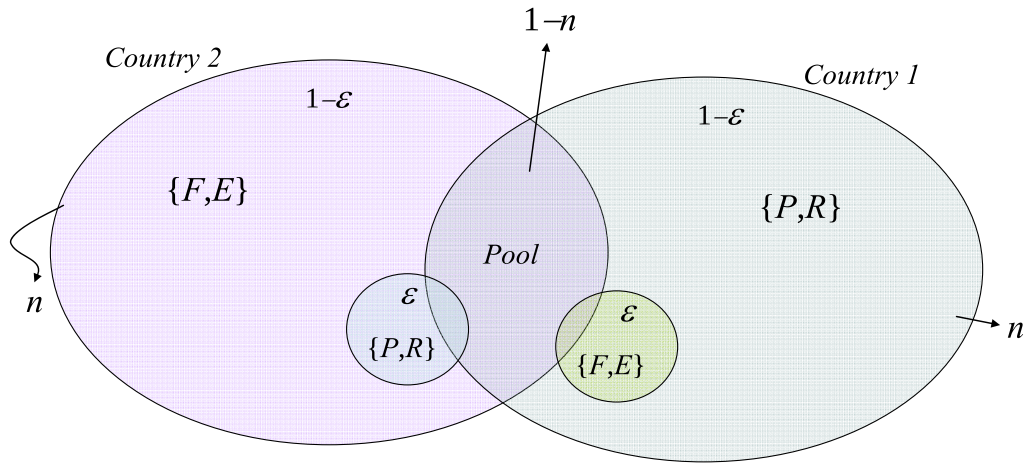

***Fig. 5*: Factor market integration.** (NOTE: ε is the expected fraction of idiosyncratic players among both employers and employees, n is the degree of national specificity of the factor markets and $1-n$ is the degree of factor market integration)

The expected payoffs equations can now be rewritten. The expected payoff after factor integration is the weighted sum of the expected payoff in the national factor market plus the expected payoff in the common pool, the weights being the relative sizes of the two pools (n and $1-n$). Notice that, as in (4) and (5), in computing the expected payoffs under factor market integration in country 2 (equations (11) and (12) below) the ω and ϕ appearing in the terms referring to own country matching are the distributions of play not the distribution of types (the two differ due to idiosyncratic play). Because we assume that all employers (employees) in country 1 are Partnership types (Reciprocators), taking account of idiosyncratic play, the country 2 agents who are matched in the pool with agents from country 1 will with probability $1-\varepsilon$ encounter employers (employees) offering P-contracts (reciprocal), while ε employers (employees) offering F-contracts (self-regarding).

Consider again country 2; the expected payoffs of respectively Partnership and Forcing contract employers under factor market integration are

$$v_P(n) = n\left[\omega_2 \frac{\rho_2^t Q_L^t}{2} + (1-\omega_2)\frac{\rho_2^t Q_N^t}{2}\right] + (1-n)\left\{\frac{1}{2}\left[\frac{\rho_2^t Q_L^t}{2}(1-\varepsilon) + \frac{\rho_2^t Q_N^t}{2}\varepsilon\right] + \frac{1}{2}\left[\omega_2 \frac{\rho_2^t Q_L^t}{2} + (1-\omega_2)\frac{\rho_2^t Q_N^t}{2}\right]\right\},$$

$$v_F(n) = n(\rho_2^t Q_N^t - w - \mu) + (1-n)\left[\frac{1}{2}(\rho_2^t Q_N^t - w - \mu) + \frac{1}{2}(\rho_2^t Q_N^t - w - \mu)\right] = \rho_2^t Q_N^t - w - \mu. \tag{11}$$

While the expected payoffs of respectively reciprocal and self-regarding employees are

$$v_R(n) = n\left\{\phi_2\left[(1+\alpha)\frac{\rho_2^t Q_L^t}{2} - \delta\right] + (1-\phi_2)[w - \eta - \alpha(\rho_2^t Q_N^t - w - \mu)]\right\} + (1-n)\frac{1}{2}\left\{\left[(1+\alpha)\frac{\rho_1^o Q_L^o}{2} - \delta\right](1-\varepsilon) + \right.$$

$$\left. + [w - \eta - \alpha(\rho_1^o Q_N^o - w - \mu)]\varepsilon + \phi_2\left[(1+\alpha)\frac{\rho_2^t Q_L^t}{2} - \delta\right] + (1-\phi_2)[w - \eta - \alpha(\rho_2^t Q_N^t - w - \mu)]\right\},$$

$$v_E(n) = n\left[\phi_2\left(\frac{\rho_2^t Q_N^t}{2} - \eta\right) + (1-\phi_2)(w-\eta)\right] + (1-n)\frac{1}{2}\left[\left(\frac{\rho_1^o Q_N^o}{2} - \eta\right)(1-\varepsilon) + (w-\eta)\varepsilon + \phi_2\left(\frac{\rho_2^t Q_N^t}{2} - \eta\right) + (1-\phi_2)(w-\eta)\right]. \tag{12}$$

Using expressions (11) and (12) we set $v_P(n) = v_F(n)$ and $v_R(n) = v_E(n)$ to calculate the new critical values, ϕ^* and ω^* in country 2 for the case of factor market integration:

$$\omega_2^*(n) = \frac{\dfrac{\rho_2^t Q_N^t - w - \mu}{1+n} - \dfrac{1-n}{1+n}\dfrac{1}{2}\left[\dfrac{\rho_2^t Q_L^t}{2}(1-\varepsilon) + \dfrac{\rho_2^t Q_N^t}{2}\varepsilon\right] - \dfrac{1}{2}\dfrac{\rho_2^t Q_N^t}{2}}{\dfrac{1}{2}\left(\dfrac{\rho_2^t Q_L^t}{2} - \dfrac{\rho_2^t Q_N^t}{2}\right)},$$

$$\phi_2^*(n) = \frac{\dfrac{1-n}{1+n}\dfrac{1}{2}\left(\dfrac{\rho_1^o Q_N^o}{2} - \eta\right)(1-\varepsilon) - \dfrac{1-n}{1+n}\dfrac{1}{2}\left\{\left[(1+\alpha)\dfrac{\rho_1^o Q_L^o}{2} - \delta\right](1-\varepsilon) - \alpha(\rho_1^o Q_N^o - w - \mu)\varepsilon\right\} + \dfrac{1}{2}\alpha(\rho_2^t Q_N^t - w - \mu)}{\dfrac{1}{2}\left\{\left[(1+\alpha)\dfrac{\rho_2^t Q_L^t}{2} - \delta\right] + \alpha(\rho_2^t Q_N^t - w - \mu)\right\} - \dfrac{1}{2}\left(\dfrac{\rho_2^t Q_N^t}{2} - \eta\right)}. \tag{13}$$

Factor market integration (a reduction in *n*) facilitates cultural-institutional transitions by having the opposite of the two effects of trade integration (the following results are valid using both autarchic and trade prices).

First, for both employers and employees in country 2, factor market integration (reducing *n*) lessens the costs of idiosyncratic play, respectively $v_F(n, \omega = 0) - v_P(n, \omega = 0)$ and $v_E(n, \phi = 0) - v_R(n, \phi = 0)$. As regards to employers this result is straightforward. *F*-type best responding employers in the $\{F,E\}$ equilibrium will be indifferent to the type of the employee they are paired with (because, as the horizontal v_F line in panel A of Figures 1 and 4 clearly shows, both reciprocal and self-regarding employees always provide quantitative labor under forcing contracting), so they are neither advantaged nor disadvantaged by factor market integration. By contrast, when $n < 1$ (factor market integration), P-type idiosyncratically playing employers will enjoy a payoff-maximizing match (with a reciprocal worker) not only with rare innovators from their own economy but also with the prevalent type of worker from the other country.

For employees, factor market integration increases the probability that both R-type idiosyncratic players and E-type individuals conforming to the convention in the $\{F,E\}$ equilibrium will make a payoff-maximizing match. However (as it is easily seen from the appendix A.2.1), the innovators' payoff advantage from market integration is greater than the benefit received by the best responders. Both idiosyncratically playing and best responding workers in country 2 additionally benefit from the higher payoffs from being matched with a country 1 producer. In this case the worker will produce the opaque good (rather than the transparent good) to be sold either at the prevailing international prices (if trade integration occurred; in which case $Q_L^o \rho_T^o > Q_L^t \rho_T^t$) or at the autarchic prices of country 1 (in the absence of trade integration; in which case $Q_L^o \rho_{1A}^o > Q_L^t \rho_{2A}^t$). But taking account of both the better matching prospects and the increase in payoffs for both best responders and idiosyncratic employees, it can be shown (see appendix A.2.1) that R-type innovators benefit from integration more than E-type best responders. In conclusion, factor market integration facilitates a transition from the $\{F,E\}$ equilibrium because it reduces the payoffs disadvantage of both idiosyncratically playing employers and employees compared to those conforming to the convention, and therefore it lessens the expected costs of innovating.

Second, for the country at the inferior $\{F,E\}$ cultural-institutional equilibrium, it can be shown (see appendix A.2.2) that

$$\frac{\mathrm{d}\omega_2^*(n)}{\mathrm{d}n} > 0 \text{ and } \frac{\mathrm{d}\phi_2^*(n)}{\mathrm{d}n} > 0$$

which show that factor market integration (reducing n) lowers the critical fraction of both employers and employees sufficient to induce a transition to the $\{P,R\}$ equilibrium.

Finally, it is possible to show that there exists a critical value, $n^* > 0$, of the degree of national specificity of the factor markets such that for $n < n^*$ a cultural-institutional transition from the {*F*,*E*} to the {*P*,*R*} convention will be induced in the absence of idiosyncratic play (see appendix A.2.3). For $n < n^*$, one of (or both) the critical values $\omega_2^*(n)$ and $\phi_2^*(n)$ is negative, so innovators do better than best responders with the result that the erstwhile {*F*,*E*} convention is no longer an absorbing state in the unperturbed dynamic. Accordingly, n_ω^* and n_ϕ^* will be the values of *n* such that (respectively) $\omega_2^*(n) = 0$ and $\phi_2^*(n) = 0$, and $n^* = \max[n_\omega^*, n_\phi^*]$.

7. Discussion

We have shown that otherwise identical economies that differ in culture and institutions may find specialization and trade welfare-enhancing, and that trade reinforces these differences by inhibiting convergence to superior cultural-institutional arrangements while factor market integration favors convergence.

Our paper is a contribution to the rapidly growing literature on institutions and trade (earlier contributions surveyed in Belloc, 2006). Comparative advantage based on institutional differences has been investigated for the following settings: financial systems (Beck, 2002; Kletzer and Bardhan, 1987; Ju and Wei, 2005; Matsuyama, 2005; Svaleryd and Vlachos, 2005), enforcement of contracts and property rights (Esfahani and Mookherjee, 1995; Levchenko, 2007; Nunn, 2007), intellectual property rights (Pagano, 2007), contracts and the division of labor (Acemoglu, Antràs and Helpman, 2009; Costinot, 2009), contractual incompleteness and the product cycle (Antràs, 2005), labor market flexibility and volatility (Cunat and Melitz, 2007), legal establishment and accounting systems (Vogel, 2007). In contrast to these papers, rather than studying the effects of exogenously given differences in institutions on comparative advantage and trade, we also consider the impact of economic integration on the endogenous dynamics of institutions. Other papers treating the effects of trade on institutions are Belloc (2009), Casella and Feinstein (2002), Dixit (2003), Do and Levchenko (2009) and Levchenko (2008). The main novelty of our approach with respect to this latter group of papers is our modeling of the complementary relationship between cultural preferences and institutions as a mechanism by which institutions associated with absolute disadvantage may persist indefinitely. In particular, our paper departs from and complements Do and Levchenko (2009) and

Levchenko (2008) in which institutional differences are a historical datum that may be modified by a cooperative lobbying game, while in our model they are implemented as an endogenously generated non-cooperative cultural-institutional equilibrium. Finally, unlike all above papers but in common with Olivier, Thoenig and Verdier (2008) and Pagano (2007) we find contrasting convergence effects of trade integration and factor market integration; but our model and these two models share little else in common, the former illustrates the dynamics of the demand for "cultural goods" that contribute to group identity while the latter concerns intellectual property.

The co-evolution of social norms and institutions is also modeled by Francois (2008). However, in contrast to our approach, in his model institutional change is implemented by an institutional designer external to the transaction (a political actor). Furthermore, while we explore the effects of economic integration on cultural-institutional equilibria, Francois (2008) studies those of increasing market competition. We share with Conconi, Legros and Newman (2008) the conclusion that liberalization need not favor the evolution of efficient institutions. In contrast to ours, in their model factor market integration may induce inefficiency, and only in conjunction with goods market integration are the effects of the two positive (in our model factor market integration has unambiguously positive effects). As in Krugman (1987)'s model of learning by doing, we show that a one time tariff may permanently alter a nation's comparative advantage and induce welfare gains.

The possibility that trade may induce institutional and cultural divergence rather than convergence is suggested by the experience of Europe in the late 19th century, when the institutional response to the import of cheap North American grain was radically different from country to country, resulting in a divergence with respect to tariffs and agrarian institutions (Gourevitch, 1977). Culture differences were also heightened, as the social solidarity of the subsidized Danish dairy cooperatives differed markedly from the nationalism associated with the German and French tariffs. Likewise, the centuries-long persistence of institutional differences among Western Hemisphere economies documented in Sokoloff and Engerman (2000) may be explained in part by the fact that trade allowed specialization in "plantation goods" such as sugar and cotton in some countries and "family farm" goods such as tobacco and wheat in others. Richard Freeman (2000) and Chiaki Moriguchi (2003) document a divergence in labor market institutions in open economies.

These cases of divergence notwithstanding, the impact of the U.S. civil war studied by

Nilsson (1994) is a reminder that cultural-institutional convergence does appear to be a powerful tendency in integrated global systems. But, like the convergence of European political institutions to the national state model over the half millennium prior to the First World War (Tilly, 1990), and the contemporaneous global diffusion of institutions and cultures of European origin, it also points to the important role of military and other political forces rather than the autonomous workings of international trade *per se* in this cultural and institutional convergence process.

References

Acemoglu, Daron, Antràs Pol, and Elhanan Helpman. 2009. Contracts and Technology Adoption, *American Economic Review* 97(3), pp. 916-943.

Acemoglu, Daron, and James Robinson. 2006. *Economic Origins of Dictatorship and Democracy*. Cambridge: University of Cambridge Press.

Antràs, Pol. 2005. Incomplete Contracts and the Product Cycle. *American Economic Review* 95(4), pp. 1054-1073.

Banerjee, Abhijit, and Lakshmi Iyer. 2005. History, Institutions and Economic Performance: The Legacy of Colonial Land Tenure Systems in India. *American Economic Review* 95(4), pp. 1190-213.

Beck, Thorsten. 2002. Financial Development and International Trade. Is there a Link? *Journal of International Economics* 57, pp. 107-131.

Belloc, Marianna. 2009. Cross-Cultural Trade and Institutional Stability. *B.E. Journal of Theoretical Economics* (*Topics*), vol. 9, n. 1: article 34

Belloc, Marianna. 2006. Institutions and International Trade: a Reconsideration of Comparative Advantage. *Journal of Economic Surveys* 20, pp. 3-26.

Bergin, James, and Barton L. Lipman. 1996. Evolution With State-Dependent Mutations. *Econometrica* 64(4), pp. 943-56.

Bewley, Truman F. 1999. *Why Wages Don't Fall During a Recession*. Cambridge: Harvard University Press.

Bisin, Alberto, and Thierry Verdier. 2001. The Economics of Cultural Transmission and the Evolution of Preferences. *Journal of Economic Theory* 97(2), pp. 298-319.

Bowles, Samuel. 1998. Endogenous Preferences: The Cultural Consequences of Markets and Other Economic Institutions. *Journal of Economic Literature* 36(1), pp. 75-111.

Bowles, Samuel. 2004. *Microeconomics: Behavior, Institutions, and Evolution*. Princeton: Princeton University Press, Princeton.

Bowles, Samuel. 2008. Policies Designed for Self Interested Citizens May Undermine ”the Moral Sentiments”: Evidence from Experiments. *Science* 320: pp. 58-83.

Bowles, Samuel, and Sandra Polania Reyes. 2009. Economic Incentives and Social Preferences: A Preference-based Lucas Critique of Public Policy. *Santa Fe Institute Working Paper*.

Cameron, Lisa. 1998. Raising the Stakes in the Ultimatum Game: Experimental Evidence from Indonesia. *Economic Inquiry* 37(1), pp. 47-59.

Casella, Alessandra, and Jonathan S. Feinstein. 2002. Public Goods in Trade on the Formation of Markets and Jurisdictions. *International Economic Review* 43(2), pp. 437-462.

Conconi, Paola, Patrick Legros, and Andrew F. Newman. 2008. Trade Liberalization and Organizational Change. *CEPR Discussion Paper* DP7028.

Costinot, Arnaud. 2009. On the Origins of Comparative Advantage. *Journal of International Economics* 77(2), pp. 255-264.

Cunat, Alejandro, and Marc Melitz. 2007. Volatility, Labour Market Flexibility, and the Pattern of Comparative Advantage, *NBER Working Paper* 13062.

Dixit, Avinash K. 2003. Trade Expansion and Contract Enforcement, *Journal of Political Economy* 111(6): pp. 1293-1317.

Do Quy-Toan, and Andrei Levchenko. 2009. Trade, Inequality and the Political Economy of Institutions, *Journal of Economic Theory* 144(4), pp. 1489-1520.

Doepke, Matthias, and Fabrizio Zilibotti. 2008. Occupational Choice and the Spirit of Capitalism, *Quarterly Journal of Economics* 123(2), pp. 747-793.

Esfahani, Hadi Salehi, and Dilip Mookherjee. 1995. Productivity, Contracting Modes, and Development. *Journal of Economic Development* 46, pp. 203-231.

Falk, Armin, and Michael Kosfeld. 2006. The Hidden Costs of Control. *American Economic Review* 96(5), pp. 1611-30.

Fehr, Ernst, and Falk, Armin. 2002. Psychological Foundations of Incentives. *European Economic Review* 46(4-5), pp. 687-724.

Fehr, Ernst, Alexander Klein, and Klaus M. Schmidt. 2007. Fairness and Contract Design.

Econometrica 75(1), pp. 121-54.

Fenoaltea, Stefano. 1984. Slavery and Supervision in Comparative Perspective: A Model, *Journal of Economic History* 44(3), pp. 635-668.

Francois, Patrick. 2008. Norms and Institutions. *Unpublished*.

Freeman, Richard. 2000. Single Peaked vs. Diversified Capitalism: the Relation Between Economic Institutions and Outcomes. *NBER Working Paper* 7556

Fershtman, Chaim, and Oren Bar-Gill. 2005. The Limit of Public Policy: Endogenous Preferences. *Journal of Public Economic Theory* 7(5), pp. 841-57.

Galor, Oded, and Omer Moav. 2002. Natural Selection and the Origin of Economic Growth. *Quarterly Journal of Economics* 117, pp. 1133-91.

Gerschenkron, Alexander. 1944. *Bread and Democracy in Germany*. Berkeley: University of California Press.

Gourevitch, Peter. 1977. International Trade, Domestic Coalitions, and Liberty: Comparative Responses to the Crisis of 1873-1896. *Journal of Interdisciplinary History* 8(2), pp. 281-313.

Greif, Avner. 1993. Contract Enforceability and Economic Institutions in Early Trade: The Maghribi Traders Coalition. *American Economic Review* 83(3), pp. 525-48.

Greif, Avner. 1994. Cultural Beliefs and the Organization of Society: An Historical and Theoretical Reflection on Collectivist and Individualist Societies. *Journal of Political Economy* 102(5), pp. 912-50.

Greif, Avner. 2002. Institutions & Impersonal Exchange: From Communal to Individual Responsibility. *Journal of Institutional and Theoretical Economics* 158(1), pp. 168-204.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2009. Cultural Biases in Economic Exchange? *Quarterly Journal of Economics*, forthcoming.

Herrmann, Benedikt, Christian Thoni, and Simon Gaechter. 2008. Antisocial Punishment Across Societies. *Science* 319 (7 March 2008), pp. 1362-67.

Henrich, Joe, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, Richard McElreath, Michael Alvard, Abigail Barr, Jean Ensminger, Natalie Smith Henrich, Kim Hill, Francisco Gil-White, Michael Gurven, Frank Marlowe, John Patton, and David Tracer. 2005. `Economic Man' in Cross-Cultural Perspective: Behavioral Experiments in 15 Small-Scale Societies. *Behavioral and Brain Sciences* 28, pp. 795-855.

Inglehart, Ronald. 1977. *The Silent Revolution: Changing Values and Political Styles among Western Publics*. Princeton: Princeton University Press.

Ju, Jiandong, and Shang-Jin Wei. 2005. Endowment Versus Finance: a Wooden Barrel Theory of International Trade. *International Monetary Fund Working Paper* 05/123.

Kandori, Michihiro G., George J. Mailath, and Rafael Rob. 1993. Learning, Mutation, and Long Run Equilibria in Games. *Econometrica* 61(1), pp. 29-56.

Keynes, John M. 1933. National Self-Sufficiency. *Yale Review* 22(4), pp. 755-769.

Krugman, Paul. 1987. The Narrow Moving Band, the Dutch Disease, and the Competitive Consequences of Mrs. Thatcher, Notes on Trade in the Presence of Dynamic Scale Economies. *Journal of Development Economics* 27, pp. 41-55.

Kindleberger, Charles P. 1962, *Foreign Trade and the National Economy*, New Haven, CT: Yale University Press.

Kletzer, Kenneth, and Pranab K. Bardhan. 1987. Credit Markets and Patterns of International Trade. *Journal of Development Economics* 27, pp. 57-70.

Levchenko, Andrei A. 2007. Institutional Quality and International Trade. *Review of Economic Studies* 74, pp. 791-819.

Levchenko, Andrei A. 2008. International Trade and Institutional Change. *Research Seminar in International Economics Discussion Paper* 579, pp. 791-819.

Levine, David K. 1998. Modeling Altruism and Spitefulness in Experiments. *Review of Economic Dynamics* 1, pp. 593-622.

Matsuyama, Kiminori. 2005. Credit Market Imperfections and Patterns of International Trade and Capital Flows. *Journal of the European Economic Association* 3, pp. 714-723.

Moriguchi, Chiaki. 2003. Implicit Contracts, the Great Depression, and Institutional Change: A Comparative Analysis of U.S. and Japanese Employment Relations, 1920-1940. *Journal of Economic History* 63(3), pp. 625-665.

Nilsson, Eric. 1994. Empirical Evidence that the Social Relations of Production Matter: the Case of the Ante-bellum U.S. South. *Cambridge Journal of Economics* 18, pp. 259-77.

Nunn, Nathan. 2007. Relation-Specificity, Incomplete Contracts and the Pattern of Trade. *Quarterly Journal of Economics* 122, pp. 569-600.

Nunn, Nathan and Leonard Wantchekon. 2008. The Slave Trade and the Origins of Mistrust in Africa. *Harvard University Department of Economics*.

Olivier, Jacques, Mathias Thoenig, and Thierry Verdier. 2008. Globalization and the Dynamics of Cultural Identity. *Journal of International Economics* 76, pp. 356-370.

Ortiz, Fernando. 1963. *Contrapunteo Cubano del Tabaco y el Azucar*. Barcelona: Editorial Ariel.

Pagano, Ugo. 2007. Cultural Globalisation, Institutional Diversity and the Unequal Accumulation of Intellectual Capital, *Cambridge Journal of Economics* 31, pp. 649–667.

Rabin, Matthew. 1993. Incorporating Fairness into Game Theory and Economics. *American Economic Review* 83(5), pp. 1281-1302.

Samuelson, Paul A. 1949. International Factor-Price Equalisation Once Again, *Economic Journal* 59, 181-197.

Siamwalla, Ammar. 1978. Farmers and Middlemen: Aspects of Agricultural Marketing in Thailand. *Economic Bulletin for Asia and the Pacific* 39(1), pp. 38-50.

Sokoloff, Kenneth L., and Stanley L. Engerman. 2000. Institutions, Factor Endowments, and Paths of Development in the New World. *Journal of Economic Perspectives* 14(3), pp. 217-32.

Spolaore, Enrico, and Romain Wacziarg. 2008. The Diffusion of Development. *NBER Working Paper* 12153.

Svaleryd, Helena, and Jonas Vlachos. 2005. Financial Markets, the Pattern of Specialization and Comparative Advantage. Evidence from OECD Countries. *European Economic Review* 49, pp. 113-44.

Tabellini, Guido. 2008. The Scope of Cooperation: Values and Incentives. *Quarterly Journal of Economics* 123(3), 905-950

Tilly, Charles. 1990. *Coercion, Capital, and European States, AD 990-1990*. Cambridge, Mass., USA: Blackwell.

Vogel, Jonathan. 2007. Institutions and Moral Hazard in Open Economies. *Journal of International Economics* 71, pp. 495-514.

Young, H. Peyton. 1993. The Evolution of Conventions. *Econometrica* 61, pp. 57-84.

Young, H. Peyton. 1998. *Individual Strategy and Social Structure: an Evolutionary Theory of Institutions*. Princeton, NJ: Princeton University Press.

Mathematical appendix

For brevity in this appendix we define $m \equiv w + \mu$. A more detailed appendix is available from the authors upon request and posted on the first author's website.

A.1.1. Trade integration increases the costs of deviation. Part A: The cost of deviation for an employer in the {*F*,*E*} equilibrium is given by $v_F(\omega = 0) - v_P(\omega = 0)$, where $v_F(\omega = 0)$ and $v_P(\omega = 0)$ are given by equations (4) in the text with ω=0. We easily obtain $v_F(\omega = 0) - v_P(\omega = 0) = (\rho^i Q_N^i)/2 - m,$ which is increasing in ρ^i. **Part B:** Similarly the corresponding cost of deviation for an employee is $v_E(\phi = 0) - v_R(\phi = 0)$, where $v_E(\phi = 0)$ and $v_R(\phi = 0)$ are given by equations (5) in the text with ϕ=0, thereby $v_E(\phi = 0) - v_R(\phi = 0) = \alpha(\rho^i Q_N^i - m),$ which is also increasing in ρ^i.

A.1.2. Trade integration decreases the critical values $\boldsymbol{\omega^*}$ and $\boldsymbol{\phi^*}$. Part A: The derivative of ω^* given in (7) in the text with respect to ρ^i is

$$\frac{\mathrm{d}\omega^*}{\mathrm{d}\rho^i} = \frac{\frac{Q_N^i}{2}\left(\frac{\rho^i Q_L^i}{2} - \frac{\rho^i Q_N^i}{2}\right) - \left(\frac{Q_L^i}{2} - \frac{Q_N^i}{2}\right)\left(\frac{\rho^i Q_N^i}{2} - m\right)}{\left(\frac{\rho^i Q_L^i}{2} - \frac{\rho^i Q_N^i}{2}\right)^2} = \frac{\left(\frac{Q_L^i}{2} - \frac{Q_N^i}{2}\right)m}{\left(\frac{\rho^i Q_L^i}{2} - \frac{\rho^i Q_N^i}{2}\right)^2} > 0.$$

Part B: The derivative of ϕ^* also given in (7) with respect to ρ^i is

$$\frac{\mathrm{d}\phi^*}{\mathrm{d}\rho^i} = \frac{\alpha Q_N^i\left[(1+\alpha)\frac{\rho^i Q_L^i}{2} - \delta - \left(\frac{\rho^i Q_N^i}{2} - \eta\right) + \alpha(\rho^i Q_N^i - m)\right] - \alpha(\rho^i Q_N^i - m)\left[(1+\alpha)\frac{Q_L^i}{2} - \frac{Q_N^i}{2} + \alpha Q_N^i\right]}{\left\{(1+\alpha)\frac{\rho^i Q_L^i}{2} - \delta - \left(\frac{\rho^i Q_N^i}{2} - \eta\right) + \alpha(\rho^i Q_N^i - m)\right\}^2}$$

$$= \frac{Q_N^i(-\delta + \eta) + \left[(1+\alpha)\frac{Q_L^i}{2} - \frac{Q_N^i}{2}\right]m}{\left\{(1+\alpha)\frac{\rho^i Q_L^i}{2} - \delta - \left(\frac{\rho^i Q_N^i}{2} - \eta\right) + \alpha(\rho^i Q_N^i - m)\right\}^2} > 0 \text{ iff } \left[(1+\alpha)\frac{Q_L^i}{2} - \frac{Q_N^i}{2}\right]m - Q_N^i(\delta - \eta) > 0.$$

A.1.3. Transition-inducing tariff rate. θ_ω^* and θ_ϕ^* are the ad-valorem tariff rates protecting the opaque good in country 2 such that, respectively, $\omega_2^* = 0$ and $\phi_2^* = 0$. **Part A:** The former is obtained by equating ω_2^* (given in (7) in the text with *i*=*t*) to zero, thereby:

$$\left[\frac{p^t}{p^t + p^o(1+\theta_\omega^*)} Q_N^t\right] / 2 - m = 0, \text{ i.e. } (1+\theta_\omega^*) = \left(\frac{Q_N^t}{2m} - 1\right)\frac{p^t}{p^o},$$

which is the first of equations (10) in the text.

Part B: Similarly the latter is obtained by equating ϕ_2^* (given in (7) in the text with *i*=*t*) to zero, thereby:

$$\frac{p^t}{p^t + p^o(1+\theta_\phi^*)} \alpha Q_N^t - m = 0, \text{i.e. } (1+\theta_\phi^*) = \left(\frac{\alpha Q_N^t}{m} - 1\right)\frac{p^t}{p^o},$$

which is the second of equations (10) in the text.

A.2.1. Factor market integration decreases the costs of deviation. Part A: The cost of deviation for an employer in the {*F*,*E*} equilibrium after factor market integration is $v_F(\omega = 0, n) - v_P(\omega = 0, n)$, where $v_F(\omega = 0, n)$ and $v_P(\omega = 0, n)$ are given by equations (11) in the text with ω = 0. This difference is smaller than the corresponding expression under factor immobility (*n*=1). This is easily seen by noting that while the expected payoff of a *F*-contract best responding employer [$v_F(\omega = 0, n)$] is unaltered, the expected payoff of a *P*-contract idiosyncratic player [$v_P(\omega = 0, n)$] is greater after factor market integration (reduction in *n*) because $\rho_2^t Q_L^t / 2 > \rho_2^t Q_N^t / 2$. **Part B:** Similarly, the cost of deviation for an employee in the {*F*,*E*} equilibrium after factor market integration is $v_E(\phi_2 = 0, n) - v_R(\phi_2 = 0, n)$, where $v_E(\varphi = 0, n)$ and $v_R(\varphi = 0, n)$ are given by equations (12) in the text with ϕ = 0. It is smaller than the corresponding expression under factor immobility (*n*=1). Indeed while both the expected payoff of *E*-type best responding employees [$v_E(\phi = 0, n)$] and the expected payoff of *R*-type idiosyncratic players [$v_R(\phi = 0, n)$] increase after factor market integration (decrease in *n*), the latter increases more than the former. This is proven to be true providing that

$$\left[(1+\alpha)\frac{\rho_1^o Q_L^o}{2} - \delta\right](1-\varepsilon) + [w - \eta - \alpha(\rho_1^o Q_N^o - m)]\varepsilon > \left(\frac{\rho_1^o Q_N^o}{2} - \eta\right)(1-\varepsilon) + (w-\eta)\varepsilon$$

which may be rewritten as

$$\varepsilon < \frac{\left[(1+\alpha)\frac{\rho_1^o Q_L^o}{2} - \delta\right] - \left(\frac{\rho_1^o Q_N^o}{2} - \eta\right)}{\left[(1+\alpha)\frac{\rho_1^o Q_L^o}{2} - \delta\right] - \left(\frac{\rho_1^o Q_N^o}{2} - \eta\right) + \alpha(\rho_1^o Q_N^o - m)}.$$

The above inequality is true because it is equivalent to $\varepsilon < 1 - \phi_1^*$ which follows from the persistence conditions (see Section 4).

A.2.2 Factor market integration increases the critical values $\omega_2^*(n)$ and $\phi_2^*(n)$. Part A: The derivative of $\omega_2^*(n)$ (given in (13) in the text with i=t) with respect to n is

$$\frac{\mathrm{d}\omega_2^*(n)}{\mathrm{d}n} = \frac{\frac{1}{2}\left[\frac{\rho_2^t Q_L^t}{2}(1-\varepsilon) + \frac{\rho_2^t Q_N^t}{2}\varepsilon\right](1+n) - \left\{(\rho_2^t Q_N^t - m) - (1-n)\frac{1}{2}\left[\frac{\rho_2^t Q_L^t}{2}(1-\varepsilon) + \frac{\rho_2^t Q_N^t}{2}\varepsilon\right]\right\}}{(1+n)^2}$$

$$= \frac{\left[\frac{\rho_2^t Q_L^t}{2}(1-\varepsilon) + \frac{\rho_2^t Q_N^t}{2}\varepsilon\right] - (\rho_2^t Q_N^t - m)}{(1+n)^2}.$$

Showing that the numerator of the last fraction is positive is equivalent to showing that the below inequality is true

$$\varepsilon < 1 - \frac{\frac{\rho_2^t Q_L^t}{2} - m}{\frac{\rho_2^t Q_L^t}{2} - \frac{\rho_2^t Q_N^t}{2}} = 1 - \omega_2^*,$$

which follows from the persistence conditions (see Section 4).

Part B: To study the sign of the derivative of $\phi_2^*(n)$ (given in (13) in the text with i=t) with respect to n, notice that the denominator of $\phi_2^*(n)$ is positive and the last fraction on the numerator does not depend on n. Hence it is easily shown that $\mathrm{d}\phi_2^*(n)/\mathrm{d}n > 0$. Indeed $\mathrm{d}[(1-n)/(1+n)]/\mathrm{d}n < 0$, and

$$\left(\frac{\rho_1^o Q_N^o}{2} - \eta\right)(1-\varepsilon) - \left[(1+\alpha)\frac{\rho_1^o Q_L^o}{2} - \delta\right](1-\varepsilon) + \alpha(\rho_1^o Q_N^o - m)\varepsilon < 0.$$

The above inequality is true because (by the same proof given in A.2.1. Part B) it is equivalent to $\varepsilon < 1 - \phi_1^*$ which follows persistence conditions (see Section 4).

A.2.3. Transition-inducing degree of national specificity. n_ω^* and n_ϕ^* are the degrees of national specificity of the factor markets such that, respectively, $\omega_2^*(n) = 0$ and $\phi_2^*(n) = 0$. **Part A:** Equating the first of expressions given in (13) to zero and solving for *n*, the former is

$$n_\omega^* = \frac{\frac{1}{2}\left[\frac{\rho_2^t Q_L^t}{2}(1-\varepsilon) + \frac{\rho_2^t Q_N^t}{2}\varepsilon\right] + \frac{1}{2}\frac{\rho_2^t Q_N^t}{2} - (\rho_2^t Q_N^t - m)}{\frac{1}{2}\left[\frac{\rho_2^t Q_L^t}{2}(1-\varepsilon) + \frac{\rho_2^t Q_N^t}{2}\varepsilon\right] - \frac{1}{2}\frac{\rho_2^t Q_N^t}{2}}.$$

Part B: Equating the second of expressions in (13) to zero and solving for *n*, the latter is

$$n_\phi^* = \frac{\left[(1+\alpha)\frac{\rho_1^o Q_L^o}{2} - \delta\right](1-\varepsilon) - \alpha(\rho_1^o Q_N^o - m)\varepsilon - \frac{1}{2}\alpha(\rho_2^t Q_N^t - m) - \left(\frac{\rho_1^o Q_N^o}{2} - \eta\right)(1-\varepsilon)}{\left[(1+\alpha)\frac{\rho_1^o Q_L^o}{2} - \delta\right](1-\varepsilon) - \alpha(\rho_1^o Q_N^o - m)\varepsilon + \frac{1}{2}\alpha(\rho_2^t Q_N^t - m) - \left(\frac{\rho_1^o Q_N^o}{2} - \eta\right)(1-\varepsilon)}.$$

DETAILED MATHEMATICAL APPENDIX (not intended for publication)

For brevity in this appendix we define $m \equiv w + \mu$.

A.1.1 Critical values $\boldsymbol{\omega^*}$ and $\boldsymbol{\phi^*}$ in autarchy. ω^* (ϕ^*) is the number of reciprocal employees (Partnership employers) in the previous period that makes an employer (employee) indifferent to offering a Partnership or a Forcing contract (to being reciprocal or self-regarding).

PART A: *Employers.* The expected payoffs to employers offering respectively *P*- and *F*-contracts, where *i* denotes the good, are:

$$\begin{aligned} v_P &= \omega \frac{\rho^i Q_L^i}{2} + (1-\omega)\frac{\rho^i Q_N^i}{2}, \\ v_F &= \omega(\rho^i Q_N^i - m) + (1-\omega)(\rho^i Q_N^i - m) = \rho^i Q_N^i - m. \end{aligned} \tag{A1}$$

ω^* is the level of ω such that $v_P(\omega) = v_F(\omega)$, i.e.

$$\omega \frac{\rho^i Q_L^i}{2} + (1-\omega)\frac{\rho^i Q_N^i}{2} = \rho^i Q_N^i - m,$$

hence

$$\omega^* = \frac{\dfrac{\rho^i Q_N^i}{2} - m}{\dfrac{\rho^i Q_L^i}{2} - \dfrac{\rho^i Q_N^i}{2}}, \tag{A2}$$

which is the first of equations (7) in the text.

PART B: *Employees.* Similarly, the expected payoffs to respectively *R*- and *E*-employees are:

$$\begin{aligned} v_R &= \phi\left[(1+\alpha)\frac{\rho^i Q_L^i}{2} - \delta\right] + (1-\phi)[w - \eta - \alpha(\rho^i Q_N^i - m)], \\ v_E &= \phi\left(\frac{\rho^i Q_N^i}{2} - \eta\right) + (1-\phi)(w-\eta). \end{aligned} \tag{A3}$$

ϕ^* is the value of ϕ such that $v_R(\phi) = v_E(\phi)$, i.e.

$$\phi\left[(1+\alpha)\frac{\rho^i Q_L^i}{2} - \delta\right] + (1-\phi)[w - \eta - \alpha(\rho^i Q_N^i - m)] = \phi\left(\frac{\rho^i Q_N^i}{2} - \eta\right) + (1-\phi)(w-\eta),$$

hence

$$\phi^* = \frac{\alpha(\rho^i Q_N^i - m)}{\left[(1+\alpha)\dfrac{\rho^i Q_L^i}{2} - \delta\right] - \left(\dfrac{\rho^i Q_N^i}{2} - \eta\right) + \alpha(\rho^i Q_N^i - m)} \tag{A4}$$

which is the second of equations (7) in the text.

A.1.2 Effects of trade integration on the costs of deviation. In this subsection, we prove that trade integration, i.e. an increase in ρ^i (with *i* = *t* in country 2, and *i* = *o* in country 1) increases the cost of deviating from the status quo cultural-institutional convention. We only consider the $\{F,E\}$ equilibrium, the extension to the $\{P,R\}$ equilibrium being straightforward. The cost of deviation is given by the difference between the expected payoff of a best responder and that of a non-best responder.

PART A: *Employers.* Rewrite the expected payoff equations for employers offering respectively *P*- and *F*-contracts when all the employees in the previous period were Homo economicus (i.e. equations (A1) with $\omega=0$):

$$\begin{aligned} v_P &= \frac{\rho^i Q_N^i}{2}, \\ v_F &= \rho^i Q_N^i - m. \end{aligned} \tag{A5}$$

The cost of deviation in the $\{F,E\}$ equilibrium is given by $v_F(\omega=0)-v_P(\omega=0)$. Using equations (A5) this is equivalent to

$$v_F(\omega=0)-v_P(\omega=0)=\frac{\rho^i Q_N^i}{2}-m, \tag{A6}$$

which is increasing in ρ^i.

PART B: *Employees.* Similarly, the expected payoff equations for respectively *R*- and *E*-employees when all the employers in the previous period were offering *F*-contracts (i.e. equations (A3) with $\phi=0$) may be written as:

$$\begin{aligned} v_R &= w-\eta-\alpha(\rho^i Q_N^i - m), \\ v_E &= w-\eta. \end{aligned} \tag{A7}$$

The cost of deviation in the $\{F,E\}$ equilibrium is thus given by $v_E(\phi=0)-v_R(\phi=0)$ which, using equations (A7), can be rewritten as

$$v_E(\phi=0)-v_R(\phi=0)=\alpha(\rho^i Q_N^i - m), \tag{A.8}$$

which is also increasing in ρ^i.

A.1.3 Effects of trade integration on the critical values ω^* and ϕ^*. In this subsection we show that trade integration, i.e. an increase in the value of the own-country-produced good ρ^i (i=t,o) in terms of *c*-good, leads to an increase in the expected number of idiosyncratic players in either class (employers and employees) sufficient to induce a transition from the $\{F,E\}$ to the $\{P,R\}$ equilibrium. To show this we study the sign of the derivative of ω^* and ϕ^* with respect to ρ^i.

PART A: Using expression (A2), the former is

$$\frac{\mathrm{d}\omega^*}{\mathrm{d}\rho^i} = \frac{\frac{Q_N^i}{2}\left(\frac{\rho^i Q_L^i}{2} - \frac{\rho^i Q_N^i}{2}\right) - \left(\frac{Q_L^i}{2} - \frac{Q_N^i}{2}\right)\left(\frac{\rho^i Q_N^i}{2} - m\right)}{\left(\frac{\rho^i Q_L^i}{2} - \frac{\rho^i Q_N^i}{2}\right)^2} = \frac{\left(\frac{Q_L^i}{2} - \frac{Q_N^i}{2}\right)m}{\left(\frac{\rho^i Q_L^i}{2} - \frac{\rho^i Q_N^i}{2}\right)^2} > 0,$$

which is equation (8) in the text and is always positive because $Q_L^i > Q_N^i$. **PART B:** Analogously, using (A4), the latter can be written as

$$\frac{\mathrm{d}\phi^*}{\mathrm{d}\rho^i} = \frac{\alpha Q_N^i\left[(1+\alpha)\frac{\rho^i Q_L^i}{2} - \delta - \left(\frac{\rho^i Q_N^i}{2} - \eta\right) + \alpha(\rho^i Q_N^i - m)\right] - \alpha(\rho^i Q_N^i - m)\left[(1+\alpha)\frac{Q_L^i}{2} - \frac{Q_N^i}{2} + \alpha Q_N^i\right]}{\left\{(1+\alpha)\frac{\rho^i Q_L^i}{2} - \delta - \left(\frac{\rho^i Q_N^i}{2} - \eta\right) + \alpha(\rho^i Q_N^i - m)\right\}^2}$$

$$= \frac{Q_N^i(-\delta+\eta) + \left[(1+\alpha)\frac{Q_L^i}{2} - \frac{Q_N^i}{2}\right]m}{\left\{(1+\alpha)\frac{\rho^i Q_L^i}{2} - \delta - \left(\frac{\rho^i Q_N^i}{2} - \eta\right) + \alpha(\rho^i Q_N^i - m)\right\}^2},$$

which is equation (9) in the text and is positive iff $\left[(1+\alpha)\frac{Q_L^i}{2} - \frac{Q_N^i}{2}\right]m - Q_N^i(\delta-\eta) > 0$.

A.1.4 Transition-inducing tariff rate. $\theta^* > 0$ is the tariff protecting the opaque good in country 2 such that a cultural-institutional transition from the {*F*,*E*} to the {*P*,*R*} convention will occur. Given the international price ratio, θ_ω^* and θ_ϕ^* are the ad-valorem tariff rates such that, respectively, $\omega_2^* = 0$ and $\phi_2^* = 0$. After a tariff on the imported *o*-good, the value of the country 2 produced *t*-good in terms of the *c*-good is $\rho^t = p^t / [p^t + p^o(1+\theta_\omega^*)]$. The transition-inducing tariff is given by $\theta^* = \min[\theta_\omega^*, \theta_\phi^*]$. **PART A:** To obtain θ_ω^*, we equate expression (A2) for country 2 to zero and solve for $1+\theta_\omega^*$:

$$\frac{\frac{p^t}{p^t + p^o(1+\theta_\omega^*)}Q_N^t}{2} - m = 0, \text{i.e.} \quad (1+\theta_\omega^*) = \left(\frac{Q_N^t}{2m} - 1\right)\frac{p^t}{p^o},$$

which is the first of equations (10) in the paper.

PART B: Whereas to obtain θ_ϕ^*, we equate expression (A4) for country 2 to zero and solve for$1+\theta_\phi^*$:

$$\frac{p^t}{p^t + p^o(1+\theta_\phi^*)}\alpha Q_N^t - m = 0, \text{i.e. } (1+\theta_\phi^*) = \left(\frac{\alpha Q_N^t}{m} - 1\right)\frac{p^t}{p^o},$$

which is the second of equations (10) in the paper.

A.2.1 Critical values $\boldsymbol{\omega_2^*(n)}$ and $\boldsymbol{\phi_2^*(n)}$ under factor market integration. The proofs contained in this subsection and in the following two are valid using both autarchic and trade prices. Hence, while we denote by subscript 1 and 2 the prices of respectively country 1 (in the {*P*,*R*} equilibrium) and country 2 (in the {*F*,*E*} equilibrium), we omit subscript "*A*" and "*T*" standing for respectively autarchy and trade in the text. Clearly, if we consider trade prices it follows that $\rho_{1T}^t = \rho_{2T}^t = \rho_T^t$ and $\rho_{1T}^o = \rho_{2T}^o = \rho_T^o$, whereas if we consider autarchic prices we have $\rho_{1A}^t > \rho_{2A}^t$ and $\rho_{1A}^o < \rho_{2A}^o$; but our conclusions do not change in substance. Again we report the proofs only considering the disadvantageous culture and institutions {*F*,*E*} country, the extension to the {*P*,*R*} country being straightforward. $\omega^*(n)$ ($\phi^*(n)$) is the number of *R*-employees (*P*-contract employers) in the previous period that, under factor market integration (*n*<1), makes an employer (employee) indifferent to offering *P*- or *F*-contracts (to being *R*- or *E*-type). The expected payoff to an individual in the {*F*,*E*} country after factor market integration is given by *n* times the expected payoff of a domestic match plus (1−*n*) times the expected payoff of a match in the common pool, the latter being given by 1/2 probability times the expected payoff from matching an individual from the {*P*,*R*} country (where everybody is best responder except ε idiosyncratic players) plus 1/2 probability times the expected payoff from matching an individual from her own country. As explained in the text, when factors of production are matched in the pool the product produced is determined by the nationality of the employer. In the case of autarchy, the prices at which the output is sold are also determined by the nationality of the employer.

PART A: *Employers.* The expected payoffs to employers offering *P*- and *F*-contracts after factor market integration (equations (11) in the text) are (notice the superscript *t* does not change because the employer determines the good produced):

$$v_P(n) = n\left[\omega_2\frac{\rho_2^t Q_L^t}{2} + (1-\omega_2)\frac{\rho_2^t Q_N^t}{2}\right] + (1-n)\left\{\frac{1}{2}\left[\frac{\rho_2^t Q_L^t}{2}(1-\varepsilon) + \frac{\rho_2^t Q_N^t}{2}\varepsilon\right] + \frac{1}{2}\left[\omega_2\frac{\rho_2^t Q_L^t}{2} + (1-\omega_2)\frac{\rho_2^t Q_N^t}{2}\right]\right\},$$

$$v_F(n) = n(\rho_2^t Q_N^t - m) + (1-n)\left[\frac{1}{2}(\rho_2^t Q_N^t - m) + \frac{1}{2}(\rho_2^t Q_N^t - m)\right] = \rho_2^t Q_N^t - m. \tag{A9}$$

To obtain $\omega_2^*(n)$ we compute the value of ω such that $v_P(n) = v_F(n)$; it follows:

$$n\left[\omega_2\frac{\rho_2^t Q_L^t}{2} + (1-\omega_2)\frac{\rho_2^t Q_N^t}{2}\right] + (1-n)\left\{\frac{1}{2}\left[\frac{\rho_2^t Q_L^t}{2}(1-\varepsilon) + \frac{\rho_2^t Q_N^t}{2}\varepsilon\right] + \frac{1}{2}\left[\omega_2\frac{\rho_2^t Q_L^t}{2} + (1-\omega_2)\frac{\rho_2^t Q_N^t}{2}\right]\right\} = \rho_2^t Q_N^t - m.$$

That, after some manipulation, becomes

$$\left[n\left(\frac{\rho_2^t Q_L^t}{2}-\frac{\rho_2^t Q_N^t}{2}\right)+(1-n)\frac{1}{2}\left(\frac{\rho_2^t Q_L^t}{2}-\frac{\rho_2^t Q_N^t}{2}\right)\right]\omega_2=(\rho_2^t Q_N^t-m)-(1-n)\frac{1}{2}\left[\frac{\rho_2^t Q_L^t}{2}(1-\varepsilon)+\frac{\rho_2^t Q_N^t}{2}\varepsilon+\frac{\rho_2^t Q_N^t}{2}\right]-n\frac{\rho_2^t Q_N^t}{2},$$

whereby

$$(1+n)\frac{1}{2}\left(\frac{\rho_2^t Q_L^t}{2}-\frac{\rho_2^t Q_N^t}{2}\right)\omega_2=(\rho_2^t Q_N^t-m)-(1-n)\frac{1}{2}\left[\frac{\rho_2^t Q_L^t}{2}(1-\varepsilon)+\frac{\rho_2^t Q_N^t}{2}\varepsilon\right]-(1+n)\frac{1}{2}\frac{\rho_2^t Q_N^t}{2}$$

and, finally, we obtain

$$\omega_2^*(n)=\frac{\dfrac{\rho_2^t Q_N^t-m}{1+n}-\dfrac{1-n}{1+n}\dfrac{1}{2}\left[\dfrac{\rho_2^t Q_L^t}{2}(1-\varepsilon)+\dfrac{\rho_2^t Q_N^t}{2}\varepsilon\right]-\dfrac{1}{2}\dfrac{\rho_2^t Q_N^t}{2}}{\dfrac{1}{2}\left(\dfrac{\rho_2^t Q_L^t}{2}-\dfrac{\rho_2^t Q_N^t}{2}\right)}, \tag{A10}$$

which is the first of equations (13) in the text. Notice that the denominator is positive.

PART B: *Employees*: The expected payoffs to *R*- and *E*-employees after factor market integration (equations (12) in the text) are:

$$v_R(n)=n\left\{\phi_2\left[(1+\alpha)\frac{\rho_2^t Q_L^t}{2}-\delta\right]+(1-\phi_2)[w-\eta-\alpha(\rho_2^t Q_N^t-m)]\right\}+(1-n)\frac{1}{2}\left\{\left[(1+\alpha)\frac{\rho_1^o Q_L^o}{2}-\delta\right](1-\varepsilon)+[w-\eta-\alpha(\rho_1^o Q_N^o-m)]\varepsilon+\right.$$
$$\left.+\phi_2\left[(1+\alpha)\frac{\rho_2^t Q_L^t}{2}-\delta\right]+(1-\phi_2)[w-\eta-\alpha(\rho_2^t Q_N^t-m)]\right\},$$
$$v_E(n)=n\left[\phi_2\left(\frac{\rho_2^t Q_N^t}{2}-\eta\right)+(1-\phi_2)(w-\eta)\right]+(1-n)\frac{1}{2}\left[\left(\frac{\rho_1^o Q_N^o}{2}-\eta\right)(1-\varepsilon)+(w-\eta)\varepsilon+\phi_2\left(\frac{\rho_2^t Q_N^t}{2}-\eta\right)+(1-\phi_2)(w-\eta)\right]. \tag{A11}$$

To obtain $\phi_2^*(n)$, we compute the value of $\phi_2(n)$ such that $v_R(n)=v_E(n)$, so we can write

$$(1+n)\frac{1}{2}\left\{\phi_2\left[(1+\alpha)\frac{\rho_2^t Q_L^t}{2}-\delta\right]-(1-\phi_2)\alpha(\rho_2^t Q_N^t-m)\right\}+(1-n)\frac{1}{2}\left\{\left[(1+\alpha)\frac{\rho_1^o Q_L^o}{2}-\delta\right](1-\varepsilon)-\alpha(\rho_1^o Q_N^o-m)\varepsilon\right\}=$$
$$=(1+n)\frac{1}{2}\phi_2\left(\frac{\rho_2^t Q_N^t}{2}-\eta\right)+(1-n)\frac{1}{2}\left[\left(\frac{\rho_1^o Q_N^o}{2}-\eta\right)(1-\varepsilon)\right],$$

which can be rewritten as

$$(1+n)\frac{1}{2}\left\{\phi_2\left[(1+\alpha)\frac{\rho_2^t Q_L^t}{2}-\delta\right]+\phi_2\alpha(\rho_2^t Q_N^t-m)\right\}-(1+n)\frac{1}{2}\phi_2\left(\frac{\rho_2^t Q_N^t}{2}-\eta\right)=$$
$$=(1-n)\frac{1}{2}\left[\left(\frac{\rho_1^o Q_N^o}{2}-\eta\right)(1-\varepsilon)\right]+(1+n)\frac{1}{2}\alpha(\rho_2^t Q_N^t-m)-(1-n)\frac{1}{2}\left\{\left[(1+\alpha)\frac{\rho_1^o Q_L^o}{2}-\delta\right](1-\varepsilon)-\alpha(\rho_1^o Q_N^o-m)\varepsilon\right\}.$$

Finally, we obtain

$$\phi_2^* = \frac{\frac{1-n}{1+n}\frac{1}{2}\left(\frac{\rho_1^o Q_N^o}{2}-\eta\right)(1-\varepsilon) - \frac{1-n}{1+n}\frac{1}{2}\left\{\left[(1+\alpha)\frac{\rho_1^o Q_L^o}{2}-\delta\right](1-\varepsilon)-\alpha(\rho_1^o Q_N^o - m)\varepsilon\right\} + \frac{1}{2}\alpha(\rho_2^t Q_N^t - m)}{\frac{1}{2}\left\{\left[(1+\alpha)\frac{\rho_2^t Q_L^t}{2}-\delta\right]+\alpha(\rho_2^t Q_N^t - m)\right\} - \frac{1}{2}\left(\frac{\rho_2^t Q_N^t}{2}-\eta\right)}, \qquad \text{(A12)}$$

which is the second of equations (13) in the text. Notice that the denominator is positive.

A.2.2 Effects of factor market integration on the costs of deviation. In this subsection we show that the cost of deviation from the best response convention in the {*F*,*E*} cultural-institutional equilibrium decreases after factor market integration.

PART A: *Employers.* First we write the expected payoff equations for employers under factor market integration when all the employees in the previous period were self-regarding. These are given by equations (A9) with ω=0,

$$\begin{aligned} v_P(n,\omega_2=0) &= n\frac{\rho_2^t Q_N^t}{2} + (1-n)\left\{\frac{1}{2}\left[\frac{\rho_2^t Q_L^t}{2}(1-\varepsilon)+\frac{\rho_2^t Q_N^t}{2}\varepsilon\right]+\frac{1}{2}\frac{\rho_2^t Q_N^t}{2}\right\}, \\ v_F(n,\omega_2=0) &= \rho_2^t Q_N^t - m. \end{aligned} \qquad \text{(A13)}$$

The cost of deviation for an employer in the {*F*,*E*} equilibrium after factor market integration is given by $v_F(\omega=0,n)-v_P(\omega=0,n)$. This difference is smaller than the corresponding expression under factor immobility (*n*=1) given in (A6). This is easily shown by the fact that while the expected payoff of a *F*-contract best responding employer [$v_F(\omega=0,n)$] is unaltered, the expected payoff of a *P*-contract idiosyncratic player [$v_P(\omega=0,n)$] is greater because $\rho_2^t Q_L^t/2 > \rho_2^t Q_N^t/2$.

PART B: *Employees.* The expected payoff equations for respectively *R*- and *E*-type employees under factor mobility when all the employers in the previous period were offering *F*-contracts, i.e. equations (A11) with $\phi_2=0$, may be written as:

$$v_R(n,\phi_2=0) = n[w-\eta-\alpha(\rho_2^t Q_N^t - m)] + (1-n)\frac{1}{2}\left\{\left[(1+\alpha)\frac{\rho_1^o Q_L^o}{2}-\delta\right](1-\varepsilon)+[w-\eta-\alpha(\rho_1^o Q_N^o - m)]\varepsilon + [w-\eta-\alpha(\rho_2^t Q_N^t - m)]\right\}$$

$$v_E(n,\phi_2=0) = n(w-\eta)+(1-n)\frac{1}{2}\left[\left(\frac{\rho_1^o Q_N^o}{2}-\eta\right)(1-\varepsilon)+(w-\eta)\varepsilon+(w-\eta)\right].$$

(A14)

The cost of deviation for an employee in the {*F*,*E*} equilibrium after factor market integration is given by $v_E(\phi_2=0,n)-v_R(\phi_2=0,n)$, which is smaller than the corresponding expression under factor immobility (*n*=1) given in (A8). Indeed while both the expected payoff of *E*-type best responding employees [$v_E(\phi_2=0,n)$] and the expected payoff of *R*-type idiosyncratic players [$v_R(\phi_2=0,n)$] increase after factor market integration, the latter increases more than the former. This is proven to be true providing that

$$\left[(1+\alpha)\frac{\rho_1^o Q_L^o}{2}-\delta\right](1-\varepsilon)+[w-\eta-\alpha(\rho_1^o Q_N^o-m)]\varepsilon>\left(\frac{\rho_1^o Q_N^o}{2}-\eta\right)(1-\varepsilon)+(w-\eta)\varepsilon,$$

which may be rewritten as

$$\varepsilon<1-\frac{\alpha(\rho^o Q_N^o-m)}{\left[(1+\alpha)\frac{\rho^o Q_L^o}{2}-\delta\right]-\left(\frac{\rho^o Q_N^o}{2}-\eta\right)+\alpha(\rho^o Q_N^o-m)}. \tag{A15}$$

Inequality (A15) is true because it is equivalent to $\varepsilon<1-\phi_1^*$, which follows from the persistence conditions (see Section 4 in the paper).

A.2.3 Effects of factor market integration on the critical values $\boldsymbol{\omega_2^*(n)}$ and $\boldsymbol{\phi_2^*(n)}$. In this subsection we show that factor market integration leads to a decrease in the expected number of idiosyncratic players in either class (employers and employees) sufficient to induce a transition from the $\{F,E\}$ to the $\{P,R\}$ cultural-institutional convention. To show this, we study the sign of the derivative of $\omega_2^*(n)$ and $\phi_2^*(n)$, given respectively by (A10) and (A12), with respect to n.

PART A: We have:

$$\frac{\mathrm{d}\omega_2^*(n)}{\mathrm{d}n}=\frac{\frac{1}{2}\left[\frac{\rho_2^t Q_L^t}{2}(1-\varepsilon)+\frac{\rho_2^t Q_N^t}{2}\varepsilon\right](1+n)-\left\{(\rho_2^t Q_N^t-m)-(1-n)\frac{1}{2}\left[\frac{\rho_2^t Q_L^t}{2}(1-\varepsilon)+\frac{\rho_2^t Q_N^t}{2}\varepsilon\right]\right\}}{(1+n)^2}$$

$$=\frac{\left[\frac{\rho_2^t Q_L^t}{2}(1-\varepsilon)+\frac{\rho_2^t Q_N^t}{2}\varepsilon\right]-(\rho_2^t Q_N^t-m)}{(1+n)^2}.$$

$\mathrm{d}\omega_2^*(n)/\mathrm{d}n$ is positive if and only if

$$\left[\frac{\rho_2^t Q_L^t}{2}(1-\varepsilon)+\frac{\rho_2^t Q_N^t}{2}\varepsilon\right]-(\rho_2^t Q_N^t-m)>0$$

which can be rewritten as

$$\varepsilon<1-\frac{\frac{\rho_2^t Q_L^t}{2}-m}{\frac{\rho_2^t Q_L^t}{2}-\frac{\rho_2^t Q_N^t}{2}}. \tag{A16}$$

Inequality (A16) is true because it is equivalent to $1-\omega_2^*$, which follows from the persistence conditions (see Section 4 in the paper).

PART B: We now turn to study the sign of $\mathrm{d}\phi_2^*(n)/\mathrm{d}n$. The denominator of (A12) is positive and the last fraction on the numerator does not depend on n. Hence it is easily shown that $\mathrm{d}\phi_2^*(n)/\mathrm{d}n>0$. Indeed $\mathrm{d}[(1-n)/(1+n)]/\mathrm{d}n<0$ and

$$\left(\frac{\rho_1^o Q_N^o}{2}-\eta\right)(1-\varepsilon)-\left[(1+\alpha)\frac{\rho_1^o Q_L^o}{2}-\delta\right](1-\varepsilon)+\alpha(\rho_1^o Q_N^o-m)\varepsilon<0,$$

which is equivalent to (A15) and is always true for the same reason that (A15) is true.

A.2.4 Transition-inducing degree of national specificity. There exists a critical value, $n^*>0$, of the degree of national specificity of the factor markets such that for $n<n^*$ one of (or both) the critical values, $\omega_2^*(n)$ and $\phi_2^*(n)$, is (are) negative, so innovators do better than those conforming to the erstwhile convention, inducing a cultural-institutional transition from the $\{F,E\}$ to the $\{P,R\}$ convention. Denote by n_ω^* the level of n such that $\omega_2^*(n)=0$, and by n_ϕ^* the level of n such that $\phi_2^*(n)=0$, the transition-inducing degree of national specificity is $n^*=\max[n_\omega^*,n_\phi^*]$. **PART A:** Equating expression (A10) to zero,

$$\omega_2^*(n)=\frac{\dfrac{\rho_2^t Q_N^t-m}{1+n}-\dfrac{1-n}{1+n}\dfrac{1}{2}\left[\dfrac{\rho_2^t Q_L^t}{2}(1-\varepsilon)+\dfrac{\rho_2^t Q_N^t}{2}\varepsilon\right]-\dfrac{1}{2}\dfrac{\rho_2^t Q_N^t}{2}}{\dfrac{1}{2}\left(\dfrac{\rho_2^t Q_L^t}{2}-\dfrac{\rho_2^t Q_N^t}{2}\right)}=0,$$

and solving for n, the former is given by

$$n_\omega^*=\frac{\dfrac{1}{2}\left[\dfrac{\rho_2^t Q_L^t}{2}(1-\varepsilon)+\dfrac{\rho_2^t Q_N^t}{2}\varepsilon\right]+\dfrac{1}{2}\dfrac{\rho_2^t Q_N^t}{2}-(\rho_2^t Q_N^t-m)}{\dfrac{1}{2}\left[\dfrac{\rho_2^t Q_L^t}{2}(1-\varepsilon)+\dfrac{\rho_2^t Q_N^t}{2}\varepsilon\right]-\dfrac{1}{2}\dfrac{\rho_2^t Q_N^t}{2}}. \qquad \text{(A17)}$$

PART B: Equating expression (A12) to zero,

$$\phi_2^*=\frac{\dfrac{1-n}{1+n}\dfrac{1}{2}\left[\left(\dfrac{\rho_1^o Q_N^o}{2}-\eta\right)(1-\varepsilon)\right]-\dfrac{1-n}{1+n}\dfrac{1}{2}\left\{\left[(1+\alpha)\dfrac{\rho_1^o Q_L^o}{2}-\delta\right](1-\varepsilon)-\alpha(\rho_1^o Q_N^o-m)\varepsilon\right\}+\dfrac{1}{2}\alpha(\rho_2^t Q_N^t-m)}{\dfrac{1}{2}\left\{\left[(1+\alpha)\dfrac{\rho_2^t Q_L^t}{2}-\delta\right]+\alpha(\rho_2^t Q_N^t-m)\right\}-\dfrac{1}{2}\left(\dfrac{\rho_2^t Q_N^t}{2}-\eta\right)}=0,$$

and solving for n, the latter is

$$n_\phi^*=\frac{\left[(1+\alpha)\dfrac{\rho_1^o Q_L^o}{2}-\delta\right](1-\varepsilon)-\alpha(\rho_1^o Q_N^o-m)\varepsilon-\dfrac{1}{2}\alpha(\rho_2^t Q_N^t-m)-\left(\dfrac{\rho_1^o Q_N^o}{2}-\eta\right)(1-\varepsilon)}{\left[(1+\alpha)\dfrac{\rho_1^o Q_L^o}{2}-\delta\right](1-\varepsilon)-\alpha(\rho_1^o Q_N^o-m)\varepsilon+\dfrac{1}{2}\alpha(\rho_2^t Q_N^t-m)-\left(\dfrac{\rho_1^o Q_N^o}{2}-\eta\right)(1-\varepsilon)}. \qquad \text{(A18)}$$