CDDRL

WORKING PAPERS

*Structural Convergence**

Romain Wacziarg

Center on Democracy, Development, and the Rule of Law
Stanford Institute for International Studies

Number 8
12 August 2004

This working paper was produced as part of CDDRL's ongoing programming on economic and political development in transitional states. Additional working papers appear on CDDRL's website: http://cddrl.stanford.edu.

Center on Democracy, Development,
and the Rule of Law
Stanford Institute for International Studies
Stanford University
Encina Hall
Stanford, CA 94305
Phone: 650-724-7197
Fax: 650-724-2996
http://cddrl.stanford.edu/

About the Center on Democracy, Development and the Rule of Law (CDDRL)

CDDRL was founded by a generous grant from the Bill and Flora Hewlett Foundation in October in 2002 as part of the Stanford Institute for International Studies at Stanford University. The Center supports analytic studies, policy relevant research, training and outreach activities to assist developing countries in the design and implementation of policies to foster growth, democracy, and the rule of law.

About the Author

Romain Wacziarg is Associate Professor of Economics at Stanford University Graduate School of Business. His research focuses on economic integration, trade and economic growth, political economy of trade, economic aspects of political separatism. His publications include *Trade Liberalization and Intersectoral Labor Movements* (forthcoming, *Journal of International Economics*, 2004); *Fractionalization, Journal of Economic Growth*, June 2003; *Stages of Diversification, American Economic Review*, March 2003; *Measuring the Dynamic Gains from Trade, World Bank Economic Review*, October 2001

Structural Convergence*

Romain Wacziarg
Stanford University

First Draft: May 2001

Abstract

This paper establishes empirically the existence of structural convergence: country pairs that converge in terms of per capita income also tend to converge in terms of their sectoral similarity, measured by the bilateral correlation of their sectoral labor shares. This is a robust feature of the data at various levels of sectoral disaggregation and data coverage. We shed light on some explanations for structural similarity, chiefly trade related determinants. Convergence in relative factor abundance accounts for approximately 1/3 of the extent of structural convergence. We argue that the existence of structural convergence has important implications for our understanding of business cycles transmission, of long-run development patterns and of the dynamics of specialization.

*Stanford Graduate School of Business, 518 Memorial Way, Stanford CA 94305-5015, wacziarg@gsb.stanford.edu. Prepared for the International Seminar on Macroeconomics in Dublin, Ireland, June 8-9, 2001. I thank Andrew Rose for generously providing his bilateral trade data, and Lanier Benkard for useful comments. All errors are mine.

1 Introduction

Do countries become more similar in terms of their sectoral structures as their incomes converge? This paper provides broad empirical support for the existence of such structural convergence. Namely, country pairs that converge in terms of per capita income also tend to converge in terms of their sectoral similarity, measured by the bilateral correlation of their sectoral labor shares.

There are several reasons for economists to be interested in the phenomenon of structural convergence. Firstly, as suggested by Imbs (2000), if shocks to the macroeconomy are sector-specific, structural convergence has implications for the international transmission of business cycles: it should give rise to increased international business cycle correlations. This is a short- to medium-run concern.

Secondly, and perhaps more importantly, a study of structural convergence can provides a novel way to examine the process of development in the longer run. The existence of structural convergence suggests that countries follow similar stages of development characterized by the rise and fall of similar types of sectors as income grows, and that countries may converge to a structural "steady-state", in which the sectoral mix of output becomes more uniform across countries (conditional on observing income convergence among them). The notion that countries grow through structural stages is consistent with recent findings in Imbs and Wacziarg (2000), who show that the sectoral concentration of labor follows a U-shaped pattern over the course of development for a broad sample of countries.

Thirdly, understanding the determinants of structural convergence can inform theoretical debates on the long-run dynamic pattern of international specialization.[1] For example, relating increased similarity in sectoral structure to changes in relative factor abundance can provide evidence on the

[1] Several caveats are in order here, however. The extent to which a finding of structural convergence can inform debates on trade induced specialization depends on the data coverage, the level of disaggregation at which structural convergence occurs and the relative importance of nontraded goods in the overall economy. The coverage of the data matters because a finding of structural convergence within manufacturing is not inconsistent with specialization, which may occur for economy-wide categories including agriculture, mining and services (or vice versa). The level of disaggregation matters because a finding of structural convergence at any given level of disaggregation is not inconsistent with the existence of specialization at narrower levels. Finally, in the context of an expansion of non-traded goods, specialization occuring within traded categories in not inconsistent with structural convergence being observed economy-wide. For further discussions of these points, see Seddon and Wacziarg (2001).

Heckscher-Ohlin model of trade. Similarly, if bilateral trade intensities are found to affect sectoral similarity negatively, this can be taken as evidence of classical (interindustry) specialization. If they affect it positively, this can be interpreted as indicating the expansion of intraindustry trade.[2]

Despite these three important reasons to study structural convergence, the concept has received very little attention in the existing literature on structural change.[3] In contrast, this paper establishes empirically the existence of structural convergence. Moreover, we explore some empirical explanations for this phenomenon, focusing mostly on the role of bilateral trade and relative factor abundance in the determination of dynamic changes in sectoral similarity.

In theory, structural change and hence structural convergence, as defined above, can result from three main forces.[4] Firstly, demand side effects, i.e. Engel effects resulting from income growth, might generate increased sectoral similarity between country pairs with converging incomes. Secondly, on the supply side, convergence in sectoral labor productivity levels across country pairs would create a tendency to allocate increasingly similar shares of labor intersectorally. Thirdly, structural convergence could be linked to trade-related considerations. In particular, if countries with converging per capita incomes also experience convergence in the determinants of comparative advantage (such as relative factor abundance), then they can be expected to structurally converge as well, because they will specialize in producing increasingly similar types of goods.[5] Additionally, the bilateral intensity of trade will be negatively related to sectoral similarity if classical (interindustry) specialization underlies the extent of trade. This study focuses mostly on the third, trade-based set of explanations for structural convergence.[6]

---

[2] See Imbs (2000, 2001) for a thorough empirical investigation of this point in the context of the OECD.

[3] A relative exception is Imbs (2000), who studies the role of sectoral similarity in the determination of international business cycles correlations for a sample of OECD countries. However, structural similarity is largely treated as an independent variable in his study, whereas the current paper seeks to explain its dynamics.

[4] See Chenery and Syrquin (1986) for a discussion of these three forces applied to structural change more broadly.

[5] See Ventura (1997) for a dynamic Heckscher-Ohlin model where such an effect can arise.

[6] We leave the consideration of the first two sets of explanations for future research. They would require the use of sectoral productivity and sectoral consumption data, which are not readily available at the level of disaggregation and for the data coverage of the present study.

This paper is organized as follows: Section 2 examines the pattern of structural convergence for a broad set of country pairs, and varying levels of disaggregation and data coverage. Section 3 examines a series of robustness issues, and focuses on some geographic features of structural convergence. Section 4 examines more closely the determinants of structural convergence, with particular attention to the role of trade intensities, factor abundance and geography. Section 5 concludes.

2 Income Convergence and Structural Similarity

2.1 Definitions and Measurement

Structural convergence is defined as follows: two countries are said to structurally converge if convergence in their per capita incomes is accompanied by convergence in their sectoral structure. Per capita income convergence occurs if the difference between the log of per capita income in the richest country and the log of per capita income in the poorest country in each pair (INCDIFF) falls.[7] The degree of similarity in sectoral structure for a pair of countries is captured by computing the correlation of sectoral labor shares at each point in time.[8] Obviously, a high correlation denotes a similar sectoral structure. The use of employment data is justified by the fact that output data in volume (i.e. deflated by sector specific price indices) is not available for most non-OECD countries in the sample.[9]

The sectoral labor data used to compute bilateral correlations comes from two sources. Firstly, we use economy-wide, 1-digit level data from the ILO (1997). These data are available for 82 countries (or 3321 country pairs), and span the period 1969 to 1997. The bilateral correlation of employment shares for the ILO data is denoted ILOCORR. Secondly, we use 3-digit manufacturing data from UNIDO (1997). Labor shares for the UNIDO data can be computed for 128 countries (or 8128 pairs) over the period 1963-1997,

[7] Alternatively, we define income convergence simply as a fall in the ratio of per capita incomes of the richest to the poorest country (INCDIFF2). This does not greatly affect the results, as discussed below.

[8] This is fairly standard. Shea (1996) used a similar measure to examine whether industry pairs tend to locate employment in the same US cities. Imbs (2000) used the correlation of sectoral labor shares across pairs of OECD countries to evaluate the degree of structural similarity.

[9] At any rate, the use of employment data to measure sector size is standard in the literature, see for example Krugman (1991) and Kim (1995).

and the bilateral correlation of these shares is denoted UNIDOCORR.[10] The Data Appendix lists the sectoral coverage, country coverage and the sources of the data used in this paper.

Tables 1 and 2 display basic summary statistics for the basic annual frequency data used throughout this section to demonstrate the existence of structural convergence. A notable feature of the dataset is the large number of observations that are available: 128,742 for the UNIDO sample, and 31,207 for the ILO sample. Pairwise correlations reveal negative and statistically significant correlations between income differences, measured either by INCDIFF or INCDIFF2, and measures of bilateral sectoral similarity. This suggests that narrower income gaps are associated with greater structural similarity. Moreover, the magnitude of the correlation is more than twice larger for the ILO (roughly 0.60) compared to the UNIDO dataset (roughly 0.25). Since these simple correlations pool between-pair and within-pair variations, however, they may not be indicative of the dynamics of structural similarity, to which we now turn.

2.2 The Evolution of Structural Similarity Through Time

A preliminary analysis of the dynamics of bilateral structural similarity can be obtained by examining its evolution through time within country pairs. To do so, we can run fixed-effects regressions of UNIDOCORR and ILOCORR on a simple time trend. Table 3 presents the results of this analysis. The ILO data demonstrated a trend towards greater structural similarity, as the coefficient on the time trend is positive and highly significant. Unconditionally, therefore, country pairs exhibit more similar economic structures through time when structure is measured using broad, economy wide sectoral categories. Results are less robust when using the UNIDO data - the data reveal a trend towards less similarity when all available observations are used ("unrestricted sample"), and more similarity when the sample is restricted to observations available in both the ILO and UNIDO samples ("common sample").

At any rate, the estimated trend is extremely weak in magnitude - it

[10] For both datasets, our panel will be unbalanced since not every country has observations for the entire time span of the data. Similarly, not every country has observations for all 9 sectors (ILO) or 28 sectors (UNIDO), although such differences will be washed away whenever fixed effects are used. Finally, the data was transformed so that each country has observations for the same number of sectors through time - if this were not the case the bilateral correlations of sector shares would not follow a consistent definition over time.

would take 10 years to *raise* the bilateral correlation of sectoral labor shares by 0.024 for the ILO dataset (the standard deviation of ILOCORR being equal to 0.329), and the same number of years to *reduce* the correlation by 0.008 for the UNIDO dataset (the standard deviation of UNIDOCORR is 0.282). In other words, there is not a sweeping tendency towards greater or lower structural similarity through time. Hence, the phenomenon of structural convergence documented below is unlikely to be accounted for by a broader trend affecting countries whether or not they converge in terms of per capita income.

2.3 Structural Similarity and Income Convergence

To assess the existence of structural convergence, we can run fixed-effects regressions of the measures of structural similarity on INCDIFF, the measure of income similarity:

$$CORR_{it} = \alpha_i + \beta \times INCDIFF_{it} + \varepsilon_{it} \tag{1}$$

where i indexes country pairs and t indexes time.

Fixed-effects estimation allows us to isolate the within-pair variation in the data - i.e. the dynamic relationship between structural similarity and income convergence, as opposed to the cross-sectional relationship. In the equation above, β can be referred to as the coefficient of structural convergence. A fixed effects estimate of β should be interpreted as a partial correlation identifying the within-pair relationship between income similarity and structural similarity (in particular, no causal meaning should be imparted to $\widehat{\beta}$). Obviously, the existence of structural convergence is associated with a negative estimate of β.

Table 4 presents the central result in this paper. Whether or not we restrict the sample to observations common to the ILO and UNIDO datasets (in order to facilitate comparisons), a narrowing of the income gap is significantly associated with greater similarity in economic structure. A significant aspect of these results is the importance of narrowing the income gap in explaining the variation in ILOCORR - the R-squared statistic varies between 0.37 and 0.47 depending on the sample - suggesting that income convergence is closely related to the dynamics of structural similarity for broad, economy-wide sectoral categories. This is consistent with an older literature on structural change which pointed out that, when considering three categories (agriculture, manufacturing and services), countries seem to go through similar development stages, characterized by the initial fall of agriculture as a share of total employment, and the concurrent rise of manufacturing and services, preceding the relative acceleration of services employment.[11] What is even more surprising, perhaps, are the results pertaining to the UNIDO dataset, where structural convergence also holds, although closing the income gap accounts for a smaller part of the overall variance in UNIDOCORR.

The magnitude of the effect is comparable for the ILO and UNIDO datasets. Indeed, the point estimate on the INCDIFF coefficient suggests that halving the income gap between the richest and the poorest country (Y_R/Y_P) should lead to a $\log 2 \times 0.0725 = 0.05$ increase in ILOCORR (or 15.3% of its standard deviation) and a $\log 2 \times 0.04 = 0.028$ increase in UNIDOCORR (or 10% of its standard deviation) in the common (ILO / UNIDO) sample. Figure 1 further displays the magnitude of the effect by plotting the fitted relationship between our measures of structural similarity and Y_R/Y_P, for the common ILO and UNIDO samples. The figure shows that the curve for the ILO is characterized by predicted bilateral sectoral labor share correlations that are always larger than for the UNIDO dataset, but that the estimated relationships are otherwise similar.

2.4 Functional Form

An alternative measure of the income gap would consist simply of Y_R/Y_P, the ratio of per capita incomes of the richest to the poorest country in each pair, labelled INCDIFF2. We can compute results based on this alternative definition to establish the functional form linking sectoral similarity measures and income differences. Table 5 presents the results of fixed effects regressions using INCDIFF2. The results are similar to those obtained using the difference in log incomes. Namely, the coefficient on INCDIFF2 is consistently negative and highly significant statistically, confirming the existence of structural convergence. The magnitude of the average effects is of the same order. At the mean of INCDIFF2 (equal to 4.291), halving the income ratio Y_R/Y_P now results in an increase of ILOCORR of 0.038 and an increase of UNIDOCORR of 0.022 (using the common sample estimates for comparability).

[11] See Chenery and Syrquin (1986) for a summary of this pattern of development. Note that the pattern uncovered in this older literature still begs for a definitive explanation. Moreover, the literature on structural transformation did not explicitly consider the dynamics of inter-country sectoral similarity - rather, it simply described what seemed to be empirical patterns holding for a variety of countries. Finally, this literature ignored intra-manufacturing dynamics captured here through the use of the UNIDO data.

To further investigate the functional form that best fits the data for the relationship between income similarity and structural similarity, we can fit polynomial functions of INCDIFF2, and plot the fitted curves against the semilog specification. If the resulting estimated curves looked similar to the log specification, this would reinforce our confidence in the functional form implied by the use of INCDIFF. Specifically, we ran fixed-effects regressions based quadratic, cubic and fourth degree polynomial functions of INCDIFF2.[12]

The quadratic specification would automatically impose non- monotonicity in the estimated relationship between sectoral similarity and income similarity, and such a non-monotonicity could conceivably characterize the true relationship as well. Estimating a quadratic specification resulted in a U-shaped curve with the minimum point occurring around INCDIFF2=25, for both the UNIDO and ILO samples. However, this was simply the result of fitting a quadratic function to the data. When isolating the subsample with INCDIFF2>25 and running a simple linear fixed-effects regression of sectoral similarity on INCDIFF2, the resulting slope was estimated to be negative. There is therefore no indication of a non-monotonic relationship between sectoral similarity and income similarity.

Higher order polynomials allow use to obtain more flexible fitted functional forms. When plotted along with the semilog specification, the cubic and fourth degree curves overlapped very closely with it (and were also similar to each other). We conclude that the use of a semilog specification approximates the underlying functional form in an appropriate fashion. Thus, for the purpose of the rest of this paper we will rely on INCDIFF as the baseline income gap measure, and on the semilog specification as the baseline functional form.[13]

[12] The estimates corresponding to the polynomial functions, as well as the fitted curves, are available upon request.

[13] Estimates using INCDIFF2 for the results presented in Sections 2 and 3 are however available upon request.

3 Robustness Analysis

In this section, we consider several robustness issues concerning the baseline results. In particular, we first investigate some geographic features of structural convergence - whether it holds for regional subsamples of the data. Secondly, we examine whether the within-pair results presented above are robust to the use of between-pair variation (which could capture longer-run phenomena), as well as other modifications of the estimation framework. Finally, we investigate whether the use of annual data might drive a spurious relationship between sectoral similarity and the income gap.

3.1 Geographic Features of Structural Convergence

We first examine whether the evidence of structural convergence presented in Section 2 is driven by specific subsets of the sample. A first split of the sample can be obtained by isolating country pairs of OECD countries. Pairs involving only OECD countries presumably entail countries of relatively similar incomes, with the result that the relationship between the income gap and sectoral similarity might be more difficult to assess. This is compounded by that fact that a fixed-effects estimator exacerbates downward bias on the slope parameter when the regressor is autocorrelated, a problem that is likely to be worse when INCDIFF exhibits smaller true variation and is of a smaller average magnitude (as is the case for the OECD subsample).

Indeed, Table 6 shows that structural convergence no longer holds in the ILO dataset when only country pairs involving OECD countries are used. INCDIFF still bears a significantly negative coefficient when considering the UNIDO dataset, suggesting evidence of structural convergence within narrowly defined manufacturing sectors within the OECD (the magnitude of the estimated coefficient is similar to that obtained in Table 4 for the full sample). Turning to country pairs involving at least one non-OECD country, we find evidence of structural convergence for both the ILO and UNIDO datasets, with estimated coefficients of similar magnitudes as for the full sample. Hence, this first sample split suggests that the finding of structural convergence is robust for the UNIDO dataset, and fragile for the ILO subsample of OECD countries.

A perhaps more interesting split of the sample would consist of a split along geographic lines. Table 7 considers sample splits requiring both countries in each pair to belong to the same region - defined as South East Asia, Latin America, Sub-Saharan Africa and Europe. Results demonstrate the presence of structural convergence almost everywhere for both the ILO and UNIDO datasets, with a particularly pronounced effect (in magnitude) in South East Asian countries, where per capita growth was high during the period under study. As expected from the OECD results presented above, structural convergence does not seem to hold for broad, economy wide sectors (ILO) in Europe. A notable feature of these results is the robustness of our finding for the UNIDO dataset.

To summarize, these simple splits of the sample leave us with several lessons. Firstly, structural convergence is particularly pronounced where structural change in general has been rapid (for example in South East Asia). Secondly, 1-digit, economy wide structural convergence does not seem to occur among rich countries. Finally, 3-digit, manufacturing sectors structural convergence seems to be almost universal - Latin America being an exception.

3.2 Estimation Issues

3.2.1 Between Variation

Turning to estimation issues, we first consider the use of some between-pair variation. In the results presented above, the use of within-pair variation was justified by the goal of characterizing the dynamic relationship between the income gap and sectoral similarity, best assessed by isolating the variation through time, within country pairs. The use of fixed-effects estimation, however, presents at least two drawbacks. Firstly, as suggested above, it exacerbates the effects of measurement error in the independent variables. This is due to the fact that, even under white noise measurement error, the error-to-truth variance ratio for an autocorrelated right-hand side variable gets larger when the variable is differenced, as is the case under fixed-effects. We would therefore expect the coefficient on INCDIFF to be biased towards zero when the within variation is isolated. Secondly, fixed-effects estimation with a time span of at most 29 years (ILO) or 34 years (UNIDO) may not be sufficient to obtain a truly long-run view of structural convergence.[14]

Table 8 presents results using the between-pair variation in the data,

[14] The same argument has often been made to justify estimating the Solow model on cross-sectional data in the empirical growth literature, even though the model's predictions refer to the within country dynamics of growth. See Mankiw, Romer and Weil (1992). See also Islam (1995) and Caselli, Esquivel and Lefort (1996) for opposing views. The cross-sectional results are interpreted as reflecting evidence of a sufficiently long-run nature, which a short time series is unable to capture.

either using a random effects estimator (which optimally weighs the within and between pair variations under the assumption that the pair-specific effects are uncorrelated with the regressor), or simply the between estimator (simple OLS on country pair means computed over time). The results are in line with expectations: the magnitude of the estimated coefficient on INCDIFF is increased compared to the benchmark results of Table 4. This is especially the case for the ILO dataset, where the absolute value of the coefficient is at least doubled. For example, in the common sample the marginal effect of INCDIFF on ILOCORR is estimated to be -0.073 under fixed effects, and -0.156 under random effects.

Hausman tests for the null hypothesis that random and fixed effects estimates do not differ significantly are also presented in Table 8. The null hypothesis is rejected for all specifications at very high levels of confidence. This provides a justification for treating fixed-effects results as a benchmark, as we have done above. However, these tests may not be particularly powerful in the presence of measurement error. Hence, we can also interpret the random effects and between estimates as providing evidence that the true extent of structural convergence was underestimated in Section 2.

To summarize, using between-pair variation results in parameter estimates that are greater in absolute value than under fixed-effects. Structural convergence is therefore a robust feature of the data even cross-sectionally, and fixed-effects estimates may understate its true extent.

3.2.2 Limited Dependent Variables

Another estimation issue has to do with the fact that the dependent variable in this study - the bilateral correlation of sector shares - is bounded below by -1 and above by 1. Firstly, this may create problems for out of sample predictions, although with the estimates obtained above and observed values of INCDIFF we never obtain predicted values of UNIDOCORR and ILOCORR that are beyond allowable bounds. Secondly, this may also result in inconsistent estimates of the parameters if a linear model is fitted to a limited dependent variable, although obviously very few observations lie at the bounds of ILOCORR and UNIDOCORR.[15] To correct for this poten-

[15] In fact, there are no country pairs in the ILO dataset with correlations equal exactly to either -1 or 1; that is, there are no countries that were ever structurally identical or diametrically different in the ILO sample. There were, however, 20 observations with UNIDOCORR=-1 and 15 with UNIDOCORR=1 (combined, these observations account for 0.03% of the UNIDO sample). This was due to the presence of some country pairs in

tial problem, Table 9 displays results for a twice-censored tobit model with random effects.[16]

The results are very similar to those obtained using the (linear) random effects estimator. Structural convergence is still observed and the magnitude of the estimated coefficient is roughly unchanged. We conclude that the results presented in Section 2 are not sensitive to an explicit consideration of the limited nature of the dependent variable in this study, and that the finding of structural convergence is not an artificial result of the use of a linear estimator.[17]

3.3 Lower Frequency Results

Lastly, we consider whether the use of annual data may have affected the results of Section 2. High frequency variations in sectoral similarity and the income gap may generate a correlation between the two variables that would vanish at lower frequencies, a problem that may be more acute under fixed-effects (within pair) estimation.[18] Another motivation for studying structural convergence at a lower, five-year interval frequency is that the results presented below in Section 4 are based on data that are only available at this frequency. Hence, we seek to establish the existence and extent of structural convergence for five year interval data as a benchmark for the estimates presented in Section 4.

Table 10 displays fixed effects estimates for the basic structural convergence equation, based on data at 5 year intervals starting in 1970 and ending in 1995 (six time periods). Compared to estimates obtained from annual data, the magnitude of the coefficient on INCDIFF is actually slightly

---

which common data for only two sectors were available (implying a bilateral correlation of sector shares equal to either -1 or 1). The results were completely unchanged when these observations were dropped.

[16] Fixed-effects tobit models are complex. As of today the parametric version of the fixed-effects tobit model has not been worked out theoretically. A semi-parametric version of the tobit model with fixed effects appears in Honoré (1992), but to my knowledge has not yet been implemented computationally. Hence, we do not present such estimates here.

[17] Twice-censored tobit - random effects estimates for other specifications presented in this paper (such as those based on INCDIFF2) are available upon request. The results there are also unchanged.

[18] However, between-pair estimates presented in Table 8 provide some preliminary indication that this in not the case - since they are the result of OLS regressions on time averages of the variables.

higher for all specifications. Hence, our last robustness check indicates that structural convergence was not an artifact of using annual data.

4 Explaining Structural Convergence: The Role of Trade

In the previous two sections, we hope to have convinced the reader that structural convergence constitutes a robust feature of the data. This new stylized fact, while it carries important implications on its own (such as those outlined in the introduction), begs for an explanation. While we probably cannot hope to provide a full account of the causes of structural convergence in a single paper, this section focuses on one set of possible explanations, based on trade-related considerations.

We start with a short conceptual discussion and then turn to empirical evidence on bilateral trade, convergence in relative factor abundance and geography. The aim is twofold. Firstly, we seek to expand the basic specification of equation (1) to improve the explained part of structural similarity. Secondly, we assess how the estimated β coefficient changes with the inclusion of additional correlates. A reduction in the magnitude of $\widehat{\beta}$ resulting from the inclusion of relative factor endowment measures in equation (1), for example, would provide evidence on the trade-based explanation for structural convergence.

4.1 Demand, Productivity and Trade

Three main non-mutually exclusive factors can contribute to the phenomenon of structural convergence. We can broadly classify them as demand explanations, productivity explanations and trade explanations.[19]

Demand. Demand-based explanations focus on the fact that the sectoral composition of demand may change in similar ways as income grows in different countries. As a very simple example isolating such an effect, consider a model where the representative consumers of two countries have identical (but non-homothetic) preferences defined over varieties of goods produced in different sectors. Suppose the countries are in autarky, and sectoral productivity is equal across sectors and across countries. Assume in addition

[19] Chenery and Syrquin (1986), chapter 3, use a similar classification to examine the determinants of structural change more broadly. In contrast, we will consider these classes of explanations as they apply to structural convergence specifically.

that the determinants of the steady-state level of income are identical across countries, and that there are diminishing returns to each of the two factors, labor and capital. If the countries start with different initial levels of capital, it is easy to see that they will converge to the same steady-state level of income through standard neoclassical income convergence. Moreover, the sectoral structure of their production will converge simply through Engel effects, since preferences are identical. Well-documented Engel effects include the relative fall of food consumption in overall expenditures as income grows, as well as the rise of health and leisure related expenditures (or services, more broadly).

A demand-side explanation for structural convergence therefore emerges as an important candidate. Moreover, it points to causality running from income convergence to structural convergence. The empirical evaluation of such an explanation would however require sectoral consumption data, which is not available at the broad, 1-digit level, and not readily available at the 3-digit level for manufacturing.[20]

Productivity. A productivity-based explanation of structural convergence would rely on the convergence of sectoral productivities across countries and sectors. Convergence in sectoral productivities can be defined as follows: define a_{ij} as the unit labor requirement of sector i in country j. Then sectoral productivity convergence between countries j and j' occurs if $a_{ij}/a_{ij'}$ approaches the same constant for all i as per capita income grows.

The following example isolates one possible mechanism whereby sectoral productivity convergence can result in structural convergence. Consider again an autarky model with two countries, and identical, Leontief preferences. Identical Leontief preferences imply that the output shares of each sector are fixed through time, and across countries.[21] In other words, we are now ruling out demand based explanations. Hence, in such an example time variation in sectoral labor shares can only result from sectorally differentiated changes in labor productivity. Several explanations can account for cross-country convergence in the schedule of labor productivities across sectors. A prime explanation would rely on technological transmissions across

[20] There is detailed international trade data for manufacturing at the sector level, which, combined with domestic output data, could be used to construct sector level consumption data for manufacturing subsectors, covering a sufficiently broad panel of countries. Moreover, the UNIDO publishes such a series at the 4-digit level. In future research we intend to use this data to evaluate the importance of demand-based explanations.

[21] See Imbs and Wacziarg (2000) for a Ricardian model with Leontief preferences. However, in contrast to the present example, this is an open economy model.

countries, which could generate sectoral productivity convergence.[22] Average (or aggregate) productivity convergence would make these countries' per capita incomes converge concurrently with convergence in the schedule of sectoral labor productivity coefficients - and mechanically the schedule of sectoral labor shares would also converge.[23]

As with demand based explanations, productivity-based explanations for structural convergence are quantifiable. Measures of value-added per worker at the sector level can be obtained at least for manufacturing subsectors, and could be used to evaluate to what extent sectoral productivity convergence contributes to structural convergence.[24]

Trade. The last main set of explanations for structural convergence relies on trade-related considerations. In classical trade models the structure of production in an open economy equilibrium is determined by the pattern of comparative advantage. If two countries' underlying pattern of comparative advantage converges, then the structure of their production can also be expected to converge.[25] Strictly speaking this is true for a two-good, two-factor Heckscher-Ohlin model of trade. In the limit of such a model if the relative abundance in capital and labor of each country become exactly identical, and the countries differ in no other way, they will produce exactly the same mix of products, and no international trade will occur in a free trade equilibrium. Moreover this process will be smooth - as relative endowments converge, so too will the structure of production (as long as the countries are not entirely specialized).

In a Ricardian model with perfectly free trade (i.e. no transportation costs), such an effect would not occur as smoothly: as relative sectoral productivities (the determinants of comparative advantage in Ricardian models)

[22] See for example Barro and Sala-i-Martin (1997). This model generates productivity convergence through technological transmissions, but in a one-sector context.

[23] The direction of causality implied by productivity convergence is not straightforward - as productivity convergence would result in both income convergence and structural convergence. This provides further justification for not interpreting the estimate of β as causal estimates, but rather as partial correlations.

[24] As it was the case for sectoral consumption data, data availability for sectoral productivity measures is more limited for the economy-wide, 1-digit sectoral classification. In future research we intend to make use of the 3-digit manufacturing data on productivity to test for this set of explanations.

[25] In what follows we will use the terms "relative factor abundance" and "relative endowments" interchangeably. We use the term "endowment" in a dynamic sense - i.e. as a state variable that can be accumulated or decumulated through time, as opposed to the static sense which is usual in pure exchange or classical trade models.

converged, countries would still remain completely specialized. They would only start producing the same mix of goods if the vectors of sectoral labor productivity parameters became linearly dependent across the two countries.[26] However, structural convergence would occur as a result of sectoral productivity convergence in a Ricardian model with trading costs. Indeed, in such a model productivity convergence, all else equal, would translate into a growing range of non-traded goods, produced in both countries.[27] Hence, the two countries would structurally converge.[28]

Other trade related considerations can have a bearing on the extent of structural convergence. Chief among those is the relationship between bilateral trade and sectoral similarity. If countries trade intersectorally, the structure of production can be expected to diverge as the volume of trade expands and countries specialize more and more. On the other hand, if the volume of trade expands mostly as a result of intraindustry trade, an expansion of trade could be positively correlated with increases in measured sectoral similarity.[29]

Changes in the extent of bilateral trade would only help explain our finding of structural convergence, i.e. the estimated magnitude of β, if the expansion of bilateral trade was somehow related to income convergence. Bergstrand and Baier (2001) show that growing income similarity (or convergence) explains virtually no part of the expansion of bilateral trade in a sample of 16 OECD countries between the late 1950s and the late 1980s. The results presented below are consistent with their finding - as the inclusion of the value of bilateral trade in the basic structural convergence equation does not modify the estimated coefficient on the income gap measure. We now turn to an empirical analysis of the relationship between trade and sectoral similarity.

[26] See Dornbusch, Fisher and Samuelson (1977) for an illustration in the context of a Ricardian model with a continuum of goods.

[27] See Imbs and Wacziarg (2000) for an illustration of this in the context of a Ricardian model with trading costs.

[28] Much in the same way as for productivity based explanations, testing for such an explanation of structural convergence would require data on sector productivity. Future research should attempt to evaluate the role of convergence in relative productivities in explaining structural convergence.

[29] See Imbs (2000) for a similar point on the relationship between trade volumes and sectoral similarity.

4.2 Structural Convergence and Bilateral Trade

We first ask whether within-pair, time variation in trade intensity may help account for structural similarity and structural convergence. Using bilateral data available at five year intervals from Rose (2001), we examine whether the inclusion of the log of the value of bilateral trade (LVALUE) in the basic structural convergence regression of Table 10 affects the coefficient on INCDIFF.

The results presented in Table 11 suggest two observations. Firstly, the coefficient estimates on INCDIFF are almost identical to those obtained when LVALUE is omitted - in all cases this is true to the third decimal. This implies that the dynamic evolution of bilateral trade intensities do not help explain structural convergence, and that income convergence is (unconditionally) unrelated to the growth of bilateral trade. The latter statement is in line with findings by Bergstrand and Baier (2001) mentioned above.

Secondly, the coefficient estimate on LVALUE itself is always negative, suggesting that a within-country pair expansion of trade is associated with less structural similarity. This can be interpreted as implying that the expansion of trade in this sample occurs mostly interindustry rather than intraindustry, although the parameter estimates on LVALUE are statistically significant at the 90% level only for the UNIDO dataset. This may not come as a surprise as the UNIDO sectoral categories are characterized by mostly tradable goods, while these are less prevalent in the ILO categories (see Appendix 1 for a list of the types of sectors included in the ILO classification).

We can conclude that the value of bilateral trade seems to have little to do with the magnitude of structural convergence as documented in Sections 2 and 3. However, bilateral trade does help in our quest for a specification explaining variations in the level of structural similarity - the negative estimated coefficient on LVALUE suggest that traditional interindustry specialization forces are at work and that countries that trade more with each other tend to look structurally more different, at least for manufacturing subsectors.[30]

[30] It should be clear to the reader that these estimates are not to be interpreted in a causal sense - they are simply conditional correlations. We do not claim to have shown that more trade "leads to" more structural dissimilarity, or that greater structural differences "lead to" more trade. Both may be true, and our estimates do not help us determine the direction of causality.

4.3 Convergence in Relative Factor Abundance

As discussed above, in a Heckscher-Ohlin framework, convergence in relative factor abundance should go hand in hand with convergence in sectoral structure. To evaluate whether this can help explain structural convergence, we construct several measures of similarity in relative factor endowments. As with the bilateral trade data above, these are computed at intervals of five years.

The first measure of similarity in relative factor abundance is denoted RELHUMS, and consists, for each country pair, of the ratio of the secondary education completion rate in the country with the lowest rate to the same measure in the country with the highest rate. This is meant to proxy for the similarity in the ratio of skilled to unskilled labor.[31] The second measure of differences in relative endowments is denoted RELKAP, and consists of the ratio of the non-residential capital stock per worker, computed by dividing the figure for the country with the lowest ratio by figure for the highest in each pair. This measure is meant to capture differences in capital-labor ratios. Thirdly, we use a measure of difference in relative land abundance (relative to labor), RELABLAND, measured by dividing the population densities of the least dense country in each pair by that of the most dense. This approximates differences in the land to labor ratio. All three measures of differences in relative factor abundance range from zero to one, with values close to one denoting greater similarity in relative endowments.

Table 12 presents results for fixed effects regressions of our baseline structural convergence equation, augmented to include the value of bilateral trade and the measures of similarity in factor abundance.[32] The table also presents F-tests for the null hypothesis that the coefficients on the factor abundance variables are jointly zero. In three of the four cases, in particular with estimates for the unrestricted samples, this null hypothesis is rejected at the 1% level. Increased similarity in capital labor ratios, captured by RELKAP, seems the most robustly significant explanatory variable among the endowment variables. Moreover, it carries the expected (positive) sign. We find

[31] We also used data on the total number of years of primary, secondary and higher education in the population aged 25 and higher, rather than secondary schooling enrollment data, to construct a ratio of relative human capital abundance. This is less appropriate theoretically, because it does not approximate the ratio of skilled to unskilled labor. The results, which are available upon request, were very similar to those using RELHUMS.

[32] The exclusion of LVALUE from these regressions did not change the estimated parameters on the other variables. Since LVALUE was found earlier to explain at least part of the changes in sectoral similarity, we keep it as a regressor in what follows.

some indication that convergence in human capital contributed to sectoral similarity for the ILO dataset, but not for the UNIDO dataset. Finally, RELABLAND does not seem to bear much of a relationship to structural similarity in either sample; this is probably due to the fact that it is a poor proxy for the differences in the land to labor ratio.[33]

A more important lesson to take from Table 12, however, is the extent to which the inclusion of variables capturing similarities in relative factor endowments affects structural convergence - i.e. the coefficient on INCDIFF. To facilitate comparisons, Table 12 also presents results for regressions which do not include measures of relative endowment similarity, but restricted to the same sample as the ones that do.[34] We find strong evidence that convergence in endowments helps account for structural convergence. In the UNIDO dataset the parameter estimate on INDCDIFF falls in absolute value from 0.0385 to 0.0246 (unrestricted sample) and from 0.0530 to 0.0371 (common sample) when the factor abundance variables are included. Hence, endowments convergence helps explain between 30% and 36.1% of structural convergence for manufacturing subsectors. In the ILO dataset, the parameter estimate on INCDIFF falls in absolute value from 0.0918 to 0.0740 (unrestricted sample) and from 0.1369 to 0.1053 (common sample). Here, endowment convergence explains between 19.4% and 23.1% of structural convergence. These results also suggest that demand and productivity based explanations probably still have a lot to contribute to our understanding of structural convergence, opening up interesting avenues for future research.

4.4 Geographic Factors: Random Effects Estimates

In the last step of our quest for an econometric specification explaining structural similarity and potentially the structural convergence coefficient β, we turn to geographic (or gravity) variables. Since these variables exhibit little or no time variation, we cannot rely on a fixed-effects estimator to estimate their incidence on structural similarity. Hence, we now turn to random effects estimates. There are several reasons for going through

[33] For example, not all of a country's surface area is usable (part of the true endowment of land) and the denominator of RELABLAND is population rather than labor.

[34] Indeed, introducing these measures of relative endowment similarity results in a fall in the number of available observations, largely due to the inclusion of RELKAP. Hence, comparing the results obtained when including the endowment measures with those in Table 11 could lead to confusing the effects of a modified sample with those of a modified specification.

this exercise. Firstly, and perhaps most importantly, geographic factors can help account for the extent of structural similarity across country pairs. Secondly, geographic features such as proximity, relative country sizes and the existence of a common border are associated with income differences, the extent of bilateral trade and probably endowments similarity as well. Hence, geographic (or gravity) variables probably belong in any estimated equation which uses between-pair variation in any proportion, to avoid omitted variables bias.[35] Finally, turning to random effects provides a check on the fixed-effects estimates which constitute most of this paper's main results.

Table 13 presents results of our structural convergence specifications using a random effects estimator.[36] To facilitate comparisons, we present random effects estimates of the specifications previously estimated using fixed-effects, in columns 1-3 and 5-7. The specifications in columns 4 and 8 include the geographic or gravity variables: indicators for the log of distance between the two countries in each pair (LDIST), a dummy variable for a common language (COMLANG), a dummy variable for a common border (BORDER) and the relative size of the countries, measured by the ratio of the population size of the smallest to the largest country in each pair (RELSIZE).[37] These are all variables thought to affect the extent of bilateral trade, and which are probably associated with both similarity in income and in relative factor abundance. Hence, they are all variables potentially omitted from the other specifications.

The results demonstrate the robustness of the fixed-effects estimates to both the use of random effects and the inclusion of the gravity variables. Most of the inferences previously derived under fixed-effects still hold: (1) The value of bilateral trade enters significantly (negatively) for the UNIDO but not the ILO dataset. (2) The inclusion of the value of bilateral trade does not affect the coefficient on INCDIFF. (3) The inclusion of relative endowment similarity measures results in a fall in the structural convergence coefficient for both the ILO and UNIDO datasets, roughly in the same proportions (one third).

[35] We should emphasize that this is not a problem when using fixed effects estimation, because gravity variables exhibit little or no time variation, and will therefore largely be accounted for by the pair-specific effects. Hence, the inclusion of geographic factors will not modify the conclusions drawn using fixed effects.

[36] The results are presented for unrestricted samples. Results for the common ILO-UNIDO sample are very similar, and available upon request.

[37] In other words, RELSIZE is decreasing in the extent of the difference in size between the two countries in a pair.

In addition to confirming these findings, Table 13 contains new results as well. Firstly, gravity variables seem to affect the level of sectoral similarity only in the UNIDO dataset. Secondly, they do so in expected ways: a longer distance between countries in a pair, and larger differences in country size imply lower structural similarity, and a common language implies greater similarity.[38] Finally, and perhaps most importantly, the geographic variables do not affect the signs, magnitude and significance of the parameter estimates for the other variables included in the regression, reinforcing our confidence in their robustness.

5 Conclusion

This paper has documented the existence and extent of structural convergence, defined as an increase in bilateral sectoral similarity through time associated with convergence in per capita income. Structural convergence is a robust feature of the data, for different levels of aggregation and data coverage. The only exception seems to occur when we restrict attention both to pairs of countries involving members of the OECD only and economy-wide 1-digit sectoral data from the ILO. This is consistent with findings in Imbs and Wacziarg (2000), who showed that rich countries seem to be in a stage of sectoral specialization, whereas most other countries are in a stage of sectoral diversification.

We proposed three explanations for structural convergence, based on convergence in demand patterns, convergence in the schedule of sectoral labor productivities, and convergence in the determinants of comparative advantage. Largely due to data limitations, which hampers the evaluation of the first two sets of explanations, we focused on a quantification of the third. We showed that changes in the intensity of bilateral trade in itself does not help explain structural convergence, but that convergence in relative factor abundance accounts for roughly a quarter to a third of its extent. These results are significant since they provide an empirical basis for viewing relative endowments as important determinants of sectoral structure. This is consistent with the Heckscher-Ohlin model of trade.

This paper should lead to interesting future research. In particular, the importance of demand and productivity convergence needs to be assessed. Unfortunately, the data is not always there, but an attempt can be

[38] The language variable is to be interpreted as another indicator of geographic closeness (a proxy for proximity).

made with whatever sectoral consumption and sectoral productivity data is available. Decomposing structural convergence into its trade, demand and productivity components appears to be the first priority of any research agenda on this topic. Evaluating the deeper causes of these three forms of convergence would be the next step.

References

Barro, Robert and Xavier Sala-i-Martin (1997), Technological Diffusion, Convergence, and Growth*, Journal of Economic Growth*, Vol. 2, No. 1.

Ben-David, Dan (1993), Equalizing Exchange: Trade Liberalization and Income Convergence, *Quarterly Journal of Economics*, vol. 108, August, pp. 653-679.

Bergstrand, Jeffrey and Scott L. Baier (2001), The Growth of World Trade: Tariffs, Transport Costs, and Income Similarity, *Journal of International Economics*, Vol. 53, No. 1, February, 1-27.

Caselli, Francesco, Gerardo Esquivel and Fernando Lefort (1996), *Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics*, Journal of Economic Growth, vol. 1, September.

Chenery, Hollis, Sherman Robinson and Moshe Syrquin (1986), *Industrialization and Growth: A Comparative Study*, Oxford: Oxford University Press for the World Bank

Honoré, Bo (1992), Trimmed LAD and Least Squares Estimation of Truncated and Censored Regression Models with Fixed Effects, *Econometrica*, vol. 60, pp. 553-565.

International Labour Office (1997), Yearbook of Labor Statistics, Geneva: ILO.

Imbs, Jean and Romain Wacziarg (2000), Stages of Diversification, *Stanford University GSB Working Paper* #1653.

Imbs, Jean (1999), Co-Fluctuations, *CEPR Discussion Paper* #2267, October.

Imbs, Jean (2000), Sectors and the OECD Business Cycle, *CEPR Discussion Paper* #2473, June.

Islam, Nazrul (1995), Growth Empirics: A Panel Data Approach, *Quarterly Journal of Economics*, Vol. CX, November 1995.

Kim, Sukkoo (1995), Expansion of Markets and the Geographic Distribution of Economic Activities: The Trends in U.S. Regional Manufacturing Structure, 1860-1987*, Quarterly Journal of Economics*,vol. 110, p. 881-908.

Krugman, Paul (1991), *Geography and Trade*, MIT Press.

Mankiw, N. Gregory, David Romer and David N. Weil (1992), A Contribution to the Empirics of Economic Growth, *Quarterly Journal of Economics*, 107-2, May, p. 407-437.

Rose, Andrew (2001), Bilateral Trade Dataset, on the following website: http://haas.berkeley.edu/~arose/frbilat.zip.

Seddon, Jessica and Romain Wacziarg (2000), Trade Liberalization and Intersectoral Labor Movements, *Stanford Graduate School of Business Research Paper* #1652 , September.

United Nations Industrial Development Organization (1997), *UNIDO Industrial Statistics Database, 3-Digit level of ISIC Code*, Vienna: UNIDO.

Ventura, Jaume (1997), Growth and Interdependence, *Quarterly Journal of Economics*, February, vol 112, pp. 57-84.

Data Appendix

A. Sectoral Coverage

1. ILO 1-Digit Classification (9 sectors)

1. Agriculture, Hunting, Forestry and Fishing
2. Mining and Quarrying
3. Manufacturing
4. Electricity, Gas and Water
5. Construction
6. Wholesale and Retail Trade and Restaurants and Hotels
7. Transport, Storage and Communication
8. Financing, Insurance, Real Estate and Business Services
9. Community, Social and Personal Services

2. UNIDO 3-Digit Classification (28 sectors)

311 Food products
313 Beverages
314 Tobacco
321 Textiles
322 Wearing apparel, exc. footwear
323 Leather products
324 Footwear, exc. rubber or plastic
331 Wood products, exc. furniture
332 Furniture, exc. metal
341 Paper and products
342 Printing and publishing
351 Industrial chemicals
352 Other chemicals
353 Petroleum refineries
354 Misc. petrol. and coal prods
355 Rubber products
356 Plastic products
361 Pottery, china, earthenware
362 Glass and products
369 Other non-metallic mineral prods
371 Iron and steel
372 Non-ferrous metals
381 Fabricated metal products
382 Machinery, except electrical
383 Machinery, electric
384 Transport equipment
385 Professional and scientific eqpt
390 Other manufactured products

B. Country Coverage

Algeria, b
Angola, b
Argentina
Australia
Austria
Azerbaijan, a
Bahamas, a
Bangladesh
Barbados
Belarus, a
Belgium
Belize, b
Benin, b
Bolivia
Botswana, b
Brazil
Bulgaria, b
Burkina Faso, b
Burundi, b
Cameroon, b
Canada
Cape Verde, b
Central Afr. Rep, b
Chile
China
Colombia
Comoros, b
Costa Rica
Cote d'Ivoire, b
Cyprus
Czechoslovakia
Denmark
Dominican Republic
Ecuador
Egypt
El Salvador
Estonia, a
Ethiopia, b
Fiji, b
Finland
France
Gabon, b
German Dem Rep, b
Germany, Fed Rep
Ghana, b
Greece
Guatemala, b
Guyana, b
Haiti, b
Honduras
Hong Kong
Hungary
Iceland
India, b
Indonesia
Iran, b
Iraq, b
Ireland
Israel
Italy
Jamaica
Japan
Jordan, b
Kenya, b
Korea
Kuwait, b
Kyrgyzstan, a
Latvia, a
Lesotho, b
Liberia, b
Luxembourg
Macao, a
Madagascar, b
Malawi, b
Malaysia
Mali, b
Malta, b
Mauritius
Mexico
Moldova, a
Morocco
Mozambique, b
Myanmar
Nepal, b
Netherlands
Neth. Antilles, a
New Zealand
Nicaragua, b
Niger, b
Nigeria, b
Norway
Oman, b
Pakistan
Panama
Papua N. Guin., b
Paraguay
Peru
Philippines
Poland
Portugal
Puerto Rico
Romania
Russian Fed, a
Rwanda, b
San Marino, a
Saudi Arabia, b
Senegal, b
Seychelles
Sierra Leone, b
Singapore
Slovakia, a
Slovenia, a
Somalia, b
South Africa, b
Spain
Sri Lanka
Suriname
Swaziland, b
Sweden
Switzerland
Syria
Taiwan, b
Tanzania, b
Thailand
The Gambia, b
Togo, b
Trinidad & Tob
Tunisia
Turkey
U.S.A
U.S.S.R., b
Uganda, b
United Kingdom
Uruguay
Venezuela
Western Samoa, b
Yemen, Arab Rep, b
Yugoslavia, b
Zaire, b
Zambia, b
Zimbabwe, b

a: not in UNIDO dataset, b: Not in ILO dataset

C. Sources and Definitions of the Variables

ILOCORR: Bilateral correlation of sector shares in a given year, ILO 1 Digit economy-wide data (9 sectors). Source: ILO sectoral employment data.

UNIDOCORR: Bilateral correlation of sector shares in a given year, UNIDO 3 Digit manufacturing data (28 sectors). Source: UNIDO sectoral manufacturing employment data

INCDIFF: Absolute value of the difference in log incomes of country 1 and 2 (i.e. the log of the ratio of incomes of the richest to the poorest country). Source: Summers and Heston v. 5.6

INCDIFF2: Ratio of incomes of the richest to the poorest country for a country pair. Source: Summers and Heston v. 5.6

LVALUE: Log of bilateral trade value. Source: Rose (originally United Nations Statistical Office).

RELHUMS: Ratio of the secondary schooling completion rate in the population aged 25 and higher, computed by dividing the figure for the country with the lowest rate by the figure for the country with the highest rate in each pair (captures the relative skilled-to-unskilled labor ratio). Source: Barro-Lee dataset.

RELKAP: Ratio of the non-residential capital stock per worker at 1985 international prices, computed by dividing the figure for the country with the lowest physical capital by the figure for the highest in each pair (captures the relative capital-labor ratio). Source: Summers-Heston, v.5.6

RELABLAND: Ratio of population densities, computed by dividing the figure of the least dense country by that of the most dense in each pair (captures the relative labor-land ratio). Source: Barro-Lee, completed using the CIA World Factbook for missing values in the land area data.

LDIST: Log of distance between countries 1 and 2. Source: Rose.

COMLANG: Dummy variable for a common language in the country pair. Source: Rose.

BORDER: Dummy variable for a common border in the country pair. Source: Rose.

RELSIZE: Ratio of population sizes, smallest to largest country in each pair. Source: Summers and Heston v. 5.6.

Table 1 - Summary statistics

Variable	# of Obs	Mean	Std. Dev.	Min	Max
unidocorr	128742	0.488	0.282	-1	1
ilocorr	31207	0.616	0.329	-0.845	0.9997
incdiff	117568	1.100	0.786	0	4.130
incdiff2	117568	4.291	4.612	1	62.168

Table 2 - Pairwise correlation matrix for the main variables

	ILOCORR	UNIDOCORR	INCDIFF
UNIDOCORR	0.385	1.000	-
	(16996)	(128742)	
INCDIFF	-0.613	-0.248	1.000
	(17621)	(114313)	(117568)
INCDIFF2	-0.569	-0.233	0.883
	(17621)	(114313)	(117568)

All displayed correlations are significant at the 99.9% level.
# of observations used to compute the correlations in parentheses.

Table 3 - Evolution of sectoral similarity through time – Fixed effects estimates.

	Unrestricted Sample		Common Sample	
	ILO	UNIDO	ILO	UNIDO
YEAR	0.0024	-0.0008	0.0026	0.0013
	(20.66)	(-20.84)	(17.37)	(13.06)
# of Obs.	31207	128742	16996	16996
(# of pairs)	(3218)	(7845)	(1892)	(1892)
R-Squared	0.009	0.002	0.009	0.004

(t-statistics in parentheses)

Table 4 – Fixed effects regressions of sectoral similarity on the absolute value of log income differences (INCDIFF)

	Unrestricted Sample		Common Sample	
	ILO	**UNIDO**	**ILO**	**UNIDO**
INCDIFF	-0.0141	-0.0462	-0.0725	-0.0400
	(-2.19)	(-30.74)	(-11.29)	(-9.81)
# of Obs.	17621	114313	14366	14366
(# of pairs)	(1940)	(7644)	(1701)	(1701)
R-Squared	0.375	0.062	0.470	0.126

(t-statistics in parentheses)

Table 5 – Fixed effects regressions of sectoral similarity on the ratio of richest to poorest incomes (INCDIFF2)

	Unrestricted Sample		Common Sample	
	ILO	**UNIDO**	**ILO**	**UNIDO**
INCDIFF2	-0.0034	-0.0096	-0.0177	-0.0103
	(-2.45)	(-37.28)	(-11.27)	(-10.35)
# of Obs	17621	114313	14366	14366
(# of pairs)	(1940)	(7644)	(1701)	(1701)
R-Squared	0.324	0.054	0.451	0.116

(t-statistics in parentheses)

Table 6 - Analysis for OECD and non-OECD countries – Fixed effects estimator

	OECD		Non-OECD	
	ILO	**UNIDO**	**ILO**	**UNIDO**
INCDIFF	0.0055	-0.0486	-0.0147	-0.0462
	(0.18)	(-7.22)	(-2.19)	(-29.76)
# of Obs.	2918	7244	14703	107069
(# of pairs)	(270)	(276)	(1670)	(7368)
R-Squared	0.586	0.051	0.304	0.055

(t-statistics in parentheses)

Table 7 - Regional analysis – Fixed effects estimator

	South East Asia		Latin America		Sub-Saharan Africa		Europe	
	ILO	**UNIDO**	**ILO**	**UNIDO**	**ILO***	**UNIDO**	**ILO**	**UNIDO**
INCDIFF	-0.2587	-0.1262	-0.1256	-0.0012	-	-0.1508	0.0323	-0.0619
	(-7.66)	(-8.51)	(-3.65)	(-0.14)		(-14.28)	(1.13)	(-12.13)
# of Obs.	277	968	1179	4924	-	4784	1992	8178
(# of pairs)	(21)	(45)	(148)	(310)	-	(588)	(218)	(351)
R-Squared	0.683	0.167	0.077	0.093	-	0.044	0.699	0.025

(t-statistics in parentheses)
* No data available

Table 8- Regressions of sectoral similarity on INCDIFF – Between and random effects results

	Unrestricted Sample		Common Sample	
	ILO	**UNIDO**	**ILO**	**UNIDO**
	Random Effects			
INCDIFF	-0.1013	-0.0518	-0.1558	-0.0647
	(-19.41)	(-36.63)	(-29.68)	(-17.25)
# of Obs.	17621	114313	14366	14366
(# of pairs)	(1940)	(7644)	(1701)	(1701)
R-Squared	0.375	0.062	0.479	0.126
Hausman Chi^2*	544.93	119.21	505.170	238.270
(p-value)	(0.00)	(0.00)	(0.00)	(0.00)
	Between Estimator			
INCDIFF	-0.2559	-0.0950	-0.3062	-0.1890
	(-29.66)	(-22.88)	(-35.39)	(-20.78)
# of Obs.	17621	114313	14366	14366
(# of pairs)	(1940)	(7644)	(1701)	(1701)
R-Squared	0.375	0.062	0.479	0.126

(t-statistics in parentheses)
* Test of the null that fixed and random effects estimates of this specification do not differ systematically.

Table 9 – Twice-censored tobit – random effects regressions of sectoral similarity on the absolute value of log income differences (INCDIFF)

	Unrestricted Sample		Common Sample	
	ILO	UNIDO	ILO	UNIDO
INCDIFF	-0.0920 .	-0.0645 (-67.82)	-0.1345 (-39.51)	-0.1118 .
# of Obs* (# of pairs)	17621 (0, 0) (1940)	114313 (20, 15) (7644)	14366 (0, 0) (1701)	14366 (0, 0) (1701)
Log likelihood	8139.838	78980.800	8052.966	13481.681

* Number of censored data points (left, right) in parentheses
(t-statistics in parentheses – not reported by STATA for some regressions)

Table 10 – Fixed effects regressions of sectoral similarity on the absolute value of log income differences (INCDIFF) – 5-year interval data

	Unrestricted Sample		Common Sample	
	ILO	UNIDO	ILO	UNIDO
INCDIFF	-0.0544 (-4.73)	-0.0531 (-12.12)	-0.1050 (-7.01)	-0.0522 (-5.16)
# of Obs. (# of pairs)	4077 (1620)	15487 (4931)	2852 (1211)	2852 (1211)
R-Squared	0.373	0.089	0.431	0.209

(t-statistics in parentheses)

Table 11 – Structural Convergence and Bilateral Trade Intensity
Fixed effects estimates – 5-year interval data

	Unrestricted Sample		Common Sample	
	ILO	UNIDO	ILO	UNIDO
INCDIFF	-0.0563 (-4.87)	-0.0552 (-12.63)	-0.1072 (-7.09)	-0.0577 (-5.67)
LVALUE	-0.0044 (-1.58)	-0.0071 (-8.85)	-0.0035 (-1.04)	-0.0088 (-3.89)
# of Obs. (# of pairs)	4077 (1620)	15487 (4931)	2852 (1211)	2852 (1211)
R-Squared	0.306	0.063	0.399	0.079

(t-statistics in parentheses)

Table 12 – Structural Convergence and Factor Endowments –
Fixed effects estimates – 5-year interval data

	Unrestricted Sample				Common Sample			
	ILO	**ILO**	**UNIDO**	**UNIDO**	**ILO**	**ILO**	**UNIDO**	**UNIDO**
INCDIFF	-0.0918	-0.0740	-0.0385	-0.0246	-0.1369	-0.1053	-0.0530	-0.0371
	(-3.94)	(-2.75)	(-5.43)	(-3.24)	(-5.39)	(-3.62)	(-3.82)	(-2.33)
LVALUE	-0.0083	-0.0086	-0.0040	-0.0043	-0.0118	-0.0123	-0.0039	-0.0041
	(-1.69)	(-1.76)	(-2.93)	(-3.17)	(-2.07)	(-2.15)	(-1.24)	(-1.32)
RELHUMS	-	0.0565	-	-0.0198	-	0.0503	-	-0.0024
		(2.54)		(-2.41)		(2.16)		(-0.19)
RELKAP	-	0.0832	-	0.0826	-	0.1356	-	0.0662
		(1.51)		(4.62)		(2.24)		(2.00)
RELABLAND	-	-0.1427	-	-0.0313	-	-0.1391	-	-0.0364
		(-1.22)		(-0.95)		(-1.11)		(-0.53)
# of Obs.	1558	1558	4909	4909	1362	1362	1362	1362
(# of pairs)	(603)	(603)	(1284)	(1284)	(557)	(557)	(557)	(557)
F test (endowments)*	-	3.30	-	9.48	-	3.64	-	1.43
(p-value)		(0.02)		(0.00)		(0.01)		(0.23)
R-Squared	0.232	0.181	0.129	0.117	0.297	0.254	0.269	0.239

(t-statistics in parentheses)
* Test of the null that the endowment variables RELHUMS, RELKAP and RELABLAND are jointly equal to zero.

Table 13 – Structural Convergence, Geography and Trade – Random effects estimates – Unrestricted sample – 5-year interval data

	ILO				UNIDO			
	1	2	3	4	5	6	7	8
INCDIFF	-0.1926 (-25.63)	-0.1922 (-25.50)	-0.1436 (-8.19)	-0.1406 (-7.91)	-0.0713 (-22.66)	-0.0721 (-22.94)	-0.0524 (-8.38)	-0.0466 (-7.49)
LVALUE	-	0.0019 (1.15)	0.0012 (0.40)	0.0008 (0.25)		-0.0051 (-7.89)	-0.0055 (-4.90)	-0.0074 (-6.56)
RELHUMS	-	-	0.0754 (3.76)	0.0732 (3.64)	-	-	-0.0125 (-1.61)	-0.0148 (-1.90)
RELKAP	-	-	0.1548 (3.88)	0.1591 (3.97)	-	-	0.1294 (8.33)	0.1321 (8.61)
RELABLAND	-	-	0.0273 (0.74)	0.0179 (0.46)	-	-	0.0796 (4.35)	0.0410 (2.25)
LDIST	-	-	-	-0.0080 (-0.55)	-	-	-	-0.0508 (-6.43)
COMLANG	-	-	-	0.0258 (0.77)	-	-	-	0.0923 (5.57)
BORDER	-	-	-	-0.0031 (-0.04)	-	-	-	0.0335 (0.91)
RELSIZE	-	-	-	0.0368 (1.00)	-	-	-	0.0977 (5.75)
# of Obs. (# of pairs)	4077 (1620)	4077 (1620)	1558 (603)	1558 (603)	15487 (4931)	15487 (4931)	4909 (1284)	4909 (1284)
R-Squared	0.373	0.375	0.343	0.340	0.089	0.078	0.177	0.233
Hausman Chi^2* (p-value)	252.52 (0.00)	263.22 (0.00)	40.29 (0.00)	42.37 (0.00)	35.9 (0.00)	53.05 (0.00)	79.09 (0.00)	71.2 (0.00)
Chi^2 (Gravity)** P value	-	-	30.52 (0.00)	1.98 (0.74)	-	-	88.35 (0.00)	137.42 (0.00)

(t-statistics in parentheses)

* Test of the null that fixed and random effects estimates of this specification do not differ systematically.

** Test of the null that the gravity variables LDIST, COMLANG, BORDER and RELSIZE are jointly equal to zero.

Figure 1 - Predicted Values of UNIDOCORR and ILOCORR, Fixed-effects, Common Sample