UNITED NATIONS CONFERENCE ON TRADE AND DEVELOPMENT

POLICY ISSUES IN INTERNATIONAL TRADE AND COMMODITIES
STUDY SERIES No. 44

REVEALED FACTOR INTENSITY INDICES AT THE PRODUCT LEVEL

by

Miho Shirotori
UNCTAD

Bolormaa Tumurchudur
UNCTAD

Olivier Cadot
University of Lausanne,
CEPR, CEPREMAP and CERDI

UNITED NATIONS
New York and Geneva, 2010

NOTE

The purpose of this series of studies is to analyse policy issues and to stimulate discussions in the area of international trade and development. The series includes studies by UNCTAD staff, as well as by distinguished researchers from academia. In keeping with the objective of the series, authors are encouraged to express their own views, which do not necessarily reflect the views of the UNCTAD secretariat or its member States.

The designations employed and the presentation of the material do not imply the expression of any opinion whatsoever on the part of the United Nations Secretariat concerning the legal status of any country, territory, city or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Material in this publication may be freely quoted or reprinted, but acknowledgement is requested, together with a reference to the document number. It would be appreciated if a copy of the publication containing the quotation or reprint were sent to the UNCTAD secretariat at the following address:

Chief
Trade Analysis Branch
Division on International Trade in Goods and Services, and Commodities
United Nations Conference on Trade and Development
Palais des Nations
CH-1211 Geneva
Switzerland

Series editor:
Khalilur Rahman
Chief, Trade Analysis Branch

UNCTAD/ITCD/TAB/46

UNITED NATIONS PUBLICATION

ISSN 1607-8291

Abstract

This paper describes a data-intensive methodology to generate indices that indicate "revealed" factor intensity of traded goods, at the most disaggregated level of product classification (SITC 5-digit or HS 6-digit). We construct the indices by calculating, for each good, a weighted average of the factor abundance of the countries that export this good, where the weights are variants of Balassa's Revealed Comparative Advantage index. In doing so, we take advantage of recent improvements in the availability of data on aggregate national endowments of primary factors (capital, education and labour force) using, inter alia, Summer and Heston's PWT (version 6.2), Barro-Lee's latest database, the World Bank and COMTRADE databases. The resulting indices are available on the UNCTAD website.

Key words: International trade, factor endowments, factor intensities, product classifications

JEL Classification: F1

Acknowledgements

The authors are grateful to Alessandro Nicita, Marco Fugazza, Sudip Ranjan Basu and Joerg Mayer for their valuable suggestions and comments.

Any mistakes or errors remain the authors' own.

Contents

List of figures

List of tables

Appendix tables:

1. Introduction

The process of export diversification has long been a major research issue in international economics. In recent years, we have seen a renewed interest in the nature and the process of export diversification.

For instance, Klinger and Lederman (2005) investigate the role of innovation in export diversification. They find that off-the-frontier innovation (e.g. the introduction of new export products) is more common among low-income countries than among high-income countries. Consistent with the U-shape pattern of the relationship between export diversification and national income, as described in Imbs and Wacziarg (2003), overall export diversification increases at low levels of development but declines with development after a middle-income point.

Hausmann and Klinger (2006) suggest that changes over time in the revealed comparative advantage of countries are associated with the pattern of relatedness (measured in terms of productive factors) determined across products. As countries change their export mix, there is a strong tendency to move towards related goods rather than to goods that are less related. They also suggest that a particular product's proximity to existing areas of comparative advantage is one of the most significant determinants of whether a country will develop an advantage in that product in the future.

These approaches provide fresh thinking into old issues and that is very welcome. However, the debate remains largely disconnected from traditional theory of factor-content of trade considerations, namely the Hecksher-Ohlin model of comparative advantage based on relative factor endowments.[1] A potential danger of such approaches is − as usual with inductive reasoning − that one may end up inferring supposedly general laws from statistical relationships that may or may not hold out of sample, and, ultimately, advising policy on the basis of these empirical findings.

This study aims to provide a tool to fill the gap between traditional, theory-based approaches and newer eclectic ones by developing a time-series database of the indices of revealed factor intensity (RFI) of export products, using a wealth of raw data accumulated in the last two decades.

The indices are constructed as follows. First, we collected (and updated whenever necessary) raw data on national factor endowments of physical capital, human capital and natural resource endowment for countries of which data was available. Combining these data gives a panel database of factor endowments at the country level, spanning close to 100 countries over three decades (our balanced panel covers 92 countries from 1971 to 2003).

Second, we calculated a "revealed" factor intensity for each product at a disaggregated level of product classification (we used both United Nations Standard International Trade Classification (SITC) 5-digit and Harmonized System (HS) 6-digit), using the factor endowments of countries exporting that product (from step 1). The idea is that a product exported predominantly by countries that are richly endowed with human capital is "revealed" to be intensive in human capital. To weigh national factor endowments in the averaging, we adopted a methodology first proposed by Hausmann, Hwang and Rodrik (2007), using as weights a slightly modified version of revealed comparative advantage (RCA). Using RCAs, instead of simple export weights, filters out scale effects.

[1] See Feenstra (2004), chapter 2 for a survey of that older literature.

Being a weighted average of factor endowments, our measure is sensitive to the country coverage of the endowments database. However, there is a trade-off between the one with a large sample size and the one which is smaller in size but without any missing values. Therefore, we propose two versions of our revealed factor intensities: (a) a "wide" one, based on the widest annual country coverage; and (b) a "consistent" one, based on a balanced panel of data.

We were also careful to weed out, as much as we could, the effect of subsidies and other trade distortions. Because these distortions are prevalent in agriculture, we used the World Bank's new Agricultural Distortions database (Anderson *et al.*, 2008) and eliminated observations where RCAs were obviously driven by policy. Without this correction, we would have high "revealed" human capital intensities for agricultural goods whose exports are subsidized by rich countries.

The resulting RFI indices are presented and analyzed in various ways in the paper. We believe the value added of the RFI indices is that it will make possible to control for Heckscher–Ohlin effects in analysis of trade diversification in a way that was not possible before.

The outline of the paper is as follows. Chapter 2 provides a detailed description of the construction of national factor endowments. Chapter 3 provides a description of the construction of the index of revealed factor intensity, and discusses caveats. It also uses cluster analysis to explore broad groupings of products on the basis of their revealed factor intensities. The explanation of the database of the indices of revealed factor intensity is attached to the paper. The database of the indices is accessible and can be downloaded from UNCTAD website (http://r0.unctad.org/ditc/tab/index.shtm).

2. National factor endowments: physical capital, human capital and natural resource endowment

2.1 Measuring national factor endowments

2.1.1 Capital stock

a. Construction of capital stock

This section describes the derivation of our database of aggregate (national) capital stock estimates. In general, two methods are available: (a) direct measurement through surveys and (b) perpetual inventory method (PIM). Because direct measures are not everywhere available, we use the PIM.

The PIM reconstructs capital stock estimates from investment flows by adding up, recursively, current investment to the previous period's capital stock, appropriately depreciated. The method raises (inter alia) two problems. One is the initial estimate of the capital stock, the other is the choice of the depreciation rate. We have followed the approach of Easterly and Levine (2001, henceforth EL)[2] and replicated their capital stock estimates using the updated version 6.2 of the Penn World Table (PWT) which provides aggregate investment figures[3] for 159 countries.

Let K_t and I_t be respectively the real capital stock and investment flow of country *i* in period *t*. The capital-accumulation equation is

$$K_{t+1} = (1-\delta)K_{i,t} + I_{it},$$

where δ is the rate of depreciation. Following EL, we assume that country i is at its steady-state capital-output ratio, which implies that $dK_t / K_t = dY_t / Y_t$. Since $dK_t = I_t - \delta K_t$, then $dK_t / K_t = I_t / K_t - \delta$. At the steady-state growth rate, be $g_i^* = dY_i^* / Y_i^* = I_i^* / K_i^* - \delta$, we can write

$$g_i^* = \frac{I_i^*}{Y_i^*}\frac{Y_i^*}{K_i^*} - \delta = \frac{\iota_i^*}{\kappa_i^*} - \delta \qquad (1)$$

[2] They used the PWT 5.6 capital stock data, based on disaggregated investment and depreciation statistics for 64 countries. They also constructed capital stock figures for more countries using aggregate investment figures.

[3] It would have been desirable to use disaggregated investment series (especially our interest is having the series for non-residential investment), but the PWT only provides with the aggregate investment series. Though the version 5.6 provides capital stock for non-residential, it covers much less countries and periods.

where ι_i^* is the investment rate and κ_i^* is the capital-output ratio. The latter can thus be written as

$$\kappa_i^* = \frac{\iota_i^*}{g_i^* + \delta} \tag{2}$$

Following EL, we construct g_i^* – the steady-state growth rate – as a weighted average of the country's average growth rate during the first 10 years for which the PWT have output and investment data and the world growth rate. That is,

$$g_i^* = \lambda \bar{g}_i + (1-\lambda)\bar{g}^W \tag{3}$$

where bars represent values averaged over the sample's first 10 years. The world growth rate is computed as 0.0423. Following Easterly et al. (1993), we set λ at 0.25. We compute ι_i^* similarly as the average investment rate during the first 10 years for which there is data. Finally, we get an estimate of the initial capital stock

$$K_0 = \kappa_i^* \tilde{Y}_0 \tag{4}$$

where $\tilde{Y}_0$ is the average real output value between 1950 and 1952 rather than simply the first observation in the sample period in order to reduce the influence of business-cycles. For countries where output and investment data do not start until 1960, everything is moved down one decade. As for δ, we again follow EL in assuming a depreciation rate of 7 per cent.

b. Description of capital stock and its reliability

Table 1a shows the coverage of our estimates of the real capital stock. We cover 159 countries, 154 of which have more than 30 years of time series. In order to construct a series for the real capital stock per worker (*K/L*), we used an indirect approach using real gross domestic product (GDP) per worker (*Y/L*), GDP per capita (*Y/P*) and population (*P*) from the PWT to infer the numbers of workers (*L*), all from the PWT.

Because the PWT has missing data for GDP per worker, the coverage was further reduced to 152 countries, 140 of which have 30 years of time series or more (see table 1b), and 136 of which have data over the common sample period 1971–2003.

Table 1. Real capital stock and real capital stock per worker estimates

a.

Time period	Number of Countries	Start of time series
14	2	1990
17	1	1988
23	1	1981
24	1	1980
30	1	1971
31	1	1973
33	38	1971
34	15	1971
43	19	1961, 62
44	11	1961, 60
45	2	1960
48	4	1956
49	2	1955
50	2	1954, 55
51	1	1954
52	3	1952, 53
53	20	1951, 52
54	35	1951
	159	

b.

Time period	Number of Countries	Start of time series
1	2	1996
14	2	1990
16	1	1988
21	1	1977
24	3	1971
24	3	1977
30	2	1971
31	1	1973
33	38	1971
42	1	1962
43	28	1961
44	3	1960
47	1	1952
48	4	1956
49	3	1955
50	2	1954
51	1	1953
52	5	1952
53	51	1951
	152	

Table 2 shows summary statistics for the 136 countries for which we have data for the common time period 1971–2003.

Table 2. Summary statistics (in 2000 United States dollars)

Variable	Obs	Mean	Std. Dev.	Min	Max
Real GDP per capita	136	8 061	8 309	521	42 419
Real Capital Stock per worker	136	31 307	37 116	376	160 177

Table 3 shows the growth of our estimates of the real capital stock per worker by income group over 1971–2003. Low-income countries have had the slowest growth, with negative growth in the 1980s and the 1990s.

Table 3. Growth rate of real capital stock per worker, by income group (annual average percentage change at 2000 US dollar)

Income group	Growth rate of Real Capital stock per worker (per cent)			
	1971-1981	1982-1992	1993-2003	1971-2003
High income: OECD	2.8	1.6	2.2	2.2
High income: nonOECD	2.0	1.0	2.0	1.6
Upper middle income	2.4	-0.4	1.7	1.2
Lower middle income	3.8	0.4	0.2	1.4
Low income	1.9	-0.6	-0.3	0.3

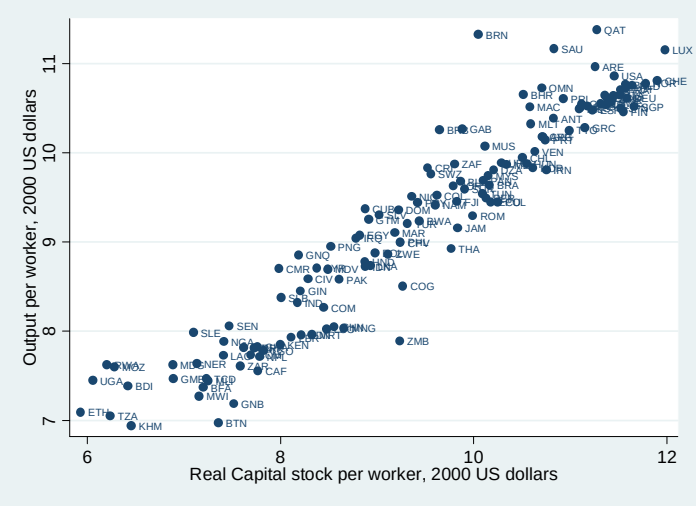
Table 4 decomposes this rate of growth in terms of income groups and regional breakdown. The general trend is a steady decline in rates of capital accumulation per worker, with some recovery in the decade between 1993 and 2003. It is notable that the growth rate of capital stock varies considerably across developing regions, as well as across periods within a developing region.

Table 4. Growth rate of real capital stock per worker, by region

	Growth rate of Real Capital stock per worker (per cent)			
	1971-1981	1982-1992	1993-2003	1971-2003
High income countries	2.6	1.4	2.1	2.0
Low and middle income				
East Asia & Pacific	4.3	3.0	2.8	3.4
East Asia & Pacific (without China)	2.6	1.4	2.1	3.2
Europe & Central Asia	6.3	1.5	0.7	2.7
Latin America & Caribbean	2.0	-1.1	1.2	0.6
Middle East & North Africa	6.0	0.1	-2.2	1.2
South Asia	3.9	3.6	3.2	3.6
Sub-Saharan Africa	1.8	-0.2	0.8	0.7

In order to check the plausibility of our estimates, we plot the sample-period average, per country, of real GDP per worker against the real capital stock per worker, both in logs (figure 1). The real GDP per-worker is a proxy to aggregate labour productivity of the country. Thus, they should be correlated if our estimates are reasonable.

Figure 1. Real GDP per worker vs. real capital stock per worker (in logs)



The scatter plot shows that they are indeed highly correlated.[4] Table 5 presents the data aggregated by income group, which again shows a plausible degree of correlation.

Table 5. Real GDP per worker vs. real capital stock per worker, by income group (in 2000 United States dollars)

Income Group	Real GDP per worker	Real Capital Stock per worker
High income: OECD	39 934	92 454
High income: nonOECD	42 842	53 368
Low income	2 715	2 470
Lower middle income	9 121	11 738
Upper middle income	16 958	23 470

[4] We have done the same figures for three different periods (1971–1981 1982–1992 and 1993–2003) to see whether the correlation was maintained over the periods. The positive correlation was observed.

As a further check, table 6 shows the correlation between our series and alternative estimates. It can be seen that the degrees of correlation are all above 90 per cent.

Table 6. Correlation of alternative capital series with our new real capital stock per worker series

	Our estimates of Real Capital Stock per worker	Replication of Klenow-Rodriguez Clare	Nehru-Dhareshwar	Larson et al.
Our estimates of Real Capital Stock per worker[a]	1			
Replication of Klenow-Rodriguez Clare	0.9854 (3 856)	1		
Nehru-Dhareshwar	0.9307 (2 382)	0.9153 (1 896)	1	
Larson *et al.*	0.7878 (1 411)	0.9381 (1 323)	0.9580 (1 138)	1

Note:
[a] As we described in the text, this is our update of Easterly and Levine (2001).
Larson *et al.* (2000) covers 62 industrial and developing countries for the years 1967–92.
Nehru–Dhareshwar (1993) covers 92 industrial and developing countries from 1960–1990.
In brackets are the number of observations.

2.1.2 Human capital stock

a. Measures for human capital stock

There are various types of proxies that have been used for measuring human capital. These include literacy rates, school enrolment ratios, educational attainment and average years of schooling. Among those, the last one – average years of schooling – is the most popular, partly because of the availability of large datasets in terms of country coverage and the length of period for which data is available.

There are several data sets on educational attainment. The available datasets can be divided into two groups depending on whether they make use of (a) census/survey data, which are the only direct numbers available together with school enrolment ratio; or (b) only the school enrolment ratio.

The first group (Kyriacou 1991[5]; and Barro and Lee, 1993, 2001) relies on census numbers whenever those are available, and fills in missing values using a regression of average years of schooling on lagged enrolment rates. However, this procedure is valid only when the relationship between these two variables is stable over time and across countries, which is not often the case. As an alternative, Barro and Lee use an accuracy test based on a sample of 30 countries with relatively

[5] Kyriacou (1991) estimated the average years of schooling of the labor force for a sample of 111 countries for the period of 1965-1985 at five-year intervals. He uses UNESCO census data and Psacharopoulos and Arriagada (1986) attainment figures to estimate average schooling years on school enrollment ratios. Psacharopoulos and Arriagada (1986) reports data on educational composition of the labor force in 99 countries and provides estimates of average years of schooling. The main drawback is that they provide only one time-series observation in most countries.

complete census numbers in order to fill in missing values.[6] As such, Barro and Lee's data may be more robust than Kyriacou's, although this is largely a matter of judgement.

The second group (Lau, Jamison and Louat, 1991; Lau, Bhalla and Louat 1991; and Nehru, Swanson and Dubey, 1995) uses only school enrolment ratios to construct human capital stock series.[7][8] Their PIM is a sophisticated version of Barro and Lee, but they ignored census data on educational attainment.

Based on Krueger and Lindahl's (2001) estimates of the reliability of the Barro and Lee and Kyriacou datasets, we chose to use Barro and Lee's data, although there are arguments in favour of both. The latest version of the dataset, described in Barro and Lee (2001), incorporates various improvements in the procedure used to fill in missing values.[9] De la Fuente and Doménech (2001) and Cohen and Soto (2000) provide useful indications on how to clean up the available census/survey data.[10] [11]

Barro and Lee estimated two sets of educational attainment rates at five-year intervals from 1960 for different levels of education for overall populations aged over 15 and over 25

---

[6] Barro and Lee use a PIM that starts with the survey numbers as benchmark stocks, and then use the school enrolment ratios to estimate the changes from the benchmarks. This method is vulnerable to inaccuracies in the underlying data on gross enrolment ratios. They assess its accuracy for the 30 countries for which they have complete census estimates for 1960, 1970 and 1980 as follows. First, they use the benchmark values for 1960 (1970) and PIM in the forward direction to estimate attainment in 1970 (1980), yielding "forward-flow" estimates. Second, they start with benchmark values in 1970 (1980) and use PIM backward to estimate attainment in 1960 (1970), yielding "backward–flow" estimates. Then they compare the accuracy of these two estimates with forecasts from simple linear trends: extrapolations from 1960 and 1970 to an estimate for 1980 and from the values for 1970 and 1980 to an estimate for 1960. They also estimated linear interpolations from the values for 1960 and 1980 to estimates for 1970 and ran several regressions of the observed values of various levels of educational attainment in 1960, 1970 and 1980 for the 30 countries on the estimates generated from forward- and backward-flow and linear extrapolation and interpolation methods. They found that linear extrapolations for 1960 and1980 were insignificant in all cases, and so was the backward-flow estimate for 1970. By contrast, the forward-flow estimate was significant in all cases for 1980, and the forward-flow and linear interpolation for 1970 were jointly significant in all cases. For more details see Barro and Lee (1993).

[7] Nehru, Swanson and Dubey (1993) introduced several improvement in Lau, Jamison and Louat's procedure. First, they collect more data on school enrolment prior to 1960 and therefore they do not have to rely on the backward extrapolation. Next, they did some adjustment for grade repetition and drop-outs.

[8] Lau, Jamison and Louat (1991) and Lau, Bhalla and Louat (1991) use a PIM and annual enrolment data to construct educational attainment series. Their PIM uses age-specific survival rates constructed for representative countries in each region.

[9] The Barro and Lee (2001) dataset improves on their earlier estimates in a number of respects. First, fill-in procedure for missing values now uses gross enrolment ratios, adjusted for repeaters. Second, in the construction of average years of schooling, they now take account of changes of school duration over time within countries.

[10] De la Fuente and Doménech (2001) construct educational attainment series for the adult population of a sample of 21 OECD countries covering the period 1960–1995. Their approach has been to collect all the information that could be found on educational attainment in each country, both from international publications and from national sources and use it to reconstruct a plausible attainment profile for each country.

[11] Cohen and Soto (2000) construct a dataset for a sample of 95 countries covering the period 1960–2000 at 10-year intervals. The key methodology is to minimize the extrapolation and keep them as close as possible to those directly available from national census. They collect census/survey data from UNESCO, the OECD's in-house educational database and websites of national statistical agencies. Their estimates refer to the 15–64 age group.

respectively.[12] For each level, the attainment rate is defined as the percentage of the relevant sub-population (over 15 or over 25) having been enrolled up to a specific level of education but no further (i.e. those who did not pursue any further education than the given level).[13]

Barro and Lee estimate the average years of schooling using attainment data as follows:

$$Av\_yrs = DUR_p\left[(1/2)h_{ip} + h_{cp}\right] + \left(DUR_p + DUR_{s1}\right)h_{is} + \left(DUR_p + DUR_{s1} + DUR_{s2}\right)h_{cs} +$$
$$+\left(DUR_p + DUR_{s1} + DUR_{s2} + (1/2)DUR_h\right)h_{ih} + \left(DUR_p + DUR_{s1} + DUR_{s2} + DUR_h\right)h_{ch}$$

where *h* represents the percentage of population with different degree of educational attainment described by subscripts. That is, for each *h*, the *j*th level of education is the highest attained: *j=ip* for incomplete primary education, *cp* for completed primary education, *is* for the first cycle of secondary education, *cs* for the second cycle of secondary, *ih* for incomplete higher education and *ch* for completed higher education. DUR_i is the duration in years of the *i*th level of schooling : *i=p* for primary, *s1* for the first cycle of secondary, *s2* for the second cycle of secondary and *h* for higher.[14]

Because the Barro and Lee dataset only gives values for each five years, we used a technique of interpolation/extrapolation to obtain yearly figures from 1960 to 2004 for 105 countries.[15] Note that the data is not adjusted for education quality. Education quality varies across countries, and available data is too fragmentary to be exploited systematically.[16]

b. Description of human capital stock

Table 7 shows the summary statistics of the estimated human capital stock series. It can be seen that there is a large variation across countries in the average number of years of schooling.

Table 7. Summary statistics, average years of schooling

Variable	Obs	Mean	Std. Dev.	Min	Max
Average Years of School	4 710	4.65	2.94	0.04	12.30

Country averages over sample period

Variable	Obs	Mean	Std. Dev.	Min	Max
Average Years of School	105	4.64	2.72	0.44	11.02

[12] Barro and Lee (2001) provide data for the population aged 25 and over and for the population aged 15 and over. The earlier version of Barro and Lee provided the data only for the population aged 25 and above in order to obtain the widest possible coverage. However, focusing only on the population aged 25 and over was ignoring the fastest growing segment of the labour force in the developing countries. Therefore, the latest version of Barro and Lee also provides the data for the population aged 15 and over which corresponds better to the labour force for many developing countries.

[13] The raw data on educational attainment come from issues of UNESCO Statistical Yearbook, which reports census and survey data by age and sex.

[14] See Barro and Lee (1993, 2001) for details.

[15] For Benin and Egypt, we extrapolated only until 1965 and 1970 respectively, since the extrapolations backward were resulting in a negative numbers. Congo, Gambia and China were extrapolated backward from 1975 until 1960 and Rwanda from 1970 until 1960.

[16] Studies by Nehru, Swanson and Dubey (1995) and Cohen and Soto (2001) show that there is a high degree of correlation between Barro and Lee estimates and other estimates of educational stocks.

Table 8 shows average years of schooling broken down both by decade (1971–1981, 1982–1992, and 1993–2003) and by income group/region. The data has 34 high-income countries (24 Organization for Economic Cooperation and Development (OECD) and 10 non-OECD), 17 upper middle-income countries, 28 lower middle-income countries, and 24 low-income countries. The low and middle-income group is further broken down into six regions: East Asia and Pacific (8 countries), Europe and Central Asia (3 countries), Latin America and Caribbean (21 countries), Middle East and North Africa (6 countries), South Asia (6 countries), and sub-Saharan Africa (25 countries). The average years of schooling rises in all regions. Among the low- and middle-income countries, sub-Saharan Africa, South Asia and Middle East and North Africa have the lowest averages but show the highest growth rates (see table 9).

Table 8. Average years of schooling by regions

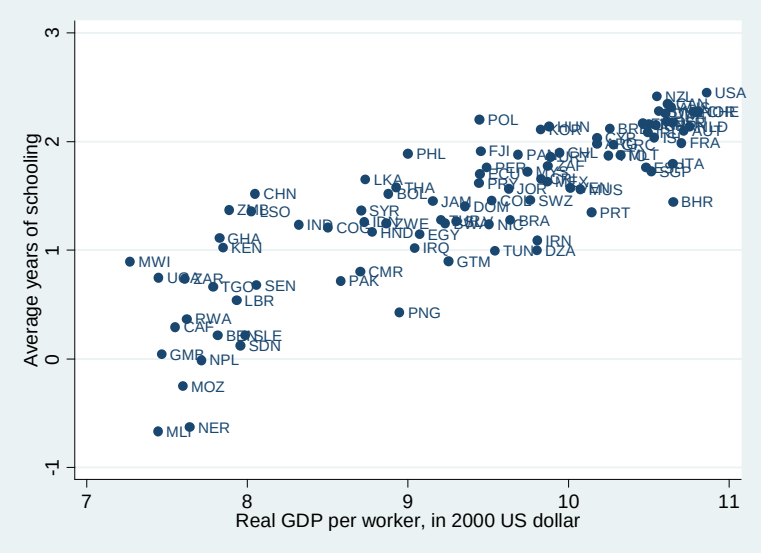
Region	1971–1981	1982–1992	1993–2003	1971–2003
High income countries	6.53	8.08	9.05	7.56
Low and middle income countries				
East Asia & Pacific	3.01	4.49	5.54	4.02
Europe & Central Asia	5.13	6.71	7.63	6.17
Latin America & Caribbean	3.45	4.80	5.64	4.35
Middle East & North Africa	1.39	3.30	5.04	2.79
South Asia	1.59	2.44	3.08	2.18
Sub-Saharan Africa	1.44	2.48	3.30	2.17

Table 9. Annual average growth rate of average years of schooling, by regions

Region	1971–1981	1982–1992	1993–2003	1971–2003
High income countries	1.3	1.1	0.8	1.1
Low and middle income countries				
East Asia & Pacific	1.8	2.4	1.3	1.9
Europe & Central Asia	1.3	1.5	0.8	1.2
Latin America & Caribbean	1.7	1.7	1.1	1.5
Middle East & North Africa	5.1	5.0	2.8	4.5
South Asia	2.6	2.0	1.9	2.3
Sub-Saharan Africa	2.8	3.2	1.7	2.6

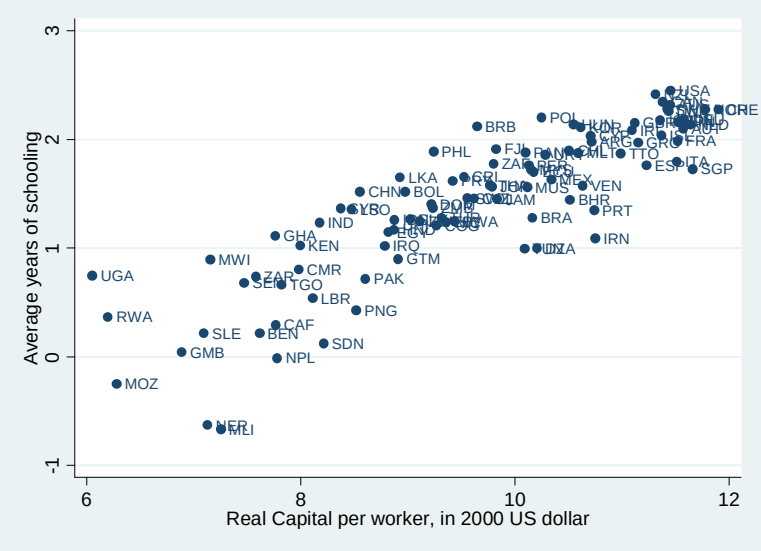
As a reliability check, figure 2 shows average output per worker at the country level (averaged over the sample period) against average years of schooling. The correlation is, as expected, quite high.

Figure 2. Real GDP per worker vs. human capital stock (in logs)



As a further check, figure 3 shows years of schooling against the real stock of capital per worker, both in logs. It can be seen that the relationship is positive, reflecting correlation with a third variable (income levels), but also concave: there is more deepening of physical capital than human capital in the north-east of the scatter plot.

Figure 3. Capital stock per worker vs. human capital stock (in logs)



2.1.3 Natural resource endowment

To measure the natural resource endowment in a country, we use the data on arable land taken from the World Bank's World Development Indicators (WDI). The series we used – arable land hectares per person – is presented in 1,000 ha per person, and covers 203 countries over the period of 1961–2005.[17] Out of those countries, 164 have 45 years or more of data (see table 10a).

Table 10a. Land database coverage

Time period	Number of countries	Start of time series
2	1	2004
3	9	2003
6	2	2000
11	2	1995
13	3	1993
14	19	1992
16	1	1990
26	1	1980
42	1	1961
45	164	1961
	203	

One justification for using arable land is that it does not stay the same over time for each country as it reflects land development or desertification. However, the availability of arable land itself is not a perfect measure of natural resource endowments of a country.

Therefore we also look into a database on natural resource capital from the World Bank's volumes "Expanding the Measure of Wealth" (1997) and "Where is the Wealth of Nations?" (2006). They offer, among others, a database on natural capital for over 100 countries. Though the database covers only two years (1994 and 2000), it provides us with the most complete measure of natural resource endowments to date and it could be used as a good indicator. Natural resource capital in the database consists of non-renewable resources (subsoil assets, including oil, natural gas, coal, and mineral resources), cropland, pastureland, forested areas (including areas used for timber extraction and non-timber forest products), and protected areas.[18] Natural capital values are given per capita and are based upon country-level data on physical stocks, and estimates of natural resource rents are based on world prices and local costs.

Table 10b presents the total values of natural capital and its components by income groups. While the value of natural capital per capita is substantially higher in high-income countries than low income ones, the percentage of cropland and pastureland in total natural capital is significantly higher in low income countries.

[17] Arable land (hectares per person) includes land defined by the Food and Agriculture Organization of the United Nations (FAO) as land under temporary crops (double-cropped areas are counted once), temporary meadows for mowing or for pasture, land under market or kitchen gardens, and land temporarily fallow. Land abandoned as a result of shifting cultivation is excluded. See WDI explanations.

[18] For details, see World Bank (2006).

Table 10b. Natural capital, year 2000 (Untied States dollars per capita)

Income group	*Subsoil assets*	*Timber resources*	*Nontimber forest resources*	*Protected areas*	*Cropland*	*Pastureland*	**Total Natural Capital**	*% of Cropland and Pastureland in total natural capital*
Low-income countries	325	109	48	111	1 143	189	1 925	69%
Middle-income countries	1 089	169	120	129	1 583	407	3 496	57%
High-income countries (OECD)	3 825	747	183	1 215	2 008	1 552	9 531	37%
World	1 302	252	104	322		536	4 011	51%

Sources: The World Bank (2006), Table 1.2.
Notes: Oil states are excluded.

Both of the data, arable land and natural capital (and its components), are given per capita. In order to make them consistent with our measures for physical capital measured per worker, we have merged the data with PWT to infer the numbers of workers and converted them into values of relative endowments per worker.

Given the data availability, we use only arable land as a measure of natural resource endowment for the panel data for the common period, that is from 1971 (1988 in HS) to 2003. But we also calculated RFI indices separately for the years 1994 and 2000, using the natural resource data.

2.2 Cluster analysis of national factor endowments

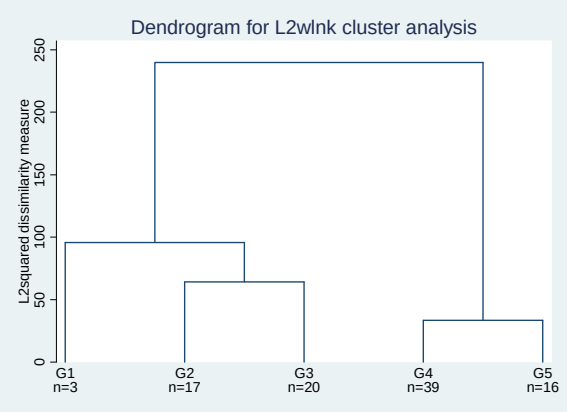
We now examine whether our data are reasonable and realistic estimates of national factor endowment by using cluster analysis. In order to avoid scale effects (the fact that the range of a variable affects its influence in the clusters' definition), we have standardized all endowment variables to a mean of zero and a standard deviation of one. This prevents the capital stock, which has a much wider range than the other two variables, from dominating the clustering procedure (see table 11).

Table 11. Summary statistics of factor endowments for the period 1971–2003

Variable	Obs	Mean	Std. Dev.	Min	Max
Capital Stock per worker	95	34 306	37 768	426	147 540
Human Capital Stock	95	5.04	2.81	0.51	11.58
Arable Land per worker	95	0.74	0.80	0.00	5.79

Two of the general types of clustering methods are hierarchical and partition. Hierarchical clustering methods create hierarchically related sets of clusters. Partition clustering methods separate the observations into mutually exclusive groups. We applied here two alternative algorithms to explore the endowment data's structure. First, we used a distance-based agglomerative clustering algorithm known as Ward's method. The resulting "dendrogram" is shown in figure 4.

Figure 4. Dendrogram of Ward's cluster analysis: countries



Long vertical lines at the top of the dendrogram indicate group of countries that are strongly dissimilar, and shorter lines indicate those that are less dissimilar. Figure 4 suggests five broad country groupings. More formally, we applied two stopping rules whose results are shown in table 12: Calinski and Harabasz' pseudo-F index, and Duda and Hart Je(2)/Je(1) index. The best stopping level is given by the maximum value of the pseudo-F index or, alternatively, by the minimum value of the Je(2)/Je(1) index. Both support a five-group structure.[19]

Table 12. Calinski and Harabasz and Duda and Hart stopping rules' result

Calinski & Harabasz		Duda & Hart		
Numbers of Clusters	**Pseudo-F**	**Numbers of Clusters**	**Je(2)/Je(1)**	**Pseudo T-squared**
2	68.8	2	0.59	25.9
3	67.5	3	0.45	43.1
4	73.6	4	0.62	32.2
5	74.3	5	0.19	4.2
6	72.1	6	0.49	38.2
7	71.3	7	0.39	22.0
8	73.9	8	0.57	11.5
9	78.1	9	0.68	8.6
10	77.7	10	0.58	7.2
11	76.6	11	0.00	.
12	75.8	12	0.63	14.1
13	76.2	13	0.58	5.0
14	76.1	14	0.63	7.6
15	76.6	15	0.14	18.0

[19] The grouping into clusters is even clearer when we exclude arable land use per capita from the endowment variables. See (appendix tables A2 and A3 and figure A1).

Based on the working hypothesis that five clusters is a "natural" partition of the data, we have used the *k*-means partition method (where *k* is specified by the user, which in our case we take as 5) to form the clusters iteratively. We base the partition on the Euclidean distance metric (also known as the Minkowski distance metric with argument 2). The general form for the distance metric between observation *i* and centroid *j* using *p* variables is given by

$$d_{ij} = \left[\sum_{m=1}^{p} \left| X_{mi} - X_{mj} \right|^{n} \right]^{1/n} \qquad (5)$$

Table 13 describes the five clusters and shows that the partition is fairly natural. Cluster 1 is made of low-income countries and lower middle-income countries (except Bahrain and Portugal, which are high-income countries but with a very low endowment of land and a relatively low capital and human capital endowment). It is characterized by low capital and human capital endowments and with the lowest endowment of arable land (all in relative to labour). Cluster 2 is made of lower middle-income countries with a few low-income countries. The difference of this cluster from the above one is that it has the second highest endowment of arable land.[20] Cluster 3 essentially consists of upper middle-income countries. Clusters 4 and 5 consist of OECD countries. The only difference between these clusters is that two countries in cluster 4 (Canada and Australia) own a large land endowment in addition to large physical and human capital endowments.

[20] Turkey, which is an upper middle-income country, is included in cluster 2, because (a) it has the lowest physical- and human-capital endowment among the upper middle income countries; and (b) its arable land endowment is one of the highest in the middle-income countries.

Table 13. Summary of clusters

Clusters	Countries in Clusters	Number of countries	Capital Stock	Human Capital Stock	Arable land	World Bank Income Group
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	Benin	41	9 366	3.08	0.47	Low income
	Bahrain					High income: nonOECD
	Bolivia					Lower middle income
	Brazil					Upper middle income
	Botswana					Upper middle income
	China					Lower middle income
	Congo, Rep.					Lower middle income
	Colombia					Lower middle income
	Costa Rica					Upper middle income
	Dominican Republic					Lower middle income
	Egypt, Arab Rep.					Lower middle income
	Ghana					Low income
	Gambia, The					Low income
	Guatemala					Lower middle income
	Honduras					Lower middle income
	Indonesia					Lower middle income
	India					Lower middle income
	Jamaica					Upper middle income
	Jordan					Lower middle income
	Kenya					Low income
	Liberia					Low income
	Sri Lanka					Lower middle income
	Lesotho					Lower middle income
	Mali					Low income
	Mozambique					Low income
	Mauritius					Upper middle income
	Malawi					Low income
	Nicaragua					Lower middle income
	Nepal					Low income
	Pakistan					Low income
	Papua New Guinea					Low income
	Portugal					High income: OECD
	Rwanda					Low income
	Senegal					Low income
	Sierra Leone					Low income
	El Salvador					Lower middle income
	Swaziland					Lower middle income
	Thailand					Lower middle income
	Uganda					Low income
	Congo, Dem. Rep.					Low income
	Zimbabwe					Low income
2	Central African Republic	13	11 942	2.68	1.50	Low income
	Cameroon					Lower middle income
	Algeria					Lower middle income
	Iran, Islamic Rep.					Lower middle income
	Iraq					Lower middle income
	Niger					Low income
	Paraguay					Lower middle income
	Sudan					Lower middle income
	Syrian Arab Republic					Lower middle income
	Togo					Low income
	Tunisia					Lower middle income
	Turkey					Upper middle income
	Zambia					Low income

Clusters	Countries in Clusters	Number of countries	Capital Stock	Human Capital Stock	Arable land	World Bank Income Group
(1)	(2)	(3)	(4)	(5)	(6)	(7)
3	Argentina	22	36 781	6.76	0.65	Upper middle income
	Barbados					High income: nonOECD
	Chile					Upper middle income
	Cyprus					High income: nonOECD
	Ecuador					Lower middle income
	Spain					High income: OECD
	Fiji					Upper middle income
	Greece					High income: OECD
	Hungary					High income: OECD
	Ireland					High income: OECD
	Korea, Rep.					High income: OECD
	Mexico					Upper middle income
	Malta					High income: nonOECD
	Malaysia					Upper middle income
	Panama					Upper middle income
	Peru					Lower middle income
	Philippines					Lower middle income
	Poland					Upper middle income
	Trinidad and Tobago					High income: nonOECD
	Uruguay					Upper middle income
	Venezuela, RB					Upper middle income
	South Africa					Upper middle income
4	Australia	2	90 764	10.32	4.59	High income: OECD
	Canada					High income: OECD
5	Austria	17	101 711	8.74	0.51	High income: OECD
	Switzerland					High income: OECD
	Germany					High income: OECD
	Denmark					High income: OECD
	Finland					High income: OECD
	France					High income: OECD
	United Kingdom					High income: OECD
	Iceland					High income: OECD
	Israel					High income: nonOECD
	Italy					High income: OECD
	Japan					High income: OECD
	Netherlands					High income: OECD
	Norway					High income: OECD
	New Zealand					High income: OECD
	Singapore					High income: nonOECD
	Sweden					High income: OECD
	United States					High income: OECD

3. Estimating the revealed factor intensity indices

We now proceed to use our endowment data to build our revealed factor intensity (RFI) indices of export products, using a methodology inspired by Hausmann, Hwang and Rodrik's index of revealed technology content (PRODY).

3.1 Methodology

Our Revealed Factor Intensity (RFI) indices for each traded good is calculated as a weighted average of the factor abundance of the countries exporting that good, with a variant of Balassa's Revealed Comparative Advantage (RCA) indices as weights. The rationale for using a variant of RCA indices as opposed to straight export shares (X_j^i / X) is to ensure that country size does not distort the ranking of goods.

For example, both China and Togo produce and export the 5 digit SITC product category 65394, "Fabrics ,woven ,of vegetable textile". In year 2000, the export value of China for this product was US$ 96 million, whereas Togo's export value was only US$ 0.1 million. However, this product constituted only 0.02% of total Chinese exports, compared to 0.05% for Togo. Therefore the index allows us to weight Togo's factor abundance more heavily than the Chinese factor abundance (37% for Togo, 17% for China) in calculating the revealed factor intensity level of the product, even though China's exports are bigger than Togo's.

Thus, the revealed capital intensity index of good *j* is calculated as

$$k_j = \sum_i \omega_j^i \frac{K^i}{L^i} \tag{6}$$

where K^i is country *i*'s capital stock, L^i is its labor force, and the weights are given by

$$\omega_j^i = \frac{X_j^i / X^i}{\sum_i \left(X_j^i / X^i \right)}. \tag{7}$$

That is, ω_j^i is a variant of Balassa's RCA for country *i* in good *j*. Balassa's index is

$$RCA_j^i = \frac{X_j^i / X^i}{X_j / X} \tag{8}$$

where X_j^i is country *i*'s exports of good *j*, $X^i = \sum_j X_j^i$ is country *i*'s aggregate exports, X_j is world exports of good *j*, and $X = \sum_j X_j$ is world aggregate exports. The denominator of ω_j^i, which is $\sum_i \left(X_j^i / X \right)$, i.e. the sum of product *j*'s shares across countries, is not identical with that of Balassa's index, which is $\sum_i X_j^i / \sum_i X^i$, i.e. the share of product *j* in world trade. In so doing, we use a trick first used by Hausmann, Hwang and Rodrik (2007) which ensures that the weights add up to one, as

$$\sum_i \omega_j^i = \sum_i \frac{X_j^i / X^i}{\sum_i \left(X_j^i / X^i \right)} = \frac{1}{\sum_i \left(X_j^i / X^i \right)} \sum_i \left(X_j^i / X^i \right) = 1 . \qquad (9)$$

This eliminates a problem of a large values of RCA indices arising from the values that are very close to zero in the denominator (a product's share in world trade) at the disaggregated level (like Hausmann *et al.*).

Similarly, the revealed human capital intensity index is given by

$$h_j = \sum_i \omega_j^i h^i \qquad (10)$$

where h^i is the average years of schooling achieved by the average person. The revealed land intensity index, finally, is calculated using arable land per person,

$$l_j = \sum_i \omega_j^i l^i \qquad (11)$$

where ℓ^i is the arable land (in hectares) per person.

Two issues are worth mentioning. Balassa indices have been criticized because (a) countries and commodities are double-counted; and (b) they are based on gross exports, whereas (as the argument goes) it should be based on net exports instead. Second, our index is potentially distorted by export subsidies, and agricultural exports are a particularly severe problem. We deal with both in turn.

3.1.1 Caveats

a. Limitations of Balassa's index

Vollrath (1987, 1989) suggested slightly amended versions of the index. One eliminates the double counting:

$$RCA_j^i \Big|_{Vollrath1} = \frac{X_j^i / X_{-j}^i}{X_j^{-i} / X^{-i}} \qquad (12)$$

where X_{-j}^i stands for country *i*'s exports net of its exports of good *j* (that is, $X_{-j}^i = X^i - X_j^i$) and X_j^{-i} and X^{-i} stand respectively for exports of good *j* by all countries except *i* and world exports net of country *i*'s. He also suggested the following version of the index, encompassing both import and export dimensions of comparative advantage:

$$RCA_j^i \Big|_{Vollrath2} = \frac{X_j^i / X_{-j}^i}{X_j^{-i} / X^{-i}} - \frac{M_j^i / M_{-j}^i}{M_j^{-i} / M^{-i}} \qquad (13)$$

We tried both specifications and decided to reject them. The first one makes little difference and is not worth the complication. The second, by contrast, introduces considerations which make it unsuitable for our purposes. To see this, consider a world of three countries: France, Germany and Ghana, with intra-industry trade in telecom equipment between Germany and France,

and no trade in that product between either of them and Ghana. Germany is a slight net exporter and France is a slight net importer. In calculating the revealed capital intensity of telecom equipment, France and Germany will cancel out each other, their aggregate weight being zero. The revealed capital intensity will then be indeterminate. This is not a far-fetched example.

Table 14 shows how using Vollrath's second correction for the RCA yields a lower revealed capital intensity for SITC 5-digit 86198 (instruments for physical or chemical analysis, traded by 96 countries) than for SITC 4-digit 4217 (rape, colza and mustard oils).

Table 14. Effect of Vollrath's correction on RCI index

Country	SITC 4-5 digit	Export	Import	$(x^a/X)/\sum(x^a/X)$	$(m^a/M)/\sum(m^a/M)$	$(x^a/X)/\sum(x^a/X)-(m^a/M)/\sum(m^a/M)$	Capital Stock per worker	RCI
Korea, Rep. of	86 198	11 275	239 806	0.17%	2.22%	-2.05%	84 821	-1'739
New Zealand	86 198	1 332	15 152	0.35%	1.61%	-1.26%	88 927	-1'123
Norway	86 198	7 351	34 717	0.38%	1.50%	-1.12%	152 748	-1'713
Canada	86 198	189 653	349 152	2.01%	2.18%	-0.17%	110 351	-188
…								
…								
Instruments for physical or chemical analysis	86 198	*8 168 913*	*7 656 878*	*100.00%*	*100.00%*			***42 041***
Rape, colza and mustard oils	4 217							***51 161***

b. Dealing with agricultural distortions

The last example raises an additional issue. Many agricultural commodities end up with high revealed capital and human capital intensities because they are exported by rich countries who subsidize them (export subsidies have been "litigated out" for most manufactured products so we ignore them). Such outcome does not arise from comparative advantage, but rather a result of direct policy intervention.

We attempt to correct for distortions in agricultural prices using a new database on agricultural distortions published by the World Bank in October 2008.[21] The database provides, among others, a nominal rate of assistance (NRA) for a number of agricultural products for developed and developing countries over the period 1955–2005. The agricultural product coverage includes 70 per cent of agricultural and food value added excluding highly processed food, beverages and tobacco, and agricultural crops, of those countries included in the sample.[22]

[21] See Anderson, *et al.* (2008) for details of the database. We would like thank Kym Anderson and Ernesto Valenzuela for their kind e-mails with very useful clarifications to our questions on the database.

[22] A similar database has been provided systematically for the last two decades by the OECD secretariat, which provides Producer Support Estimates (PSEs) and Consumer Support Estimates (CSEs). However, these estimates are given only for a few key products, and for a much smaller number of countries (only for high income countries and five non-European Union developing countries) for the years from 1986–2005.

The NRA is measured as the unit value of production at the distorted price less its value at the undistorted free market price expressed as a fraction of the undistorted price.[23]

$$NRA_{ik}^{t} = \frac{E * P(1 + t_{ik}^{t}) - E * P}{E * P} = \tau_{ik}^{t} \qquad (14)$$

where E is the domestic currency price of foreign exchange and P is the foreign currency price of good k in the international market and τ_{kt}^{i} is the ad-valorem equivalent of the array of tariffs and domestic tax and subsidies affecting good k in country i in year t, which can be positive or negative. It is typically positive in high-income countries subsidizing and protecting agriculture, and negative in low-income countries taxing theirs (Anderson, 2008). Our correction consisted of weeding out observations (country × product pairs) characterized by nonzero NRAs in order to keep only undistorted RCAs.

3.1.2. Data coverage

In constructing the database, we faced a trade-off between "width" and "consistency" in country-endowment data. On the one hand, we are interested in having indices for as many countries as possible, to give a width to the database. On the other hand, to track the evolution of RFI indices for each good over several years, we need to have a complete (i.e. balanced) panel of data on endowments of a given set of countries for the same length of years, to ensure that the indices are constructed in comparable ways.

However, if there is systematic bias in the selection of countries in the panel (say, if low-income countries are underrepresented in the data), RFI indices will be biased against factors of which low-income countries are poorly endowed. This may not necessarily alter the ranking of goods by RFI, but will affect the relative intensities. In order to minimize this bias, the wide-coverage (unbalanced) panel includes, each year, all the countries for which data are available in that year.

Table 15 gives the largest number of countries with a common set of time periods for all trade and endowment data (the balanced panel). The resulting 92 countries are tracked over a sample period of 33 years (1971–2003).

Table 15. Balanced data coverage

	Number of years	Time period	Number of Countries
Capital Stock	33	1971–2003	136
Human Capital Stock	45	1960–2004	105[a]
Land	45	1961–2005	165[b]
This study coverage	**33**	**1971–2003**	**92[c]**

Notes:
[a] Egypt has 35 years of time series over the period 1970–2004.
[b] Benin has 40 years of time series over the period 1965–2004.
[c] WITS does not provide trade data for all these 33 years for 3 countries such as Lesotho, Swaziland and Botswana.

[23] OECD's PSEs are calculated as a fraction of the distorted value; that is, $PSE = t_m / (1 + t_m)$ and for a positive t_m it is smaller than NRA and is necessarily less than 100 per cent, which is not the case for the NRA. See Anderson, Kurzweil, Martin, Sandri and Valenzuela (2008) and OECD (2007).

Table 16 shows the widest range of countries for each factor endowment (the unbalanced panel). For example we have the capital stock data starting from 1951, but the number of countries which have the data varies from year to year. The number of countries having all three endowments' data varies between 76 and 99, depending on the years (bottom line of the table 16).

Table 16. Wide (unbalanced) coverage for each endowment

	Range of years	Range of number of countries
Capital Stock	1951–2003	51–141
Human Capital Stock	1960–2004	103–105
Arable Land	1961–2005	165–203
All three endowments	***1961–2003***	***76–99***

Table 17 presents the number of countries covered in the unbalanced data sets for each year.

Table 17. Wide (unbalanced) coverage

Year	Number of countries	Year	Number of countries
1961	76	1983	99
1962	77	1984	99
1963	77	1985	99
1964	77	1986	99
1965	78	1987	99
1966	78	1988	99
1967	78	1989	99
1968	78	1990	99
1969	78	1991	99
1970	79	1992	98
1971	98	1993	98
1972	98	1994	98
1973	99	1995	98
1974	99	1996	98
1975	99	1997	98
1976	99	1998	98
1977	99	1999	98
1978	99	2000	99
1979	99	2001	98
1980	99	2002	98
1981	99	2003	98
1982	99		

Finally, as regards product classification, we calculated the indices using two different classification schemes: Revision 1 of the United Nations Standard International Trade Classification (SITC 5-digit) and the Harmonized System (HS88/92 6-digit). Each has its own advantages and disadvantages. SITC provides longer years of trade statistics (since 1962) with fewer revisions than the HS, thus has the advantage of giving maximum comparability over the sample period. HS gives us a more disaggregated product classification, at the 6-digit level, than SITC. Whereas there are only over 1,000 products at the 4-5 digit products of the SITC classification, there are over 5,000 products at the HS 6-digit level.

Our SITC database covers 1971–2003, and our HS6 covers 1988–2003, with few countries until 1992.

3.2 Results – revealed factor intensity indices

We now illustrate the results of our RFI indices for the year 2000. Table 18 shows the summary statistics of the RFI indices for each good in the SITC classification (corresponding tables of results using the HS classification are given in appendix A).

Table 18. Summary statistics of revealed factor intensity indices, year 2000 (SITC classification)

Variable	Obs	Mean	Std. Dev.	Min	Max
rhci	1166	7.07	1.62	1.52	11.21
rci	1166	60 257	30 726	2 608	149 916
rnri_land	1166	0.61	0.35	0.07	4.18
rnri_nc	1166	14 768	8 998	2 028	73 993
rnri_sa	1166	4 826	5 428	31	61 315
rnri_pc	1166	6 909	4 011	1 087	43 272

rhci	Revealed human capital intensity
rci	Revealed (physical) capital intensity
rnri_land	Revealed natural resource intensity – land
rnri_nc	Revealed natural resource intensity – natural capital
rnri_sa	Revealed natural resource intensity – sub-oil assets
rnri_pc	Revealed natural resource intensity – pastured and crop land

Table 19 shows simple averages of RFI indices for 10 industries at the SITC-1 aggregation level.

Table 19. Simple averages of factor intensity indices, by SITC 1 digit industries

sitc1	SITC 1 digit description	RHCI	RCI	RNRI_land	RNRI_nc	RNRI_sa	RNRI_pc
0	Food and live animals	6.27	39 067	0.79	15 428	4 422	8 314
1	Beverages and tobacco	6.95	52 538	0.61	15 070	4 614	7 704
2	Crude materials, inedible	6.37	42 159	0.74	16 382	5 256	7 640
3	Mineral fuels, lubricants	6.94	47 869	0.69	20 925	12 070	6 187
4	Animal and vegetable oils and fats	5.67	34 756	0.74	12 748	3 795	6 709
5	Chemicals	7.66	72 169	0.59	16 641	6 119	7 354
6	Manufact goods classified chiefly	7.06	62 059	0.55	13 667	4 342	6 380
7	Machinery and transport equipment	8.23	87 231	0.56	15 474	4 705	6 998
8	Miscellaneous manufactured articles	7.04	60 941	0.46	11 818	3 607	5 814
9	Commod. & transacts. not class. acc	7.60	75 250	0.79	18 288	7 253	6 115

The revealed capital intensity (RCI) indices and the revealed human capital intensity (RHCI) Indices appear highly correlated.

Factor intensity rankings are reported in table 20 (a-f) for all 10 industries and three factors (also for factors that are calculated using the World Bank data on natural capital). Resulting rankings are plausible. For instance, machinery and transport equipment or chemicals are revealed as intensive in capital and human capital. By contrast, food and live animals, animal and vegetable oils and fats, or crude materials have the lowest RFI indices for capital and human capital, but rank near the top in terms of land intensity.

Table 20. Ranking of industries in terms of revealed factor intensity indices

a. RHCI

Rank	SITC 1 digit description	RHCI
1	Machinery and transport equipment	8.23
2	Chemicals	7.66
3	Commod. & transacts. not class. acc	7.60
4	Manufact goods classified chiefly	7.06
5	Miscellaneous manufactured articles	7.04
6	Beverages and tobacco	6.95
7	Mineral fuels, lubricants	6.94
8	Crude materials, inedible	6.37
9	Food and live animals	6.27
10	Animal and vegetable oils and fats	5.67

b. RCI

Rank	SITC 1 digit description	RCI
1	Machinery and transport equipment	87 231
2	Commod. & transacts. not class. acc	75 250
3	Chemicals	72 169
4	Manufact goods classified chiefly	62 059
5	Miscellaneous manufactured articles	60 941
6	Beverages and tobacco	52 538
7	Mineral fuels, lubricants	47 869
8	Crude materials, inedible	42 159
9	Food and live animals	39 067
10	Animal and vegetable oils and fats	34 756

c. RNRI (Arable Land)

Rank	SITC 1 digit description	RNRI_land
1	Commod. & transacts. not class. acc	0.79
2	Food and live animals	0.79
3	Animal and vegetable oils and fats	0.74
4	Crude materials, inedible	0.74
5	Mineral fuels, lubricants	0.69
6	Beverages and tobacco	0.61
7	Chemicals	0.59
8	Machinery and transport equipment	0.56
9	Manufact goods classified chiefly	0.55
10	Miscellaneous manufactured articles	0.46

d. RNRI (Total Natural Capital)

Rank	SITC 1 digit description	RNRI_nc
1	Mineral fuels, lubricants	20 925
2	Commod. & transacts. not class. acc	18 288
3	Chemicals	16 641
4	Crude materials, inedible	16 382
5	Machinery and transport equipment	15 474
6	Food and live animals	15 428
7	Beverages and tobacco	15 070
8	Manufact goods classified chiefly	13 667
9	Animal and vegetable oils and fats	12 748
10	Miscellaneous manufactured articles	11 818

e. TNRI (Subsoil Assets)

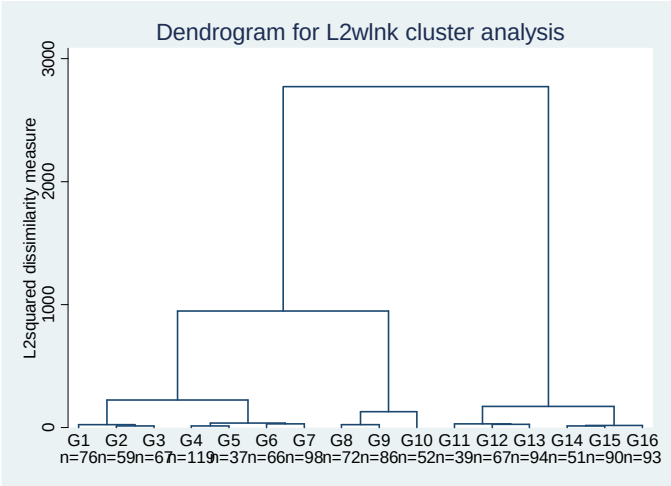
Rank	SITC 1 digit description	RNRI_sa
1	Mineral fuels, lubricants	12 070
2	Commod. & transacts. not class. acc	7 253
3	Chemicals	6 119
4	Crude materials, inedible	5 256
5	Machinery and transport equipment	4 705
6	Beverages and tobacco	4 614
7	Food and live animals	4 422
8	Manufact goods classified chiefly	4 342
9	Animal and vegetable oils and fats	3 795
10	Miscellaneous manufactured articles	3 607

f. RNRI (Pastureland and Cropland)

Rank	SITC 1 digit description	RNRI_pc
1	Food and live animals	8 314
2	Beverages and tobacco	7 704
3	Crude materials, inedible	7 640
4	Chemicals	7 354
5	Machinery and transport equipment	6 998
6	Animal and vegetable oils and fats	6 709
7	Manufact goods classified chiefly	6 380
8	Mineral fuels, lubricants	6 187
9	Commod. & transacts. not class. acc	6 115
10	Miscellaneous manufactured articles	5 814

We now turn to cluster analysis to explore whether industries can be clustered into naturally homogenous groups in terms of the RCI and RHCI indices, i.e. factor intensity, using the same algorithms as in the previous section. Figure 5 shows that Ward's dendrogram gives six well-identified clusters of products at the finest disaggregated level of the SITC Rev 1 (4-5 digits).

Figure 5. Dendrogram of Ward's cluster analysis: industrial sectors



The number of clusters is validated by the stopping-rule results shown in table 21. The six-cluster solution is the most favourable under Calinski and Harabasz Pseudo-F indices, and to a lesser extent under Duda and Hart indices.

Table 21. Stopping rule result

Calinski & Harabasz		Duda & Hart		
Numbers of Clusters	**Pseudo-F**	**Numbers of Clusters**	**Je(2)/Je(1)**	**Pseudo T-squared**
2	1 712	2	0.39	1 144
3	2 308	3	0.45	630
4	2 137	4	0.49	456
5	2 195	5	0.35	388
6	***2 374***	***6***	***0.69***	***142***
7	2 186	7	0.71	83
8	2 053	8	0.40	244
9	1 964	9	0.54	138
10	1 919	10	0.65	105
11	1 880	11	0.62	96
12	1 864	12	0.70	99
13	1 840	13	0.60	101
14	1 812	14	0.52	127
15	1 799	15	0.58	89

Table 22 shows summary statistics for the clusters just identified. They are ordered from the least intensive in capital and human capital (cluster 1), to the most intensive in both capital and human capital (cluster 6).

Table 22. Summary statistics of the clusters

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
Revealed Capital Intensity Index	10 783	23 386	39 601	58 076	81 636	102 341
Revealed Human Capital Intensity Index	3.29	5.20	6.37	7.38	8.16	9.05
Number of Goods	72	167	224	273	237	193

Tables 23 and 24 show the industry composition of the clusters at the SITC-1 and SITC-2 levels respectively. Two industries account for 50 per cent or more for all clusters except cluster 3 (49 per cent), with the highest proportion accounted for by the top 2 industries in clusters 4, 5 and 6 (54 per cent in each).

Table 23

a. Percentages of SITC 1 digit industries, by clusters

SITC sectors at 1 digit	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
0 Food and live animals	26	17	18	10	3	4
1 Beverages and tobacco	1	1	0	2	1	0
2 Crude materials, inedible	31	26	17	12	5	7
3 Mineral fuels, lubricants	0	1	5	3	0	1
4 Animal and vegetable oils and fats	8	4	3	1	1	0
5 Chemicals	4	11	8	13	25	18
6 Manufact goods classified chiefly	17	25	31	36	29	26
7 Machinery and transport equipment	0	1	4	7	22	28
8 Miscellaneous manufactured articles	10	15	14	18	13	14
9 Commod. & transacts. not class. acc	3	0	0	0	1	2
	100	100	100	100	100	100

b. Percentages of clusters, by SITC 1 digit industries

SITC sectors at 1 digit	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	
0 Food and live animals	15	22	31	20	6	6	100
1 Beverages and tobacco	8	8	8	50	25	0	100
2 Crude materials, inedible	14	27	23	20	7	9	100
3 Mineral fuels, lubricants	0	5	55	32	0	9	100
4 Animal and vegetable oils and fats	25	29	25	13	8	0	100
5 Chemicals	2	11	10	21	35	21	100
6 Manufact goods classified chiefly	4	12	21	28	20	15	100
7 Machinery and transport equipment	0	1	7	14	39	39	100
8 Miscellaneous manufactured articles	4	15	18	29	18	16	100
9 Commod. & transacts. not class. acc	25	0	0	0	38	38	100

Examination of tables 23 and 24 (particularly the latter) shows that cluster composition is far from perfectly overlapping with industry composition. Factor intensities vary substantially not just between, but also within industries, and this pattern remains at all levels of disaggregation.

For instance, an industry sector, SITC 65 (textile yarns, fabrics, made-up articles) covers a wide variety of goods whose factor contents vary from the least human/physical capital intensive to the most capital intensive (table 24b).

This suggests that analyses of the factor content of trade, whether motivated by the empirical validation of trade models or by policy advice, should best be carried out at high degrees of disaggregation.

Table 24

a. Percentages of SITC 2 digit industries, by clusters

SITC 1 digit description	SITC 2 digit	SITC 2 digit description	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6
0 Food and live animals	0	Live animals	1.4	0.0	1.3	0.0	0.0	0.5
	1	Meat and meat preparations	1.4	0.6	2.2	1.5	0.4	0.5
	2	Dairy products and eggs	0.0	0.6	0.9	1.1	0.0	0.0
	3	Fish and fish preparations	1.4	0.6	0.4	0.7	0.0	0.0
	4	Cereals and cereal preparations	1.4	3.0	1.8	1.5	1.3	1.0
	5	Fruit and vegetables	9.7	4.2	7.1	2.2	0.4	1.0
	6	Sugar, sugar preparations and honey	1.4	0.6	0.9	0.4	0.0	0.0
	7	Coffee, tea, cocoa, spices	8.3	4.2	0.9	0.0	0.4	0.0
	8	Feed.-stuff for animals	1.4	1.2	0.4	1.5	0.4	0.5
	9	Miscellaneous food preparations	0.0	1.8	1.8	0.7	0.0	0.0
1 Beverages and tobacco	11	Beverages	0.0	0.0	0.4	1.5	1.3	0.0
	12	Tobacco and tobacco manufactures	1.4	0.6	0.0	0.7	0.0	0.0
2 Crude materials, inedible, except f	21	Hides, skins and furskins, raw	2.8	1.2	0.4	0.7	0.0	0.5
	22	Oil-seeds, oil nuts and oil kernels	5.6	2.4	0.0	0.4	0.0	0.0
	23	Crude rubber	1.4	0.0	0.0	0.4	0.8	0.0
	24	Wood, lumber and cork	4.2	1.2	1.3	0.4	0.4	1.0
	25	Pulp and paper	0.0	0.6	0.9	1.5	0.0	1.0
	26	Textile fibres, not manufactured	5.6	6.0	4.0	1.8	0.8	1.0
	27	Crude fertilizers and crude minerals	2.8	5.4	4.0	2.9	2.1	1.0
	28	Metalliferous ores and metal scrap	5.6	4.8	3.1	1.5	0.4	1.6
	29	Crude animal and vegetable materials	2.8	4.2	2.7	2.2	0.0	1.0
3 Mineral fuels, lubricants and relat	32	Coal, coke and briquettes	0.0	0.0	0.9	0.7	0.0	0.5
	33	Petroleum and petroleum products	0.0	0.6	3.1	1.8	0.0	0.5
	34	Gas, natural and manufactured	0.0	0.0	0.9	0.0	0.0	0.0
	35	Electric energy	0.0	0.0	0.4	0.0	0.0	0.0
4 Animal and vegetable oils and fats	41	Animal oils and fats	0.0	0.6	0.4	1.1	0.4	0.0
	42	Fixed vegetable oils and fats	5.6	2.4	1.3	0.0	0.4	0.0
	43	Animal and vegetable oils and fats,	2.8	1.2	0.9	0.0	0.0	0.0
5 Chemicals	51	Organic chemicals	2.8	6.6	3.1	9.2	13.1	10.4
	52	Inorganic chemicals	0.0	0.0	0.9	0.0	0.0	0.0
	53	Dyeing, tanning and colouring materials	0.0	1.2	0.4	0.7	2.1	1.0
	54	Medicinal and pharmaceutical products	0.0	0.6	0.0	0.0	3.0	1.0
	55	Perfume materials, toilet & cleansing preparations	1.4	0.6	0.9	0.7	0.0	0.0
	56	Fertilizers, manufactured	0.0	1.8	0.9	0.0	0.0	0.0
	57	Explosives and pyrotechnic products	0.0	0.0	0.4	0.4	0.4	0.5
	58	Plastic materials, etc.	0.0	0.0	0.0	0.4	2.1	0.5
	59	Chemical materials and products	0.0	0.6	0.9	1.5	4.2	4.7
6 Manufact goods classified chiefly b	61	Leather and leather manufactures, nes	2.8	1.8	1.3	1.5	0.4	0.0
	62	Rubber manufactures, nes	0.0	0.0	1.3	1.5	1.3	0.0
	63	Wood and cork manufactures	1.4	1.2	2.7	1.5	0.8	0.5
	64	Paper, paperboard and manufactures	0.0	0.0	0.4	1.8	3.0	3.1
	65	Textile yarn, fabrics, made-up articles	9.7	12.6	7.6	8.1	4.6	1.6
	66	Non-metallic mineral manufactures, nes	0.0	4.8	3.6	4.8	6.3	4.7
	67	Iron and steel	0.0	1.2	3.6	5.5	4.2	5.7
	68	Non-ferrous metals	2.8	1.2	5.4	3.7	3.0	6.7
	69	Manufactures of metal, nes	0.0	2.4	5.4	7.3	5.5	4.1
7 Machinery and transport equipment	71	Machinery, other than electric	0.0	0.0	0.4	2.9	13.1	16.6
	72	Electrical machinery	0.0	0.0	1.8	2.9	5.1	5.7
	73	Transport equipment	0.0	0.6	2.2	1.1	4.2	5.7
8 Miscellaneous manufactured articles	81	Sanitary, plumbing, heating and lighting fixtures	0.0	0.0	0.9	0.7	0.8	0.0
	82	Furniture	0.0	0.6	0.0	0.7	0.4	0.0
	83	Travel goods and handbags	0.0	0.0	0.4	0.0	0.0	0.0
	84	Clothing	4.2	7.8	1.3	1.8	0.0	0.0
	85	Footwear	0.0	0.6	1.3	0.4	0.0	0.0
	86	Scientif & control instrum, photographic apparatus	0.0	0.0	0.4	4.0	5.9	8.8
	89	Miscellaneous manufactured articles	5.6	6.0	9.4	9.9	5.5	5.2
9 Commod. & transacts. not class. acc	94	Animals, nes, incl. zoo animals	1.4	0.0	0.0	0.0	0.0	0.0
	95	Firearms of war and ammunition	1.4	0.0	0.0	0.0	1.3	1.0
	96	Coin, other than gold coin	0.0	0.0	0.0	0.0	0.0	0.5
			100.0	100.0	100.0	100.0	100.0	100.0

b. Percentages of clusters, by SITC 2 digit industries

SITC 1 digit description	SITC 2 digit	SITC 2 digit description	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	
0 Food and live animals	0	Live animals	20.0	0.0	60.0	0.0	0.0	20.0	100.0
	1	Meat and meat preparations	7.7	7.7	38.5	30.8	7.7	7.7	100.0
	2	Dairy products and eggs	0.0	16.7	33.3	50.0	0.0	0.0	100.0
	3	Fish and fish preparations	20.0	20.0	20.0	40.0	0.0	0.0	100.0
	4	Cereals and cereal preparations	5.3	26.3	21.1	21.1	15.8	10.5	100.0
	5	Fruit and vegetables	17.9	17.9	41.0	15.4	2.6	5.1	100.0
	6	Sugar, sugar preparations and honey	20.0	20.0	40.0	20.0	0.0	0.0	100.0
	7	Coffee, tea, cocoa, spices	37.5	43.8	12.5	0.0	6.3	0.0	100.0
	8	Feed.-stuff for animals	10.0	20.0	10.0	40.0	10.0	10.0	100.0
	9	Miscellaneous food preparations	0.0	33.3	44.4	22.2	0.0	0.0	100.0
1 Beverages and tobacco	11	Beverages	0.0	0.0	12.5	50.0	37.5	0.0	100.0
	12	Tobacco and tobacco manufactures	25.0	25.0	0.0	50.0	0.0	0.0	100.0
2 Crude materials, inedible, except f	21	Hides, skins and furskins, raw	25.0	25.0	12.5	25.0	0.0	12.5	100.0
	22	Oil-seeds, oil nuts and oil kernels	44.4	44.4	0.0	11.1	0.0	0.0	100.0
	23	Crude rubber	25.0	0.0	0.0	25.0	50.0	0.0	100.0
	24	Wood, lumber and cork	25.0	16.7	25.0	8.3	8.3	16.7	100.0
	25	Pulp and paper	0.0	11.1	22.2	44.4	0.0	22.2	100.0
	26	Textile fibres, not manufactured	12.5	31.3	28.1	15.6	6.3	6.3	100.0
	27	Crude fertilizers and crude minerals	5.7	25.7	25.7	22.9	14.3	5.7	100.0
	28	Metalliferous ores and metal scrap	14.8	29.6	25.9	14.8	3.7	11.1	100.0
	29	Crude animal and vegetable materials	8.7	30.4	26.1	26.1	0.0	8.7	100.0
3 Mineral fuels, lubricants and relat	32	Coal, coke and briquettes	0.0	0.0	40.0	40.0	0.0	20.0	100.0
	33	Petroleum and petroleum products	0.0	7.1	50.0	35.7	0.0	7.1	100.0
	34	Gas, natural and manufactured	0.0	0.0	100.0	0.0	0.0	0.0	100.0
	35	Electric energy	0.0	0.0	100.0	0.0	0.0	0.0	100.0
4 Animal and vegetable oils and fats	41	Animal oils and fats	0.0	16.7	16.7	50.0	16.7	0.0	100.0
	42	Fixed vegetable oils and fats	33.3	33.3	25.0	0.0	8.3	0.0	100.0
	43	Animal and vegetable oils and fats,	33.3	33.3	33.3	0.0	0.0	0.0	100.0
5 Chemicals	51	Organic chemicals	2.1	11.5	7.3	26.0	32.3	20.8	100.0
	52	Inorganic chemicals	0.0	0.0	100.0	0.0	0.0	0.0	100.0
	53	Dyeing, tanning and colouring materials	0.0	16.7	8.3	16.7	41.7	16.7	100.0
	54	Medicinal and pharmaceutical products	0.0	10.0	0.0	0.0	70.0	20.0	100.0
	55	Perfume materials, toilet & cleansing preparations	16.7	16.7	33.3	33.3	0.0	0.0	100.0
	56	Fertilizers, manufactured	0.0	60.0	40.0	0.0	0.0	0.0	100.0
	57	Explosives and pyrotechnic products	0.0	0.0	25.0	25.0	25.0	25.0	100.0
	58	Plastic materials, etc.	0.0	0.0	0.0	14.3	71.4	14.3	100.0
	59	Chemical materials and products	0.0	3.8	7.7	15.4	38.5	34.6	100.0
6 Manufact goods classified chiefly b	61	Leather and leather manufactures, nes	15.4	23.1	23.1	30.8	7.7	0.0	100.0
	62	Rubber manufactures, nes	0.0	0.0	30.0	40.0	30.0	0.0	100.0
	63	Wood and cork manufactures	6.3	12.5	37.5	25.0	12.5	6.3	100.0
	64	Paper, paperboard and manufactures	0.0	0.0	5.3	26.3	36.8	31.6	100.0
	65	Textile yarn, fabrics, made-up articles	8.6	25.9	21.0	27.2	13.6	3.7	100.0
	66	Non-metallic mineral manufactures, nes	0.0	15.1	15.1	24.5	28.3	17.0	100.0
	67	Iron and steel	0.0	4.3	17.4	32.6	21.7	23.9	100.0
	68	Non-ferrous metals	4.3	4.3	26.1	21.7	15.2	28.3	100.0
	69	Manufactures of metal, nes	0.0	7.0	21.1	35.1	22.8	14.0	100.0
7 Machinery and transport equipment	71	Machinery, other than electric	0.0	0.0	1.4	11.1	43.1	44.4	100.0
	72	Electrical machinery	0.0	0.0	11.4	22.9	34.3	31.4	100.0
	73	Transport equipment	0.0	3.3	16.7	10.0	33.3	36.7	100.0
8 Miscellaneous manufactured articles	81	Sanitary, plumbing, heating and lighting fixtures	0.0	0.0	33.3	33.3	33.3	0.0	100.0
	82	Furniture	0.0	25.0	0.0	50.0	25.0	0.0	100.0
	83	Travel goods and handbags	0.0	0.0	100.0	0.0	0.0	0.0	100.0
	84	Clothing	12.5	54.2	12.5	20.8	0.0	0.0	100.0
	85	Footwear	0.0	20.0	60.0	20.0	0.0	0.0	100.0
	86	Scientif & control instrum, photographic apparatus	0.0	0.0	2.3	25.6	32.6	39.5	100.0
	89	Miscellaneous manufactured articles	4.7	11.8	24.7	31.8	15.3	11.8	100.0
9 Commod. & transacts. not class. acc	94	Animals, nes, incl. zoo animals	100.0	0.0	0.0	0.0	0.0	0.0	100.0
	95	Firearms of war and ammunition	16.7	0.0	0.0	0.0	50.0	33.3	100.0
	96	Coin, other than gold coin	0.0	0.0	0.0	0.0	0.0	100.0	100.0

4. Conclusion

We constructed a database of revealed factor intensity (RFI) indices for each export good at a very detailed disaggregation level, using data from up to 99 countries for the period between 1961 and 2003. To calculate the indices on factor endowments, we used (and in some cases updated or made estimates) data from different sources: (a) Barro and Lee's dataset on educational achievements; (b) Easterly and Levine's estimates on national capital stock; and (c) the World Bank's World Development Indicators.

First, we constructed two country-endowment datasets: a "wide" one with the maximum number of countries in each year, and a "consistent" one with 92 countries with full data over 33 years (1971–2003).

Second, using these data, we followed Hausmann, Hwang and Rodrik's (2007) methodology to construct the RFI indices for all export goods at the finest disaggregation level available in harmonized trade data: SITC-5 and HS-6. For each good and relative factor (capital/labour, human capital and land/labour), the RFI indices is calculated as a weighted average of the relative factor abundances of the countries exporting that good, using slightly modified versions of Balassa's RCA indices as weights.

Our RFI indices allow us to systematically classify products according to their factor intensities, at the most disaggregated level of product classification. This is an advantage over other ad hoc attempts, as the degree of factor intensity can widely vary within an industry, e.g. as classified at the HS 2-digit level. Thus, we believe the RFI indices generate a more economically meaningful categorization of products that can be used for policy advice as well as positive trade analysis.

The RFI indices could be used for many purposes. As mentioned in the introduction, it will enable us to revisit the issue of export diversification with a more standard, theory-based approach, i.e. taking into account the effect of changes in relative factor endowment, than the recent eclectic approaches with inductive reasoning. For instance, one could explore to which extent export diversification proceeds from changes in comparative advantage. This could have interesting implications to policymakers and export-promotion agencies when they need to identify or prioritize sectors for export diversification.

A recent study by Cadot, Carrère and Strauss-Kahn (version March 2009) has used our RFI indices to verify a conjecture that diversification in middle to high income countries may simply reflect a slow adjustment to changes in its comparative advantage. The paper confirms a robust hump-shaped relationship between export diversification and the level of income, i.e. export diversification continues up until a certain level of income, but then stops and moves to export specialization as income increase. Then, using the RFI indices, they were able to suggest the reason behind the hump shape was because countries fail to close a tail of export lines that no longer belong to their comparative advantage. Their export bundles are therefore artificially inflated. That is, the slow adjustment of production/export lines may explain the hump-shaped relationship between diversification and development.

In addition, issues that can be explored using the RFI indices would include: (a) how does the capital content of exports evolve with income levels? (b) are there systematic deviations linked e.g. to governance failures (an "anti-capital" bias)? and (c) does the factor content of trade vary with its destination (e.g. Southern countries could export more capital-intensive goods to other Southern countries than to Northern ones)?

Appendix A. Appendix tables and figures

Table A1. Countries included in the sample

	World Bank Country Code	Country Name
1	ARG	Argentina
2	AUS	Australia
3	AUT	Austria
4	BEN	Benin
5	BHR	Bahrain
6	BOL	Bolivia
7	BRA	Brazil
8	BRB	Barbados
9	CAF	Central African Republic
10	CAN	Canada
11	CHE	Switzerland
12	CHL	Chile
13	CHN	China
14	CMR	Cameroon
15	COG	Congo
16	COL	Colombia
17	CRI	Costa Rica
18	CYP	Cyprus
19	DEU	Germany
20	DNK	Denmark
21	DOM	Dominican Republic
22	DZA	Algeria
23	ECU	Ecuador
24	EGY	Egypt
25	ESP	Spain
26	FIN	Finland
27	FJI	Fiji
28	FRA	France
29	GBR	United Kingdom
30	GHA	Ghana
31	GMB	Gambia, The
32	GRC	Greece
33	GTM	Guatemala
34	HND	Honduras
35	HUN	Hungary
36	IDN	Indonesia
37	IND	India
38	IRL	Ireland
39	IRN	Iran, Islamic Republic of
40	IRQ	Iraq
41	ISL	Iceland
42	ISR	Israel
43	ITA	Italy
44	JAM	Jamaica
45	JOR	Jordan
46	JPN	Japan
47	KEN	Kenya

	World Bank Country Code	Country Name
48	KOR	Korea, Republic of
49	LBR	Liberia
50	LKA	Sri Lanka
51	MEX	Mexico
52	MLI	Mali
53	MLT	Malta
54	MOZ	Mozambique
55	MUS	Mauritius
56	MWI	Malawi
57	MYS	Malaysia
58	NER	Niger
59	NIC	Nicaragua
60	NLD	Netherlands
61	NOR	Norway
62	NPL	Nepal
63	NZL	New Zealand
64	PAK	Pakistan
65	PAN	Panama
66	PER	Peru
67	PHL	Philippines
68	PNG	Papua New Guinea
69	POL	Poland
70	PRT	Portugal
71	PRY	Paraguay
72	RWA	Rwanda
73	SDN	Sudan
74	SEN	Senegal
75	SGP	Singapore
76	SLE	Sierra Leone
77	SLV	El Salvador
78	SWE	Sweden
79	SYR	Syrian Arab Republic
80	TGO	Togo
81	THA	Thailand
82	TTO	Trinidad and Tobago
83	TUN	Tunisia
84	TUR	Turkey
85	UGA	Uganda
86	URY	Uruguay
87	USA	United States
88	VEN	Venezuela, Bolivarian Rep. of
89	ZAF	South Africa
90	ZAR	Congo, Dem. Rep. of
91	ZMB	Zambia
92	ZWE	Zimbabwe

Figure A1. Dendrogram of Ward's cluster (natural resource excluded from the variables)

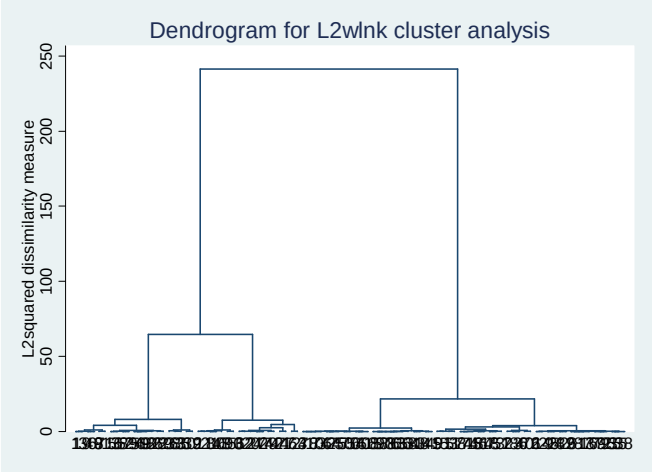


Table A2. Calinski and Harabasz and Duda and Hart stopping rules' result (arable land per capita excluded from the variable lists)

Calinski & Harabasz		Duda & Hart		
Numbers of Clusters	**Pseudo-F**	**Numbers of Clusters**	**Je(2)/Je(1)**	**Pseudo T-squared**
2	166.93	2	0.34	70.42
3	201.84	3	0.39	83.75
4	***207.08***	***4***	***0.52***	***17.82***
5	188.53	5	0.57	12.11
6	188.06	6	0.47	10.33
7	182.31	7	0.42	19.40
8	182.86	8	0.64	17.14
9	191.07	9	0.50	14.88
10	200.20	10	0.35	12.96
11	209.29	11	0.28	54.34
12	223.14	12	0.36	17.87
13	227.71	13	0.31	6.53
14	229.34	14	0.33	8.09
15	233.77	15	0.38	8.18

Table A3. Summary of clusters (arable land per capita excluded from the variable lists)

	Countries in Cluster	Number of countries	Capital Index	Human Capital Index	World Bank Income Group	World Bank Regional Group
	Benin				Low income	Sub-Saharan Africa
	Central African Republic				Low income	Sub-Saharan Africa
	Cameroon				Lower middle income	Sub-Saharan Africa
	Gambia, The				Low income	Sub-Saharan Africa
	Guatemala				Lower middle income	Latin America & Caribbean
	Iraq				Lower middle income	Middle East & North Africa
	Kenya				Low income	Sub-Saharan Africa
	Liberia				Low income	Sub-Saharan Africa
	Mali				Low income	Sub-Saharan Africa
	Mozambique				Low income	Sub-Saharan Africa
Cluster 1	Malawi	22	2 632	1.65	Low income	Sub-Saharan Africa
	Niger				Low income	Sub-Saharan Africa
	Nepal				Low income	South Asia
	Pakistan				Low income	South Asia
	Papua New Guinea				Low income	East Asia & Pacific
	Rwanda				Low income	Sub-Saharan Africa
	Sudan				Lower middle income	Sub-Saharan Africa
	Senegal				Low income	Sub-Saharan Africa
	Sierra Leone				Low income	Sub-Saharan Africa
	Togo				Low income	Sub-Saharan Africa
	Uganda				Low income	Sub-Saharan Africa
	Dem. Rep. of the Congo				Low income	Sub-Saharan Africa
	Australia				High income: OECD	..
	Austria				High income: OECD	..
	Canada				High income: OECD	..
	Switzerland				High income: OECD	..
	Germany				High income: OECD	..
	Denmark				High income: OECD	..
	Finland				High income: OECD	..
	France				High income: OECD	..
	United Kingdom				High income: OECD	..
Cluster 2	Ireland	20	98 827	8.86	High income: OECD	..
	Iceland				High income: OECD	..
	Israel				High income: non-OECD	..
	Italy				High income: OECD	..
	Japan				High income: OECD	..
	Netherlands				High income: OECD	..
	Norway				High income: OECD	..
	New Zealand				High income: OECD	..
	Singapore				High income: non-OECD	..
	Sweden				High income: OECD	..
	United States				High income: OECD	..
	Bahrain				High income: non-OECD	..
	Bolivia				Lower middle income	Latin America & Caribbean
	Brazil				Upper middle income	Latin America & Caribbean
	Botswana				Upper middle income	Sub-Saharan Africa
	China				Lower middle income	East Asia & Pacific
	Congo, Rep.				Lower middle income	Sub-Saharan Africa
	Colombia				Lower middle income	Latin America & Caribbean
	Costa Rica				Upper middle income	Latin America & Caribbean
	Dominican Republic				Lower middle income	Latin America & Caribbean
	Algeria				Lower middle income	Middle East & North Africa
	Egypt, Arab Rep.				Lower middle income	Middle East & North Africa
	Ghana				Low income	Sub-Saharan Africa
	Honduras				Lower middle income	Latin America & Caribbean

	Countries in Cluster	Number of countries	Capital Index	Human Capital Index	World Bank Income Group	World Bank Regional Group
	Indonesia				Lower middle income	East Asia & Pacific
	India				Lower middle income	South Asia
	Iran, Islamic Rep. of				Lower middle income	Middle East & North Africa
Cluster 3	Jamaica	32	15 043	3.90	Upper middle income	Latin America & Caribbean
	Jordan				Lower middle income	Middle East & North Africa
	Sri Lanka				Lower middle income	South Asia
	Lesotho				Lower middle income	Sub-Saharan Africa
	Mauritius				Upper middle income	Sub-Saharan Africa
	Nicaragua				Lower middle income	Latin America & Caribbean
	Portugal				High income: OECD	..
	Paraguay				Lower middle income	Latin America & Caribbean
	El Salvador				Lower middle income	Latin America & Caribbean
	Swaziland				Lower middle income	Sub-Saharan Africa
	Syrian Arab Republic				Lower middle income	Middle East & North Africa
	Thailand				Lower middle income	East Asia & Pacific
	Tunisia				Lower middle income	Middle East & North Africa
	Turkey				Upper middle income	Europe and Central Asia
	Zambia				Low income	Sub-Saharan Africa
	Zimbabwe				Low income	Sub-Saharan Africa
	Argentina				Upper middle income	Latin America & Caribbean
	Barbados				High income: non-OECD	..
	Chile				Upper middle income	Latin America & Caribbean
	Cyprus				High income: non-OECD	..
	Ecuador				Lower middle income	Latin America & Caribbean
	Spain				High income: OECD	..
	Fiji				Upper middle income	East Asia & Pacific
	Greece				High income: OECD	..
	Hungary				High income: OECD	..
	Korea, Rep.				High income: OECD	..
Cluster 4	Mexico	21	35 393	6.70	Upper middle income	Latin America & Caribbean
	Malta				High income: non-OECD	..
	Malaysia				Upper middle income	East Asia & Pacific
	Panama				Upper middle income	Latin America & Caribbean
	Peru				Lower middle income	Latin America & Caribbean
	Philippines				Lower middle income	East Asia & Pacific
	Poland				Upper middle income	Europe and Central Asia
	Trinidad and Tobago				High income: non-OECD	..
	Uruguay				Upper middle income	Latin America & Caribbean
	Venezuela (Bolivarian Rep. of)				Upper middle income	Latin America & Caribbean
	South Africa				Upper middle income	Sub-Saharan Africa

Table A4. Summary statistics of revealed factor intensity indices, year 2000 (HS classification)

Variable	Obs	Mean	Std. Dev.	Min	Max
rhci	5009	7.32	1.68	0.79	11.57
rci	5009	66 948	33 406	1 407	165 297
rnri_land	5009	0.57	0.32	0.10	4.84
rnri_nc	5009	14 380	8 823	1 859	89 591
rnri_sa	5009	4 679	5 858	7	74 197
rnri_pc	5009	6 744	3 549	854	46 780

Figure A2. Dendrogram of Ward's cluster analysis (HS classification)

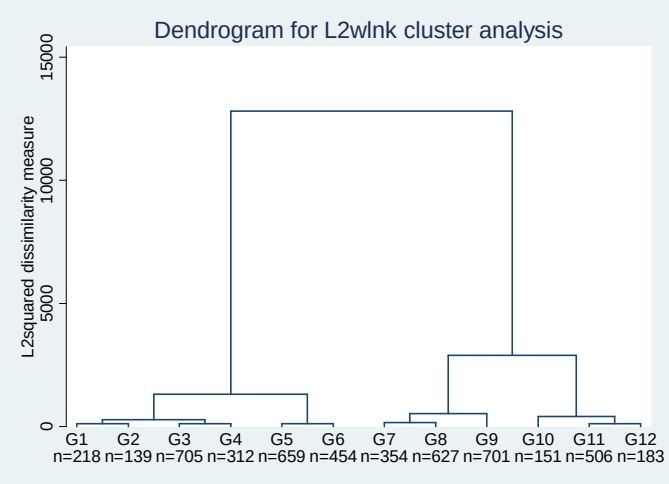


Table A5. Calinski and Harabasz and Duda and Hart stopping rules' result (HS classification)

Calinski & Harabasz		Duda & Hart		
Numbers of Clusters	Pseudo-F	Numbers of Clusters	Je(2)/Je(1)	Pseudo T-squared
2	8 881	2	0	3 903
3	9 068	3	0	2 821
4	***9 395***	***4***	***1***	***1 500***
5	8 807	5	0	1 068
6	8 658	6	1	709
7	8 493	7	1	673
8	8 118	8	1	555
9	7 698	9	1	606
10	7 443	10	1	442
11	7 328	11	1	353
12	7 336	12	1	489
13	7 191	13	1	630
14	7 085	14	1	540
15	7 044	15	1	296

Table A6. Summary statistics of the clusters
(HS classification)

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Revealed Capital Intensity Index	18 619	42 873	72 971	105 368
Revealed Human Capital Intensity Index	4.45	6.51	7.91	9.04
Number of Goods	769	1 333	1 559	1 348

Table A7. Percentages of HS sections' industries, by clusters
(HS classification)

HS sections	Section Descriptions	Cluster 1	Cluster 2	Cluster 3	Cluster 4
1	Live animals	4.7	4.7	3.8	2.7
2	Veg.	13.5	6.1	3.6	2.1
3	Fats and Oils	2.6	1.5	0.6	0.2
4	Bev. & Tobac.	4.6	5.6	3.6	1.1
5	Mineral	4.4	4.9	2.2	1.3
6	Chemical	5.9	9.6	17.5	22.9
7	Plastics	0.7	2.9	5.4	4.5
8	Leather	3.1	2.3	0.8	0.5
9	Wood	2.9	2.1	1.2	0.6
10	Paper	0.7	2.5	3.7	4.0
11	Textile	42.0	22.7	9.7	2.4
12	Footwear	1.6	2.9	0.3	0.0
13	Stone & Glass	1.0	3.9	2.6	2.8
14	Preciuos Stones	0.9	1.5	0.8	0.8
15	Base Metal	6.1	11.4	13.3	13.4
16	Machinery	2.6	6.6	18.8	26.8
17	Trans.Eq	0.8	1.7	3.5	3.6
18	Optical	0.4	3.0	4.7	8.4
19	Arms	0.1	0.1	0.5	0.5
20	Misc.	1.3	4.1	3.3	1.2
21	Works of Arts	0.3	0.2	0.2	0.0
		100.0	100.0	100.0	100.0

Table A8. Percentages of HS chapters' industries, by clusters (HS classification)

Sections	Chapters	Chapter Descriptions	Cluster 1	Cluster 2	Cluster 3	Cluster 4
1 Live animals	1	Live animals	0	0	0	0
	2	Meat and edible meat offal	1	1	1	1
	3	Fish and crustaceans, molluscs	3	2	1	2
	4	Dairy produce; birds eggs; natural honey	0	1	1	0
	5	Products of animal origin	0	1	0	0
2 Veg.	6	Live trees and other plants	1	0	0	0
	7	Edible vegetables	2	2	1	0
	8	Edible fruits and nuts	2	1	1	0
	9	Coffee, tea, maté and spices	3	1	0	0
	10	Cereals	1	0	0	0
	11	Products of milling industry	1	1	0	1
	12	Oil seeds and oleaginous fruits	2	1	1	1
	13	Lac; gums, resins	1	0	0	0
	14	Veg. planting materials	1	0	0	0
3 Fats and Oils	15	Animal or vegetable fats and oils	3	2	1	0
4 Bev. & Tobac.	16	Preparations of meat, of fish	0	1	0	0
	17	Sugars and sugar confectionery	0	1	0	0
	18	Cocoa and cocoa preparations	1	0	0	0
	19	Prep. of cereals, flour	0	0	1	0
	20	Prep. of vegetables, fruits, nuts	1	2	1	0
	21	Misc. edible preparations	0	0	1	0
	22	Beverages, spirits and vinegar	0	1	1	0
	23	Waste from food industries	1	1	0	0
	24	Tobacco	1	0	0	0
5 Mineral	25	Salt, sulfur, earths and stone	3	2	1	1
	26	Ores, slag and ash	1	1	1	0
	27	Mineral fuels, mineral oils	1	2	0	0
6 Chemical	28	Inorganic chemicals	2	3	4	4
	29	Organic chemicals	1	3	6	12
	30	Pharmaceutical products	0	0	1	1
	31	Fertilizers	1	1	0	0
	32	Tanning and dyeing extracts	0	1	1	1
	33	Essential oils and resinoids	1	1	1	0
	34	Soap, organic surface-active agents	1	0	1	0
	35	Albuminoidal substances	0	0	0	0
	36	Explosives; pyrotechnic products	0	0	0	0
	37	Photographic and cinematographic goods	0	0	1	2
	38	Miscellaneous chemical products	0	1	1	2
7 Plastics	39	Plastics and articles thereof	0	1	4	3
	40	Rubber and articles thereof	1	2	1	1
8 Leather	41	Raw hides and skins	2	1	0	0
	42	Leader of leather	1	1	0	0
	43	Furskins and artificial fur	0	0	1	1
9 Wood	44	Wood and articles of wood; wood charcoal	2	1	1	1
	45	Cork and articles of cork	0	1	0	0
	46	Manufactures of straw, of esparto	1	0	0	0
10 Paper	47	Pulp of wood	0	1	0	0
	48	Paper and paperboard	0	2	3	3
	49	Printed books, newspapers, pictures	0	0	1	0
11 Textile	50	Silk	0	1	0	0
	51	Wool, fine or coarse animal hair	0	1	1	0
	52	Cotton	10	3	1	0
	53	Other vegetable textile fibers	2	1	0	0
	54	Man-made filaments	1	2	1	0

Sections	Chapters	Chapter Descriptions	Cluster 1	Cluster 2	Cluster 3	Cluster 4
	55	Man-made staple fibers	4	3	2	1
	56	Wadding, felt and nonwovens	1	1	1	0
	57	Carpets	1	1	0	0
	58	Special woven fabrics	1	1	1	0
	59	Impregnated, coated, covered textile fabrics	0	0	1	1
	60	Knitted and crocheted fabrics	0	1	1	0
	61	Apparel and clothing	7	4	0	0
	62	Apparel and clothing, not knitted and crocheted	9	4	0	0
	63	Other textiles	5	1	0	0
	64	Footwear	1	2	0	0
12 Footwear	65	Headgear	0	0	0	0
	66	Umbrellas	0	1	0	0
	67	Prepared feathers	1	0	0	0
	68	Stone, plaster, cement	1	2	1	1
13 Stone & Glass	69	Ceramic	0	1	0	1
	70	Glass	0	1	2	1
14 Precious Stones	71	Precious stones	1	2	1	1
	72	Iron and steel	2	4	4	5
	73	Articles of iron or steel	1	2	3	2
	74	Copper	1	1	2	1
	75	Nickel	0	0	0	1
	76	Aluminum	0	1	1	0
15 Base Metal	78	Lead	0	0	0	0
	79	Zinc	0	0	0	0
	80	Tin	0	0	0	0
	81	Other base metals	0	0	0	2
	82	Tools, implements, cutlery	1	2	1	1
	83	Misc. articles of base metal	1	0	1	0
16 Machinery	84	Nuclear reactors	2	3	10	21
	85	Electrical MAshinery	0	3	9	5
	86	Railway and Tramway	0	0	1	1
17 Trans.Eq	87	Vehicles other than railway and tramway	0	1	3	1
	88	Aircraft, spacecraft	0	0	0	1
	89	Ships and boats	0	0	0	1
	90	Optical	0	1	3	6
18 Optical	91	Clocks and watches	0	1	1	2
	92	Musical instruments	0	1	1	0
19 Arms	93	Arms	0	0	1	1
	94	Furniture	1	1	1	0
20 Misc.	95	Toys and games	0	2	1	1
	96	Misc. manu.articles	0	2	1	0
21 Works of Arts	97	Works of art	0	0	0	0
			100	100	100	100

Appendix B. Dataset of index of revealed factor intensity

The dataset that contains the revealed factor intensity (RFI) indices is available on the UNCTAD website (http://r0.unctad.org/ditc/tab/index.shtm).

The dataset is available both in Stata and Excel. For each format, the dataset consists of three folders, titled (a) SITC; (ii) HS; and (c) RFII_1994and2000.

Folders – SITC and HS

The SITC folder contains the RFI indices for each good calculated at the finest disaggregated level of the SITC. Though the finest level of the SITC is a 5-digit level, not every 4-digit level is divided into 5 digits. Therefore our data consists of a mix of four and five digits. The indices are given in a separate file for each year (called year_sitc_indices.dta if in Stata, and year_sitc_indices.xml). The year coverage is from 1971 to 2003.

The HS folder contains the RFI indices for products classified at the HS 6-digit level. The year coverage is from 1988 to 2003.

For those indices calculated from a wide or "unbalanced" dataset, the filename is "unb_year_sitc_indices.dta".

Table B.1 (a and b) provide the description of all the variables included in these SITC and HS folders.

Table B1
a. Revealed factor intensity indices

a. SITC (file name = "year"_sitc_indices.dta or "year"_sitc_indices.xml)

	Variable name	Variable description
1	product	SITC code at either 4 or 5 digit level
2	productname	Corresponding product description
3	digit	Number of digit (either 4 of 5)
4	sitc45	SITC code at either 4 or 5 digit level, in string form
5	sitc1	Corresponding SITC 1 digit code
6	sitc1_desc	SITC 1 digit product description
7	sitc2	Corresponding SITC 2 digit code
8	sitc2_desc	SITC 2 digit product description
9	sitc3	Corresponding SITC 3 digit code
10	sitc3_desc	SITC 3 digit product description
11	Export	World Export of the product
12	Import	World Import of the product
13	rnci	Revealed Human Capital Intesity Index
14	rci	Revealed Psycical Capital Intesity Index
15	rnri_land	Revealed Natural Resource Intensity Index
16	percentage	Percentage of excluded exports, due to the lack of data, in total exports

Table B1
b. Revealed factor intensity indices database

b. HS (file name = hs_"year"_indices.dta or hs_"year"_indices.xml)

	Variable name	Variable description
1	year	year
2	product	HS code at six digit level
3	h0productname	Corresponding product description
4	sect	Corresponding HS section code
5	sect_desc	HS section description
6	hs2	Corresponding HS 2 digit code
7	Export	World Export of the product
8	Import	World Import of the product
9	rhci	Revealed Human Capital Intesity Index
10	rci	Revealed Psycical Capital Intesity Index
11	rnri_land	Revealed Natural Resource Intensity Index
12	percentage	Percentage of excluded exports, due to the lack of data, in total exports

Folder-RFI indices for the years 1994 and 2000

The RFI indices calculated using additional data on natural resources from the World Bank are given in a folder called "RFII_1994and2000". The World Bank data is available only for the years 1994 and 2000. The list of the variables and their descriptions are given in Table B.2 (a-b).

Table B2. Revealed factor intensity indices database (1994, 2000)

a. SITC

	Variable name	Variable description
1	product	SITC code at either 4 or 5 digit level
2	productname	Corresponding product description
3	digit	Number of digit (either 4 of 5)
4	sitc45	SITC code at either 4 or 5 digit level, in string form
5	sitc1	Corresponding SITC 1 digit code
6	sitc1_desc	SITC 1 digit product description
7	sitc2	Corresponding SITC 2 digit code
8	sitc2_desc	SITC 2 digit product description
9	sitc3	Corresponding SITC 3 digit code
10	sitc3_desc	SITC 3 digit product description
11	Export	World Export of the product
12	Import	World Import of the product
13	rnci	Revealed Human Capital Intesity Index
14	rci	Revealed Psycical Capital Intesity Index
15	rnri_land	Revealed Natural Resource Intensity Index (Arable Land)
16	rnri_nc	Revealed Natural Resource Intensity Index (Total Natural Capital)
17	rnri_sa	Revealed Natural Resource Intensity Index (Subsoil Assets)
18	rnri_pc	Revealed Natural Resource Intensity Index (Pastureland and Cropland)

(file name = 1994 _sitc_indices.dta or 1994_sitc_indices.xml)

(file name = 2000 _sitc_indices.dta or 2000_sitc_indices.xml)

b. HS (file name = hs_1994_indices.dta or hs_1994_indices.xml)

	Variable name	Variable description
1	year	year
2	product	HS code at six digit level
3	h0productname	Corresponding product description
4	sect	Corresponding HS section code
5	sect_desc	HS section description
6	hs2	Corresponding HS 2 digit code
7	Export	World Export of the product
8	Import	World Import of the product
9	rnci	Revealed Human Capital Intesity Index
10	rci	Revealed Psycical Capital Intesity Index
11	rnri_land	Revealed Natural Resource Intensity Index (Arable Land)
12	rnri_nc	Revealed Natural Resource Intensity Index (Total Natural Capital)
13	rnri_sa	Revealed Natural Resource Intensity Index (Subsoil Assets)
14	rnri_pc	Revealed Natural Resource Intensity Index (Pastureland and Cropland)

(file name = hs_1994_indices.dta or hs_1994_indices.xml)

(file name = hs_2000_indices.dta or hs_2000_indices.xml)

Files – country endowments

In addition, we have attached our newly constructed database on countries endowments (called “endowments_all.dta” and “endowments1994_2000.dta”; the same file name for Excel). Table B.3 (a-b) presents the list of variables and the descriptions.

Table B3

a. Endowment database, by country

	Variable name	Variable description
1	isocode	PWT 6.2: Country Code
2	countryname	Country name
3	year	Year
4	phys_cap_pw	Physical Capital Stock per Worker
5	hum_cap	Average Years of Schooling
6	land_pw	Arable Land hectares per worker
7	workers	Number of Workers
8	phys_cap	Physical Capital Stock
9	land	Arable Land hectares
10	region	World Bank Region Classification
11	income	World Bank Income Classification
12	group	World Bank Income Classification

b. Endowment database, by country (1994 and 2000)
(with data on natural capital and its components from the World Bank)

	Variable name	Variable description
1	isocode	PWT 6.2: Country Code
2	countryname	Country name
3	year	Year
4	workers	Number of workers
5	phys_cap_pw	Physical Capital Stock per Worker
6	hum_cap	Average Years of Schooling
7	land_pw	Arable Land hectares per worker
8	nc	Natural Capital, $ per worker
9	sa	Subsoil Assets, $ per worker
10	tr	Timber Resources, $ per worker
11	ntr	Non Timber Resources, $ per worker
12	pa	Protected Areas, $ per worker
13	p	Pastureland, $ per worker
14	c	Cropland, $ per worker
15	pc	Pastureland and Cropland, $ per worker

References

Anderson K (2008). Distorted agricultural incentives and economic development: Asia's experience. CEPR DP 6914.

Anderson K, Kurzweil M, Martin W, Sandri D and Valenzuela E (2008). Measuring distortions to agricultural incentives, revisited. *World Trade Review*. 7(4): 675–704.

Balassa B (1965). Trade liberalization and "revealed" comparative advantage. *The Manchester School of Economic and Social Studies*. 33: 99–123.

Barro RJ and Lee JW (1993). International comparisons of educational attainment. *Journal of Monetary Economics*. 32: 5–23.

Barro RJ and Lee JW (2000). International data on educational attainment: updates and implications. *CID Working Paper*. No. 42. Harvard University.

Barro RJ and Lee JW (2001). International data on educational attainment: updates and implications. *Oxford Economic Papers*. 53(3): 541–563.

Carrère C, Strauss-Kahn V and Cadot O (2007). Export diversification: what's behind the hump? Mimeo, University of Lausanne.

Cohen D and Soto M (2001). Growth and human capital: good data, good results. OECD Development Centre Working Papers. 179, OECD Development Centre.

De La Fuente A and Doménech R (2001). Educational attainment in the OECD, 1960-1995. CEPR Discussion Paper, 3416.

Easterly W, Kremer M, Pritchett L and Summers LH (1993). Good policy or good luck: country growth performance and temporary shocks. *Journal of Monetary Economics,* 32(3), 459-483.

Easterly W and Levine R (2001). It's not factor accumulation: stylized facts and growth models. *The World Bank Economic Review*. 15(2): 177–219.

Feenstra R (2004). *Advanced International Trade*. MIT Press.

Hausmann R, Hwang J and Rodrik D (2007). What you export matters. *Journal of Economic Growth.* 12: 1–25.

Hausmann R and Klinger B (2006). Structural transformation and patterns of comparative advantage in the product space. Mimeo.

Imbs J and Wacziarg R (2003). Stages of Diversification. American Economic Review. 93(1): 63-86.

Nehru V and Dhareshwar A (1993). A new database on physical capital stock: sources, methodology and results. *Revista de Analisis Economico*. 8(1): 37–59.

Nehru V, Swanson E and Dubey A (1993). A new database on human capital stock: sources, methodology, and results. Policy Research Working Paper 1124. World Bank. Washington, D.C.

Klenow P and Rodriguez-Clare A (1997). The neoclassical revival in growth economics: has it gone too far? In *National Bureau of Economic Research Macroeconomics Annual 1997*, edited by Ben S.Bernanke and Julio Rotemberg: 73–103, Cambridge, MA: MIT Press.

Klinger B and Lederman D (2005). Diversification, innovation and imitation off the global technological frontier. Mimeo.

Krueger A and Lindahl M (2001). Education for growth: why and for whom? *Journal of Economic Literature*. American Economic Association. 39(4): 1101–1136.

Kyriacou GA (1991). Level and growth effects of human capital: a cross-country study of the convergence hypothesis. Unpublished. New York University.

Larson DF, Butzer R, Mundlak Y and Crego A (2000). A cross-country database for sector investment and capital. *World Bank Economic Review*. 14: 371–391.

Lau L, Jamison D and Louat F (1991). Education and productivity in developing countries: an aggregate production function approach. Working Paper Series, 612, World Bank.

Lau L, Bhalla S and Louat F (1991). Human and physical capital stock in developing countries: construction of data and trends. Draft mimeo. World Development Report, World Bank.

OECD (2007). Producer and consumer support estimates, OECD database 1986–2006. Available at www.oecd.org/document/59/0,3343,en _2649_33727_39551355_1_1_1_1,00.html.

OECD (2008). Agricultural policies in OECD countries: at a glance.

Psacharopoulos G and Ariagada AM (1986). The educational composition of the labor force: an international comparison. *International Labor Review*. 125: 561–574.

Vollrath TL (1987). Revealed competitiveness for wheat. Economic Research Service Staff Report, No AGES861030, U.S. Dept. of Agriculture, Economic Research Service. Washington, D.C.

Vollrath TL (1989). Competitiveness and protection in world agriculture. Agricultural Information Bulletin, 567, United States Dept. of Agriculture, Economic Research Service. Washington, D.C.

World Bank (1997). Expanding the measure of wealth: indicators of environmentally sustainable development.

World Bank (2006). Where is the wealth of nations? Measuring capital for the twenty-first century.

UNCTAD Study Series on

POLICY ISSUES IN INTERNATIONAL TRADE AND COMMODITIES

No. 1 Erich Supper, **Is there effectively a level playing field for developing country exports?**, 2001, 138 p. Sales No. E.00.II.D.22.

No. 2 Arvind Panagariya, **E-commerce, WTO and developing countries**, 2000, 24 p. Sales No. E.00.II.D.23.

No. 3 Joseph Francois, **Assessing the results of general equilibrium studies of multilateral trade negotiations**, 2000, 26 p. Sales No. E.00.II.D.24.

No. 4 John Whalley, **What can the developing countries infer from the Uruguay Round models for future negotiations?**, 2000, 29 p. Sales No. E.00.II.D.25.

No. 5 Susan Teltscher, **Tariffs, taxes and electronic commerce: Revenue implications for developing countries**, 2000, 57 p. Sales No. E.00.II.D.36.

No. 6 Bijit Bora, Peter J. Lloyd, Mari Pangestu, **Industrial policy and the WTO**, 2000, 47 p. Sales No. E.00.II.D.26.

No. 7 Emilio J. Medina-Smith, **Is the export-led growth hypothesis valid for developing countries? A case study of Costa Rica**, 2001, 49 p. Sales No. E.01.II.D.8.

No. 8 Christopher Findlay, **Service sector reform and development strategies: Issues and research priorities**, 2001, 24 p. Sales No. E.01.II.D.7.

No. 9 Inge Nora Neufeld, **Anti-dumping and countervailing procedures – Use or abuse? Implications for developing countries**, 2001, 33 p. Sales No. E.01.II.D.6.

No. 10 Robert Scollay, **Regional trade agreements and developing countries: The case of the Pacific Islands' proposed free trade agreement**, 2001, 45 p. Sales No. E.01.II.D.16.

No. 11 Robert Scollay and John Gilbert, **An integrated approach to agricultural trade and development issues: Exploring the welfare and distribution issues**, 2001, 43 p. Sales No. E.01.II.D.15.

No. 12 Marc Bacchetta and Bijit Bora, **Post-Uruguay round market access barriers for industrial products**, 2001, 50 p. Sales No. E.01.II.D.23.

No. 13 Bijit Bora and Inge Nora Neufeld, **Tariffs and the East Asian financial crisis**, 2001, 30 p. Sales No. E.01.II.D.27.

No. 14 Bijit Bora, Lucian Cernat, Alessandro Turrini, **Duty and quota-free access for LDCs: Further evidence from CGE modelling**, 2002, 130 p. Sales No. E.01.II.D.22.

No. 15 Bijit Bora, John Gilbert, Robert Scollay, **Assessing regional trading arrangements in the Asia-Pacific**, 2001, 29 p. Sales No. E.01.II.D.21.

No. 16 Lucian Cernat, **Assessing regional trade arrangements: Are South-South RTAs more trade diverting?**, 2001, 24 p. Sales No. E.01.II.D.32.

No. 17 Bijit Bora, **Trade related investment measures and the WTO: 1995-2001**, 2002.

No. 18 Bijit Bora, Aki Kuwahara, Sam Laird, **Quantification of non-tariff measures**, 2002, 42 p. Sales No. E.02.II.D.8.

No. 19 Greg McGuire, **Trade in services – Market access opportunities and the benefits of liberalization for developing economies**, 2002, 45 p. Sales No. E.02.II.D.9.

No. 20 Alessandro Turrini, **International trade and labour market performance: Major findings and open questions**, 2002, 30 p. Sales No. E.02.II.D.10.

No. 21 Lucian Cernat, **Assessing south-south regional integration: Same issues, many metrics**, 2003, 32 p. Sales No. E.02.II.D.11.

No. 22 Kym Anderson, **Agriculture, trade reform and poverty reduction: Implications for Sub-Saharan Africa**, 2004, 30 p. Sales No. E.04.II.D.5.

No. 23 Ralf Peters and David Vanzetti, **Shifting sands: Searching for a compromise in the WTO negotiations on agriculture**, 2004, 46 p. Sales No. E.04.II.D.4.

No. 24 Ralf Peters and David Vanzetti, **User manual and handbook on Agricultural Trade Policy Simulation Model (ATPSM)**, 2004, 45 p. Sales No. E.04.II.D.3.

No. 25 Khalil Rahman, **Crawling out of snake pit: Special and differential treatment and post-Cancun imperatives**, 2004.

No. 26 Marco Fugazza, **Export performance and its determinants: Supply and demand constraints**, 2004, 57 p. Sales No. E.04.II.D.20.

No. 27 Luis Abugattas, **Swimming in the spaghetti bowl: Challenges for developing countries under the "New Regionalism"**, 2004, 30 p. Sales No. E.04.II.D.38.

No. 28 David Vanzetti, Greg McGuire and Prabowo, **Trade policy at the crossroads – The Indonesian story**, 2005, 40 p. Sales No. E.04.II.D.40.

No. 29 Simonetta Zarrilli, **International trade in GMOs and GM products: National and multilateral legal frameworks**, 2005, 57 p. Sales No. E.04.II.D.41.

No. 30 Sam Laird, David Vanzetti and Santiago Fernández de Córdoba, **Smoke and mirrors: Making sense of the WTO industrial tariff negotiations**, 2006, Sales No. E.05.II.D.16.

No. 31 David Vanzetti, Santiago Fernandez de Córdoba and Veronica Chau, **Banana split: How EU policies divide global producers**, 2005, 27 p. Sales No. E.05.II.D.17.

No. 32 Ralf Peters, **Roadblock to reform: The persistence of agricultural export subsidies**, 2006, 43 p. Sales No. E.05.II.D.18.

No. 33 Marco Fugazza and David Vanzetti, **A South-South survival strategy: The potential for trade among developing countries**, 2006, 25 p.

No. 34 Andrew Cornford, **The global implementation of Basel II: Prospects and outstanding problems**, 2006, 30 p.

No. 35 Lakshmi Puri, IBSA: **An emerging trinity in the new geography of international trade**, 2007, 50 p.

No. 36 Craig VanGrasstek, **The challenges of trade policymaking: Analysis, communication and representation**, 2008, 45 p.

No. 37 Sudip Ranjan Basu, **A new way to link development to institutions, policies and geography**, 2008, 50 p.

No. 38 Marco Fugazza and Jean-Christophe Maur, **Non-tariff barriers in computable general equilibrium modelling**, 2008, 25 p.

No. 39 Alberto Portugal-Perez, **The costs of rules of origin in apparel: African preferential exports to the United States and the European Union**, 2008, 35 p.

No. 40 Bailey Klinger, **Is South-South trade a testing ground for structural transformation?**, 2009, 30 p.

No. 41 Sudip Ranjan Basu, Victor Ognivtsev and Miho Shirotori, **Building trade-relating institutions and WTO accession**, 2009, 50 p.

No. 42 Sudip Ranjan Basu and Monica Das, **Institution and development revisited: A nonparametric approach**, 2010, 26 p.

No. 43 Marco Fugazza and Norbert Fiess, **Trade liberalization and informality: New stylized facts**, 2010, 45 p.

No. 44 Miho Shirotori, Bolormaa Tumurchudur and Olivier Cadot, **Revealed factor intensity indices at the product level,** 2010, 55 p.

United Nations Publications

All orders from **North America, Latin America, the Caribbean and Asia and the Pacific** should be sent to:

United Nations Publications
Room DC2-853, 2 UN Plaza
New York, NY 10017, USA
Telephone: (212) 963-8302, Toll Free 1-800-253-9646 (North America only)
Fax: (212) 963-3489
E-mail: publications@un.org

Customers in **Europe, Africa and the Middle East** should send their orders to:

Section des Ventes et Commercialisation
Bureau E-4, CH-1211
Geneva 10, Switzerland
Telephone: 41 (22) 917-2613/2614
Fax: 41 (22) 917-0027
E-mail: unpubli@unog.ch

For further information, please visit: *http://www.unctad.org/tab*

QUESTIONNAIRE

UNCTAD Study series on

POLICY ISSUES IN INTERNATIONAL TRADE AND COMMODITIES

(Study series no. 44: Revealed factor intensity indices at the product level)

Readership Survey

Since 1999, the Trade Analysis Branch of the Division on International Trade in Goods and Services, and Commodities of UNCTAD has been carrying out policy-oriented analytical work aimed at improving the understanding of current and emerging issues in international trade and development. In order to improve the quality of the work of the Branch, it would be useful to receive the views of readers on this and other similar publications. It would therefore be greatly appreciated if you could complete the following questionnaire and return to:

Trade Analysis Branch, DITC
Rm. E-8076
United Nations Conference on Trade and Development
Palais des Nations
CH-1211 Geneva 10, Switzerland

1. Name and address of respondent (optional):

_____

_____

_____

_____

2. Which of the following describes your area of work?

Government	☐	Public enterprise	☐
Private enterprise institution	☐	Academic or research	☐
International organization	☐	Media	☐
Not-for-profit organization	☐	Other (specify) _____	

3. In which country do you work? _____

4. Did you find this publication to your work? ☐ Very useful ☐ Of some use ☐ Little use

5. What is your assessment of the contents of this publication?
☐ Excellent ☐ Good ☐ Adequate ☐ Poor

6. Other comments:

_____

_____

_____

_____