Cultural-Institutional Persistence under Autarchy, International Trade, and Factor Mobility

Marianna Belloc
Samuel Bowles

SFI WORKING PAPER: 2013-01-003

SANTA FE INSTITUTE

Cultural-Institutional Persistence under Autarchy, International Trade, and Factor Mobility*

Marianna Belloc† and Samuel Bowles‡

January 24, 2013

Abstract

We address two seemingly unrelated empirical anomalies: the remarkable historical persistence of cultural and institutional differences affecting production and distribution even among nations and regions engaged in extensive trading, and the shortcomings of the standard model that predicts international specialization and trade on the basis of differences in factor endowments or technologies. We model the endogenous evolution of both culture (the distribution of preferences affecting individual behavior) and institutions (the distribution of contracts among employers and employees), showing that in otherwise identical economies, different cultural-institutional conventions can persist over long periods. Transitions between cultural-institutional conventions occur as a result of decentralized and un-coordinated contractual or behavioral innovations by firms or workers. In a two-good/two-factor/two-country trade model, we then show that: (*i*) because goods differ in the kinds of contracts and preferences that are appropriate for their production, cultural-institutional differences support differing competitive prices in autarchy, and so provide the basis for specialization and comparative advantage; (*ii*) the resulting gains from trade raise the cost of deviations from the prevailing culture and institutions and, as a result, trade will impede transitions to the superior convention; and (*iii*) by contrast, by reducing the cost of innovating, international mobility of factors of production facilitates convergence to superior cultural-institutional conventions. Our model thus provides a possible unified resolution of the anomalies concerning patterns of specialization and trade, and cultural-institutional persistence.

Keywords: endogenous institutions, endogenous social preferences, evolutionary game theory, culture, trade integration, factor mobility, globalization

JEL Classification: D02, D03, F15, F16

*We would like to thank Alberto Alesina, Pranab K. Bardhan, Paolo Bartoletti, Tanmoy Bhattacharya, Ennio Bilancini, Robert Boyd, Eric Brousseau, Wendy Carlin, Roy Cerqueti, Jeremy van Cleve, Arnaud Costinot, Simone D'Alessandro, Michele Di Maio, Francesco Drago, Duncan Foley, Maurizio Franzini, Roberto Galbiati, Pierre Garrouste, Francesco Giavazzi, Peter Grajzl, Sung-Ha Hwang, Ricardo Hausmann, Guido Lorenzoni, Eric Nilsson, Nathan Nunn, Ugo Pagano, Emmanuel Raynaud, Arslan Razmi, Larry Samuelson, Pietro Vertova, Peyton Young, and Giulio Zanella for helpful comments. Thanks also to the Behavioral Science Program of the Santa Fe Institute, the U.S. National Science Foundation, the University of Siena and the European Science Foundation for support of this project.

†Sapienza University of Rome, via del Castro Laurenziano 9, 00161, Rome (Italy). Email: marianna.belloc@uniroma1.it.

‡(Corresponding author) Santa Fe Institute, 1399 Hyde Park Rd, Santa Fe, 87501, NM (USA). Email: bowles@santafe.edu.

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. Latin American institutions and elite cultures that owed their origin to the 16th century exploitation of slaves and coerced Native American labor on plantations and in mining persisted long after sugar and gold had lost their central role in these economies (Sokoloff and Engerman, 2000). Current levels of trust in distinct African populations vary inversely with the exposure to the slave trade that ended two centuries ago (Nunn and Wantchekon, 2011). 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). The system of coerced supply of mining labor in 17th and 18th century colonial Peru had effects on household consumption, land tenure and stature centuries later (Dell, 2010).

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 for a globally integrated world economy in which competition is sometimes thought to induce institutional and cultural convergence. Our explanation of the persistence of cultural and institutional differences is based on the endogenous co-determination 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 not heretofore formally modeled. We model the decentralized updating of both preferences and contractual choices and show how this process can support durable cultural and institutional differences. Because the impact of culture and institutions on the cost of production differs across commodity groups, these cultural-institutional differences could provide a basis for comparative advantage. The resulting specialization and trade, we show, need not obliterate cultural and institutional diversity, and this may help to account both for the long term persistence of cultural and institutional differences and for the ways in which trade patterns deviate from predictions based solely on factor endowments (see, e.g., Trefler, 1995; Helpman, 1999, for a survey).

Suggestive evidence that goods differ in the kinds of institutions and culture that are best adapted for their production is provided by economic historians. 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 this 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 'effort intensive' and 'care intensive' productive activities. A similar distinction between sugar and tobacco was made by Fernando Ortiz (1963) who contrasted the coerced labor and hierarchical and authoritarian culture of the sugar plantation regions of Cuba with the self-motivated labor and liberal culture of the tobacco family-farming areas.

We refer to differences across economies in the distribution of preferences (including social norms) as cultural differences, while differences in the distribution of employment contracts are referred to as institutional differences. Thus we might say that Norway has a trusting culture because most Norwegians exhibit high levels of interpersonal trust, or that farming by former slaves in the post-bellum U.S. south was governed by the institution of share cropping because this was by far the most common contract. Note that an institution – the prevalence of a particular share contract in farming, for example, or primigeniture as a wealth inheritance practice – need not reflect deliberate public policy – the explicit prohibition of alternative contracts, for example. Instead, like cultures, institutions may persist as the result of decentralized actions in non-cooperative settings (Young, 1998).

We develop a two-country/two-factor/two-good model in which countries may differ in their institutions and cultures. Production and distribution are governed by employers' choice between two contracts, either joint residual claimancy under partnerships or fixed wage and monitoring contracts with the employer as the sole residual claimant. The relevant preference differences are captured by assuming that employees may be reciprocal — responding positively to employers' trust in them, and conversely — or self-regarding — simply maximizing material payoffs. Finally, goods differ in the extent to which their production depends on qualitative labor, namely labor which is prohibitively costly to verify and hence cannot be cost-effectively secured by a fixed wage and monitoring contract requiring an explicit level of labor input. Where non-verifiable aspects of work are important to production, social norms such as reciprocity or a positive work ethic are required for high levels of productivity.

The main novelty of our approach and of our related paper (Belloc and Bowles, 2013) is that, rather than treating institutions and preferences as exogenous, or as determined by a national-level constitutional bargain, we use evolutionary game theory to model the interacting dynamics of both as the result of decentralized non-cooperative interactions among economic agents (Young, 1998; Binmore, Samuelson and Young, 2003). In our model, the contract that maximizes employers' profits depends on the preferences that prevail in a given country. Partnership contracts without monitoring, for example, are more profitable where social preferences such as the work ethic or reciprocity are common; purely self-regarding employees would take advantage of the lack of monitoring, reducing profits. The distribution of preferences in turn is based on a cultural updating process in which the payoffs associated with different preferences (and the behaviors they support) depend on the distribution of contracts in the economy.

As a result of the complementarity between distinct preferences and contracts, there may exist a number of distinct distributions of both preferences and contracts that are stationary because each is a mutual best response to the other. The subset of these stationary states that are evolutionarily stable are termed cultural-institutional conventions. It is this mutual dependence of preferences and contracts and the differences among goods

in the importance of non-verifiable qualitative labor that provides the basis for national specialization in our model, thus playing a role analogous to technology-based economies of scale in Paul Krugman's (1987) model of trade among countries with identical factor endowments and technologies. Transitions occur among these cultural-institutional conventions when sufficiently many innovators deviate from the status quo convention (adopting non-best response preferences or contracts) 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.

Second, economic integration may reinforce rather than destabilize institutional and cultural diversity and impede transitions, even to Pareto-improving conventions. This result contradicts the view, popular among critics of trade liberalization since John Maynard Keynes (1933), that trade will lead to institutional and cultural convergence and thus defeat attempts by nations that, as he put it (p. 762), would prefer to "have a try at working out our own salvation". Keynes' view is especially thought to be true when one nation's culture and institutions confer absolute advantage in all or most products. But since trade allows countries to specialize in the goods in which they are relatively more advantaged (or less disadvantaged) given their institutions and culture, it increases the joint surplus in the cultural-institutional status quo, even in an absolutely disadvantaged country. These gains from trade increase the returns available to employers and employees and, hence, raise the cost of an employer-employee mismatch that is likely to occur as the result of deviations from the prevalent preferences and contracts. This trade induced increase in the "innovation penalty" acts as an impediment to cultural-institutional transitions.

Our third result is that, in contrast to trade, factor market integration facilitates convergence between cultural-institutional conventions. The reason is that factor mobility provides a kind of 'innovation insurance' as it lowers the expected costs of deviating from the status quo.

The paper is organized as follows. After a discussion of related literature (section 2), we begin with the behavioral and institutional features of our model and the empirical evidence motivating them (section 3). In sections 4 and 5 we develop a model of endogenous preferences and contracts and transitions among cultural-institutional conventions under autarchy, which we extend to the two-country case in section 6 to account for specialization and trade. We then explore the persistence of cultural and institutional differences following trade integration (section 7), and factor mobility (section 8). Section 9 addresses the robustness of our results to variations in our modeling choices. Section 10 concludes and considers some broader implications of our approach.

2 Related literature

Our paper is a contribution to two distinct literatures: the economic theory of culture, institutions and their evolution, and international trade theory. We provide a conceptual framework addressing two puzzles: the shortcomings of standard trade models in the prediction of specialization and trade, and the long term persistence of economically important cultural-institutional differences among nations in a liberalized trading system.

We extend the stochastic evolutionary game approach pioneered by Cavalli-Sforza and Feldman (1981), Boyd and Richerson (1985), Foster and Young (1990), Young (1993) and Kandori, Mailath and Rob (1993) to model the joint evolution of institutions and culture. The complementarity between preferences and contracts, and the resulting cultural-institutional divergence among economies, we demonstrate, provide a model of the mechanisms underlying both cultural persistence and divergence in recent historical studies. An example is the historical account by Greif and Tabellini (2010) of the societal bifurcation between China and Europe in which "the diffusion of specific values in the society explains the emergence of one form of organization over another", while "the emergence of one moral system or another is explained by the distinct initial distribution of individuals across organizations" (p. 2). Like Alesina and Giuliano (2010), Guiso, Sapienza and Zingales (2009), Greif (1994, 2006), Tabellini (2010), Spolaore and Wacziarg (2009) and Brugger, Lalive and Zweimuller (2010), we study the economic importance of cultural differences. Unlike all above papers but in common with Bisin and Verdier (2001, 2011), Bowles (1998), Fershtman and Bar-Gill (2005), and Galor and Maov (2002), we model cultural evolution. The co-evolution of social norms and institutions is also modeled by Bidner and Francois (2011); however, in contrast to our approach, in their model institutional change is implemented by an institutional designer external to the transaction. Like Acemoglu (2003), we ask why inefficient institutions persist, but in our model persistence results not from the concentrated political power of those who benefit from the institutions, but rather from a coordination failure among decentralized agents in a non-cooperative interaction.

Our paper also contributes to the literature on international trade and institutions. Previous works on institutional comparative advantage have shown, both theoretically and empirically, that cross-country differences in institutional settings, e.g. in the quality of contracting institutions (Acemoglu, Antràs and Helpman, 2009; Costinot, 2009; Nunn, 2007; Levchenko, 2007), of credit market institutions (Kletzer and Bardhan, 1987) or in the accounting systems (Vogel, 2007), may be an independent source of gains from specialization and trade (earlier contributions are surveyed in Belloc, 2006). Our contribution to this literature is a novel explanation for the persistence of differences in the labor productivity across countries (Clark and Feenstra, 2003) even with identical technologies (narrowly defined as physical input-output relationships). These productivity differences may contribute to the explanation of Leontief's paradoxical result half a century ago that exports of the U.S. – a capital abundant economy – were more labor intensive than its imports. This could occur because seemingly capital abundant countries may actually be labor abundant once their labor endowment is measured in efficiency units (as Leontief himself suggested (1953, pp. 344-345), and Trefler (1993) and Davis and Weinstein (2001), among others, demonstrated empirically). Our model shows how, due to cultural-institutional differences, labor produc-

tivity may differ across countries even with identical technologies, an explanation very much in the spirit of Leontief's own reference to America's superior "organization".

Because in our model institutions as well as culture are endogenous and the two are jointly determined, we are able to explore the impact of economic integration on the persistence of institutions, rather than confining ourselves to the effects of given differences in institutions on comparative advantage and relative labor productivity Other papers treating the effects of trade expansion on institutional change are Belloc (2009), Greif (2006), Do and Levchenko (2009) and Levchenko (2012). The main novelty of our approach with respect to this group of papers is our modeling of the complementary relationship between preferences and institutions as a mechanism by which institutions associated with absolute disadvantage may persist indefinitely. In particular, our paper departs from and complements the work of Do and Levchenko (2009) and Levchenko (2012), in which institutional differences are a historical datum that may be modified by a cooperative lobbying game; in our model they are implemented as an endogenously generated non-cooperative cultural-institutional convention.

In common with Olivier, Thoenig and Verdier (2008), Karabay and McLaren (2010) and Pagano (2007), we find contrasting convergence effects of trade integration and factor market integration. But our paper and these three papers share little else in common: the first illustrates the dynamics of the demand for 'cultural goods' that contribute to group identity; the second is about the effects of globalization on wages and workers' welfare; the third concerns intellectual property. We share with Conconi, Legros and Newman (2012) the conclusion that liberalization need not favor the evolution of efficient institutions. However, in contrast to ours, in their model factor market integration induces international convergence of organizational choices only in conjunction with goods market integration.

The idea that trade integration may lead to divergence is also maintained by Zeira (2011) and Galor and Mountford (2008). In their models, divergence occurs in human capital investment, technology, or population, rather than in culture and institutions as in our model.

3 Overview

An economy is populated by employers and employees. Employers hire employees (each employer can hire just one employee and each employee can be hired only by one employer), the employment relationship being a random employee-employer match for a single interaction in which the employer implements a contract under which the employee works. Labor is perfectly mobile across industries but (until section 8) immobile across countries. There are no other factors of production. Our model is based on five distinctive features that we believe are of broad empirical relevance.

First, there are two aspects of labor. Qua$N$titative labor (denoted by N) includes time at work, compliance with explicit directions, simple effort readily measured either by input or output, and other aspects of work that are observable at low cost, either directly or that may be inferred from the associated outputs. Because it is readily observable, quantitative labor is cost-effectively verifiable and can be enforced by contracts. By contrast, qua$L$itative work (qualitative labor provided in addition to quantitative labor is denoted by L) consists

of care, creativity, problem solving and other non-routine aspects of work that are more difficult to verify, and hence not cost-effectively subject to explicit contracts conditional on individual performance. Production requires quantitative labor and is also enhanced by qualitative labor. Each employee may provide either quantitative labor alone or both quantitative and qualitative labor.

Second, there are two contracts, a Fixed-wage-monitoring contract (denoted by F) and a Partnership (denoted by P); the two contracts are chosen to represent the extreme points of a continuum in the degree of gain-sharing. In the former, the employer is the residual claimant, while the employee is paid a fixed wage w and works under close supervision by the employer (as in many secondary labor market jobs). In the latter, work is motivated by gain-sharing with the employer based on joint residual claimancy (as is the case in many legal practices, financial consulting, and software design – e.g. at Google, Microsoft). Under the Partnership the 'employee' is granted a positive fraction b of the revenue of the Partnership, and selects the type of labor (quantitative alone or both quantitative and qualitative) without supervision. We do not model the determination of b, but rather regard it is a determined custom as is the case with many share contracts (Young and Burke, 2001). Similarly, the (real) wage is exogenously determined by the bargaining power of employees and employers. For example, it could be determined by the no-shirking condition of an efficiency wage model (Shapiro and Stiglitz, 1984), and so would vary with the worker's fallback position, which could depend on the unemployment benefit (fixed in real terms). An alternative wage determination process consistent with our assumptions is the classical Arthur Lewis (1954) type model of a surplus labor economy, where the wage depends on the average income in subsistence farming or some other non-traded goods sector. What matters for the model is that the customs and bargaining power or other influences on b and w are unaffected by the terms of trade. This does not mean that workers do not benefit from the gains from trade: but whether they do depends (as we think it should on empirical grounds) on the institutions in force. As a result, employees in the P-contract share gains from trade, should they exist, while those in the F-contracts do not.

The third feature of our model builds on the idea modelled by Huck, Kubler, and Weibull (2012) and shown empirically by Bandiera, Barankay, and Resul (2005), that the same social norm may have different effects on productivity depending on the kind of incentives that are implemented. To model this effect we assume 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 these 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; Houser, Xiao, MacCabe, and Smith, 2008; Fehr and Rockenbach, 2003, surveyed in Bowles, 2008, and Bowles and Polania, 2012). We simplify by assuming just two kinds of preferences. Those who we term Reciprocators (denoted by R) care about the form of the contract *per se*: their utility is increasing in their own payoffs and may be either increasing or decreasing in the payoffs of the employer depending on the Reciprocator's belief about the type of the employer, in the spirit of Rabin (1993), Levine (1998) and Fehr and Falk (2002). Thus, the utility of employee h who is matched with employer -h depends on his

own material payoff (π_h), including the disutility of labor, and on the payoff of the employer ($\pi_{\text{-}h}$):

$$U_{h,\text{-}h} = \pi_h + \alpha_h \gamma_{h,\text{-}h} \pi_{\text{-}h} \tag{1}$$

where α (> 0 for the Reciprocator, denoted by R, and $= 0$ otherwise) is the strength of h's reciprocity preferences and $\gamma_{h,\text{-}h}$ ($= -1, 1$) is h's belief about -h's type, the latter depending on the form of contract that -h implements when interacting with h. In the model below a Partnership signals the good will and trust of the employer, leading to $\gamma_{h,\text{-}h} = 1$; while the employer's close surveillance under a Fixed-wage-monitoring contract signals distrust with $\gamma_{h,\text{-}h} = -1$ as a result. Other employees, who we will term Homo Economicus (denoted by E), care only about their own material payoffs (that is, $\gamma_{h,\text{-}h} = 0$) so that $U_h = \pi_h$ regardless the partner. We refer to preferences of this kind as self-regarding.

Though unconventional in the trade literature, the resulting complementarity between contracts and preferences is consistent with extensive empirical evidence (Fehr, Gachter and Schmidt, 1997; Brown, Falk and Fehr, 2004). Falk and Kosfeld (2006) used a gift exchange game to explore the idea that 'control aversion' may make close supervision counterproductive. Experimental agents in a role similar to an employee in our model 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 he 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 to impose a lower bound on the agent's production. The principal could infer that a self-regarding agent would perform at the lower bound or, in the absence of the bound, at zero, and thus the imposition of the bound would maximize the principal's payoffs. But in the experiment agents offered a lower level of production when the principal imposed the bound than when the principal elected to impose no bound. In post-experiment debriefings employees reported that they interpreted the imposition of the lower bound as a signal of the employer's distrust.

In a related experiment by Houser et al. (2008) a principal could offer an agent a fixed wage and monitoring contract that would induce a self-regarding agent to transfer all of the surplus to the principal but for which a reciprocal agent might retaliate by transferring nothing and enduring a fine by the principal. There was an alternative unmonitored partnership like contract in which a self-regarding agent would garner the entire surplus, but a reciprocal one would share the surplus with the principal. All four of the possible pairings occurred in the experiment, with both principals and agents gaining the highest payoffs when partnerships were offered to reciprocators, and reciprocators faced with the forcing contract doing the worst (they lost money).

The fourth feature of our model is that while both culture and institutions are endogenous, neither is the result of instantaneous individual maximization, nor are they implemented by a mechanism designer or other kind of collective choice. Rather both evolve inertially 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 instruction, schooling and other effects operating at the national level, we represent this process of cultural-institutional evolution by a society-wide dynamic operating prior to economic matching for production. Thus individuals acquire their preferences in much the same manner as they come to have a particular accent, that is to say unwittingly, not by a conscious choice. For this reason, our employees do not condition their preferences on the kind of contract (Partnership, Fixed-wage-monitoring) they experience in any period; rather they periodically update by (almost always) best responding to the distribution of contracts in the past. (For evidence of institutional and cultural inertia see Hannan and Freeman, 1989, Luttmer and Singhal, 2011, and the references therein.) Similarly firms do not condition their contracts on the preferences of the employee (Reciprocator, Homo economicus) with whom they are paired in a given period; rather they periodically update by (almost always) best responding to the past distribution of employee preferences.

Thus firms offering a forcing contract in a population composed almost entirely of reciprocators will make lower profits and either fail and be replaced by a firm offering a partnership, or in order to avoid failure will restructure the firm and offer a partnership contract. A similar process leads a reciprocal employee – or more realistically his or her children – facing only firms offering forcing contracts to adopt self-regarding preferences. Other than the standard assumption that updating is a standard payoff-monotonic best response process, our model does not specify how these adjustments take place. In both examples above updating would certainly involve both psychological influences (such as employees attempts to reduce cognitive dissonance by adjusting their types to be able to garner higher payoffs) and societal forces (schools and parents seeking to socialize the young so to function well the kinds of jobs that are currently available). It is this sociological and psycological processes, not some assumed meta-preference function of the individuals, that account for long term preference updating.

A key assumption here is here inertia in updating. The conflictual labor relations at Fiat, for example, could not be abandoned overnight were the employees suddenly transformed into Volkswagen's workers, accustomed to a more cooperative approach. Introducing Google-style work organization at General Motors would take years. For analogous reasons we assume that production requires some product specific kinds of capital (which we do not model here) so that firms cannot switch their product mix costlessly (General Motors could not costlessly shift to producing software if by chance it found itself with a self-motivated workforce for one period).

Fifth, there are two goods that differ in which cultural-institutional settings are appropriate for their production. The first is intensive in quantitative labor and termed transparent (the t-good) because it is more intensive in the labor activities that are readily observed. Transparent goods include standardized manufactured goods (exemplified by most goods produced on an assembly line and any good the production of which is cost effectively compensated by piece rates), most grains and sugar. 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), complex and quality-variable manufactured goods (such as wine and cigars), personal services ranging from legal advice to preparing meals, and care-sensitive agricultural products (many vegetables and fruits). Hence, denoting by Q_L^i the quantity of good $i = o, t$ 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, we have:

$$\frac{Q_L^o}{Q_N^o} > \frac{Q_L^t}{Q_N^t} \tag{2}$$

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

These five features of our model are intended to capture real distinctions among countries engaged in international exchange. Norms and preferences influencing economic behavior as well as institutions differ significantly among economies (Inglehart, 1977; Tabellini, 2010; Henrich, Boyd, Bowles et al., 2005; Bisin and Verdier, 2011; Knack and Keefer, 1997). In particular, reciprocal social preferences appear to be more prevalent in the higher income countries. Among subjects in 15 countries, for example, 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).

Similarly Bloom, Sedun, and Van Reenen (2012) provide evidence that among 4,000 firms in Europe, Asia and the U.S. norms of interpersonal trust affect firm structure, promoting greater decentralization and allowing for scale increases and productivity gains. Likewise institutions that favor gain-sharing, such as democratic governance, trade unions and the rule of law are more highly developed in the European, North American and other richer nations (World Bank, 2012a; Polity IV Project, 2010; Przeworski et al., 2000). For these reasons we represent an economy whose cultural-institutional convention is characterized by partnerships and reciprocal preferences as having a Pareto-superior cultural-institutional environment and, as a result, enjoying absolute advantage with respect to countries in which fixed wage contracts and high levels of monitoring may elicit quantitative (but not qualitative) labor services from entirely self-regarding economic agents.

Our model leads us to expect a positive association between a country's comparative advantage in opaque goods, on the one hand, and both a measure of the prevalence of partnership-like employment contracts and the level of reciprocity and related social norms, on the other. Cross nationally comparable measures of the prevalence of partnership contracts do not exist. As a measure of reciprocity, we use the level of cooperation sustained in the already mentioned behavioral experiment conducted by Herrmann et al. (2008). In a sample of 15 countries, they implemented a multi-period public goods game with an option for members (anonymously) to impose a cost on other members of the group (at a cost to themselves) once their contributions to the public good were revealed. We regard this as the best available cross national measure of the behavioral pattern we have termed reciprocity for two reasons: it is based on actual behavior with real monetary costs and benefits (rather than a survey) and it captures both good will towards fellow contributors and hostility towards those who would exploit the cooperativeness of others.

To locate sectors on the opaque-transparent continuum, we use the index computed by Costinot, Oldenski and Rauch (2011) that captures a sector's intensity in routine tasks. To do this, they employ measures developed by Autor, Levy and Murnane (2003), who classify occupations on the basis of the importance of "solving problems" and routine activities (in our terms, respectively qualitative and quantitative labor). To obtain a measure of countries'

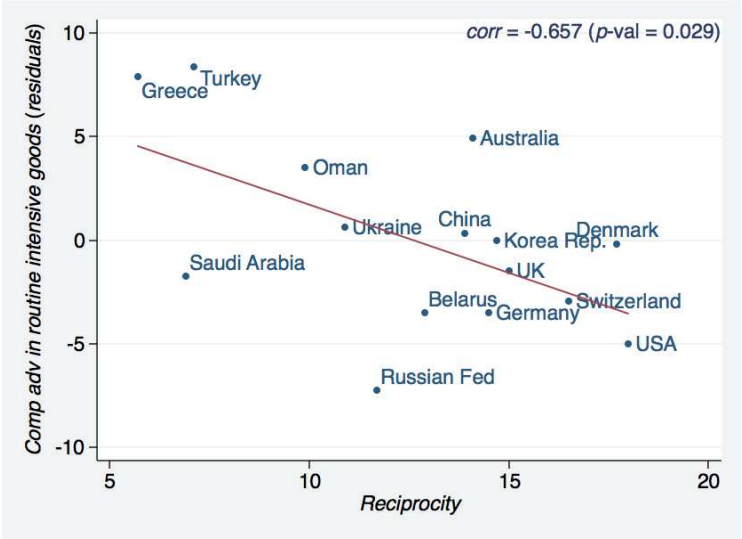

Figure 1: *Comparative advantage in routine intensive goods and reciprocity.* Note: The figure reports the relationship between the reciprocity index and the residuals from a regression of comparative advantage in routine intensive goods on GDP pro capite. Reciprocity (Herrmann et al., 2008) is the mean of contribution of a public good (20 is the maximum contribution possible, zero the minimum). Comparative advantage is computed as explained in the text. The correlation coefficient between the measure of comparative advantage and the reciprocity index obtained not controlling for GDP pro capite is 0.892 (p-val=0.005).

comparative advantage in routine intensive goods, we match these data with international bilateral trade data from Feenstra et al. (2005) and, following Costinot (2009), estimate the equation: $x_{ij}^k = \theta_{ij} + \lambda_j^k + v_i z^k + \varepsilon_{ij}^k$, where x_{ij}^k is logarithm of exports from country i to country j in sector k (3-digit NAICS), θ_{ij} is the exporter-importer fixed effect, λ_j^k is the importer-sector fixed effect, v_i is the exporter fixed effect, z^k is the intensity in routine tasks of sector k, and ε_{ij}^k is the idiosyncratic error term. Our measure of comparative advantage in routine intensive goods is given by the OLS estimate of v_i (data details are reported in appendix A).

In Figure 1 we report the relationship between the reciprocity index and the residuals from a regression of the comparative advantage measure on a constant and GDP pro capite (World Bank, 2012b). From the figure we can see two things. First, there are substantial cultural differences among nations. And, second, these differences are associated with patterns of trade: countries with higher levels of reciprocity have a comparative advantage in the production of goods whose labor inputs are less routine and more oriented toward problem solving, even after controlling for GDP pro capite.

In the next sections we present a model that captures the above five aspects. For tractability and to provide a transparent illustration of the processes underlying our results, we adopt specific production and utility functions, and present just two polar extreme sets of preferences and contracts. These assumptions could be modified – allowing some gain sharing in the F-contract for example – without altering the qualitative results. Our results are driven entirely by the gains from trade and the resulting specialization of countries and do not depend in any way on the particular description of the workers preferences or

contractual alternatives. All that is required for our three main results (Theorems 1-3) is that both cultural-institutional configurations are Nash equilibria.

4 Cultural-institutional complementarities in autarchy

In this section, before introducing specialization and trade (section 6), we describe production and distribution in a single-good autarchic economy, identifying the relevant Nash equilibria. Institutional and cultural updating and the resulting dynamics of transitions among cultural-institutional equilibria are illustrated in section 5.

Employers maximize profits, while employees maximize utility. Payoffs (profits and utility respectively) are measured in the number of units of the single good. The (risk-neutral) utility function of employees is additive in consumption, 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 greater costs δ (> 0). Table 1 gives the payoffs of the eight possible payoffs.

	Worker/preference	
Firm/contract	*Reciprocator*	*Homo economicus*
Partnership	$\boldsymbol{(1-b)Q_L, [b+\alpha(1-b)]Q_L-\delta}$	$(1-b)Q_N$, $bQ_N-\eta$
Fixed-wage-monitoring	$Q_N-w-m, w-\eta-\alpha(Q_N-w-m)$	$\boldsymbol{Q_N-w-m, w-\eta}$

Table 1: *Matrix of payoffs*. Note: Payoffs in bold type indicate pure stable Nash equilibria. A third Nash equilibrium in mixed strategies exists.

We are interested in the case where the Partnership-Reciprocator and Fixed wage-Homo economicus matches are Nash equilibria so that two distinct cultural-institutional configurations exist. To avoid uninteresting cases, we study the case in which workers with different preferences respond differently to the partnership contract, the E-worker providing quantitative labor only, and the R-worker providing both quantitative and qualitative labor. Thus, we assume that for the Homo economicus working under a Partnership the share of increased output associated with qualitative labor is less than the additional disutility required. But for reciprocal employees, who have preferences over the kind of contract that is implemented by the employer *per se*, we assume that the positive valuation of the payoff to the partner is sufficient to offset the greater disutility of labor δ ($\gamma = 1$ in equation (1)). These two assumptions guarantee that[1]

a1. $[b+\alpha(1-b)]Q_L - \delta > bQ_N - \eta > bQ_L - \delta$,

where Q_L is the amount of production obtained using one unit of both qualitative and quantitative labor, and Q_N the amount obtained with a single unit of quantitative labor.

Under the Fixed-wage-monitoring contract the employee is given a fixed compensation ($w > 0$) and is closely monitored (at a cost m). As a consequence, the E-worker provides

[1] Inserting the payoffs in equation (1) with $\alpha > 0$ and $\gamma = 1$ one obtains: $(bQ_L-\delta)+\alpha(1-b)Q_L = [b+\alpha(1-b)]Q_L-\delta$. The subscript h and $-h$ for the individuals are hereafter omitted with no loss of clarity.

quantitative labor only, incurring the associated disutility η. Similarly, the R-worker in the F-contract provides quantitative labor only at a cost η. However, unlike the Homo economicus, the Reciprocator values the payoff of the employer negatively ($\gamma = -1$ in equation (1)), subtracting $\alpha(Q_N - w - m)$ from his utility (where the expression in brackets is the profit of the employer).

To exclude cases with a single Nash equilibrium, we further assume that the workers' compensation under the Partnership is greater than the cost of labor under a Fixed-wage-monitoring contract, so that

a2. $bQ_N > w + m$,

and that qualitative (in addition to quantitative) labor is sufficiently more productive than just quantitative labor, so that

a3. $(1 - b)Q_L > Q_N - w - m$.

Then two Nash equilibria in pure strategies exist, namely $\{P,R\}$, that is the Partnership contract matched with the Reciprocal employee, and $\{F,E\}$, that is the Fixed-wage-monitoring contract matched with the Homo economicus. Assumptions **a1-a3** insure that two Nash equilibria exist and induce an unambiguous ranking of the two, with $\{P,R\}$ Pareto-dominating $\{F,E\}$.

To model the mutual dynamics of preferences and contracts, suppose that there are z members of each of the two populations (firms and workers), where z is a large and finite real number. In every period members of each population are paired randomly with members of the other population to play the 2x2 game with payoffs given in Table 1. The state of play in each period is the number of firms and employees offering partnerships and adopting reciprocal preferences respectively, denoted by $z\omega$ and $z\phi$, where both ω and ϕ are elements of $(0,\ 1/z,\ 2/z,\ \ldots (z-1)/z,\ 1)$. We will see that transitions between the inferior and superior outcomes will occur given that individuals are boundedly rational and occasionally do not best respond.

Using the payoffs in Table 1, the expected payoffs to employers implementing the P- and F-contracts are respectively

$$\begin{aligned} v^P(\omega) &= \omega(1-b)Q_L + (1-\omega)(1-b)Q_N, \\ v^F(\omega) &= \omega[Q_N - (w+m)] + (1-\omega)[Q_N - (w+m)] = Q_N - (w+m); \end{aligned} \tag{3}$$

whereas the expected payoffs to the R- and E-employees are

$$\begin{aligned} v^R(\phi) &= \phi\{[b + \alpha(1-b)]Q_L - \delta\} + (1-\phi)\{w - \eta - \alpha[Q_N - (w+m)]\}, \\ v^E(\phi) &= \phi(bQ_N - \eta) + (1-\phi)(w - \eta). \end{aligned} \tag{4}$$

These expected payoff functions are illustrated in Figure 2.

The state of play that equates the expected payoffs of the two strategies is given by ($z\phi^*$, $z\omega^*$), where:

$$\begin{aligned} \omega^* &= \frac{bQ_N - w - m}{(1-b)(Q_L - Q_N)}, \\ \phi^* &= \frac{\alpha(Q_N - w - m)}{\{[b + \alpha(1-b)]Q_L - \delta\} - (bQ_N - \eta) + \alpha(Q_N - w - m)}. \end{aligned} \tag{5}$$

Assumption **a1** ensures that ϕ^* is interior to the unit interval because $w - \eta > w - \eta - \alpha(Q_N - w - m)$ is always true for positive profits; assumptions **a2** and **a3** ensure that ω^*

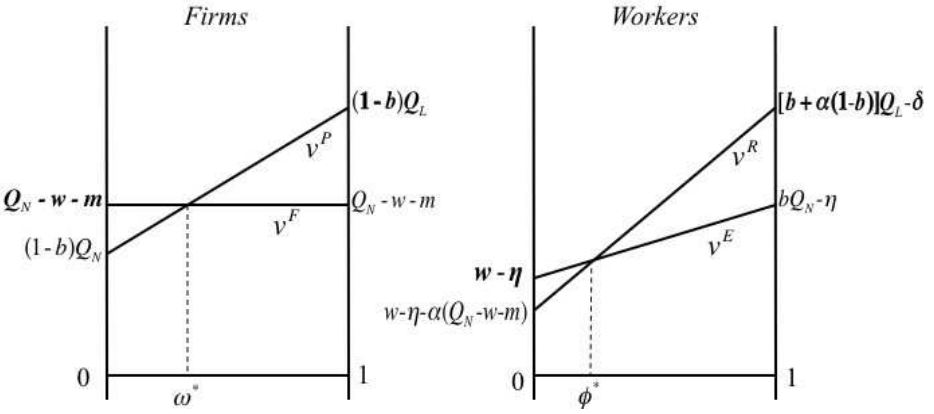

Figure 2: *Expected payoffs under autarchy to P- and F-firms (left) and to R- and E-workers (right).* Note: ω is the fraction of the workers that are reciprocal and ϕ is the fraction of the firms that are partnerships. The vertical intercepts are from Table 1. Payoffs in bold type refer to the payoffs in the pure Nash equilibria.

is also interior. Due to the finiteness of population size and the fact that the state of play must take integer values, the interior equilibrium $(z\phi^*, z\omega^*)$ will occur with a sufficiently small probability that can be ignored (as we will see, should $(z\phi^*, z\omega^*)$ occur it would not be stable in the revision dynamics). We define $z\phi^+$ $(\geq z\phi^*)$ and $z\omega^+$ $(\geq z\omega^*)$ as the least integers such that P-contracts and R-preferences for the population in question is a strict best response. Figure 3 gives the state space of this process.

5 Cultural-institutional persistence and transitions in autarchy

We now provide an explicit dynamics to show that the stationary states in this setting are those labeled (z, z) and (0,0) in Figure 3 – all employers offering partnership contracts to reciprocal workers $\{P,R\}$ or all employers offer fixed wage contracts to self-regarding workers $\{F,E\}$. Cultural and institutional updating is a standard payoff-monotonic myopic best response process perturbed by infrequent idiosyncratic (non-best response) play (Foster and Young, 1990; Weibull, 1995; Young, 1998; Bowles, 2004). Following matching each member of the two populations has a type revision opportunity. To ensure that the process is acyclic, we let the revision process be asynchronous (Binmore, Samuelson, and Young, 2003). Thus in even numbered periods one population has the opportunity to revise its type first, after which the other population revises; in odd numbered periods the other revises first.

Individuals are boundedly rational and best respond to the state of play with probability $(1-\sigma)$. But with probability σ they, respectively workers and firms, adopt the preferences and contracts that are not a best response. On empirical grounds, we would like the likelihood of non-best response play to be sensitive to the expected costs (rather than being simply random as in many perturbed Markov process models). Hence, we let σ be strictly decreasing in both the cost of deviating from the best response and the agents' degree of rationality (defined below) as in Blume (2003), Myatt and Wallace (2004) and Kreindler and Young (2011).

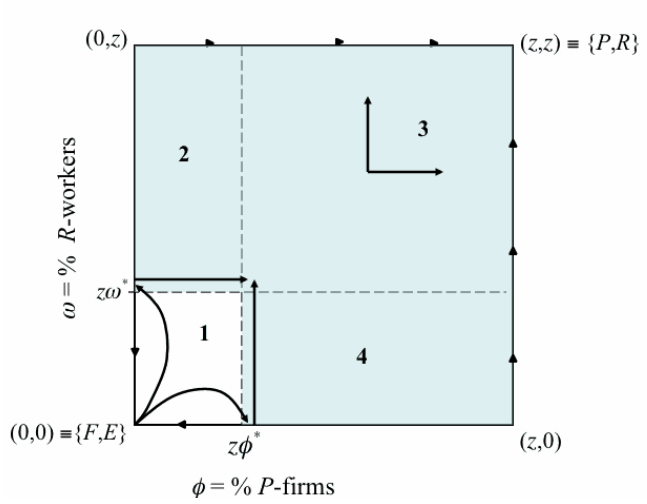

Figure 3: *Transitions from* $\{F,E\}$ *to* $\{P,R\}$. Note: In any period, if the number of deviants from the $\{F,E\}$ convention exceeds $z\phi^*$ or $z\omega^*$, subsequent best responses will lead the population to the $\{P,R\}$ convention.

The cost of deviating from the *status quo* $\{F,E\}$ equilibrium, represented by differences between the left axis vertical intercepts in Figure 2, for firms and workers respectively, are defined by

$$\begin{aligned} \Delta^{firm} &\equiv v^F(\omega = 0) - v^P(\omega = 0), \\ \Delta^{worker} &\equiv v^E(\phi = 0) - v^R(\phi = 0). \end{aligned} \tag{6}$$

Then, using the log-linear revision process in Blume (2003), we have the following expressions for the probability of conforming to the equilibrium and deviating from it, where $k = firm, worker$,

$$\begin{aligned} P(best\ response) &= 1 - \sigma(\Delta^k, \beta) = \frac{e^{\beta\Delta^k}}{1 + e^{\beta\Delta^k}}, \\ P(non\text{-}best\ response) &= \sigma(\Delta^k, \beta) = \frac{1}{1 + e^{\beta\Delta^k}}. \end{aligned} \tag{7}$$

We interpret β as a measure of rationality of the agent when making a choice because $\sigma(\Delta^k, \beta)$ is decreasing in β: the more rational is the agent (the larger is β), the smaller the probability that he will make the non-best response. When $\beta = 0$ the agent chooses randomly between the two options, and as β goes to infinity (with positive Δ^k), the agent never deviates, that is $\sigma(\Delta^k, \beta)$ tends to zero. Agents are identical with respect to β.

As shown in Figure 3, under this dynamics the only evolutionary stable states are (z, z) and (0,0), that is states in which all employer-employee pairs are at one or the other of the Nash equilibria in Figure 2. What is important in what follows is that, for sufficiently rational agents, once a population is in the neighborhood of either of these two states the associated convention, namely $\{F, E\}$ and $\{P, R\}$, will persist over very long periods. The reason is that non-best response updating will be rare, so that, as in Kandori, Mailath and Rob (1993: 52), "an equilibrium is upset by large jumps (from the equilibrium to the basin of attraction of the other equilibrium)" or as Binmore, Samuelson and Young (2003: 309) put it, by a "single burst of mutations". This allows us to focus on what they term

"one-step transitions".[2]

As we are interested in one-step transitions from the inferior to the superior equilibrium, we study the minimum numbers of deviant firms and workers such that, with sufficiently rational agents, the population will enter the basin of attraction of the superior $\{P, R\}$ equilibrium. The basin of attraction of a particular stationary state is the set of states that, for the above dynamics and sufficiently rational agents, leads to that state. For sufficiently rational individuals, the state following idiosyncratic play will almost always lie within the basin of attraction of the *status quo* convention. Then even if by chance this perturbed state is a substantial displacement from uniform adherence to the convention, in the next period the population will with high probability return to the vicinity of the *status quo* state (0,0), and because agents have only a one period memory the excursion from the convention will have no lasting effect.

Suppose that in period 1 firms previously offering F-contracts update and that $z\phi^+$ of them offer P-contracts instead of best responding with an F-contract. In the subsequent period workers update, and because the workers' best response to the state $z\phi^+$ is to adopt R- rather than E-preferences, each will adopt R-preferences with probability $1-\sigma$. But as β goes to infinity σ goes to zero, so there exists some finite β such that, as a result of workers updating, with virtual certainty we will have $z\omega \geq z\omega^+$, and the population will be in the set of states for which both firms' and workers' best responses will lead to $\{P, R\}$. Thus the minimum number of P-contracting firms sufficient to escape from the inferior equilibrium is $z\phi^+$. Analogous reasoning shows that the minimum number of R-workers sufficient to escape the inferior equilibrium is $z\omega^+$. It follows that from the initial state $\{F, E\}$ the basin of attraction of the superior equilibrium is composed of quadrants 2, 3 and 4 in Figure 3.[3]

We use the term conventions to describe the economy wide adoption of the cultural and institutional profiles $\{P, R\}$ and $\{F, E\}$ when these profiles are evolutionarily stable given the endogenous nature of culture and institutions.

Definition 1: The two pure Nash equilibria are termed cultural-institutional conventions when they represent mutual best responses conditional on virtually all other agents doing the same.

This definition ensures that for each cultural-institutional convention there is a neighboring set of states for which the dynamics described by equations (3)-(7) carry the population back to the convention, if agents are sufficiently rational.

Because deviations from best response are independent, waiting times for a transition from one stationary state to the other are approximated by the inverse of the probability (μ) that in a given period the number of non-best responses will be sufficient to enter the basin of attraction of the other equilibrium. Because we assume a large population and a small rate of idiosyncratic play (β sufficiently large), the probability of escaping the $\{F, E\}$ convention is approximated by the probability that exactly the minimum number

[2] This is a simplification, the results below apply to any path of any number of steps from any state in the basin of attraction of the $\{F, E\}$ convention to a state not contained in it.

[3] From the initial state $\{P, R\}$, the basin of attraction of the $\{F, E\}$ convention is constituted by the quadrants not containing $\{P, R\}$, the two basins of attraction not being disjoint in this case, as in Binmore, Samuelson, and Young (2003).

of deviants in each population ($z\phi^+$,$z\omega^+$) will occur (the probability that larger numbers deviate being sufficiently small to be ignored in this approximation (Binmore, Samuelson, and Young, 2003)). Thus, we have

$$\begin{aligned}\mu^{firm} &= \left(\begin{array}{c} z \\ z\phi^+ \end{array}\right)(\sigma^{firm})^{z\phi^+}(1-\sigma^{firm})^{z-z\phi^+}, \\ \mu^{worker} &= \left(\begin{array}{c} z \\ z\omega^+ \end{array}\right)(\sigma^{worker})^{z\omega^+}(1-\sigma^{worker})^{z-z\omega^+}.\end{aligned} \tag{8}$$

Since we are interested in the effect of the integration of goods and factor markets on the probability of an exit from the inferior equilibrium induced by firms and workers respectively, that is μ^{firm} and μ^{worker}, in the next sections we replace the autarchic one-good economy with a two-good and two-country model of specialization and trade. Of course we are interested in the joint probability that a transition takes place induced by either the employers or the employees but, in the view of the fact that the results below hold for each of the transition probabilities independently, they hold *a fortiori* for the joint probability.

6 Cultural-institutional differentiation and comparative advantage

We consider two otherwise identical countries (same technologies, no differences in workers' skills, identical demand functions), one in which P-contracts and R-workers are matched and the other in which F-contracts and E-workers are matched, and determine the conditions under which these cultural-institutional configurations will be evolutionarily stable in the dynamics described in the previous section, and whether specialization and the gains from trade will be supported as a result.

Agents consume a given composite bundle (indicated by c) of the two goods produced; thus prices have no effect on consumption proportions. For simplicity, we assume that the composite good is made up of one half of the transparent (t) and one half of the opaque (o) good. We denote by p^o the price 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 by p^t the price 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). In the remainder of the paper payoffs (profits and utility respectively) will be measured in the number of units of the composite good commanded. Markets are competitive in the sense that employers take the price of the good as exogenously given.

Autarchy prices will differ between the two countries because the relative cost of production of the two (opaque and transparent) goods depends on whether both qualitative and quantitative labor is used ($\{P,R\}$) or quantitative labor only ($\{F,E\}$). Without loss of generality, we associate country A with the $\{F,E\}$ convention and country B with the $\{P,R\}$ convention, and so hereafter the subscripts A and B will denote the autarchic prices in the two economies. Denoting by C_N^A and C_L^B the quantity of good c that can be obtained

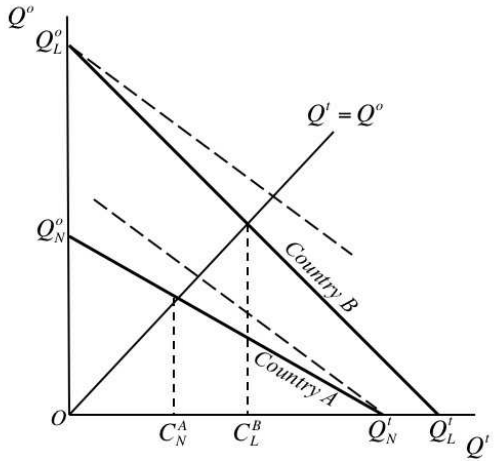

Figure 4: *Production possibility frontiers and the gains from trade in the two countries.* Note: Labor endowment normalized to 1. Sloped dashed lines indicate the international terms of trade. C^A_N and C^B_L are the quantities of the composite good available under autarchy in country A and country B respectively.

using one unit of labor in respectively country A and B in autarchic equilibrium, we have[4] $C^A_N = 2Q^o_N Q^t_N/(Q^o_N + Q^t_N)$ and $C^B_L = 2Q^o_L Q^t_L/(Q^o_L + Q^t_L)$. Since under autarchy both transparent and opaque goods (the composite goods) are produced in competitive equilibrium, producers must be indifferent to which product to produce so (given that costs are the same) employer's revenue must be equal across sectors, thus: $C^A_N = Q^o_N p^o_A = Q^t_N p^t_A$ and $C^B_L = Q^o_L p^o_B = Q^t_L p^t_B$ with $i = t, o$ in the two countries respectively. From this it follows that $p^t_A = 2Q^o_N/(Q^o_N + Q^t_N)$ and $p^o_A = 2Q^t_N/(Q^o_N + Q^t_N)$ in country A, and $p^t_B = 2Q^o_L/(Q^o_L + Q^t_L)$ and $p^o_B = 2Q^t_L/(Q^o_L + Q^t_L)$ in country B. So (as one would expect) for the composite good to be produced in each country the relative price of transparent and opaque goods must be equal to the relative marginal costs of the two goods (the marginal rate of transformation), i.e. $p^t_A/p^o_A = Q^o_N/Q^t_N$ in the $\{F, E\}$ economy (country A) and $p^t_B/p^o_B = Q^o_L/Q^t_L$ in the $\{P, R\}$ economy (country B).

In Figure 4, the two solid lines represent the production possibility frontiers of the two countries. Because $Q^o_L > Q^o_N$ and $Q^t_L > Q^t_N$, country B enjoys an absolute advantage in the production of both goods, and so its production possibility frontier is all above that of country A. However, the cultural and institutional differences across countries (like differences in endowments or technologies in the standard models) result in differences in the ratios of marginal costs of goods in autarchy and, as a result, confer different comparative advantages to the two economies considered. Indeed, given (2), it follows that:

$$\frac{p^t_B}{p^o_B} = \frac{Q^o_L}{Q^t_L} > \frac{Q^o_N}{Q^t_N} = \frac{p^t_A}{p^o_A} \tag{9}$$

[4] The quantity of N-labor needed to produce one unit of the composite good is: $(1/Q^o_N + 1/Q^t_N)/2 = (Q^o_N + Q^t_N)/(2Q^o_N Q^t_N)$. The productivity of one unit of N-labor in terms of composite goods will be just the inverse of this ratio. Similarly we can obtain the productivity of one unit of L-labor.

or, given the definition of autarchic prices in the two countries, $p_B^t > p_A^t$ ($p_B^o < p_A^o$). Country B, where the established cultural-institutional convention is able to elicit qualitative (in addition to quantitative) labor in all the employment relations, is superior in the production of both commodities, but has a relatively greater advantage in the production of the o-good where qualitative aspects of work are relatively more important. By contrast, country A 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. Thus, providing that the international terms of trade, represented in Figure 4 by the slope of the dashed lines and denoted by p_T^t/p_T^o (the subscript T referring to trade), falls strictly between the autarchic relative prices of the two countries, specialization and trade will take place.[5]

Theorem 1 (Specialization and trade): After the two economies open up to trade, country A will completely specialize in the production of (and will export) the transparent good, while country B will specialize in the production of (and will export) the opaque good.

Proof: Complete specialization follows directly from the linearity of the two production possibility frontiers and the fact that $p_T^t > p_A^t$ and $p_T^o > p_B^o$.[6] Trade follows from complete specialization and the assumption of identical demand function across countries.

But the fact that the culture and institutions prevailing in each country are a source of comparative advantage, and that opening up to trade enables the two otherwise identical countries to trade and enjoy welfare gains, would be of little interest if trade were to erode the differences upon which cultural-institutional comparative advantage depends. Because both culture and institutions are endogenous in our model, we must confirm that the cultural and institutional differences that underlie comparative advantage persist after the two countries open up to international exchange.

To demonstrate this, the assumptions ensuring that the two Nash equilibria exist in the single good case illustrated in Table 1 (**a1-a3**) must be extended to take account of the fact that in the two-good, two-country case the prices that will obtain in autarchy in the two countries will differ. Given definition 1 we have the following lemma:

Lemma 1. Both $\{F, E\}$ and $\{P, R\}$ are cultural-institutional conventions if the following three assumptions hold, where $j = A, B, T$ refers respectively to autarchy in country A, autarchy in country B and trade (C_N^T and C_L^T are the amounts of composite good that can be purchased respectively in country A and country B under specialization and trade),

a1'. $[b + \alpha(1-b)]C_L^j - \delta > bC_N^j - \eta > bC_L^j - \delta$,

a2'. $bC_N^j > w + m$,

[5] 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 notational clutter associated with joint production in one country.

[6] Notice that p^t and p^o are relative prices (the price of respectively the transparent and the opaque good in terms of the composite good).

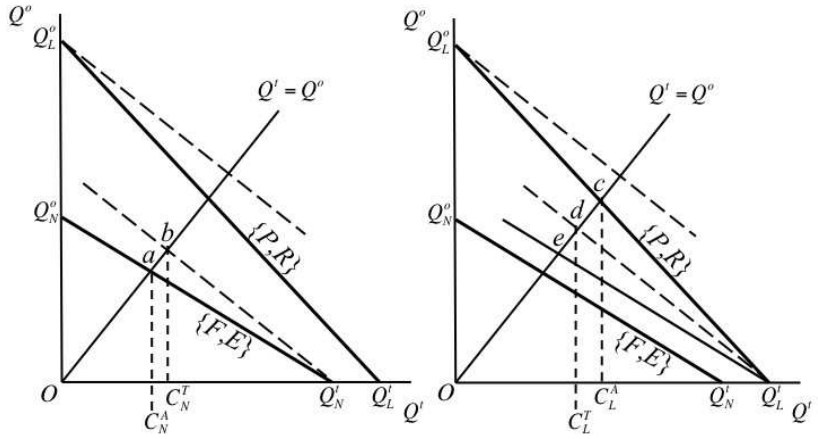

Figure 5: *Production possibility frontiers and effects of trade for two matches, the status quo* $\{F, E\}$ *(left) and the joint idiosyncratic match* $\{P, R\}$ *(right).* Note: Labor endowment normalized to 1. Sloped dashed lines indicate the international terms of trade. For the idiosyncratic match the value of production (units of the composite good commanded) is less after trade because the negative effect of specialization in the transparent good (shift from point c to point e) is not fully offset by the favorable price effect (shift from point e to point d).

a3'. $(1-b)C^j_L > C^j_N - w - m$.

Proof. These assumptions ensure that ϕ^* and ω^* are within the unit interval after trade so that there exists a neighborhood of $\{F, E\}$ in which the best response by firms and workers is respectively F-contracts and E-preferences, and there exist a neighborhood of $\{P, R\}$ in which the best response by firms and workers is respectively P-contracts and R-preferences.

7 Trade and transitions to superior cultural-institutional conventions

We can now investigate whether trade favors cultural and institutional convergence to the superior convention (so we only consider the point of view of the disadvantaged country A). To do this we consider agents at the $\{F, E\}$ convention facing the prices prevailing at that state under autarchy and under trade and explore how the difference in prices and the resulting specialization will affect the likelihood that a sufficient number of them will deviate from the *status quo*, inducing a transition to the superior convention. The revision process of preferences and contracts is based on the *status quo* prices (rather than the prices that would obtain were a transition to occur) because the revision process in (7) – perturbed myopic best response – is backward rather than forward looking. From (7) we know that for $\beta > 0$ the realized number of deviants will depend on the cost of deviating, that is the difference in the expected payoffs of best responding and innovating defined in (6).

The effects of specialization and trade in country A are shown in Figure 5, where we represent the production possibility frontiers for respectively the *status quo* $\{F, E\}$ match (left panel) and the joint idiosyncratic $\{P, R\}$ match (right panel).

Under autarchy the employer-employee pair produces equal quantities of the two goods

with the resulting production of the composite good being given by the intersection of the production possibility frontier and the composite good ray from the origin (point a). After trade integration the country specializes in the production of the transparent good in which it has a comparative advantage, because virtually all the contracts are such that only N labor is provided (with the only exception of the rare jointly idiosyncratic $\{P, R\}$ pairs). Some of the produced quantity of the transparent good is then exchanged for the opaque good in which the other country specializes in order to obtain the composite good. These specialization and trade effects for the *status quo* $\{F, E\}$ match are jointly represented by the movement from a to b in Figure 5. As can be easily verified, the value of production of this match in terms of the composite good increases, $C_N^T > C_N^A$.

In the idiosyncratic $\{P, R\}$ match L- in addition to N-labor is provided. Under autarchy both goods are produced in equal quantity resulting in a quantity of the composite good (point c) superior to the $\{F, E\}$ match because the $\{P, R\}$ pair has an absolute advantage in the production of both goods. Here trade has two effects. The first is the specialization in the t-production at which the $\{P, R\}$ pair is relatively less productive, which were the autarchy prices to obtain would be represented by the movement from c to e. But there is a second offsetting effect: the increase in the relative price of the t-good, represented by the movement from e to d. The positive price effect only partially compensates the negative specialization effect because the terms of trade is strictly between the slopes of the two production possibility frontiers. It follows that $C_L^T < C_L^A$.

All the other terms (δ, η, w, m and γ) entering the agents' expected payoffs are measured in units of the composite goods, and so remain unaltered.

To study the effects on the probability of a transition, we rewrite the expected payoff equations, (3) and (4), in terms of units of the composite good commanded, where $j = A, T$ indicates either autarchy or trade in country A, respectively for firms

$$\begin{aligned} v^P(\omega) &= \omega(1-b)C_L^j + (1-\omega)(1-b)C_N^j, \\ v^F(\omega) &= \omega[C_N^j - (w+m)] + (1-\omega)[C_N^j - (w+m)] = C_N^j - (w+m), \end{aligned} \tag{3'}$$

and workers

$$\begin{aligned} v^R(\phi) &= \phi\{[b+\alpha(1-b)]C_L^j - \delta\} + (1-\phi)\{w - \eta - \alpha[C_N^j - (w+m)]\}, \\ v^E(\phi) &= \phi(bC_N^j - \eta) + (1-\phi)(w-\eta). \end{aligned} \tag{4'}$$

which are depicted in Figure 6.

It is easily shown that the two critical population fractions of individuals that equate the expected payoffs are now

$$\begin{aligned} \omega_j^* &= \frac{bC_N^j - w - m}{(1-b)(C_L^j - C_N^j)}, \\ \phi_j^* &= \frac{\alpha(C_N^j - w - m)}{\{[b+\alpha(1-b)]C_L^j - \delta\} - (bC_N^j - \eta) + \alpha(C_N^j - w - m)}. \end{aligned} \tag{5'}$$

Differences between left vertical intercepts in Figure 6 illustrate the cost of deviating from the *status quo* $\{F, E\}$ convention under autarchy and trade for firms and workers, which are defined as in (6) using equations (3') and (4'). Inspection of the figure motivates:

Lemma 2: Trade integration increases the cost of deviating from the $\{F, E\}$ convention for firms and workers respectively, namely, $\Delta_T^{firm} > \Delta_A^{firm}$ and $\Delta_T^{worker} > \Delta_A^{worker}$.

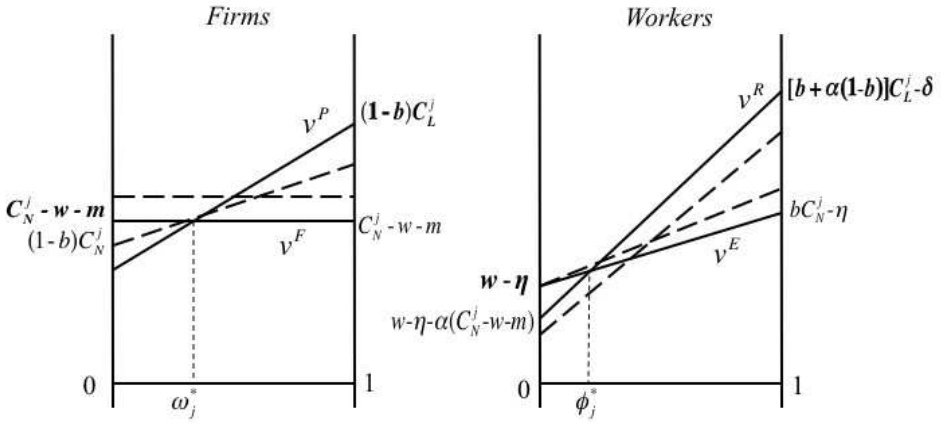

Figure 6: *Payoffs to P- and F-firms (left) and R- and E-workers (right) in country A before and after trade integration.* Note: ϕ is the fraction of the Partnership employers and ω the fraction of Reciprocator employees. Payoffs in bold type refer to the payoffs in the stable pure Nash equilibria. Dashed lines represent expected payoff lines after trade integration ($j = A, T$).

Proof: It follows from the fact that the costs of deviation for both firms and workers are increasing in C_N^j $(j = A, T)$[7] that is greater after trade.

The economic intuition behind this result is that deviating from the convention almost always entails a mismatch and hence forgoing some of the surplus, the value of which is higher after trade integration as a result of the gains from specialization and trade.

In addition to increasing the incentive not to innovate, trade will alter the critical fraction of innovators to induce a transition from the inferior to the superior convention:

Lemma 3: Trade integration increases the critical fraction of innovating R-workers and P-contracting firms sufficient to induce, respectively, the F-contracting firms to best respond by adopting P-contracts, namely, $\omega_T^* > \omega_A^*$ (so that $\omega_T^+ \geq \omega_A^+$) and the E-workers to best responding by adopting R-preferences, namely, $\phi_T^* > \phi_A^*$ (so that $\phi_T^* \geq \phi_A^*$).

Proof: The critical fractions ω_j^* and ϕ_j^* are given by the ratio of the cost (for respectively firms and workers) of deviating from the $\{F, E\}$ convention to the sum of this cost and the cost of deviation from the $\{P, R\}$ convention, conditional on all matches specializing in the production of the transparent good. As we have seen in Lemma 2, after trade and specialization the former is greater than under autarchy, because $C_N^T > C_N^A$ (cf. Figure 5, left panel). By contrast the latter is lessened by specialization in the t-good production at which the $\{P, R\}$ match is relatively less productive so that $C_L^A > C_L^T$ (cf. Figure 5, right panel).

Theorem 2 gives the effects of trade on the expected waiting time of a transition from the inferior to the superior convention.

Theorem 2 (Effect of trade integration on transitions): If agents are sufficiently rational, trade integration decreases the probability that in a given period the number of

[7] Costs of deviation, for firms and workers respectively, are $\Delta_j^{firm} = bC_N^j - (w+m)$ and $\Delta_j^{worker} = \alpha[C_N^j - (w+m)]$.

non-best responses will be sufficient to escape the $\{F, E\}$ convention, that is, for transitions induced by respectively workers and firms, $\mu_T^{worker} < \mu_A^{worker}$ and $\mu_T^{firm} < \mu_A^{firm}$.

Proof: See appendix B.1.

We know, by Lemma 2, that trade increases the cost to both firms and workers of deviating from the inferior convention. Since by Lemma 3 we also know that trade increases the critical fraction of idiosyncratically playing agents to induce a transition, it follows from (8) that the probability that a transition in a given period will occur must be reduced by trade and hence the expected waiting time of escaping the basin of attraction of the $\{F, E\}$ convention is increased for sufficiently rational agents. Thus, removing impediments to international exchange fortifies the preexisting cultural and institutional differences upon which specialization and trade are based. This is true even if there exists an alternative cultural-institutional convention to which a transition would be Pareto-improving.

8 Factor market integration and transitions to superior cultural-institutional conventions

Many of the effects of international economic integration – like factor price equalization in Paul Samuelson's theorem (Samuelson, 1948) – are independent of whether integration is accomplished through the elimination of barriers to trade in commodities or through the mobility of factors of production. However, where comparative advantage is based on country differences in culture and institutions, as in our model, this is not the case.

In contrast to goods market integration, factor market integration facilitates transitions from inferior to superior cultural-institutional conventions. In contrast to trade integration: in the neighborhood of the inferior convention, factor market integration lessens the costs of deviating from the $\{F, E\}$ convention. The probability of an optimal match for idiosyncratic R-workers and P-firms is not just the (vanishingly small) likelihood of meeting an idiosyncratic type from one's own economy but also the substantial chance for the 'right' match which occurs when paired with best responding individuals from the other country in the pool. Thus factor market integration has what we will call an 'innovation insurance' effect. This is in contrast to commodity market integration, which imposes an 'innovation penalty' because, as we have seen, the gains from trade heighten the opportunity costs of the frequent mismatches that innovators may expect when paired with agents from their own country.

We could model the movement of firms and workers from country A to country B and the converse, but a more parsimonious way to represent factor market integration, and one that captures the essentials, is to simply posit a distinct cosmopolitan matching process in addition to the within economy matching we have assumed thus far. 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 7, 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 we have:

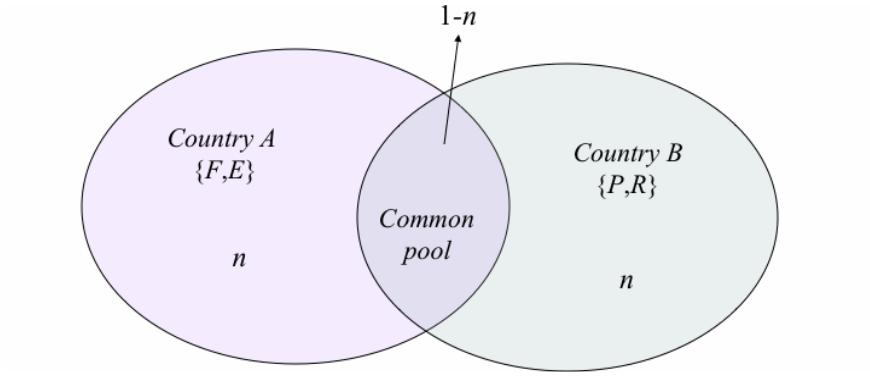

Figure 7: *Factor market integration.* Note: n is the degree of national specificity of the factor markets and $1-n$ is the degree of factor market integration.

Definition 2: The degree of factor market integration is equal to $1-n$, where with probability n an individual is matched with an individual from his own nation and with probability $1-n$ he is matched with an individual randomly drawn from the common pool.

When factors are immobile, one may imagine the two countries as two 'villages' within which all production takes place; 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 who they encounter there. 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. 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.

When the factors of production are matched in the pool the product produced is determined by the nationality of the employer, consistent with a degree of product specificity of the physical assets of the employer, while the skills of the worker are less so.[8] The results below can be obtained using either autarchic or trade prices, meaning that factor market integration facilitates transitions to the Pareto-superior convention independently of whether the *status quo* is nationally specific factor markets under autarchy or trade integration. In the case of autarchy, the prices at which the output is sold are also determined by the nationality of the employer.[9]

The expected payoff after factor market 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 $1-n$. The expected payoff in the common pool, in turn, is the weighted sum of the expected payoffs from matching an

[8] This assumption may easily be relaxed without altering the conclusions in any relevant way.

[9] For example, when an employee from country A is matched with an employer from country B, the pair will produce the opaque good to be sold at the autarchic prices of country B. This assumption may easily be relaxed without altering the conclusions in any relevant way.

individual resident in country A and in country B with weights given by the relative sizes of the two countries. (Expected payoff equations and critical values are reported in the appendix B.2.2 for reasons of space). In what follows the subscript $n = 1$ indicates that the relevant quantity refers to factor immobility, while $n < 1$ denotes factor market integration.

In contrast to trade integration (Lemma 2), factor market integration (reducing n) lessens the expected cost of innovation for both employers and employees at the $\{F, E\}$ equilibrium. We have:

Lemma 4: Factor market integration decreases the cost of deviating from the $\{F, E\}$ convention for respectively firms and workers, namely, $\Delta_{n<1}^{firm} < \Delta_{n=1}^{firm}$ and $\Delta_{n<1}^{worker} < \Delta_{n=1}^{worker}$.

Proof: See appendix B.2.3.

The case of employers is straightforward. The best responding F-contracting employer will be unaffected by factor market integration, because his profits do not depend on the type of employee with whom he is matched. By contrast, employers who idiosyncratically implement P-contracts when $n < 1$ will enjoy a payoff-enhancing match with a reciprocal worker not only with the rare innovators from their own economy, but also with the prevalent type of worker from the other country, who will constitute a sizeable fraction of the workers in the cosmopolitan pool. So while the expected payoff to the best responder is unchanged, the expected payoff to the idiosyncratic agent increases, leading to a lessened cost of innovation.

The case of employees is less straightforward but easily demonstrated. Because the worker's share in a partnership is larger than the wage in a fixed wage contract even when only quantitative labor is provided, factor market integration increases the expected payoff to best responding workers who adopt E-preferences. But, by increasing the probability of meeting a P-contracting firm, factor market integration raises the expected payoff to workers idiosyncratically adopting reciprocal preferences even more (the payoff of a reciprocal employee matched with a partnership firm is larger than that of a Homo Economicus working under a partnership), and thereby lessens the cost of deviating from the $\{F, E\}$ convention.

In addition to the reduced cost of innovating, it is easily shown that:

Lemma 5: Factor market integration decreases the critical fraction of innovating R-workers and P-contracting firms sufficient to induce, respectively, the F-contracting firms to best respond by adopting P-contracts, namely, $\omega_{n<1}^{*} < \omega_{n=1}^{*}$ (so that $\omega_{n<1}^{+} \leq \omega_{n=1}^{+}$) and the E-workers to best responding by adopting R-preferences, namely, $\phi_{n<1}^{*} < \phi_{n=1}^{*}$ (so that $\phi_{n<1}^{+} \leq \phi_{n=1}^{+}$).

Proof: See appendix B.2.4.

The economic intuition behind this result is transparent. Recall that, each of the two critical fractions sufficient to induce a transition is given by the cost of deviating from the *status quo* $\{F, E\}$ convention divided by the sum of this cost and the cost of deviation from the $\{P, R\}$ convention. We know from Lemma 4 that, for both workers and firms the cost of deviating from the *status quo* $\{F, E\}$ match decreases after factor market integration because of the 'innovation insurance' effect. As a result the critical fractions (ω^* and ϕ^*) must decrease because for neither firms nor workers can the cost of deviating from a $\{P, R\}$

match be altered by integration of the factor market with country B where, in equilibrium, all the firms are Partnerships and all the workers are Homo economicus.

Summing up, Lemma 4 and Lemma 5 show that factor market integration reduces the costs of deviation from the inferior status-quo convention while also reducing the critical fraction of deviants. Finally, it follows from these two results that:

Theorem 3 (Effect of factor market integration on transitions): If agents are sufficiently rational, factor market integration increases the probability of escaping the $\{F, E\}$ convention for transitions induced by both innovating workers and firms, that is $\mu_{n<1}^{firm} > \mu_{n=1}^{firm}$ and $\mu_{n<1}^{worker} > \mu_{n=1}^{worker}$.

Proof: See appendix B.2.5.

9 Robustness and extensions

Is our model robust to plausible variations in assumptions? For reasons of tractability and transparency we have modeled two specific contracts and behavioral types, but the underlying mechanisms work under quite general assumptions. Neither the strong complementarity in consumption, nor the linearity of the production functions is necessary for our results.

Concerning the latter, adopting concave production functions might entail less than complete specialization, but would not affect the mechanisms underlying our theorems. Relaxing the strong complementarity in consumption would increase the gains from trade. This would increase the cost of deviating from the *status quo* inferior equilibrium and, *ceteris paribus*, by equations (3)-(8) reduce the likelihood of transition to the Pareto-superior outcome. Allowing substitution in consumption increases the gains from trade because in our set up the gains from trade are restricted to income effects and do not allow substitution effects which, if included, will increase the gains from trade due to the price change following integration.

Our revision process for preferences and contractual choices and the resulting dynamics could also be modified without changing our results. While the way that our model takes account of the cost of idiosyncratic play is plausible and the fact that trade increases this cost contributes to the trade effect on transitions, even this feature could be jettisoned without altering our conclusions. For example, were we to adopt a revision process with a fixed probability of idiosyncratic play independent of payoffs (as in Young, 1993, and Kandori, Mailath and Rob, 1993), our theorems 2 and 3 would remain unchanged because trade increases the critical amount of idiosyncratic play necessary to induce a transition which, for a given rate of idiosyncratic play, necessarily entails a reduced probability of escape from the inferior convention in a given period.

We can even replace the evolutionary model of transition with a centralized bargaining model without changing our major result. In appendix C we show that trade impedes transitions to Pareto-superior conventions even if (implausibly, we believe) employers and workers could agree on the contracts and behavioral patterns that would govern the production process, but faced an adjustment cost of making the transition from a nation of fixed wage contracts with self-regarding workers to a nation of partnership contracts and reciprocal workers. The bargaining model does not, however, reproduce the result that factor market integration favors convergence, suggesting that the way the transition is modeled

does make a difference, at least for this result.

The mechanisms underlying our main conclusions impose two generic requirements about the workings of the economy, each of which we find quite plausible (in light of the evidence presented in section 3). First is the complementarity of preferences and contracts, according to which different contracts work better with different individual motivations. Second is the distinction between transparent and opaque goods and the insight that different contract-preference matches will differ in the degree to which they are appropriate for the production of each. Where these two conditions occur, there may be gains from trade made possible by specialization according to cultural-institutional comparative advantage and, as a result, contract-preference mismatches are more costly under trade than autarchy, the key condition for our major result.

Our model is clearly just one way to represent these underlying mechanisms, which, if we are correct, are applicable quite generally.

10 Discussion

We have shown that cultural-institutional differentiation between otherwise identical econo mies provides the basis for comparative advantage, specialization and gains from trade and that these differences will persist in the presence of trade because if agents are sufficiently rational, trade integration impedes convergence even to Pareto-superior cultural-institutional conventions. By contrast, factor market integration favors convergence.

By making behavioral or contractual experimentation more costly, the gains from trade thus increase the impediments to cultural-institutional transitions. Thus, under trade liberalization a nation's cultural-institutional convention may persist over very long periods, even when a Pareto-superior convention exists and when the *status quo* convention confers absolute disadvantage with respect to other countries in all goods. Our model of how endogenous institutions and culture may lock an economy into a set of preferences and contracts that precludes the use of high powered incentives thus provides an answer to what Esfahani and Mookherjee (1995) consider to be "the fundamental question" in the study of economic differences among nations, namely "why inefficient legal or cultural institutions persist" (p. 205).

In light of our results it is certainly tempting to think that a half a millennium ago cultural-institutional advantages – interpersonal trust and the rule of law for example – may have contributed to the specialization of Firenze and Siena in finance (Goldthwaite, 2009). Likewise international trade may have contributed to the persistence of the regional cultural and institutional differences in Cuba in the 19th century, allowing the slave regions of Cuba to specialize in the export of sugar (a prime example of the transparent good), while the non-slave Pinar del Rio region exported tobacco and high quality cigars (exemplifying opaqueness, as in Ortiz, 1963).

Our model may hold insights beyond the study of comparative advantage. For example, 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. 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). Cultural 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. Freeman (2000) and Moriguchi (2003) also document a divergence in labor market institutions in open economies. The "cultural and institutional bifurcation" of China and Europe studied by Greif and Tabellini (2010) could persist even in the presence of exchange (which would favor Europe's specialization in goods in which economies of scale were more pronounced). Finally, the first decade of the Euro-zone saw a quite dramatic institutional divergence among its members; between-country differences in the World Bank's measure of the strength of the rule of law increased substantially and significantly over this period, primarily along North-South lines (see appendix A.2, and Boltho and Carlin, 2012).

Appendix

A. Data appendix

Reciprocity (Herrmann et al., 2008): the indicator is the level of cooperation sustained in a public goods experiment with altruistic punishment of free riders.

Intensity in routine tasks (Costinot et al., 2011): The index corresponds to the average task routineness in each 3-digit NAICS sector. Task routineness is measured using the "importance of making decision and solving problems" in each task according to the U.S. Department of Labor's Occupational Information Network (O*NET). The intensity in tasks across sectors is measured using the share of employment of 6-digit occupations in the Bureau of Labor Statistics (BLS) Occupational Employment Statistics (2006).

Trade data (Feenstra et al., 2005): Taken from the World Trade Flows Database 2000.

Comparative advantage (authors' computation): Computed as explained in the text. Trade data are originally in SITC Rev 2 classification, and converted to NAICS using Feenstra and Lipsey (http://www.nber.org/lipsey/sitc22naics97/) concordance tables.

GDP pro capite (World Bank, 2012b): GDP pro capite in current US dollars in 2000, taken from the World Bank Indicators.

Sample: *Exporters* include all countries for which the Herrmann et al. (2008) index is available: Australia, Belarus, China, Denmark, Germany, Greece, Korea Rep., Oman, Russian Fed., Saudi Arabia, Switzerland, Turkey, Ukraine, UK, USA. *Importers* include the 50 largest importing countries accounting for more than 90% world imports: Areas NES, Argentina, Australia, Austria, Belgium-Lux, Bermuda, Brazil, Canada, Chile, China, China HK SAR, Czech Rep, Denmark, Egypt, Finland, France, Germany, Greece, Hungary, Indonesia, Iran, Ireland, Israel, Italy, Japan, Korea Rep., Kuwait, Malaysia, Mexico, Netherlands, New Zealand, Nigeria, Norway, Philippines, Poland, Portugal, Russian Fed., Saudi Arabia, Singapore, South Africa, Spain, Sweden, Switzerland, Taiwan, Thailand, Turkey, UK, USA, United Arab Em., Venezuela. *Sectors* include all the 3-digit NAICS manufacturing categories (with the exclusion of petroleum and coal products): food products (311); beverage and tobacco products (312); textile mills (313); textile product mills (314); ap-

parel (315); leather and allied products (316); wood products (321); paper products (322); printing and related support activities (323); chemicals (325); plastics and rubber products (326); nonmetallic mineral products (327); primary metals (331); fabricated metal products (332); machinery, computer and electronic products (333); electrical equipment, appliance, and components (335); transport equipment (336); furniture and related products (337); miscellaneous products - medical equipment and other supplies, other miscellaneous (339).

Rule of law (Kaufmann, Kraay and Mastruzzi, 2010). The index is taken from the World Governance Indicators (World Bank, 2012a). The rule of law index "reflects perceptions of the extent to which agents have confidence in and abide by the rules of society, and in particular the quality of contract enforcement, property rights, the police, and the courts, as well as the likelihood of crime and violence". The index ranges from -2.5 (weak) to 2.5 (strong governance performance). Over the period for which these data are available, the Euro-zone nations range from a maximum of 1.98 for Finland in 1998 to a minimum of 0.28 for Italy in 2006. The variances of the log of the index in 2000 and 2010 are 0.0869 and 0.2537 respectively; the p-value for the variance-comparison test (null hypothesis that the variance ratio is equal to 1) is 0.0448. *Sample (Euro-zone in 2001)*: Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, and Spain.

B. Mathematical appendix

B.1 (Theorem 2). Trade integration decreases the probability of escaping the $\{F, E\}$ convention. Using equations (8), we have:

$$\lim_{\beta\to\infty}\frac{\mu_T^k}{\mu_A^k}=\frac{\frac{z!}{(zq_T^k)!(z-zq_T^k)!}}{\frac{z!}{(zq_A^k)!(z-zq_A^k)!}}\lim_{\beta\to\infty}\frac{[\sigma(\beta,\Delta_T^k)]^{zq_T^k}}{[\sigma(\beta,\Delta_A^k)]^{zq_A^k}}\frac{[1-\sigma(\beta,\Delta_T^k)]^{z-zq_T^k}}{[1-\sigma(\beta,\Delta_A^k)]^{z-zq_A^k}}$$

where q_j^k (with $j=A,T$ and $k=worker,firm$) is the fraction of deviants in population k sufficient to escape the basin of attraction of the $\{F,E\}$ convention as explained in section 5 in the paper ($q_j^{worker}=\omega_j^+$ and $q_j^{firm}=\phi_j^+$). Omitting the constant term, and using equations (7), the above limit can be rewritten as:

$$\lim_{\beta\to\infty}\frac{\left(\frac{1}{1+e^{\beta\Delta_T^k}}\right)^{zq_T^k}\left(1-\frac{1}{1+e^{\beta\Delta_T^k}}\right)^{z-zq_T^k}}{\left(\frac{1}{1+e^{\beta\Delta_A^k}}\right)^{zq_A^k}\left(1-\frac{1}{1+e^{\beta\Delta_A^k}}\right)^{z-zq_A^k}}=\lim_{\beta\to\infty}\frac{(1+e^{\beta\Delta_A^k})^{zq_A^k}}{(1+e^{\beta\Delta_T^k})^{zq_T^k}}=\lim_{\beta\to\infty}\frac{(e^{\beta\Delta_A^k})^{zq_A^k}}{(e^{\beta\Delta_T^k})^{zq_T^k}}$$

that, after defining $y\equiv e^{\beta}$ and for finite z, can be solved as

$$\lim_{\beta\to\infty}\frac{\mu_T^k}{\mu_A^k}=\lim_{y\to\infty}\frac{[y^{\Delta_A^k}]^{zq_A^k}}{[y^{\Delta_T^k}]^{zq_T^k}}=\begin{cases}0 \text{ iff } \Delta_A^k q_A^k<\Delta_T^k q_T^k\\ 1 \text{ iff } \Delta_A^k q_A^k=\Delta_T^k q_T^k\\ \infty \text{ iff } \Delta_A^k q_A^k>\Delta_T^k q_T^k\end{cases} \tag{B1}$$

Given Lemma 2 and Lemma 3, we know that, for respectively workers and firms, $\Delta_A^{worker}\omega_A^+<\Delta_T^{worker}\omega_T^+$ and $\Delta_A^{firm}\phi_A^+<\Delta_T^{firm}\phi_T^+$. Hence, we can conclude that there exists $\bar{\beta}$ such that for $\beta>\bar{\beta}$, it must be that for respectively workers and firms $\mu_T^{worker}<\mu_A^{worker}$ and $\mu_T^{firm}<\mu_A^{firm}$.

B.2.1. Country probability difference of matching P-firms and R-workers. In the neighborhood of the $\{F, E\}$ convention, the probability of an employer being paired with a reciprocal employee conditional to being resident, respectively, in country B and in country A is $n(1-\sigma)+(1-n)[s_A\sigma+s_B(1-\sigma)]$ and $n\sigma+(1-n)[s_A\sigma+s_B(1-\sigma)]$, where s_A and s_B are the relative sizes of the two countries. It is straightforward to see that the difference between the two is $n(1-2\sigma)$. If β is sufficiently large, σ goes to zero and these expressions can be approximated by n. Similar expressions are readily found for the corresponding country difference in the probability of an employee being paired with a Partnership.

B.2.2. Expected payoffs and critical fractions of deviants under factor market integration. The proofs contained in this subsection and in the following ones are obtained using autarchic prices. Notice that the nationality of the employer determines the good produced and the price at which the output is sold. Conclusions would not change were this hypothesis relaxed (and/or trade prices were considered instead).

Part 1. The expected payoffs to employers implementing P- and F-contracts are

$$\begin{aligned} v^P_{n<1}(\omega) &= n(1\text{-}b)[\omega C^A_L+(1\text{-}\omega)C^A_N]+(1\text{-}n)(1\text{-}b)\{s_B C^A_L+s_A[\omega C^A_L+(1\text{-}\omega)C^A_N]\}, \\ v^F_{n<1}(\omega) &= n[C^A_N\text{-}(w+m)]+(1\text{-}n)\{s_B[C^A_N\text{-}(w+m)]+s_A[C^A_N\text{-}(w+m)]\}=C^A_N\text{-}(w+m). \end{aligned} \tag{B2}$$

By equating $v^P_{n<1}(\omega)$ and $v^F_{n<1}(\omega)$ and solving for ω, we obtain

$$\omega^*_{n<1}=\frac{\frac{(1-n)s_B}{ns_B+s_A}\{[C^A_N-(w+m)]-(1-b)C^A_L\}+[bC^A_N-(w+m)]}{(1-b)(C^A_L-C^A_N)}. \tag{B3}$$

Part 2. The expected payoffs to R- and E-employees are

$$\begin{aligned} v^R_{n<1}(\phi)=&n\{\phi\{[b+\alpha(1\text{-}b)]C^A_L\text{-}\delta\}+(1\text{-}\phi)[w\text{-}\eta\text{-}\alpha(C^A_N\text{-}w\text{-}m)]\}+(1\text{-}n)\{s_B\{[b+\alpha(1\text{-}b)]C^B_L\text{-}\delta\}+ \\ &+s_A\phi\{[b+\alpha(1\text{-}b)]C^A_L\text{-}\delta\}+(1\text{-}\phi)[w\text{-}\eta\text{-}\alpha(C^A_N\text{-}w\text{-}m)]\}, \\ v^E_{n<1}(\phi)=&n[\phi(bC^A_N\text{-}\eta)+(1\text{-}\phi)(w\text{-}\eta)]+(1\text{-}n)\{s_B(bC^B_N\text{-}\eta)+s_A[\phi(bC^A_N\text{-}\eta)+(1\text{-}\phi)(w\text{-}\eta)]\}. \end{aligned} \tag{B4}$$

By equating $v^R_{n<1}(\phi)$ and $v^E_{n<1}(\phi)$ and solving for ϕ, we obtain

$$\phi^*_{n<1}=\frac{\frac{(1-n)s_B}{ns_B+s_A}\{(bC^B_N-\eta)-\{[b+\alpha(1-b)]C^B_L-\delta\}\}+\alpha(C^j_N-w-m)}{\{[b+\alpha(1-b)]C^A_L-\delta\}-(bC^A_N-\eta)+\alpha(C^A_N-w-m)}. \tag{B5}$$

B.2.3 (Lemma 4). Factor market integration decreases the costs of deviation. *Part 1.* The cost of deviation from the $\{F, E\}$ convention for employers, $\Delta^{firm}_{n<1}$, is given by $v^F_{n<1}(\omega=0)-v^P_{n<1}(\omega=0)$, where $v^F_{n<1}(\omega=0)$ and $v^P_{n<1}(\omega=0)$ are defined by equations (B2) with $\omega=0$. This difference is decreasing in the degree of factor market integration $(1-n)$, because $C^A_L>C^A_N$ (the intuition is that $v^F_{n<1}(\omega=0)$ is unaltered by factor market integration $(n<1)$, while $v^P_{n<1}(\omega=0)$ is increased).

Part 2. The cost of deviation from the $\{F, E\}$ convention for workers, $\Delta^{worker}_{n<1}$, is given by $v^E_{n<1}(\phi=0)-v^R_{n<1}(\phi=0)$, where $v^E_{n<1}(\phi=0)$ and $v^R_{n<1}(\phi=0)$ are defined by equations (B3) with $\phi=0$. This difference is decreasing in the degree of factor market integration $(1-n)$, because $[b+\alpha(1-b)]C^B_L-\delta>bC^B_L-\eta$ (the intuition is that both $v^E_{n<1}(\phi=0)$ and $v^R_{n<1}(\phi=0)$ are increased by factor market integration $(n<1)$, but $v^R_{n<1}(\phi=0)$ is increased more).

B.2.4 (Lemma 5). Factor market integration decreases the critical fractions of deviants. *Part 1.* $\omega^*_{n<1}$ is decreasing in the degree of factor market integration because $[C^A_N - (w+m)] - (1-b)C^A_L < 0$. When $n = 1$, (B3) is equal to the first of expressions (5). *Part 2.* $\phi^*_{n<1}$ is decreasing in the degree of factor market integration because $\{(bC^B_N - \eta) - [b + \alpha(1-b)]C^B_L - \delta\} < 0$. When $n = 1$ (B5) is equal to the second of expressions (5).

B.2.5 (Theorem 3). Trade integration increases the probability of escaping the $\{F, E\}$ convention. Using the same methodology as in B.1, it can be shown that

$$\lim_{\beta \to \infty} \frac{\mu^k_{n<1}}{\mu^k_{n<1}} = \begin{cases} 0 \text{ iff } \Delta^k_{n=1} q^k_{n=1} > \Delta^k_{n<1} q^k_{n<1} \\ 1 \text{ iff } \Delta^k_{n=1} q^k_{n=1} = \Delta^k_{n<1} q^k_{n<1} \\ \infty \text{ iff } \Delta^k_{n=1} q^k_{n=1} < \Delta^k_{n<1} q^k_{n<1} \end{cases} \tag{B6}$$

Given Lemma 4 and Lemma 5, we know that, for respectively workers and firms, $\Delta^{worker}_{n=1}\omega^+_{n=1} > \Delta^{worker}_{n<1}\omega^+_{n<1}$ and $\Delta^{firm}_{n=1}\phi^+_{n=1} > \Delta^{firm}_{n<1}\phi^+_{n<1}$. Hence, we can conclude that there exists $\bar{\beta}_n$ such that for $\beta > \bar{\beta}_n$, it must be that for respectively workers and firms $\mu^{worker}_{n<1} > \mu^{worker}_{n=1}$ and $\mu^{firm}_{n<1} > \mu^{firm}_{n=1}$.

C. Extension: A centralized model of bargaining

Suppose that employers and workers could agree on the contracts and behavioral patterns that would govern the production process, but faced an adjustment cost of making the transition from a nation of fixed wage contracts with self-regarding workers to a nation of partnership contracts and reciprocal workers. Following the transition the two economies would have identical production possibilities, and hence would reap no gains from specialization. But both employers and workers would receive higher payoffs following a transition (as long as the pre-transition terms of trade were strictly between the relative cost ratios of the two countries, that is as long as the gains from trade were shared between countries). The two classes thus would readily agree to implement a transition if for each the present value of the resulting gains exceeded their share of the cost of transition.

We can ask if trade will facilitate a transition to the Pareto-superior cultural-institutional equilibrium by determining the effect of trade liberalization on the greatest rate of time discount (common to both classes) such that both classes would not prefer to remain at the inferior equilibrium. This critical rate of time discount (denoted by θ^*) will depend on the payoffs at the inferior equilibrium under both autarchy and trade, the payoffs at the superior equilibrium under autarchy, and on the distribution of the costs of transition among the two groups, which we assume sum to a total τ but can be divided in any way that allocates to the firms a non-negative share f ($\leqslant 1$). Because in this setup agreement between the two classes is necessary for a transition, we determine for each class the greatest degree of impatience (discount rate) consistent with that class not preferring the *status quo*, as a function of the cost of transition borne by that class. We then allocate the costs of transition so as to maximize the minimum of these two critical rates (so as to allow the maximum degree of impatience among employees and firms alike that is consistent with a transition). This requires that the critical rates be equalized (if they differ, then shifting some of the cost of transition to the class with the higher critical rate will raise the minimum of the two).

Trade will impede transitions if θ^* is greater under autarchy than under trade, namely $\theta^*_A > \theta^*_T$. This indeed will be the case. While the gains to workers of making a transition to the superior equilibrium are not affected by trade, this is not the case for firms; as a result of gains from trade (allowing country A to specialize in its least disadvantaged good), the increase in the payoffs of firms induced by the transition to the superior cultural-institutional equilibrium will be less after trade liberalization than under autarchy. As a result it is not difficult to show that

Theorem 4 (Trade impedes bargained transitions to Pareto-superior equilibria). Trade decreases the greatest rate of time discount such that both workers and firms do not prefer to remain at the inferior $\{F, E\}$ equilibrium, that is $\theta^*_A > \theta^*_T$.

Proof. Assume a given total cost τ. Trade does not affect employees' gain from a transition, $\{[b+\alpha(1-b)]C^A_L - \delta\} - (w-\eta)$. The maximum impatience of a worker consistent with her not preferring the *status quo* is that which equates the present value of the transition to its costs, or

$$\frac{\{[b+\alpha(1-b)]C^A_L - \delta\} - (w-\eta)}{\tau(1-f)} = \bar{\theta}^{worker}(\tau, f), \tag{C1}$$

with $\bar{\theta}^{worker}(\tau, f)$ increasing in f because the smaller the cost of transition borne by the worker the more impatient can she be and yet still prefer a transition. Under autarchy and trade respectively firms will be indifferent to a transition if their rate of time preference is equal to $\bar{\theta}^{firm}_A(\tau, f)$ and $\bar{\theta}^{firm}_T(\tau, f)$ respectively, given by

$$\bar{\theta}^{firm}_A(\tau, f) = \frac{C^A_L(1-b) - [C^A_N - (w+m)]}{\tau f} > \frac{C^A_L(1-b) - [C^T_N - (w+m)]}{\tau f} = \bar{\theta}^{firm}_T(\tau, f) \tag{C2}$$

with $\bar{\theta}^{firm}(\tau, f)$ decreasing in f. Inequality (C2) is true because as a result of gains from trade, $C^T_N > C^A_N$.

To determine the greatest level of impatience consistent with a transition under autarchy and trade respectively, we find f_j with $j = A, T$, the f that solves $\bar{\theta}^{worker}(\tau, f_j) = \bar{\theta}^{firm}_j(\tau, f_j)$, that is

$$\frac{\{[b+\alpha(1-b)]C^A_L - \delta\} - (w-\eta)}{\tau(1-f_j)} = \frac{C^A_L(1-b) - [C^j_N - (w+m)]}{\tau f_j}, \tag{C3}$$

which gives us $\theta^*_j(\tau, f_j)$. But we know that $f_A > f_T$ because of (C2) above (under autarchy employers can bear a larger cost without opposing a transition than under trade). Thus to induce firms' participation in the transition, under trade workers must bear a larger cost of the transition, so their (and the common) greatest level of time preference consistent with a bargained transition is lower under trade than under autarchy, that is, $\theta^*_A(\tau, f_A) < \theta^*_T(\tau, f_T)$ because $\bar{\theta}^{worker}(\tau, f)$ is increasing in f.

The intuition behind the result is illustrated in Figure 8.

Observation 1: If (contrary to our assumption) the gains from trade were to some extent shared with workers, so that under trade in the $\{F, E\}$ equilibrium workers payoff exceeded $w-\eta$ that they receive under autarchy, then the $\bar{\theta}^{worker}(\tau, f)$ function downwards, as it did the $\bar{\theta}^{firm}(\tau, f)$ function. This would reduce θ^*_T and thus increase the effect of trade as an impediment to a transition ($\theta^*_A(\tau, f) - \theta^*_T(\tau, f)$ would be larger).

Observation 2: The decentralized and top down models of transition differ in important ways, however. This is best seen by noting that the model of centralized transitions does

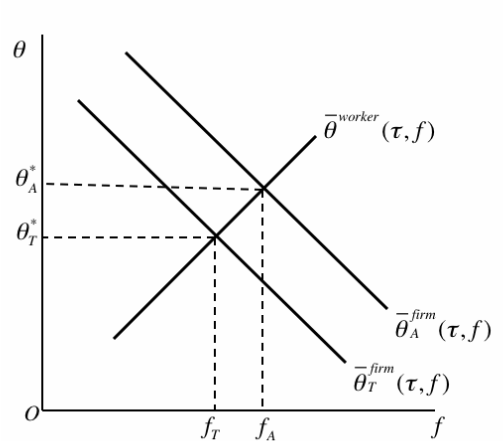

Figure 8: *Transition to a Pareto-superior equilibrium requires more far-sighted employers and workers under trade than under autarchy.* Note: θ is the rate of time preference, f is the share of transition costs borne by the firms. $\bar{\theta}^{worker}$ and $\bar{\theta}^{firm}$ are the maximum level of impatience consistent with firms and workers respectively not preferring the *status quo* rather than a transition.

not replicate the result of Theorem 3 (factor market integration facilitates transitions to the superior equilibrium). This is because factor market integration, like trade reduces the payoff gains associated with a transition. Firms in the *status quo* in country A are not affected (their workers will perform N labor only in any case), while workers' expected payoffs at the *status quo* equilibrium improve (R and E-workers alike do better when they are paired with a P-firm in the pool).

References

Acemoglu, D. 2003. Why not a Political Coase Theorem? Social Conflict, Commitment and Politics. *Journal of Comparative Economics*, 31: 620-52.

Acemoglu, D., P. Antràs, E. Helpman. 2007. Contracts and Technology Adoption. *American Economic Review*, 97(3): 916-943.

Aghion, P., Y. Algan, P. Cahuc, A. Shleifer. 2010. Regulation and Distrust. *Quarterly Journal of Economics,* 125(3): 1015-1049.

Alesina, A., P. Giuliano. 2010. The Power of the Family. *Journal of Economic Growth*, 15(2): 93-125.

Autor, D.H., F. Levy, R.J. Murnane. 2003. The Skill Content of Recent Technological Change: An Empirical Exploration. *Quarterly Journal of Economics*, 118(4): 1279-1333.

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

Bandiera, O., I. Barankay, I. Rasul. 2005. Preferences and the Response to Incentives: Evidence from Personnel Data. *Quarterly Journal of Economics*, 120: 917-62.

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

Belloc, M. 2006. Institutions and International Trade: a Reconsideration of Comparative Advantage. *Journal of Economic Surveys*, 20(1): 3-26.

Belloc, M., S. Bowles. 2013. The Persistence of Inferior Cultural-Institutional Conventions. *American Economic Review: Papers and Proceedings*, forthcoming.

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

Bidner, C., P. Francois. 2011. Cultivating Trust: Norms, Institutions and the Implications of Scale. *Economic Journal*, 121(555): 1097-1129.

Binmore, K., L. Samuelson, H.P. Young. 2003. Equilibrium Selection in Bargaining Models. *Games and Economic Behavior*, 45(2): 296-328.

Bisin, A., T. Verdier. 2011. The Economics of Cultural Transmission and Socialization. In *Handbook of Social Economics* 1A, ed. J. Benhabib, A. Bisin, and M. Jackson. The Netherlands: North-Holland.

Bisin, A., T. Verdier. 2001. The Economics of Cultural Transition and the Evolution of Preferences. *Journal of Economic Theory*, 97(2): 298-319.

Bloom, N., R. Sadun, J. Van Reenen. 2012. The Organization of Firms across Countries. *Quarterly Journal of Economics*, forthcoming.

Blume, L.E. 2003. How Noise Matters. *Games and Economic Behavior*, 44: 251-271.

Boltho, A., W. Carlin. 2012. *The Problems of European Monetary Union - Asymmetric Shocks or Asymmetric Behavior.* Vox: http://www.voxeu.org/index.php?q=node/7808 (March 31, 2012).

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

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

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

Bowles, S., S. Polania Reyes. 2012. Economic Incentives and Social Preferences: Complements or Substitute. *Journal of Economic Literature*, 50: 1-57.

Boyd, R., P.J. Richerson. 1985. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Brown, M., A. Falk, E. Fehr. 2004. Relational Contracts and the Nature of Market Interactions. *Econometrica*, 72(3): 747-80.

Brugger, B., R. Lalive, J. Zweimuller. 2010. Does Culture Affect Unemployment? Evidence from the Rostigraben. *University of Laussane*.

Cavalli-Sforza, L.L., M.W. Feldman. 1981. *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton (NJ): Princeton University Press.

Clark, G., R.C. Feenstra. 2003. Technology and the Great Divergence. In *Globalization in Historical Perspective*, ed. M.D. Bordo, A.M. Taylor, and J.G. Williamson. NBER Inc. Chicago: University of Chicago Press: 277-322.

Conconi, P., P. Legros, A.F. Newman. 2012. Trade Liberalization and Organizational Change. *Journal of International Economics*, 86(2): 197-208.

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

Costinot, A., L. Oldenski, J.E. Rauch. 2011. Adaptation and the Boundary of Multinational Firms. *Review of Economics and Statistics*, 93(1): 298-308.

Davis, D.R., D.E. Weinstein. 2001. An Account of Global Factor Trade. *American Economic Review*, 91(5): 1423-1453.

Dell, M. 2010. The Persistent Effect of Peru's Mining Mita. *Econometrica*, 78: 1863-903.

Do, Q-T., A. Levchenko. 2009. Trade, Inequality and the Political Economy of Institutions. *Journal of Economic Theory*, 144(4): 1489-1520.

Esfahani, H.S., D. Mookherjee. 1995. Productivity, Contracting Modes, and Development. *Journal of Economic Development*, 46: 203-231.

Falk, A., M. Kosfeld. 2006. The Hidden Costs of Control. *American Economic Review*, 96(5): 1611-30.

Fehr, E., A. Falk. 2002. Psychological Foundations of Incentives. *European Economic Review*, 46(4-5): 687-724.

Fehr, E., S. Gachter, G. Kirchsteiger. 1997. Reciprocity as a Contract Enforcement Device: Experimental Evidence. *Econometrica*, 65(4): 833-60.

Fehr, E., A. Klein, K.M. Schmidt. 2007. Fairness and Contract Design. *Econometrica*, 75(1): 121-54.

Fehr, E., B. Rockenbach, 2003. Detrimental Effects of Sanctions on Human Altruism. *Nature*, 422: 137-140.

Feenstra, R.C., R.E. Lipsey, H. Deng, A.C. Ma, H. Mo (2005) World Trade Flows: 1962-2000. *NBER WP* 11040.

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

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

Foster, D., H.P. Young. 1990. Stochastic Evolutionary Game Dynamics. *Theoretical Population Biology*, 38: 219-232.

Freeman, R.B. 2000. Single Peaked vs. Diversified Capitalism: The Relation Between Economic Institutions and Outcomes. In *Advances in Macroeconomic Theory*, ed. J. Dreze, 139-170. London, UK: Palgrave.

Galor, O., O. Moav. 2002. Natural Selection and the Origin of Economic Growth. *Quarterly Journal of Economics*, 117(4): 1133-91.

Galor, O., A. Mountford. 2008. Trading Population for Productivity: Theory and Evidence. *Review of Economic Studies*, 75(4): 1143-79.

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

Goldthwaite, R. 2009. *The Economy of Renaissance Florence.* Baltimore: Johns Hopkins.

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

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

Greif, A. 2006. *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade.* Cambridge Mass: Cambridge University Press.

Greif, A., G. Tabellini. 2010. Cultural and Institutional Bifurcation: China and Europe Compared. *American Economic Review: Papers & Proceedings*, 100(2): 1-10.

Guiso, L., P. Sapienza, L. Zingales. 2009. Cultural Biases in Economic Exchange? *Quarterly Journal of Economics*, 124(3): 1095-1131.

Hannan, M., J. Freeman. 1989. *Organizational Ecology.* Cambridge, MA: Harvard University Press.

Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, R. McElreath, M. Alvard, A. Barr, J. Ensminger, N. Smith Henrich, K. Hill, F. Gil-White, M. Gurven, F. Marlowe, J. Patton, D. Tracer. 2005. 'Economic Man' in Cross-Cultural Perspective: Behavioral Experiments in 15 Small-Scale Societies. *Behavioral and Brain Sciences*, 28: 795-855.

Helpman, E. 1999. The Structure of Foreign Trade. *Journal of Economic Perspectives*, 13(2): 121-144.

Herrmann, B., C. Thoni, S. Gaechter. 2008. Antisocial Punishment Across Societies. *Science*, 319: 1362-67.

Houser, D., E. Xiao, K. McCabe, V. Smith. 2008. When Punishment Fails: Research on Sanctions, Intentions and Non-cooperation. *Games and Economic Behavior*, 62(2): 509-532.

Huck, S., D. Kubler, J. Weibull. 2012. Social Norms and Economic Incentives in Firms. *Journal of Economic Behavior & Organization*, 83, 173-85.

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

Kandori, M.G., G.J. Mailath, R. Rob. 1993. Learning, Mutation, and Long Run Equilibria in Games. *Econometrica*, 61(1): 29-56.

Karabay, B., J. McLaren. 2010. Trade, Offshoring and the Invisible Handshake. *Journal of International Economics*, 82(1): 26-34.

Kaufmann, D., A. Kraay, M. Mastruzzi. 2010. The Worldwide Governance Indicators: Methodology and Analytical Issues. *World Bank, Policy Research Working Paper*, 5430.

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

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

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

Kreindler, G.H, H.P. Young. 2011. Fast Convergence in Evolutionary Equilibrium Selection. *Discussion Paper 569, Department of Economics, University of Oxford*.

Krugman, P. 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: 41-55.

Leontief, W. 1953. Domestic Production and Foreign Trade; the American Capital Position Re-Examined. *Proceedings of the American Philosophical Society*, 97(4): 332-349.

Levchenko, A.A. 2007. Institutional Quality and International Trade. *Review of Economic Studies*, 74(3): 791-819.

Levchenko, A.A. 2012. International Trade and Institutional Change. *Journal of Law, Economics, and Organization*, forthcoming

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

Lewis, W.A. 1954. Economic Development with Unlimited Supplies of Labor. *Machester School*, 22: 139-91.

Luttmer, E.F.P., M. Singhal. 2011. Culture, Context, and the Taste for Redistribution. *American Economic Journal: Economic Policy*, 3: 157-179.

Moriguchi, C. 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): 625-665.

Myatt, D.P., C.C. Wallace. 2004. Adaptive Play by Idiosyncratic Agents. *Games and Economic Behavior*, 48: 124-138.

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

Nunn, N. 2007. Relation-Specificity, Incomplete Contracts and the Pattern of Trade. *Quarterly Journal of Economics*, 122(2): 569-600.

Nunn, N., L. Wantchekon. 2011. The Slave Trade and the Origins of Mistrust in Africa. *American Economic Review*, 101: 3221-3252.

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

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

Polity IV Project. 2010. *Political Regime Characteristics and Transitions*, 1800-2009. http://www.systemicpeace.org/polity/polity4.htm.

Przeworski, A, M.E. Alvarez, J.A. Cheibub, F. Limongi. 2000. *Democracy and Development: Political Institutions and Well-being in the World*, 1950-1990. Cambridge, Mass: Cambridge University Press.

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

Samuelson, P. 1948. International Trade and the Equalization of Factor Prices. *Economic Journal*, 58(230): 163-84.

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

Spolaore, E., R. Wacziarg 2009. The Diffusion of Development. *Quarterly Journal of Economics*, 124(2): 469-530.

Tabellini, G. 2010. Culture and Institutions: Economic Development in the Regions of Europe. *Journal of The European Economic Association*, 8(4): 677-716.

Trefler, D. 1993. International Factor Price Differences: Leontief was Right! *Journal of Political Economy*, 101(6): 961-987.

Trefler, D. 1995. The Case of Missing Trade and Other Mysteries. *American Economic Review*, 85(5): 1029-1046.

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

Weibull, J. W. 1995. *Evolutionary Game Theory*. Cambridge, MA: MIT Press.

World Bank. 2012a. *Worldwide Governance Indicators*. info.worldbank.org/governance/wgi/.

World Bank. 2012b. http://data.worldbank.org/indicator/NY.GDP.PCAP.CD.

Young, H.P. 1993. An Evolutionary Model of Bargaining. *Journal of Economic Theory* 59(1): 145-168.

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

Young, H.P., M. Burke. 2001. Competition and Custom in Economic Contracts: A Case Study of Illinois Agriculture. *American Economic Review*, 91(3): 559-73.

Zeira, J. 2011. Globalization and Divergence. *Hebrew University of Jerusalem and LUISS*.