International Trade and the Persistence of Cultural-Institutional Diversity

Marianna Belloc
Samuel Bowles

SFI WORKING PAPER: 2009-03-005

SFI Working Papers contain accounts of scientific work of the author(s) and do not necessarily represent the views of the Santa Fe Institute. We accept papers intended for publication in peer-reviewed journals or proceedings volumes, but not papers that have already appeared in print. Except for papers by our external faculty, papers must be based on work done at SFI, inspired by an invited visit to or collaboration at SFI, or funded by an SFI grant.

©NOTICE: This working paper is included by permission of the contributing author(s) as a means to ensure timely distribution of the scholarly and technical work on a non-commercial basis. Copyright and all rights therein are maintained by the author(s). It is understood that all persons copying this information will adhere to the terms and constraints invoked by each author's copyright. These works may be reposted only with the explicit permission of the copyright holder.

www.santafe.edu
Abstract

Cultural and institutional differences among nations may affect 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 experimental and historical evidence motivating a model of the joint dynamics of preferences and institutions under autarchy and trade. 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) economic integration 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) the distributional impact of convergence to superior culture and (domestic) institutions differs from the impact of the removal of barriers to trade such that under plausible conditions the welfare of workers is enhanced by a cultural institutional transition but not by trade integration, while the reverse is true for employers.

JEL CODES: D02, F15, F16

KEYWORDS: institutions, incomplete contracts, culture, trade integration
1. Introduction

Because the verifiability of the quality of inputs and output differ among goods, goods also differ in the extent to which their production and distribution can be cost-effectively governed by complete contracts, ranging from those few commodities where complete contracts are feasible (allowing use of piece rates in production, for example) to others (such as the production and distribution of information, labor effort, and many other services) whose quality and quantity is highly variable and difficult to assess, and where as a result complete contracts generally cannot be written and enforced. We call the former kind of goods transparent and the latter opaque.

It follows that institutions, particularly those involving the contractual environment, differentially affect production costs depending on where the good lies along the opaque – transparent continuum. For opaque goods, social norms such as the work ethic and truth-telling are widely recognized to facilitate exchange and production (Arrow, 1971). Thus countries where these and other social preferences are common are able to make cost-effective use of incomplete contracts and, as a result, are likely to have a comparative advantage in the production of opaque goods. The implied correspondence between the goods in which a country specializes and its institutions and cultures has been much commented upon by historians and economists with a historical orientation (Kindleberger, 1962; Sokoloff and Engerman, 2000).

An example is the study by Eric Nilsson (1994) of the effects of the emancipation of slaves at the time of the U.S. Civil War on comparative advantage and specialization. Cotton, according to Nilsson, was a “slave commodity” for which effort levels beyond that which could be coerced from the worker were of little importance. For other commodities – manufacturing and tobacco in Nilsson’s empirical study –, variations in the quality of effort were more important, and difficult to secure by coercion. A similar distinction between sugar and tobacco was made in the much earlier study of Cuba by Fernando Ortiz (1963) in which the contrasting hierarchical and authoritarian culture of the sugar plantation regions and the liberal culture of the tobacco family farming areas were described. To study the effect of an exogenous institutional shock on specialization, Nilsson exploited the natural experiment provided by the end of slavery following the victory of the Union forces in the U.S. civil war. His study of production specialization in 169 counties in the Confederacy found that the end of slavery brought about a significant change in specialization, away from the “slave commodity” (cotton) and towards other commodities (manufacturing and tobacco). In both cases the effects are large and highly statistically significant. Until recently, these and other examples of institutional effects on trade and specialization have been given little attention by economic theorists.

We develop a two-good/two-country model with endogenous preferences and institutions where employee-employer relations are shaped by social norms and governed by (possibly incomplete) contracts. We refer to differences across economies in the kinds of contracts that are offered as institutional differences, while 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 (2008), Tabellini (2008) and Spolaore and Wacziarg (2008), we study the economic importance of cultural differences. In our model, institutions are conventions (mutual best responses), and which of the many possible conventions will occur depends on the preferences which prevail in a given country. In similar manner, the distribution of preferences is based on a cultural updating process (similar to that modeled by Bowles, 1998, Bisin and Verdier, 2001, and Fershtman and Bar-Gill, 2005), 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 institutions that supports the multiplicity of equilibria in our model. Transitions may occur among these cultural-institutional conventions as a result of individual experimentation and other forms of idiosyncratic play. (We briefly consider the case of bargained cooperative transitions in Section 6).

We derive three key results. First, for historical reasons two otherwise identical countries may support different cultural-institutional equilibria and these cross-country differences in the institutional and cultural environment are an independent source of comparative advantage, even in the absence of differences in technologies or factor endowments. Second, using our model of endogenous institutions and preferences we show that economic integration reinforces rather than destabilizes institutional and cultural diversity and may even impede transitions to Pareto-improving conventions. The second result contradicts the view, popular 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 countries specialize in the goods that are relatively more advantaged by their institutions and preferences, trade increases the joint surplus in the institutional and cultural status quo, thereby also raising the cost of local deviations from the prevalent preferences and contracts. It also may increase the number of innovators or other deviants required to induce a transition to the superior convention. Third, employers and employees may have different interests in cultural-institutional transitions and the removal of barriers to trade, employers benefiting from the latter but not necessarily from the former, while employers benefit from the former but not from the latter.

Opaque goods make up a substantial fraction of the output of the more advanced economies. Examples are the production and distribution of information-intensive goods and many services ranging from health care to entertainment and other recreational services. By contrast poorer nations produce large shares of agricultural and manufactured goods that are closer to the transparent pole of the opaque-transparent continuum. Cultural comparisons across nations are more difficult but 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 incomplete contracts and extensive social preferences such as trust and the work ethic as having a “good” cultural-institutional environment and, as a result, enjoying absolute advantage with respect to other countries in which complete contracts may elicit low (but not high) quality contributions from entirely self-interested economic agents. Here we depart from the common practice in the institutions and trade literature (e.g. Bardhan, Mookherjee, and Tsumagari, 2007) where complete contracts are characteristic of the rich “North” while incomplete contracts prevail in the “South”. The difference arises from our representation of incomplete contracts and social preferences as complements rather than substitutes, which is empirically motivated both by observations in natural settings and by behavioral experiments in which complete contracts often crowd out trust, the work ethic, and reciprocity (Bewley, 1999; Bowles, 2008). Thus the use of incomplete contracts need not reflect a nation’s deficient institutional environment; rather it may be a profit maximizing choice in a society where reciprocal and other social preferences are common.

We begin with empirical evidence for economically relevant cultural differences, and the mutual dependence of these with institutional differences (Section 2). We then develop a model of endogenous preferences and contractual choice showing that asymptotically stable cultural-institutional equilibria may exist and are not generally unique (Section 3). We embed this model of the co-evolution of preferences and institutions in a model of international exchange, illustrating cultural-institutional comparative advantage (Section 4). We explore the persistence of cultural and institutional differences following trade integration in a context of non-cooperative decentralized interactions (Section 5), briefly considering the case of a bargained cooperative game (Section 6). Finally, we conclude with some historical cases of trade induced cultural and institutional divergence (Section 7).

2. Endogenous preferences and contracts

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 longstanding 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.

Evidence of the importance of exchange-supporting norms is provided by both historical and other case studies (Greif, 1994; Baker, 1984) and behavioral experiments (Fehr, Klein and Schmidt, 2007; Fehr and Gaechter, 2000). Experimental subjects, for example, frequently choose actions resulting in zero payoffs for both parties, rather than accept what they consider to be unfair terms of exchange with their experimental partner. The extent of norms supporting exchange under incomplete contracts and other preferences influencing economic behavior
appears to differ significantly among societies (Inglehart, 1977; Herrmann, Thoni, and Gaechter, 2008; Henrich, Boyd, Bowles et al., 2005). Where such social norms are absent, the costs of producing goods for which complete contracting is infeasible will be elevated. Thus one may expect the distribution of preferences to influence the product composition of output via its influence on the distribution of cost-effective contracts.

The converse is also true: the contracts in force influence the evolution of preferences. Martin Brown, Armin Falk, and Ernst Fehr (2004) designed an experiment to explore the effects of contractual incompleteness on the nature of labor market interactions. In their complete contracting experimental condition, the level of effort promised by the “employee” was enforced by the experimenter, while in their incomplete contracting condition the employee could provide any level of effort (irrespective of any promise or agreement with the employer). Employers and employees knew the identification numbers of those they were interacting with, so they could use information they had acquired in previous rounds as a guide to with whom they would like to interact, the wage and level of effort to offer, and the nature of the contract they would offer. Employers had the opportunity to make a private contract offer (rather than broadcasting a public offer) to the same employee in the next period, thus attempting to initiate an ongoing relationship with the employee.

Very different relations between employers and employees emerged under the complete and incomplete contracting conditions. In the first, 90 percent of the employment relationships lasted less than three periods (most of them were single-shot). By contrast, only 40 percent of the relationships were this brief under the incomplete contracting condition and most agents formed trusting relationships with their partners. “Employers” in the incomplete contracting condition offered wages considerably in excess the employee's cost of providing effort. When employers were disappointed by the effort provided by the employees, they terminated the relationship, thereby withdrawing the implied rent from the employee. The differences in behavior under the two treatments were particularly pronounced in later rounds of the game, suggesting that the agents learned from their experiences, and updated their behaviors accordingly.

Relatedly, Peter Kollock (1992) investigated the structural origins of trust in a system of exchange that was analogous to the Thai rice and raw rubber markets described above, with similar results. Using an experimental design based on the exchange of goods of variable quality, he found that trust in and commitment to trading partners as well as a concern for ones own and others' reputations emerges when product quality is variable and non-contractible but not when quality is contractible.

These experiments suggest that reciprocal or trusting agents might respond negatively to the imposition of a more complete contract. Armin Falk and Michael Kosfeld (2006) explored the idea that “control aversion” may be a mechanism underlying this counter-productive aspect of incentives. Experimental agents in a role similar to an employee chose a level of “production” that was costly to them and beneficial to the principal (the employer). The agent's choice effectively determined the distribution of gains between the two, with the agent's maximum
payoff occurring if she produced nothing. Before the agent's decision, the principal could elect to leave the choice of the level of production completely to the agent's discretion, or impose a lower bound on the agent's production (three bounds were varied by the experimenter across treatments, the principal’s choice was whether or not to impose it). The principal could infer that a self-regarding agent would perform at the lower bound and thus imposition of the bound would maximize the principal’s payoffs.

But in the experiment, agents chose a lower level of production when the principal imposed the bound. Apparently anticipating this response, fewer than a third of the principals opted for its imposition in the moderate or low bound treatments. The minority of “untrusting” principals earned on average half of the profits of those who did not seek to control the agents’ choice in the low bound treatment, and a third less in the intermediate bound condition. In post-play interviews, most agents agreed with the statement that the imposition of the lower bound was a signal of distrust.

These and other experiments (Bowles, 2008) suggest that the distribution of preferences in a population depends in part on the distribution of contracts. To model this mechanism, the acquisition of new preferences may be represented as a now-standard cultural evolution process in which individuals periodically update their behavioral norms (perhaps frequently, perhaps only during adolescence) 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).

Equilibrium (that is, stationary) preferences will depend on the nature of the updating rules and the structure of social interactions given by the society’s institutions. The latter are important as they determine who meets whom to do what tasks and with what benefits. Among the institutions making up this cultural environment are the structure of markets, contracts, legislation and other aspects of society affected by public policy. A consequence is that differing economic institutions support different equilibrium preferences.

These observations along with the experimental evidence suggest that the norms that proliferate and become common in a society may depend on the prevalent forms of contract, with contractual incompleteness supporting the evolution of trusting and reciprocal behaviors. The converse is also true: lower levels of trust and reciprocity would lead those offering contracts to be willing to pay for more complete contracts. If levels of trust and reciprocity on the one hand and contractual incompleteness on the other are mutually determining, one may define a joint equilibrium set of norms and contracts. For a given population there may be any number of these equilibria, some with high levels of trust and relatively incomplete contracts and others with the converse. In the high trust-incomplete contract equilibrium, goods whose production and distribution entail non-verifiable information and other sources of contractual incompleteness will have lower marginal costs (relative to the marginal costs of goods for which complete contracting is less costly) than in the latter.

To explore the effects of endogenous preferences and institutions in a globally integrated
setting we provide the following model.

3. Cultural-institutional equilibrium under autarchy

An economy is populated by employers and employees. Employers hire employees to produce one of two goods: the opaque good (denoted by the superscript \( o \)) and the transparent good (denoted by \( t \)). Markets are competitive in the sense that employers take the price of the good as exogenously given. There is only one factor of production, labor, which is perfectly mobile across industries but immobile across countries. Each employee is endowed with one unit of labor which can be provided in production with two different qualities: low quality effort (\( e_L \)) and high quality effort (\( e_H \)), one unit of high effort producing more (in either good) than one unit of low effort. The employment relationship is a random employee-employer pair for a single interaction and is regulated by a contract. High quality effort cannot be contracted for because it is not verifiable due to information asymmetries or other reasons.

The production process of the opaque good is more intensive in effort quality: the increase in production obtained employing high rather than low quality effort is relatively greater in the \( o \)-sector than in the \( t \)-sector. Denoting by \( Q^i(e_H) \) and \( Q^i(e_L) \) the quantity of good \( i \) (\( i = t, o \)) obtained using one unit of respectively high quality effort and low quality effort, we therefore have:

\[
\frac{Q^o(e_H)}{Q^o(e_L)} > \frac{Q^t(e_H)}{Q^t(e_L)}.
\]

Employers maximize profits, while employees maximize utility. The employer may offer the employee one of two contracts: complete (C) or incomplete (I). If the complete contract is offered, the employee receives a fixed compensation (\( w > 0 \)) just sufficient to offset the cost of providing low quality effort (\( \eta > 0 \)). These are \( C \)-type employers. Employers offering a \( C \)-contract must also pay a cost of \( \mu (> 0) \) for monitoring and contractual enforcement. According to the incomplete contract, the employer pays the employee half of the output resulting from the transaction and does not monitor the employee. These are \( I \)-type employers.

Employees are also of two types: reciprocator (R) or selfish (S). S-type employees are completely self-regarding and provide low effort irrespective of the contract. R-type employees interpret the \( I \)-contract as a sign of trust on the part of the employer, and reciprocate by providing high effort, incurring a cost of \( \delta (> \eta) \) as a result. However, as in the Falk and Kosfeld (2006) experiment above, when offered a \( C \)-contract (in which case the employer garners the whole of the joint surplus) R-type employees feel that the employer is distrusting and seeking to exploit them. As a result, they experience a subjective cost and reject the contract, no production taking place. The extent of the subjective cost is proportional to the profits the employer would have made had the contract been accepted. For example, if competitive conditions constrain the employer to a zero-surplus payoff, the R-type worker would feel no insult if she also were to
receive a wage that just compensated the subjective cost of her effort. This is consistent with experimental play in the Ultimatum Game (Cameron, 1998): second movers are willing to forego a larger payoff to avoid participating in a transaction in which the first mover would make a lot of money. Thus we assume that for the reciprocal workers the cost for being offered a complete contract is proportional (with a proportionality factor $\gamma$) to the profits that the employer would make if the employee were to accept the contract.

Agents consume a composite bundle (indicated by $c$) 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 $p^i$ and $p^o$ the price of the $t$-good and the price of the $o$-good, we define $\rho^i = p^i/(p^i+p^o)$ and $\rho^o = p^o/(p^i+p^o)$ respectively as 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 unit of the $c$-good one can purchase with one unit of the $t$-good). The utility function is additive in consumption of the composite good and subjective utility associated with the contract and effort quality.

Table 1, where $\pi^i$ and $u^i$ denote respectively the profits of the employers and the utility of the employees in sector $i = t,o$, reports the matrix of payoffs, which are measured in number of units of the composite good commanded. Since in autarchic equilibrium both goods are produced and labor is mobile among sectors within a country, $\rho^i_Q^o(e_H)$ must be equal to $\rho^i_Q^i(e_H)$ and $\rho^o_Q^o(e_L)$ must be equal to $\rho^o_Q^i(e_L)$, the subscript “A” referring to autarchy. Thus the entries in Table 1 are invariant across sectors. We assume that employers and employees update their contracts and preferences with regard to these payoffs (that is, on the basis of inter-sectoral competitive equilibrium prices).

<table>
<thead>
<tr>
<th>Employer/Contract</th>
<th>Reciprocator</th>
<th>Selfish</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incomplete</td>
<td>$\pi^i(e_H) = \rho^i_Q^i(e_H)/2$</td>
<td>$\pi^i(e_L) = \rho^i_Q^i(e_L)/2$</td>
</tr>
<tr>
<td></td>
<td>$u^i(e_H) = \rho^i_Q^i(e_H)/2 - \delta$</td>
<td>$u^i(e_L) = \rho^i_Q^i(e_L)/2 - \eta$</td>
</tr>
<tr>
<td>Complete</td>
<td>$\pi^i(0) = 0$</td>
<td>$\pi^i(e_L) = \rho^i_Q^i(e_L) - w - \mu$</td>
</tr>
<tr>
<td></td>
<td>$u^i(0) = -\gamma[\rho^i_Q^i(e_L) - w - \mu]$</td>
<td>$u^i(e_L) = w - \eta = 0$</td>
</tr>
</tbody>
</table>

*Table 1: Matrix of payoffs. (NOTE: Payoffs in bold type indicate pure stable Nash equilibria)*
in a dynamic setting because the “inferior” convention \{C,S\} is also asymptotically stable.

Writing the fraction of the employers offering incomplete contracts in the previous period as \(\phi\) and recalling that in competitive equilibrium the payoffs in Table 1 are invariant across sectors, the expected payoffs to the \(R\)- and \(S\)-employees are respectively:

\[
v_R = \phi \left[ \frac{\rho^i Q^i(e_H)}{2} - \delta \right] + (1 - \phi)(-\gamma)[\rho^i Q^i(e_L) - (w + \mu)],
\]
\[
v_S = \phi \left[ \frac{\rho^i Q^i(e_L)}{2} - \eta \right] + (1 - \phi) \times 0.
\]

Similarly, writing the fraction of the employees who were reciprocators in the previous period as \(\omega\), the expected payoffs to employers offering the \(I\)- and \(C\)-contracts are:

\[
v_I = \omega \frac{\rho^i Q^i(e_H)}{2} + (1 - \omega) \frac{\rho^i Q^i(e_L)}{2},
\]
\[
v_C = \omega \times 0 + (1 - \omega)[\rho^i Q^i(e_L) - (w + \mu)].
\]

These expected payoff functions are illustrated in Figure 1, the vertical intercepts being taken from Table 1.

Figure 1: Expected payoffs under autarchy to \(R\)- and \(S\)-employees (panel A) and to \(I\)- and \(C\)-employers (panel B). (Note: \(\phi\) is the fraction of the employers offering incomplete contracts and \(\omega\) the fraction of the employees being reciprocators in the previous period. Payoffs in bold type refer to the stable pure Nash equilibria in Table 1)

Suppose that both employers and employees periodically update their preferences and the behavioral strategies they support by best responding to the distribution of play in the other class in the previous period. 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 (people do not condition their updating on an already known kind of employment contract in which they will engage). 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 has the same strategy as the employee A, A does not update. If, on the contrary, B follows a different strategy, A compares the two payoffs: if B has greater payoff, A switches to B’s strategy with a probability equal to $\beta (>0)$ times the payoff difference, whereas she retains her own strategy otherwise. It is easily shown that this process gives the replicator dynamic equations:

$$\frac{d\phi}{d\tau} = \phi(1-\phi)\beta[v_B(\omega) - v_C(\omega)],$$

$$\frac{d\omega}{d\tau} = \omega(1-\omega)\beta[v_R(\phi) - v_S(\phi)],$$

where $\tau$ denotes time. We are now interested in the stationary states, such that $d\phi/d\tau=0$ and $d\omega/d\tau=0$. It is easy to find that:

$$\frac{d\phi}{d\tau} = 0 \text{ for } \phi = 0, \phi = 1 \text{ and } \omega^* = \frac{\rho^i Q^i(e_L) - 2(w + \mu)}{\rho^i[Q^i(e_H) + Q^i(e_L)] - 2(w + \mu)},$$

$$\frac{d\omega}{d\tau} = 0 \text{ for } \omega = 0, \omega = 1 \text{ and } \phi^* = \frac{2\gamma[\rho^i Q^i(e_L) - (w + \mu)]}{\rho^i[Q^i(e_H) - Q^i(e_L)] - 2(\delta - \eta) + 2\gamma[\rho^i Q^i(e_L) - (w + \mu)]}.$$  

The resulting dynamical system is illustrated in Figure 2 where the arrows indicate the out of equilibrium adjustment given by the replicator dynamic equations. The states where $d\phi/d\tau=0$ and $d\omega/d\tau=0$ are institutional-cultural equilibria. The state $(\phi^*, \omega^*)$ is stationary, but it is a saddle: small movements away from $\phi^*$ or $\omega^*$ are not self-correcting. (Two additional unstable stationary states, namely $\{\phi = 1, \omega = 0\}$ and $\{\phi = 0, \omega = 1\}$ are of no interest.) The asymptotically stable states are $\{I\} \text{ and } \{C\} \text{ (that is, } \{1\} \text{ and } \{0,0\})$. In what follows it will be helpful to note (from equation (5)) that the critical values $\phi^*$ and $\omega^*$ are simply given by the cost (for respectively employees and employers) of deviating from the $\{0,0\}$ equilibrium divided by the sum of this cost and the cost of deviating from the $\{1,1\}$ equilibrium.

In this deterministic setting, the initial state determines which of these two asymptotically stable states occurs. Of course institutions may be altered by a joint decision of the representatives of one or both classes to adopt a new set of rules of the game (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, 1998). Suppose that with probability $1-\varepsilon$ myopic best response updating occurs as described above, but with probability $\varepsilon$ the employee chooses randomly from the two behavioral
traits and the employer likewise randomizes her contractual offer. Idiosyncratic play may be due to deliberate experimentation, error, or any other reason for non-best-response play. The rate of non-best-response play \( (\varepsilon) \) may be exogenously given, or may be state-dependent (Bergin and Lipman, 1996). For example the likelihood of experimentation or error may be the smaller, the greater is the cost of deviating from the status quo best response.

![Figure 2: Co-evolution of institutions and preferences, and persistence of two institutional-cultural equilibria in a given country.](image)

In a plausible version of this process (e.g. Young, 1998), the resulting perturbed Markov process is ergodic so over the long run both \( \{I,R\} \) and \( \{C,S\} \) will occur, with infrequent transitions (assuming that \( \varepsilon \) is small) 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 institutional-cultural 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 \( \{I,R\} \) equilibrium will be more persistent if \( \phi^* \omega^* < (1-\phi^*)(1-\omega^*) \) that is, if \( \{I,R\} \) is the risk-dominant equilibrium, and conversely for the \( \{C,S\} \) equilibrium.

4. Cultural-institutional comparative advantage

Assume now that the world economy comprises two countries, 1 and 2, that are identical in all relevant respects (same labor endowment, same technology, same demand function), except for their recent histories which have given them different cultural and institutional conventions. As a result their production possibility frontiers differ, and the two countries enjoy comparative advantage in the production of different goods.

Let us suppose that country 1 is near the \( \{I,R\} \) equilibrium (virtually all pairs except those playing idiosyncratically) are reciprocal types working under incomplete contracts, while country
2 is near the \{C,S\} equilibrium. As it is evident from Figure 3 (where the slope of the dashed lines indicates the international terms of trade), because \(Q^o(e_H) > Q^o(e_L)\) and \(Q^i(e_H) > Q^i(e_L)\), the \{I,R\} country enjoys an absolute advantage in the production of both goods. In autarchic equilibrium there will be only one relative price in each country such that both goods are produced equal to the marginal rate of transformation (MRT) in the two countries, namely: \(p^o_{1A} / p^o_{1A} = Q^o(e_H) / Q^o(e_H) = MRT_1\) and \(p^o_{2A} / p^2_{2A} = Q^o(e_L) / Q^o(e_L) = MRT_2\), where \(p^o_{1A} / p^o_{1A}\) and \(p^o_{2A} / p^2_{2A}\) are the autarchic relative prices in the two countries. Therefore, given (1), we have:

\[
\frac{p^o_{1A}}{p^o_{1A}} = \frac{Q^o(e_H)}{Q^o(e_H)} < \frac{Q^o(e_L)}{Q^o(e_L)} = \frac{p^o_{2A}}{p^2_{2A}},
\]

so that country 1 has a comparative advantage in the production of the opaque good, whereas country 2 has a comparative advantage in the production of the transparent good.

\[\text{Figure 3: Production possibility frontiers in the two countries.}\]

(NOTE: Each country has a normalized labor endowment of 1)

Providing that the international terms of trade, \(p^T_2 / p^T_1\) (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. Furthermore, 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 transport 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 $\rho_{o}^t > \rho_{o}^{\text{A}}$ and $\rho_{t}^{t} > \rho_{t}^{\text{A}}$, where $\rho_{o}^t = p_o^t / (p_o^t + p_o^t)$ and $\rho_{t}^{t} = p_t^t / (p_t^t + p_t^t)$. 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, $\rho_{o}^c Q^o(e_H) > \rho_{o}^{\text{A}} Q^o(e_H)$ and $\rho_{t}^c Q^t(e_L) > \rho_{t}^{\text{A}} Q^t(e_L)$: the c-good value of output in the two countries increases. All the other terms ($\delta, \eta, \mu$ and $\gamma$) in the payoff matrix (Table 1), which are measured in units of the composite goods, remain unaltered.

5. Effects of trade on culture and institutions

Thus 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? First, will the two countries different cultural-institutional equilibria persist after the two countries open up to international exchange? Second, 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, $\phi^*$ and $\omega^*$? 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 in the two countries as a result of trade. The circle (square) around the one (zero) on the horizontal axis in panel A signifies that the entire population of employers (except idiosyncratic players) in the given country is of the I-type (C-type); a circle (square) around the one (zero) on the horizontal axis in panel B indicates that the entire population of employees (except idiosyncratic players) in the given country is of the R-type (S-type). Payoffs received by the individuals in equilibrium before and after trade are emphasized in bold fonts in the relevant panel.

It is easy to verify that trade integration does not destroy the cultural institutional differences upon which specialization is based. After trade, the critical values of $\phi$ and $\omega$ remain within the unit interval and trade increases the cost of deviating from the status quo cultural-institutional convention for both groups in both countries. This is because deviating almost always entails a mismatch with the result being that some or all of the surplus is forgone, the value of which is higher after trade integration. Thus the cost of deviating is increasing in $\rho^i (i=t,o)$, and trade increases the amount of the composite good that may be purchased with one unit of the good in which the country specializes (i.e. increases $\rho^i$, where $i=o$ in country 1 and $i=t$ in country 2).
Fig. 4: Payoff changes to R- and S-employees (panel A) and to I- and C-employers (panel B) after trade openness. (NOTE: $\phi$ is the fraction of the employers offering incomplete contracts and $\omega$ the fraction of the employees being reciprocators in the previous period)
Table 2 shows the percentage increases in the cost of deviating from the status quo induced by trade, $\Delta \rho_i (i=t,o)$ denoting the change in $\rho$ from autarchy to trade.

<table>
<thead>
<tr>
<th>Equilibrium</th>
<th>Individual {I,R}</th>
<th>{C,S}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employee</td>
<td>$\frac{\Delta \rho_i [Q^i(e_H) - Q^i(e_L)]}{\rho_A^i [Q^i(e_H) - Q^i(e_L)] - 2(\delta - \eta)}$</td>
<td>$\frac{\Delta \rho_i Q^i(e_L)}{\rho_A^i Q^i(e_L) - (w + \mu)}$</td>
</tr>
<tr>
<td>Employer</td>
<td>$\frac{\Delta \rho_i}{\rho_A^i}$</td>
<td>$\frac{\Delta \rho_i Q^i(e_L)}{\rho_A^i Q^i(e_L) - 2(w + \mu)}$</td>
</tr>
</tbody>
</table>

Table 2: Trade induced percentage increases in the cost of deviation from the status quo cultural institutional convention.

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 critical values $\phi^*$ and $\omega^*$. This can be seen from equations (4), along with the fact that trade increases both $(v_I - v_C)$ and $(v_R - v_S)$ when $\omega = 1 = \phi$ and increases both $(v_C - v_I)$ and $(v_S - v_R)$ when $\omega = 0 = \phi$. Thus trade will not induce a non-cooperative transition from \{C,S\} to the \{I,R\} equilibrium despite the fact that country 2 welfare would thereby be enhanced because the \{I,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 \{C,S\} to the \{I,R\} equilibrium. To see this we study the effect of trade (that is, the increase in $\rho$ with $i=t,o$) on $\phi^*$ and $\omega^*$. In the case of $\omega^*$ the result is unambiguous: trade increases the critical fraction of reciprocal workers necessary to induce the C-type employers to best respond by adopting I-contracts

$$\frac{d \omega^*}{d \rho^i} = \frac{2(w + \mu)Q^i(e_H)}{\{\rho^i [Q^i(e_H) + Q^i(e_L)] - 2(w + \mu)\}^2} > 0.$$ (8)

The reason is that while the costs of deviating from both equilibria increase for the employers, trade increases the cost of deviating from the \{C,S\} equilibrium of country 2 proportionally more (in Table 2, the bottom-right entry exceeds the bottom-left entry).

The economics underlying this surprising result is as follows. Because in country 2 the cost of hiring labor under complete contracting $(w + \mu)$ is unaffected by the increase in the value of output induced by trade, the effect of trade on profits of a C-type employer matched with an S-employee is proportionally greater than the effect of trade on the payoff of an I-type employer
matched with a reciprocal employee. As a result, employers who offer the C-contract in anticipation of being matched with S-type workers are more reluctant following trade to abandon complete contracting and will switch to the I-contract only if they have encountered a larger fraction of reciprocal workers the previous period. For the same reason in country 1, employers offering the I-contract expecting matches with R-type workers will be induced to switch to the C-contract by fewer idiosyncratically playing S workers following trade than before.

The effect of trade on $\phi^*$ cannot be signed in general, but (under plausible conditions) it too may increase following integration. We have

$$\frac{d\phi^*}{d\rho} = \frac{2\gamma [Q^i(e_H) - Q^i(e_L)](w + \mu) - 2(\delta - \eta)Q^i(e_L)}{\rho^i [Q^i(e_H) - Q^i(e_L)] - 2(\delta - \eta) + 2\gamma [\rho^i Q(e_L) - (w + \mu)]^2},$$

from which we see that $d\phi^*/d\rho > 0$ iff $[Q^i(e_H) - Q^i(e_L)](w + \mu) > 2(\delta - \eta)Q^i(e_L)$. This will be the case if the increased contribution to output associated with high quality effort is sufficiently great compared to the increased disutility associated with providing it and/or the output obtained employing low quality effort is sufficiently small compared to the cost of hiring labor in complete contracting (the subjective cost of potential exploitation for a reciprocal employee being offered a C-contract is sufficiently small). Here we cannot sign the expression because we cannot determine which of the proportional effects of trade on the foregone utility of the worker should a mismatch occur is greater (compare top-left and top-right entries in Table 2).

6. Trade, distributional conflict, and the persistence of inefficient cultural-institutional equilibria

Thus economic integration need not destabilize and, indeed, may fortify the preexisting cultural and economic differences on which specialization and trade is 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.

The source of persistent inefficiency in this model is the coordination failure arising from the decentralized nature of preference formation and contractual choice coupled with the impossibility of matching behavioral types to contractual offers. But the persistence of cultural-institutional diversity and the associated inefficiency may be a more general result, occurring even if we abandon our assumption that culture and institutions evolve by a non-cooperative decentralized dynamic.

Suppose, by contrast, that at some cost of coordination and transition, the entire
memberships of the two classes could simultaneously bind themselves to switch from one Nash equilibrium to the other. It is easily shown that for the country with inferior institutions (country 2), the welfare gains associated with such a transition are greater under autarchy than following trade integration. The reason is that the best that country 2 could do by implementing such a transition is to attain the standard of living experienced by country 1 in autarchy, because such a transition would eliminate cultural-institutional diversity and, as a result, the gains from trade would also vanish. Because the income gap between country 2 and country 1 under autarchy necessarily exceeds the difference between country 2’s income under trade and country 1’s income under autarchy, the removal of barriers to trade thereby diminishes the prospective gains from such a cultural-institutional transition.

Note also that the gains from such a cultural-institutional transition to a superior equilibrium would be unequally distributed between economic agents. Because their utility is zero in the \{C,S\} equilibrium irrespective of the presence or absence of barriers to trade, whereas they have positive utility in the \{I,R\} convention, country 2 employees would necessarily gain from a transition to the \{I,R\} convention. But country 2 employers benefit from trade and specialization and hence may have lower incomes under the \{I,R\} cultural-institutional equilibrium. The reason is that in the latter specialization and trade do not occur (because the cultural-institutional transition eliminates the country differences on which specialization is based). In the absence of barriers to trade they may thus oppose a cultural institutional transition, even if they would have benefited by the transition under autarchy. To see this compare vertical intercepts in panel B.2 of Figure 4 drawn to illustrate the possibility of just this case. The employer’s payoff at the \{C,S\} equilibrium under trade \((ρ^i_Q(we) - w - μ)t\) may be greater than the employer’s payoff at the \{I,R\} equilibrium under autarchy \((ρ^i_A(we) / 2\), even if the latter are greater than the autarchy payoffs at the \{C,S\} equilibrium \((ρ^i_A(we) - w - μ)\).

Moreover to the extent that by increasing competition and for other reasons not explicitly treated in our model trade increases productivity levels, it may bring into existence de novo the conditions mentioned above for cultural-institutional diversity. Recall that the conditions for the existence of more than one asymptotically stable cultural-institutional equilibrium in our model are that \(ρ^i Q^i(e_L) > 2(w + μ)\) and \(ρ^i [Q^i(e_H) - Q^i(e_L)] > 2(\delta - \eta)\). These conditions will not be fulfilled by economies at very low levels of productivity (low output levels at either quality level relative to the additional disutility of high quality effort and to the cost of hiring labor under complete contracting). But if more conventional endowments-related sources of comparative advantage exist, trade-induced increases in the value of output might eventually satisfy this condition for cultural-institutional diversity, such that once established, distinct cultures and institutions could emerge and endure, even if (as in our model) there existed a convention that conferred absolute advantage in both goods.
7. Discussion

Our paper is a contribution to the rapidly growing literature (surveyed in Belloc, 2006) on institutions and trade. The existence of institutional comparative advantage has been investigated by a number of recent contributions which focus on different aspects of the institutional setting: 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), contracts and the division of labor (Acemoglu, Antrás, and Helpman, 2005; Antrás, 2005; Costinot, 2007), labor market flexibility and volatility (Cunat and Melitz, 2007), legal establishment and accounting systems (Vogel, 2007). Our paper differs in important ways from these contributions: rather than taking institutional differences as exogenous, we consider institutional comparative advantage in a model where endogenous cultural-institutional dynamics are the result of a decentralized non-cooperative game among economic agents. Like Belloc (2005), Casella and Feinstein (2002), Dixit (2003), Do and Levchenko (2006) and Levchenko (2008), we also study the impact of economic integration on domestic institutions. The main novelty of our approach with respect to these contributions consists in identifying the complementary relationship between cultural preferences and institutions as the mechanism by which institutions associated with absolute disadvantage may persist indefinitely. In particular, our paper departs from and complements Levchenko (2008) and Do and Levchenko (2006)’s papers because in their models institutional differences are a historical datum that may be modified by a cooperative lobbying game, while in our model they are an endogenously generated non-cooperative cultural-institutional equilibrium.

The co-evolution of social norms and institutions is also modeled by Francois (2008). However, differently from our paper, in his model institutional change occurs via the effort of an institutional designer which is external to the transaction (a political actor). Furthermore, while we explore the effects of trade openness on cultural-institutional equilibria, Francois (2008) studies those of increasing market competition. Like Conconi et al. (2008), we show that liberalization need not contribute to allocative efficiency. Yet in our model this occurs because inferior institutions are thereby reinforced, while in theirs the price changes associated with liberalization induce organizational changes that may reduce consumer welfare.

In this paper, we have shown that otherwise identical economies that differ in culture and institutions may find specialization and trade welfare-enhancing and that trade need not induce cultural-institutional convergence to superior cultural-institutional arrangements.

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 also diverged, as the social solidarity of the 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 by 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. Furthermore, Pagano (2007) shows that, in an integrated global economy, the distribution of intellectual assets across countries can be a source of institutional comparative advantage and that increasing economic integration is compatible with persistent diversity in domestic systems of intellectual property rights. Freeman (2000) and 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), 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 process.

References

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


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


