Statistical Appendix

This Appendix contains two sections. Section A presents the calculations of the relative skill intensity of sectors mentioned in the paper. Section B provides supplementary information on the figures and tables in the paper, including a list of the countries included. The data for the figures and tables are available in an Excel workbook discussed at the end of section B.

A. Sectoral skill intensity calculations

The calculations make use of the Socioeconomic Accounts of the World Input-Output Database 2013 (WIOD), http://www.wiod.org/database/seas13, which contain data on employment (of employees and the self-employed, not distinguished by gender), measured in hours, for three skill categories in 35 industries in 40 countries. WIOD data run from 1995 to 2009, but these calculations refer to a single year, 2000, in the middle of the period covered by the paper.

Skill, measured by schooling, follows the International Standard Classification of Education (ISCED). 'Low skilled' workers are ISCED categories 0, 1 and 2 (everything below completed upper secondary). 'Medium skilled' are ISCED categories 3 and 4 (complete upper secondary and some tertiary, but below a college degree), and 'high skilled' are categories 5 and 6 (a 2–4 year college degree, or its vocational equivalent, and above). The present calculations use two measures of skill intensity: the ratio of high to medium-plus-low (H/(M + L)), and the ratio of high-plus-medium to low (denoted by (H + M)/L).

The 35 WIOD industries are aggregated into a smaller number of sectors that correspond as closely as possible with the classification used in the paper. The three broad sectors are:

Primary (NP, where N stands for 'narrow', excluding processed primary products): agriculture, hunting, forestry and fishing; mining and quarrying (2 WIOD industries).

Manufacturing (BM, where B stands for 'broad', including processed primary products): all 14 WIOD manufacturing industries (from 15t16 to 36t37).

Services (SVS, which are everything except NP and BM, referred to collectively as GOODS): the other 19 WIOD industries.

Broad manufacturing is divided into two subsectors:

Processed primary (PP): food, beverages and tobacco; coke, refined petroleum and nuclear fuel (2 WIOD industries).

Narrow manufacturing (NM), which is further divided into:

Skill intensive (NMH, where H stands for 'high'): chemicals; machinery n.e.c.; electrical and optical equipment; transport equipment (4 WIOD industries).

Labour-intensive (more precisely, unskilled-labour intensive, NML, where L stands for 'low'): textiles; leather and footwear; wood and cork; pulp, paper, printing and publishing; rubber and plastics; other non-metallic minerals; basic metals and fabricated metal; manufacturing n.e.c., recycling (8 WIOD industries)

The two measures of skill intensity are compared for pairs of sectors relevant to the paper by calculating five ratios: BM/NP, NM/PP, PP/NP, NMH/NML, SVS/GOODS. This is done for each of the 40 countries individually, and for regional aggregates of countries that correspond as closely as possible to those used in the paper:

Land-abundant OECD. Australia*, Canada, Finland, Sweden, United States.

Land-scarce OECD. Austria, Belgium, Denmark, France, Germany, Greece*, Ireland*, Italy*, Japan, Netherlands, Portugal*, Spain, United Kingdom (Cyprus, Luxemburg and Malta are omitted because they are below the size limit for inclusion in the paper).

China. China.

Other East Asia. Indonesia*, South Korea, Taiwan.

Former Soviet sphere. Bulgaria*, Czech Republic, Estonia*, Hungary, Latvia*, Lithuania*, Poland, Romania*, Russia, Slovak Republic, Slovenia.

Latin America. Brazil, Mexico.

Middle East & North Africa. Turkey*.

India. India.

Other South Asia. None.

Sub-Saharan Africa. None.

The coverage of the WIOD data is biased strongly towards more developed (including FSS) countries, with only 8 developing countries included, none of which is in Other South Asia or sub-Saharan Africa. In the 12 countries with a *, moreover, there is no recorded variation of skill intensity among individual manufacturing (BM) industries, so these countries are omitted from the regional calculations of the three ratios that use intra-manufacturing data: NM/PP, PP/NP, NMH/NML.[1]

[1] This lack of variation is likely to be due to the data being derived from censuses or labour force surveys in which all manufacturing industries are classified as a single sector.

The regional results are in Table A.1 (the country-level results are in Relative skill intensities WIOD.xlsx). Manufacturing is more skill-intensive than primary production in every region, on both measures of skill intensity, though the difference is smaller in more developed regions (a ratio of about two) than in developing regions (where in most cases the ratio is above five).[2] Services, which include industries such as education and health, are more skill-intensive than goods in all countries, but again the difference is larger in developing regions (where much of goods sector employment consists of peasant agriculture).

Within manufacturing, narrow manufacturing is in a majority of cases somewhat more skill-intensive than primary processing, but not in the land-scarce OECD, China and Other East Asia (or in 15 out of 35 individual countries with the necessary data). In all regions, though, primary processing is more skill-intensive than (narrow) primary production, with the differences again being smaller in more developed regions than in developing regions.[3]

The measured relative skill intensity of 'skill-intensive' narrow manufacturing exceeds that of 'labour-intensive' narrow manufacturing in all regions, but the regional pattern is complicated, with the differences being larger in the land-abundant OECD, China, and India than in the land-scarce OECD, Other East Asia and Latin America. Which of the two measures of skill intensity shows the greater difference also varies among regions: they are similar in both OECD regions and Other East Asia, but the difference is larger for H/(M + L) than (H + M)/L in China and India, and vice versa in Latin America.

In 15 individual countries, 14 of which are European, 'skill-intensive' manufacturing is similar in skill intensity to 'labour-intensive' manufacturing or (in four cases) of lower skill intensity.[4] This is consistent with the fragmentation of manufacturing discussed in the paper: in developed countries, even by 2000, the most labour-intensive goods and activities in 'labour-intensive' manufacturing sectors had disappeared, especially in Europe, whose labour market and social security institutions kept unskilled wages relatively high.

[2] The only exceptions at country level are Germany, where there is little difference, and South Korea, where primary production is much more skill-intensive than manufacturing (because most workers in agriculture are recorded as having a college degree or its vocational equivalent).
[3] The exceptions at country level, as with the BM/NP ratio, are Germany and South Korea.
[4] The countries with little difference in skill intensity between NMH and NML are Austria, Belgium, Germany, Denmark, Finland, France, Hungary, Poland, Slovenia, Sweden and United Kingdom. Those where NMH is less skill intensive than NML are the Netherlands, Spain and Slovakia, as well as Taiwan, where the pattern probably mainly reflects concentration within NMH on activities of relatively low skill intensity.

Table A.1 Relative skill-intensity of sectors by region, 2000

Land-abundant OECD	H/(M+L)	(H+M)/L
BM/NP	1.9	1.9
NM/PP	1.3	1.6
PP/NP	1.5	1.5
NMH/NML	2.0	1.9
SVS/GOODS	1.9	1.6

Land-scarce OECD	H/(M+L)	(H+M)/L
BM/NP	1.9	2.2
NM/PP	1.0	1.0
PP/NP	2.1	2.9
NMH/NML	1.2	1.2
SVS/GOODS	1.9	1.8

China	H/(M+L)	(H+M)/L
BM/NP	17.7	11.8
NM/PP	0.9	0.8
PP/NP	19.8	14.1
NMH/NML	3.5	2.1
SVS/GOODS	9.2	5.9

Other East Asia	H/(M+L)	(H+M)/L
BM/NP	2.3	4.7
NM/PP	1.0	0.9
PP/NP	4.8	12.4
NMH/NML	1.2	1.2
SVS/GOODS	3.3	2.8

Former Soviet sphere	H/(M+L)	(H+M)/L
BM/NP	1.5	1.9
NM/PP	1.4	1.2
PP/NP	1.1	3.0
NMH/NML	1.2	1.3
SVS/GOODS	4.0	3.5

Latin America	H/(M+L)	(H+M)/L
BM/NP	7.7	8.1
NM/PP	1.4	1.4
PP/NP	5.8	6.1
NMH/NML	1.3	1.9
SVS/GOODS	4.2	3.5

Middle East & North Africa	H/(M+L)	(H+M)/L
BM/NP	13.2	6.5
NM/PP	na	na
PP/NP	na	na
NMH/NML	na	na
SVS/GOODS	9.3	5.6

India	H/(M+L)	(H+M)/L
BM/NP	5.2	2.6
NM/PP	2.3	1.8
PP/NP	2.7	1.6
NMH/NML	4.8	2.9
SVS/GOODS	6.5	3.3

Source: calculations using WIOD socioeconomic accounts described in appendix text. As in the paper, the regions are aggregates of the countries within them.

B. Supplementary information on figures and tables

This section lists the countries included in each region, provides supplementary information on the sources and calculations used for each figure and table in the paper, and describes the data in an accompanying Excel workbook for future users.

Country coverage of regions

The data cover only countries whose total populations exceeded 1 million in 1990. During the period, some countries split up or united, so to make comparisons between the start and end of the period, the data for these countries had to be added together (on the same principles as for the addition of countries into regional aggregates), yielding a total of 130 countries.[5]

The 130 'comparable' countries are grouped into ten regions, according to the membership of these regions in 1985. The 'OECD', however, omits Turkey (shifted to MENA), and is divided between land-abundant and land-scarce countries (with 'abundance' defined as ≥ 3 km^2 of land area per 100 adults (15+) in 1990). The Former Soviet Sphere includes the USSR and Eastern Europe. The developing regions follow the World Bank's classification, except that China and India are separated from their respective regions.[6] The membership of the regions (subject to the availability of data, discussed later) is as follows.

1. *Land-abundant OECD* (7). Australia, Canada, Finland, New Zealand, Norway, Sweden, United States.

2. *Land-scarce OECD* (14). Austria, Belgium, Denmark, France, Germany (adding together in 1985 East Germany and West Germany), Greece, Ireland, Italy, Japan, Netherlands, Portugal, Spain, Switzerland, United Kingdom.

3. *China* (1).

4. *Other East Asia* (15). Cambodia, Hong Kong, Indonesia, North Korea, South Korea, Laos, Malaysia, Mongolia, Myanmar, Papua New Guinea, Philippines, Singapore, Taiwan, Thailand, Vietnam.

5. *Former Soviet sphere* (8). Albania, Bulgaria, Czechoslovakia (adding together in 2014 the Czech Republic and the Slovak Republic), Hungary, Poland, Romania, USSR, Yugoslavia. In 2014, the 'USSR' adds together Armenia, Azerbaijan, Belarus, Estonia, Georgia, Kazakhstan, Kyrgyz Republic, Latvia, Lithuania, Moldova, Russian Federation, Tajikistan, Turkmenistan, Ukraine and

[5] At the start, there were 132 non-small countries, of which the two Germanys and the two Yemens were united by the end of the period, reducing the number to 130. At the end, the number had risen to 152, with the addition of three countries through division of Czechoslovakia, Ethiopia and Sudan, five countries through division of Yugoslavia (excluding small Montenegro), and 14 countries though division of the Soviet Union.

[6] The main difference is that developing regions here include some countries that the World Bank now classifies as high-income: Hong Kong, South Korea, Singapore (Other East Asia); Chile, Puerto Rico, Uruguay, Trinidad and Tobago (Latin America); Israel, Kuwait, Oman, Saudi Arabia, United Arab Emirates (MENA).

Uzbekistan; while 'Yugoslavia' adds together Bosnia and Herzegovina, Croatia, Kosovo, Macedonia, Serbia and Slovenia.

6. *Latin America* (23). Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Dominican Republic, Ecuador, El Salvador, Guatemala, Haiti, Honduras, Jamaica, Mexico, Nicaragua, Panama, Paraguay, Peru, Puerto Rico, Trinidad and Tobago, Uruguay, Venezuela.

7. *Middle East & North Africa* (18). Algeria, Egypt, Iran, Iraq, Israel, Jordan, Kuwait, Lebanon, Libya, Morocco, Oman, Saudi Arabia, Syrian Arab Republic, Tunisia, Turkey, United Arab Emirates, West Bank and Gaza, Yemen (adding together in 1985 the Yemen Arab Republic and the People's Democratic Republic of Yemen).

8. *India* (1).

9. *Other South Asia* (5). Afghanistan, Bangladesh, Nepal, Pakistan, Sri Lanka.

10. *Sub-Saharan Africa* (38). Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Central African Republic, Chad, Congo, Dem. Rep., Congo, Rep., Cote d'Ivoire, Ethiopia (adding together in 2014 Eritrea and Ethiopia), Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Mozambique, Namibia, Niger, Nigeria, Rwanda, Senegal, Sierra Leone, Somalia, South Africa, Sudan (adding together in 2014 Sudan and South Sudan), Tanzania, Togo, Uganda, Zambia, Zimbabwe.

In the following discussion of the coverage of data for each region, the 'omission' of countries is with reference to the 130 countries listed above, and takes for granted the omission of those not included in this list (either because they are small or because they are aggregated with other countries for comparability between 1985 and 2014).

The figures and tables are discussed below in the order of their appearance in the paper. In all cases, data for Taiwan are obtained from Taiwan's own statistical publications rather than from the standard international sources cited in the notes to the figures and tables.

Figure 1. Regional endowment ratios, 1985
Figure 2. Changes in regional endowment ratios, 1985-2010

The nine countries for which gaps in the Barro and Lee average years of schooling data are filled with estimates from UNESCO literacy data are: Oman (MENA); Angola, Burkina Faso, Chad, Ethiopia, Guinea, Guinea-Bissau, Madagascar and Nigeria (sub-Saharan Africa). The method is described in section 2.1 of Caselli.pdf. After filling these gaps, the regional aggregates in these figures omit for lack of education data the following five countries: North Korea (Other East Asia); Puerto Rico (Latin America); Lebanon and West Bank and Gaza (MENA); and Somalia (sub-Saharan Africa).[7]

[7] They also omit, in the FSS region in both years, Azerbaijan, Belarus, Georgia, Turkmenistan and Uzbekistan (within the USSR) and Bosnia, Kosovo and Macedonia (within Yugoslavia).

Figure 3. Trade share of GDP at constant prices, 1985 and 2014 (%)

The regional aggregates in this figure omit no countries. The UN source was supplemented for Czechoslovakia and the USSR in 1985 and for Taiwan in both years with estimates from other sources based on trade shares at current prices.

Table 1. Exports by broad sector, 2014 levels and 1985-2014 changes

To divide SITC data on merchandise exports between manufactures and primary products in a way that is consistent with the ISIC definitions of manufacturing and primary in output and employment data, each three-digit category in SITC 0-4 is classified as either 'unprocessed' or 'processed'. The unprocessed items are 'primary' exports, while the processed items are added to SITC 5-9 as 'manufactured' exports. The code classification is as in section 4 of Caselli.pdf, with the exception of India's cut-diamond exports, which accounted for 16% of its merchandise exports in 2000 and are classified as labour-intensive manufactures rather than primary.[8] The share of services was calculated from WTO data on merchandise exports and services exports, supplemented for some countries by national accounts and balance of payments data.

Sectoral shares of the domestic value added content of exports (goods and services) are derived by combining (a) sectoral shares of gross exports (calculated from UNCTAD and WTO data) with (b) the share of value added in the gross exports of each sector (using OECD/WTO TiVA data). The estimates for 1985 use TiVA data for 1995, the earliest available year.

The gross export data cover almost all countries (as detailed below), but the TiVA data cover only about 50 countries (the omissions are detailed below). To estimate regional aggregates, calculations of (a) and (b) are first done for each country with the necessary data. Estimates of (a) and (b) for each region are then derived separately from the country-level values, and finally combined to obtain regional-level estimates of sectoral shares (on the assumption that countries with data for (a) or (b) are, when aggregated, representative of their regions).

Gross export data are missing for: North Korea (Other East Asia), Albania (FSS), Puerto Rico (Latin America), Lebanon, West Bank & Gaza, and Yemen (MENA), and Angola, Botswana, Lesotho, Namibia, and Somalia (Sub-Saharan Africa).[9] In all cases these gaps are due to lack of UNCTAD data on merchandise exports.

TiVA data are missing for: North Korea, Laos, Mongolia, Myanmar, Papua New Guinea (Other East Asia); Albania and Yugoslavia (FSS); in Latin America, all countries except Argentina, Brazil, Chile, Colombia, Costa Rica and Mexico; in MENA, all countries except Israel, Saudi Arabia,

[8] This exception is as in Wood and Calandrino (2000), and differs from Wood and Mayer (2011). Cut-diamond exports are SITC 66729. SITC 667 exports are classified as (narrow) primary in all other cases.
[9] They are also missing within the FSS in 2014 for Azerbaijan, Georgia, Kyrgyz Republic and Turkmenistan (USSR) and Bosnia, Kosovo, Macedonia and Serbia (Yugoslavia).

Tunisia and Turkey; in Other South Asia, all countries (the regional aggregate uses the data for India); and in Sub-Saharan Africa, all countries except South Africa.[10]

Figure 4. Manufactured/primary exports and skill/land endowments, 2014

The omissions from the regional endowment aggregates are as specified in the discussion of figures 1 and 2 above. The omissions from the regional export (domestic value added content) aggregates are as specified in the discussion of table 1 above. The cross-country regression is based on data that split Czechoslovakia, the USSR and Yugoslavia into separate countries.

Table 2. Sectoral export ratios regressed on endowment ratios

Eleven countries are omitted from all the regressions in this table: they all lack (gross) export data, as detailed in the discussion of table 1 above, and five of them also lack schooling data, as detailed in the discussion of figures 1 and 2 above. A further 37 countries with low narrow manufactured export shares are omitted from the regressions in panel D.

Table 3. Manufactured export composition, 2014 levels and 1985-2014 changes

The regional aggregates in this table omit the countries listed in the discussion of table 1 above that lack gross export data.

Following Wood and Mayer (2011) and as in Caselli.pdf, SITC categories are assigned to skill-intensive and labour-intensive narrow manufactures in the following way:

Labour-intensive (NML)	SITC-Rev 2 categories
Leather and rubber products	61–62
Wood and paper products	63–64
Textiles, clothing, travel goods and footwear	65, 83–85
Non-metallic mineral products, excl. precious stones	66 less 667
Iron and steel and metal products	67, 69
Furniture and plumbing equipment	81–82
Ships, bicycles and trains	78 (less 781-784), 79 (less 792)
Miscellaneous	89, 9 (less 941, 971)
Plus for all except OECD and FSS countries, Hong Kong, South Korea, Singapore, Taiwan:	
Computers, office and communications equipment	75–77
Skill-intensive (NMH)	
Chemicals	5 (less 524)
Non-electrical machinery	71–74
Motor vehicles and aircraft	781–784, 792

[10] They are also missing within the USSR in both years for all countries except Estonia, Latvia, Lithuania and the Russian Federation. An aggregate for Yugoslavia was not calculated because TiVA data are available only for Croatia and Slovenia.

Scientific instruments, watches and cameras	87, 88
Plus for OECD and FSS countries, Hong Kong, South Korea, Singapore, Taiwan:	
Computers, office and communications equipment	75–77

The different allocation of SITC 75-77 in different countries is a rough attempt to allow for the fact that the production which generates this category of exports is a mixture of skill-intensive component manufacturing, mainly in more educated countries, and labour-intensive assembly operations, mainly in less educated countries.

Figure 5. Skill-intensity of manufactured exports and skill/labour endowments, 2014

The omissions from the regional endowment aggregates are as specified in the discussion of figures 1 and 2 above. The omissions from the regional gross export aggregates are as specified in the discussion of table 1 above, and the division of SITC categories between 'skill-intensive' and 'labour-intensive' is as in the discussion of table 3 above. The cross-country regression is based on data that split Czechoslovakia, the USSR and Yugoslavia into separate countries.

The regression data omit countries where narrow manufactures are under 10% of total exports, which are included in the regional aggregates. However, the omission of such countries from the aggregates for Sub-Saharan Africa, which contains the highest share of them, only slightly alters the region's ratio of skill-intensive to labour-intensive manufactured exports.

Table 4. Output by broad sector, 2014 levels and 1985-2014 changes

The main source is the UN National Accounts Main Aggregates (UNMA) database, which has data for almost every country and year (http://unstats.un.org/unsd/snaama/methodology.pdf). As a result, only North Korea (Other East Asia) is missing from this table. The sectoral shares in the World Development Indicators database are far less complete, but are strongly correlated with UNMA (pooling manufacturing shares in 1985, 2000 and 2014, R = 0.94).

For China and India in 1985, the share of manufacturing was adjusted using other sources, as explained in the notes to the table. For China between 1970 and 2010 the UNMA industry share was split in proportion to the shares of manufacturing, mining and utilities in the GGDC database (http://www.rug.nl/ggdc/productivity/10-sector/). A similar split was made for the USSR in 1985 using the data in http://src-h.slav.hokudai.ac.jp/database/SESS.html#USSR-S1. The UNMA share of manufacturing was adjusted also in Bulgaria 1985, Chad 1985, Ghana 1985, Mexico 1985, Papua New Guinea 2014, and Togo 1985, often to resolve anomalies that emerged in the estimation of the mining share described below.

Mining is part of the primary sector (with agriculture), but UNMA merges it into 'industry', together with manufacturing and utilities. In almost all cases, however, manufacturing is also shown separately, so the sum of mining and utilities can be derived. In the great majority of cases, the share of mining in GDP can then be obtained by subtracting the share of utilities in the same (or a nearby) year, using more detailed information from the UN National Accounts database, Eurostat and OECD data. As a check, and to assist where the subtraction of utilities yielded

implausible results (most often negative mining) or data on utilities were unavailable, the mining share of GDP was also estimated by multiplying the manufacturing share of GDP by the ratio of mining value added to manufacturing value added (if the necessary more detailed data were available). In a few cases, anomalies were resolved or gaps filled by using data for other years or other countries.

Figure 6. Manufactured/primary output and skill/land endowments, 2014

The omissions from the regional endowment aggregates are as specified in the discussion of figures 1 and 2 above; and the only omission from the regional output (GDP share) aggregates is North Korea, as explained in the discussion of table 4 above. The cross-country regression is based on data that split Czechoslovakia, the USSR and Yugoslavia into separate countries.

Table 5. Sectoral output ratios regressed on endowment ratios

Five countries are omitted from the regressions in this table, in all cases for lack of schooling data: Lebanon, North Korea (which also lacks output data), Puerto Rico, Somalia, and West Bank & Gaza.

Table 6. Employment by broad sector, 2014 levels and 1985-2014 changes

Only North Korea and Somalia are missing from the data for primary, all manufacturing, and services employment. The 20 countries where gaps in the Haraguchi, Cheng and Smeets (Haraguchi et al) dataset were filled with WESO data are: Laos, Papua New Guinea (Other East Asia); Lebanon, West Bank & Gaza, Yemen (MENA); Afghanistan (South Asia); Benin, Burundi, Central African Republic, Chad, Democratic Republic of the Congo, Guinea, Guinea-Bissau, Lesotho, Liberia, Mauritania, Namibia, Sierra Leone, Sudan, Togo (S-S Africa).

The Haraguchi et al dataset does not separate out mining, for which (in order to calculate the primary share of employment) WESO data were used for all countries and years. The share of services was calculated as a residual (100 minus the shares of primary and manufacturing). For the 20 missing countries and for mining, sectoral employment shares in 1985 were assumed to be the same as in 1991, the first year of the WESO data. The only gaps in the WESO data are for North Korea and Somalia (missing also in Haraguchi et al) and for Syria.

The data for formal manufacturing employment are far less complete: even taking data for 1985 and 2014 from nearby (or in some cases distant) years, 24 countries are missing in 1985, 38 in 2014, and 47 countries in one or other of these years (though only 15 in both years). Regional shares of formal manufacturing in total employment were thus calculated as follows.

Gaps for countries with formal manufacturing employment data in at least one year were filled by extrapolation in proportion to total manufacturing employment.[11] Then for each region the share of formal manufacturing employment in total manufacturing employment was calculated as a weighted average across all countries with formal manufacturing employment data. This regional share was multiplied by total manufacturing employment for all countries in the region to estimate of total formal manufacturing employment in the region. The assumption is that countries with formal manufacturing employment data are in aggregate representative of all countries in the region.

Haraguchi et al, WESO and China

World-wide long-term datasets on employment are afflicted by large gaps and inconsistencies in the country and year coverage of primary sources, and are therefore necessarily incomplete and/or based on adjustment, splicing and extrapolation. The WESO dataset achieves almost complete country coverage by extrapolation, but starts only in 1991, making it less suitable for the present paper than Haraguchi et al, whose country coverage, however, is less complete.[12]

Where they overlap, the Haraguchi et al and WESO country-level data are strongly correlated. Pooling the 1985 (for WESO, 1991), 2000 and 2014 data on sectoral shares, for primary R = 0.97, for manufacturing R = 0.92, and for services R = 0.96. Running the regressions in table 7 with WESO rather than (WESO-augmented) Haraguchi et al data only slightly alters the sizes and significance levels of coefficients that matter for the analysis in the paper.

At the regional level, if table 6 is calculated with WESO rather than Haraguchi data, the only important difference is for China, discussed below. For other regions, the levels and changes of sectoral shares do not differ in ways that affect the conclusions. The all-manufacturing share in FSS falls less, because the WESO series start after the initial phase of transition. In most regions in 2014, primary sector shares are somewhat lower and service sector shares somewhat higher than in Haraguchi et al, and the service share for Other South Asia rises (as in all other regions) rather than falling.

The differences for China are massive, and of special importance for manufacturing: the WESO 2014 (all) manufacturing share of 11.7% is far below the Haraguchi et al share of 18.2% (an absolute difference of 50 million manufacturing workers), and the WESO manufacturing share falls by 2.2 percentage points over the period, rather than rising by 3.4 percentage points, as in the Haraguchi et al series. The WESO primary share for China in 2014 is 9 percentage points lower, and the WESO service share 15 percentage points higher, than in Haraguchi et al.

Sectoral employment data in China, particularly for manufacturing, are highly problematic. No single government agency is responsible for collecting the information on an annual basis, the

[11] This step was undertaken using formal employment data for 2000, as well as for 1985 and 2014, so there were even fewer countries missing than the 15 mentioned in the previous paragraph. Extrapolating values in this way assumes no change over the relevant period in the formal share of total manufacturing employment.
[12] The employment dataset of Felipe et al (2016) covers 63 countries during 1970-2010, but was not available for use in the present paper.

institutional structure of manufacturing has changed radically over the period, and the coverage of relevant statistical series has also changed (Banister 2005, 2013). As a result, although it is universally agreed that there are far more manufacturing workers in China than in any other country, there are widely differing estimates of their numbers.

The Haraguchi et al series for China is from the GGDC 10-sector database, whose derivation is explained in de Vries et al (2015). Its all-manufacturing 2014 employment share of 18.2% is at the top of the range of recent estimates, though not far above the 2010 population census estimate of 16.9% cited by Felipe et al (2016, table 8). More importantly for the present paper, its movement over time fits with other evidence: a slight (one percentage point) rise in the share between 1985 and 1996, a fall of more than two percentage points between 1996 and 2002, as a result of job-shedding from state-owned enterprises, and a rise of almost five percentage points (43 million workers) after 2002. The rise over this last period, though higher than the 2003-09 rise of 15 million estimated by Banister (2013), is similar to the 2003-14 rise of 49 million reported by Lardy (2015), the 2000-10 rise of 43 million between the population census figures cited by Banister (2005, table3) and Felipe et al (2016, table 8), and the 2002-2010 rise of 39 million in the UNIDO series used in this paper to measure formal employment.

The WESO series is less credible. From 1991 (the first year) to 1997, its all-manufacturing employment share is only about one percentage point below the Haraguchi et al series, as is the case also in 2003 (after a widening during 1998-2001). From 2003, however, the WESO share gradually declines, so the two series diverge, and by 2012 (when the Haraguchi et al series ends) the difference between them is over six percentage points. The decline implied by the WESO series in the absolute number of manufacturing workers in China during 2003-12 is inconsistent with all other evidence.

Table 7. Sectoral employment ratios regressed on endowment ratios

Five countries are omitted from all regressions in this table for lack of schooling data: Lebanon, North Korea, Puerto Rico, Somalia, and West Bank & Gaza (North Korea and Somalia also lack employment data). The regressions involving formal manufacturing employment omit a further 43 countries that do not have formal manufacturing employment data in both 1985 and 2014 (four of the five countries that lack schooling data are among the 47 countries without formal employment data in both years, Puerto Rico being the exception).

Table 8. Manufacturing share changes 1985-2014 *(percentage points)*

The data in this table are copied from those in table 1 (exports), table 4 (output), and table 6 (employment) the notes to which thus apply also to this table.

Figure 7. GDP per capita, 1985 and 2014 (thousand 2005 US dollars)

Data are missing only for the West Bank & Gaza.

Data for figures, tables and regressions

The underlying data, plus other series not used in the paper, are in a single workbook called Wood source data.xlsx. It contains the following four sheets, the first three of which all have the same variables in the same column order for 1985, 2000 and the most recent available year (called '2014').

'Country-level data' contains data (with some gaps) for the 130 'comparable' countries listed earlier, and in 2000 and 2014 also for the 23 countries that emerged from Czechoslovakia (2), USSR (15) and Yugoslavia (6, excluding small Montenegro).

'Regional averages' contains aggregate (weighted average) and unweighted average values of all variables for the ten regions listed above plus East Asia (China plus Other East Asia), South Asia (India plus Other South Asia) and the World. It also includes the country-level data used to calculate these averages, with the countries grouped by region.

'Regional summary' is a simplification of the previous sheet that contains only the aggregates (weighted averages) for the ten regions in the paper and for the world.

'Constant-price trade' contains the data on constant-price exports and imports used to construct figure 3, both at country level and aggregated to regional level, for 1985 and 2014.

References

Banister, J. (2005). 'Manufacturing employment in China', *Monthly Labour Review*, July https://www.bls.gov/opub/mlr/2005/07/art2full.pdf, accessed 4 February 2017.

Banister, J. (2013). 'China's manufacturing employment and hourly compensation, 2002-2009', Bureau of Labour Statistics, https://www.bls.gov/fls/china_method.pdf, accessed 4 February 2017.

De Vries, G., K. de Vries, R. Gouma, S. Pahl and M. Timmer (2015). 'GGDC 10-Sector Database: Contents, Sources and Methods', http://www.rug.nl/ggdc/docs/10sector_sm_jan2015.pdf, accessed 4 February 2017.

Felipe, J., A. Mehta and C. Rhee (2016). 'Manufacturing Matters … but it's the Jobs that Count'. Update of Working Paper 420, 2014, Asian Development Bank.

Lardy, N. (2015). 'Manufacturing Employment in China'. Peterson Institute for International Economics, https://piie.com/blogs/realtime-economic-issues-watch/manufacturing-employment-china, accessed 4 February 2017.