Data appendix for economic growth in the long run

Tamura, Robert and Dwyer, Gerald P and Devereux, John and Baier, Scott

Clemson University, Clemson University, Queens College, CUNY, Clemson University

14 September 2012

Data Appendix for Economic Growth in the Long Run*

Robert Tamura, Gerald P Dwyer, John Devereux, Scott Baier†

November 2016

Abstract

This extended data appendix describes the sources and methods used to construct the data used in our paper *Economic Growth in the Long Run*.

1 Introduction

This data appendix describes the sources and procedures used to construct Figures 1-4 in the text, as well as the data for all of the empirical analyses. Almost all of the real output data come from Maddison. We used two different sources, which occasionally differ, but generally are quite similar: Maddison (1995) *Monitoring the World Economy: 1820-1992*, and Maddison (2003) *The World Economy: Historical Statistics*. All of Maddison source data is listed in 1990 Geary-Khamis dollars. We converted these into 2000 international dollars using the US GDP deflator. This keeps the data roughly consistent with our earlier work, Baier, Dwyer and Tamura (2006), except for change of base year. We generally used Maddison for all real PPP per capita income. Unless noted otherwise, all schooling data, historical age distribution data, investment data come from B. R. Mitchell (2003). I abbreviate this source as Mam, Maa, Meu for *International Historical Statistics: the Americas, 1750-2000*, *International Historical Statistics: Africa, Asia and Oceania, 1750-2000*, and *International Historical Statistics: Europe, 1750-2000*. For some population figures in 1980 and 1990 we used those from Summers and Heston online, hereafter abbreviated as S & H online. For 2000 population we used data from the Time Almanac 2001. For 2010 population we used data from *Wikipedia*. The share of the population aged 15-64 for all countries in 2010 come from *World Bank Development Indicators*. Some age distribution data is not available from Mitchell. These are typically smaller undeveloped countries of Africa and Asia. Typically we do not have data prior to 1950. We used data from Keyfitz and Flieger (1990), which provide age distribution data in quinquennial manner from 1950-2000.

1.1 Physical Capital

Physical capital investment rates prior to 1992 were measured using Mam, Maa, and Meu. Mam, Maa and Meu provide annual information on gross physical capital formation. Between the census years t-1

*We thank Kevin M. Murphy, Richard Rogerson, Todd Schoellman, Curtis Simon, Paula Tkac, Kei-Mu Yi for helpful comments and suggestions. We also thank the seminar participants at the 2016 Las Vegas APEE Meetings, University of North Carolina at Chapel Hill, 2009 NBER Time and Space Conference at the Philadelphia FRB, 2009 Midwest Macroeconomics Meetings at Indiana University, Human Capital and Economic Development Conference at the Korean Labor Institute

†Clemson University, Clemson University, Queens College of City University of New York, Clemson University

and t, we calculate the mean investment rate. Most of the data comes from real capital formation as a fraction of real income, but when this data was missing we substituted nominal values of capital formation relative to nominal income. For some countries we used investment rates from S & H online. In order to calculate physical capital per worker, we created average investment rates over the coverage years. For example suppose we have two observations on income per worker in 1990 and 2000, respectively. In order to calculate the physical capital per worker in 2000 we use the average investment rate from 1990 to 1999 inclusive. We use a perpetual inventory method of calculation. In order to illustrate this method, assume that we observe output per worker in 1990 and 2000, y_{1990} and y_{2000}, respectively. Let i_{2000} be the average investment rate for years 1990 to 1999 inclusive, and i_{1990} be the investment rate in 1990. Finally assume that k_{1990} is the physical capital per worker in 1990. Physical capital per worker in 1991 would be given by:

$$k_{1991} = \frac{k_{1990}(1-\delta) + y_{1990}i_{1990}}{g_w} \tag{1}$$

where g_w is the growth rate of the labor force between 1990 and 1991, and δ is the annual depreciation rate on capital. Now let g_y be the annualized growth rate of output per worker from 1990 to 2000 and redefine g_w to be the annualized growth rate of labor force between 1990 and 2000. Repeated substitution of the above relation produces:

$$k_{2000} = i_{2000}y_{1990}\sum_{i=0}^{9}\frac{(1-\delta)^i g_y^{9-i}}{g_w^{i+1}} + \frac{(1-\delta)^{10}k_{1990}}{g_w^{10}} \tag{2}$$

The first term on the right hand side is, by assumption, a finite geometric sum and hence finite. The last term is an exponentially decaying term of the previous periods physical capital per worker. Thus we can rewrite the above expression as:

$$k_{2000} = i_{2000}y_{1990}\frac{\frac{g_y^9}{g_w}\left[1 - \frac{(1-\delta)^{10}}{g_y g_w^{11}}\right]}{1 - \frac{(1-\delta)}{g_y g_w}} + \frac{(1-\delta)^{10}k_{1990}}{g_w^{10}} \tag{3}$$

For initial conditions we typically use information on sectoral output shares and apply the US capital - sectoral output ratios to these to construct initial capital estimates. Otherwise, we have information for other sources on the initial capital stock value.

What values of δ_t to use? An earlier version of this paper assumed that δ_t was a step function in the output per worker, y_t. However the work of Piketty and and Zucman (2013) convinced us that it was simpler and better to use a constant depreciation rate, independent of income. Furthermore Piketty and Zucman find that the physical capital output ratio is generally in the range of 3 or greater, whether in the 19th century, or 20th century or 21st century! For a capital output ratio of 3, a $\frac{1}{30}$ depreciation rate would produce a capital charge on income equal to 10% of income. We assumed for any country i:

$$\delta_i = \frac{1}{30} \tag{4}$$

These rates are consistent with producing $\frac{k}{y}$ ratios for rich countries that are around 3 in 2010. In the final section, we detail the special cases of depreciation rates due to the switch from central planning in *Central & Eastern European* countries to market economies after the fall of the Soviet Union/Russia.

Table 1: Regressions of PPP investment rates on nominal investment rates

variable	decade average	yearly observations
nominal investment rate	1.6127	1.5749
	(0.1017)	(0.0344)
lny	0.0091	0.0093
	(0.0025)	(0.0009)
nominal investment rate * lny	-0.0838	-0.0800
	(0.0127)	(0.0043)
constant	-0.0687	-0.0697
	(0.0186)	(0.0065)
number of observations	610	5842
$\bar{R}^2$	.9065	.8989

Notes: Table reports our estimates of PPP investment PWT rates on nominal investment rates from Mitchell.

The next question is what investment rate to use. The investment rates from Mitchell are nominal investment rates. That is investment rates assuming that the price of capital is equal to the price of consumption for all countries. However Summers Heston provide PPP adjusted as well as nominal investment rates for countries. We ran the regression of PPP adjusted investment rates against nominal investment rates, log of real output per capita in 1985 dollars, and an interaction between nominal investment rates and log of real output per capita in 1985 dollars and a constant. However before we ran the regressions we calculated decade averages of PPP investment rates, nominal investment rates, log of real output per capita in 1985 dollars and the interaction of the average nominal investment rate and log of real output per capita in 1985 dollars. Column one of Table 1 produces these regression results. Column two of Table 1 provides a robustness check on the regressions by examining the results of a regression using all of the years of the data, instead of averaging them by decade, there are no differences. For the paper we used the results in column 1 to produce our PPP investment rates. The range of the variables are given in the following Table. The first half of Table 2 provides the mean, standard deviation, minimum and maximum for the Summers Heston 1950-2000 data. The second half of the table provides the same information for our Mitchell data. As can be seen, almost all of the values of the variables in the years prior to 1950 are completely contained in the range from 1950-2000. The notable exception is the negative investment rate in the Netherlands during the 1930s. However since this is not an initial year, it does not cause any problems in the calculations of real physical capital per worker, although obviously real capital per worker falls between 1930 and 1940! Hence our estimates of PPP investment rates are in fact projections and not extrapolations.

1.2 Sectoral Computations

There are times when we do not have information on investment rates. Previous editions of this work assumed that the missing years were replaced with average investment rates typically for the first 30-50 years (3-5 observations) that are available. However this can be quite the stretch if we are extrapolating

Table 2: Descriptive statistics of Summers & Heston (1950-2000) and Mitchell prior to 1950

	Summers & Heston 1950-2000				Mitchell prior to 1950			
variable	mean	std dev	min	max	mean	std dev	min	max
nominal investment rate	0.1656	0.0936	0.0121	0.5440	0.1200	0.0437	-0.008	0.3586
lny	7.852	1.025	5.658	10.134	7.364	0.644	5.623	8.755
nominal investment rate * lny	1.3532	0.8505	0.0735	4.733	0.8913	0.358	-0.0646	2.854

a century! So we have modified our approach using sectoral output shares. Typically Mitchell provides sectoral output shares of GDP, which we aggregate into three categories: farming, manufacturing and services. Manufacturing is a catch-all category that includes manufacturing, mining and construction.[1] We generally have data from 1950-2010, where for the richest countries we often have sectoral output shares back into the 19th century. In order to estimate sectoral shares where no data was found, we used Sabillon (2005). He provides decade growth rates for the 1950s, 1960s, 1970s, 1980s and 1990s, half century growth rates 1900-1950, and 19th century growth rates of GDP, agriculture and most times manufacturing and occasionally services. His coverage is almost comprehensive with our countries. In order to compute past sectoral shares we used the following typical formula:

$$s_{ag,t-X} = s_{ag,t} exp(-X[\gamma_{ag,t-X} - \gamma_{y,t-X}]) \tag{5}$$

$$s_{man,t-X} = s_{man,t} exp(-X[\gamma_{man,t-X} - \gamma_{y,t-X}]) \tag{6}$$

$$s_{services,t-X} = 1 - s_{ag,t-X} - s_{man,t-X} \tag{7}$$

where X is the number of years prior to the first observed value of sectoral output share, $\gamma_{ag,t-X}$ is the annual growth rate of agriculture in the decade containing year $t-X$, $\gamma_{y,t-X}$ is the annual growth rate of GDP in the decade containing year $t-X$. When we do not have growth rate observations on say agriculture or manufacturing, for some countries, we almost always have growth rate observations on services. Below we list in Table 3 the countries that Sabillon does not have: [2] Table 4 provides the sectoral capital-output

Table 3: Missing Countries from Sabillon (2005) Sectoral Data

Region	N	Missing Countries
wc	18	none
se	8	Cyprus
cee[1]	24	Slovak Republic
nic	5	none
asia	20	Afghanistan, Fiji, North Korea
ssa	48	Djibouti, Eritrea, Gabon, Gambia, Guinea-Bissau, Lesotho, Malawi, Mauritius, Reunion, Seychelles, Swaziland
la	28	Barbados, Belize, Guyana, Trinidad
me	12	none
na	5	none

[1] Often this is categorized as industrial or secondary sector. This typically fits better the industrial composition that is used by the *WDI*.

ratios for the U.S. by time period that are used to produce aggregate capital for a country in which real investment rates are not available, and other papers have not constructed capital stock estimates. In this

Table 4: U.S. sectoral capital output ratios

year	farming	manufacturing	services
≤ 1854	3.06	0.46	3.77
1855-1864	3.76	0.47	3.76
1865-1874	3.13	0.47	3.76
1875-1884	4.52	0.55	3.76
1885-1894	5.01	0.58	3.77
1895-1904	4.19	0.82	3.76
1905-1914	4.69	0.91	3.76
1915-1924	4.42	0.71	3.76
1925-1934	4.54	1.40	3.87
1935-1944	2.52	1.48	3.92
1945-1954	3.34	0.83	4.40
1955-1964	3.22	0.76	3.87
1965-1974	3.19	0.81	3.53
1975-1984	4.71	0.97	4.01
1985-1994	3.36	0.97	3.25
1995-2004	2.73	0.88	2.87

subsection we demonstrate that the physical capital stock measures derived from sectoral output shares are reasonable. There are two concerns that arise from our method of estimating physical capital using sectoral output shares. The first is that we only have sectoral physical capital - output ratios for the United States, and hence applying these to countries outside of the United States maybe problematic. The second concern is that our measure of sectoral physical capital - output ratios are only available for the United States back to 1850, and the use of the 1850 value for all years prior to 1850 maybe problematic. Here we believe that we can convince the reader that neither of these worries are too severe to prevent the use of this method. We show this by regressing the log of physical capital taken from other individuals, or from our perpetual inventory calculation against the log of physical capital per worker from the sector method. Of course when we use the sectoral method to produce the estimate, we do not include these in the regression. For all years, except 2010, we have estimates from the sectoral method. Table 5 presents the results of these regressions. We run four samples, all years, only years before 1975, only years before 1955 and only years before 1925. The samples decline from 1038 observations, to 574, 348 and 211. We include region dummies and decade dummies. The results are robust to inclusion of these region and decade controls. We find that the sectoral capital per worker estimates are highly and significantly correlated with the measures of capital per worker arising from other researchers like Piketty and Zucman, as well as by perpetual inventory. The typical coefficient on log sectoral physical capital per worker ranges from 0.76 to 1.079. Our view is that the sectoral capital stock is a good estimate of physical capital when other researchers have not produced

[2]None of the former Soviet Republics are included in Sabillon (2005). However all observations for these countries begin in 1970, and Easterly and Fischer (1995) provide capital values for 1970, 1980 and 1990. The information from Sabillon is similar for Sub-Saharan Africa, but broader in scope to van Waijenburg. In van Waijenburg information is available on real wage growth from 1880-1945 for a few select Africa countries, and regions of Africa. While van Waijenburg's data typically predate our coverage, they are consistent with the relatively slow growth or output in Sub-Saharan Africa from 1950-2010. Also quite similar to Sabillon is Smits.

estimates of physical capital.

Table 5: Log Physical Capital per Worker Stock Regressions

Variable	$\ln k$	$\ln k$	$\ln k$	$\ln k$	$\ln k$	$\ln k$	$\ln k$	$\ln k$
$\ln k_{sector}$	1.0787	1.0921	1.0408	0.9764	1.0250	1.0625	0.9408	0.9180
	(0.0167)	(0.0162)	(0.0203)	(0.0258)	(0.0253)	(0.0260)	(0.0306	0.0419)
constant	-0.7664	-0.6547	-0.4634	0.4271	-0.3272	-0.6981	0.4458	0.7161
	(0.1619)	(0.1605)	(0.2142)	(0.2653)	(0.2404)	(0.2539)	(0.3106)	(0.4168)
sample	all years	all years	all years	all years	yr $\leq$ 1974	yr $\leq$ 1974	yr $\leq$ 1974	yr $\leq$ 1974
decade dummies	no	yes	no	yes	no	yes	no	yes
region dummies	no	no	yes	yes	no	no	yes	yes
N	1038	1038	1038	1038	574	574	574	574
R^2	.8011	.8327	.8152	.8474	.7418	.7584	.7651	.7772
Variable	$\ln k$	$\ln k$	$\ln k$	$\ln k$	$\ln k$	$\ln k$	$\ln k$	$\ln k$
$\ln k_{sector}$	0.9810	1.0563	0.8090	0.8125	1.0299	1.1483	0.7623	0.8110
	(0.0414)	(0.0441)	(0.0491)	(0.0713)	(0.0666)	(0.0715)	(0.0770	0.1149)
constant	0.0961	-0.6823	1.6991	1.7023	-0.3120	-1.5316	2.1624	1.7203
	(0.3893)	(0.4251)	(0.5165)	(0.7081)	(0.6169)	(0.6795)	(0.8695)	(1.1705)
sample	yr $\leq$ 1954	yr $\leq$ 1954	yr $\leq$ 1954	yr $\leq$ 1954	yr $\leq$ 1924	yr $\leq$ 1924	yr $\leq$ 1924	yr $\leq$ 1924
decade dummies	no	yes	no	yes	no	yes	no	yes
region dummies	no	no	yes	yes	no	no	yes	yes
N	348	348	348	348	211	211	211	211
R^2	.6189	.6496	.6742	.6848	.5334	.5817	.6487	.6578

1.3 Labor Force

While most of our labor force data comes from Mitchell's three volumes, and supplemented by modern sources like WDR and *WDI*, there are periods, particularly in the 19th century when our labor force data comes differently. We used data on the age structure of the population and the urban-rural population distribution to estimate labor force. The urban-rural population shares come from essentially four sources, Clio infra (an internet site that provides urbanization shares for countries from 1500 to 2000), WorldOMeters (an internet site with urbanization rates from 1950-2010), Banks (1971) and *WDI*. We have urban-rural population shares for every country, with the exception of Hong Kong (1820-1910) and Taiwan (1820-1890). For Hong Kong we back project for the 19th century, and for Taiwan we used the historic ratio of urbanization in Taiwan compared with mainland China for the 19th century. We assume the age structure of the national population is common between urban and rural locations. We then solve for the labor force participation rates of adults 15 and older for both urban and rural locations in order to fit the data on labor force. We choose not to estimate a constant participation rate for each type, but rather to find a pair of participation rates for each year in order to match the data on labor force. Typically we have higher labor force participation rates in rural population compared with urban, but not always. If year τ if the earliest year we observe data on labor force, then we compute labor force for years $t < \tau$, by typically applying the year τ labor force participation rates for rural and urban populations for each year. In our write up for

each country, we list explicitly these participation rates, as well as the fit of the data using this method in overlapping years.

1.4 Schooling and Human Capital

Enrollment rates for years prior to 2010 generally come from Mitchell. When reported primary and secondary school enrollments are reported separately, we used the age distribution of the population 0-4, 5-9, 10-14, 15-19 to produce the relevant enrollment rates. When they are not reported separately, we typically assumed that the vast majority of enrollments are primary school attendees. There is typically an overlapping observation year, with enrollment rates reported by *World Development Reports* or *Human Development Reports*, that allow us to assign the proportion enrolled in each school category in the overlapping years, and then gradually reduce the secondary share of total enrollments back in time. The details are given in each country description. Finally for years prior to data from Mitchell on school enrollment, we used estimates of literacy from various sources, and other estimates of enrollments, again detailed within each country's description. Enrollment rates for 2010 come from *World Development Indicators*. After the penultimate draft of this paper was written, we became aware of Lee & Lee (2016) in working paper form. Their original contribution produces years of schooling for men and women 15-64, from 1870-2010. They have enrollment data prior to 1870, but mostly concentrate on 1870 onward. Their work provides years of schooling estimates for 111 countries.

To calculate the stock of human capital of each type, primary school stock, secondary school stock and higher education stock, we used a perpetual inventory method. We focused on males, but we typically used information on total enrollments, not enrollments of men, as this typically is not available in Mitchell. Although this induces a downward bias in the measure of male enrollments, we feel that the information is still valuable. We used the same method as in BDT (2006), which is a variant of TTMB (2007). The following example will illustrate the nature of our calculations. In period $t+1$, the stock of adults, $H_{t+1]^i}$, aged 25 and older, with exposure to education level i, i =none, primary (but no more), secondary (but no more) and higher education (but no more) is given by:

$$H_{t+1}^i = H_t^i(1-\delta_t^{hc}) + I_t^i \tag{8}$$

where δ_t^{hc} is the *death rate* and I_t^i is the flow of new adults with exposure to education level i and no more. We assumed that δ_t^{hc} does not vary by education class. It is useful to put the human capital measure as a fraction of the labor force. Thus we normalize and produce:

$$\frac{H_{t+1}^i}{L_{t+1}} = \frac{H_t^i}{L_t}\frac{L_t}{L_{t+1}}(1-\delta_t^{hc}) + \frac{I_t^i}{L_{t+1}} \tag{9}$$

$$h_{t+1}^i = h_t^i\frac{L_t}{L_{t+1}}(1-\delta_t^{hc}) + \frac{I_t^i}{L_{t+1}} \tag{10}$$

In order to proceed we need a measure of the death rate of adults. We constructed this for each country using the information provided below about the ages of schooling for primary and secondary education. We illustrate our calculations by presenting the case where primary school lasts from ages 6-13 and secondary school lasts from ages 14-17. Assume that the time gap between period t and t+1 is 10 years. Then we

used the following equation:

$$L_{t+1} = L_t(1-\delta_t^{hc}) + r_t^{hi}\ell\,[9-24]_t + (r_t^{sec} - r_t^{hi})\ell\,[8-17]_t + (r_t^{el} - r_t^{sec})\ell\,[0-13]_t + (1-r_t^{el})\ell\,[0-13]_t \quad (11)$$

where r_t^{hi} is the higher education enrollment rate, r_t^{sec} is the secondary school enrollment rate, r_t^{el} is the primary school enrollment rate, and $\ell[i-j_t]$ is the number of males between the ages of i and j, inclusive in period t. Notice that this definition allows for the calculation of the common term in all equations, $\frac{L_t}{L_{t+1}}(1-\delta_t^{hc})$ in terms of obsevables:

$$\frac{L_t}{L_{t+1}}(1-\delta_t^{hc}) = 1 - r_t^{hi}\ell\,[9-24]_t - (r_t^{sec} - r_t^{hi})\ell\,[8-17]_t - (r_t^{el} - r_t^{sec})\ell\,[0-13]_t - (1-r_t^{el})\ell\,[0-13]_t \quad (12)$$

To complete the analysis we use the information on enrollments for I_t^i. In this example, these are:

$$I_t^{hi} = r_t^{hi}\ell\,[9-24]_t \quad (13)$$

$$I_t^{sec} = (r_t^{sec} - r_t^{hi})\ell\,[8-17]_t \quad (14)$$

$$I_t^{el} = (r_t^{el} - r_t^{sec})\ell\,[0-13]_t \quad (15)$$

$$I_t^{none} = (1-r_t^{el})\ell\,[0-13]_t \quad (16)$$

1.5 Extent of Data Imputation

In this section we detail the degree of data imputation. We believe that the level of imputation is small, arguably less than 5 percent of the observations, and insignificant. We do not impute real output. So all imputed data occur in order to use the data on output per worker. Population is never imputed for any country. Data for the age distribution, labor force, education enrollments and investment rates are possible candidates for imputation. Our main concern is with schooling data.

We supplement the Mitchell *Historical Statistics* volumes with data from *World Development Indicators*, *WDI* for 2000 and 2010 when needed. For historical data prior to Mitchell, we used a variety of sources, each country entry details these. The most important data sources were Benavot & Riddle (1988), Morrisson & Murtin (2009) and Easterlin (1981, 1998). When our early values are able to match the 1870 stock education shares of Morrisson & Murtin, we consider the information validated.[3] As a result, out of 2044 observations only 85 primary school enrollment rates were extrapolated. In addition to these sources we used information on literacy contained in: Morris and Adelman (1988), Steckel and Floud (1997), and Benavot and Riddle (1988). We follow the rule that it takes 3 years of schooling to become literate, see Harman (1970), Mitch (1984,2004) and Resnick and Resnick (1977). Out of 168 countries only 12 countries have extrapolated primary school enrollment rates. Of these 12, 9 have multiple years of extrapolation, three of these only extrapolate the first two or three observation years, and three more extrapolate the first observation. Thus only six countries have primary school enrollment rates that are extrapolated for more than three observations: Hong Kong (1820-1880), Nepal (1820-1900), Vietnam (1820-1890), Venezuela (1820-1880), Jordan (1820-1890), and Lebanon (1820-1890).

For secondary school enrollment rates, 6.5 percent of the observations are extrapolated, 133 out of 2044. Here 18 countries out of 55 countries in *Western Countries*, *Southern Europe*, *Central & Eastern Europe* and

[3]The path is almost certainly not unique, but we typically assume that enrollment rates are weakly increasing from the first year of observation to the earliest information available from a source.

Newly Industrialized Countries have extrapolated values for secondary school enrollments. These countries are responsible for 78 out of 133 extrapolated values. Typically these are the earliest observations, and even more than with primary school enrollment rates, the first observations from data are typically very small values. In *Asia* only 6 out of 20 countries have extrapolated values. *Sub-Saharan Africa* has only 1 out of 48 countries requiring data extrapolation. Only 1 out of 28 *Latin America* countries have extrapolated values, Argentina (1800-1860). Four out of 12 *Middle East* countries have extrapolated values. Jordan (1820-1890), Lebanon (1820-1890) are 16 of the 19 extrapolated values. Morrison & Murtin do not provide any education measures for Jordan and Lebanon. *North Africa* did not have any extrapolated values.

For higher education enrollment rates, only 109 of the 2044 observations are extrapolated. Two of 18 countries in the *Western Countries* region have extrapolated values, but only 3 observations out of 355 are extrapolated. All but one of these are first observations, and the other is a second observation. No extrapolated observations come from the *Southern Europe* region. Four of 28 countries in *Central & Eastern Europe* have extrapolated values. The Czech Republic (1820-1900), Poland (1870-1910) and Romania (1870-1890) are responsible for 17 out of 18 extrapolated values. None of these countries are in Morrisson & Murtin database, and thus cannot be benchmarked. In the *Newly Industrialized Countries* region Hong Kong (1820-1940), Singapore (1820-1870) and Taiwan (1820-1905) are the countries that have extrapolated values. As with the *Central & Eastern Europe* region, none of these countries are in the Morrisson & Murtin database. They constitute 29 extrapolated observations. Thus Three of the 20 countries in *Asia* have extrapolated values. However only Nepal (1820-1900) and Vietnam (1820-1890) have extrapolated observations other than first observations. These two countries are not in the Morrisson & Murtin database. They are 17 of the 18 extrapolated values. Of the 48 *Sub-Saharan Africa* countries, 11 have extrapolated values. However only Namibia has extrapolated values that are not first observations. Only 2 of 373 observations from *Latin America* are extrapolated, and only 2 out of 28 countries have extrapolated values. In the *Middle East* region 5 of 12 countries have extrapolated values. Three of these five countries have extrapolated first or second observations. Only Jordan (1820-1910) and Lebanon (1820-1930) have extrapolated values for observations beyond the first or second year. As before, we have missing values because these two countries are not in the Morrisson & Murtin database. Finally *North Africa* has no extrapolated values.

Thus for the schooling data, we have at most 181 observations with extrapolated values. Hence less than 9% of the observations required extrapolations in schooling. Typically these are first observations, where schooling enrollment rates are typically very low and we benchmark the observation with the closest data point from the country.

Table 6: Data Imputation: Non Interpolated Values

Country	obs.	age dist.	labor force	invest rate	primary	secondary	higher
Western Countries							
Australia	21	4: 1820-1850	none	none	none	none	none
Austria	20	5: 1820-1860	none	none	none	1: 1820	1: 1820
Belgium	20	3: 1820-1840	none	none	none	1: 1820	none
Canada	20	3: 1820-1840	none	none	none	5: 1820-1860	none
Denmark	20	2: 1820-1830	none	none	none	6: 1820-1870	none
Finland	19	3: 1820-1840	none	none	none	none	none

Country	obs.	age dist.	labor force	invest rate	primary	secondary	higher
France	24	5: 1800-1840	none	none	none	none	none
Germany	23	7: 1800-1860	none	none	none	3: 1800-1820	none
Iceland	7	none	none	none	none	none	none
Ireland	20	4: 1820-1850	none	none	none	5: 1820-1860	none
Luxembourg	7	none	none	none	none	none	none
Netherlands	22	4: 1800-1830	none	none	none	4: 1800-1830	2: 1800-1810
New Zealand	25	5: 1820-1860	none	none	none	6: 1820-1874	none
Norway	20	3: 1820-1840	none	none	none	2: 1820-1830	none
Sweden	22	none	none	none	none	none	none
Switzerland	20	5: 1820-1860	none	none	1: 1820	1: 1820	none
UK	22	2: 1801-1811	none	none	none	none	none
USA	23	1: 1790	none	none	none	none	none
Southern Europe							
Cyprus	7	none	none	none	none	none	none
Greece	20	5: 1820-1860	none	none	none	none	none
Israel	8	none	none	none	none	none	none
Italy	20	4: 1820-1850	none	none	none	none	none
Malta	6	none	none	none	none	none	none
Portugal	22	6: 1800-1849	none	none	none	none	none
Spain	22	5: 1800-1840	none	none	none	5: 1800-1840	none
Turkey	21	12: 1820-1927	none	none	none	none	none
Central and Eastern Europe							
Albania	15	8: 1870-1940	none	none	none	none	none
Armenia	5	none	none	none	none	none	none
Azerbaijan	5	none	none	none	none	none	none
Belarus	5	none	none	none	none	none	none
Bulgaria	16	2: 1870-1880	none	none	none	1: 1870	none
Czech	20	9: 1820-1900	none	none	none	none	9: 1820-1900
E Germany	5	none	none	none	none	none	none
Estonia	5	none	none	none	none	1: 1970	1: 1970
Georgia	5	none	none	none	none	none	none
Hungary	15	none	none	none	none	none	none
Kazakhstan	5	none	none	none	none	none	none
Kyrgyzstan	5	none	none	none	none	none	none
Latvia	5	none	none	none	none	none	none
Lithuania	5	none	none	none	none	none	none
Moldova	5	none	none	none	none	none	none
Poland	15	5: 1870-1910	none	none	none	5: 1870-1910	5: 1870-1910
Romania	16	3: 1870-1890	none	none	none	none	3: 1870-1890
Russia	20	8: 1820-1890	none	none	none	3: 1820-1840	none
Slovak	3	none	none	none	none	none	none

Country	obs.	age dist.	labor force	invest rate	primary	secondary	higher
Tajikistan	5	none	none	none	none	none	none
Turkmenistan	5	none	none	none	none	none	none
Ukraine	5	none	none	none	none	none	none
Uzbekistan	5	none	none	none	none	none	none
Yugoslavia	11	1: 1910	none	none	none	none	none
Newly Industrialized Countries							
Hong Kong	20	13: 1820-1940	none	none	7: 1820-1880	13: 1820-1940	13: 1820-1940
Japan	22	9: 1800-1880	none	none	none	none	none
Singapore	21	13: 1820-1940	none	none	none	6: 1820-1870	6: 1820-1870
S. Korea	21	11: 1820-1920	none	none	none	none	none
Taiwan	22	8: 1820-1890	none	none	none	10: 1820-1905	10: 1820-1905
Asia							
Afghanistan	7	none	none	none	none	none	none
Bangladesh	7	none	none	none	none	none	none
Bhutan	4	none	none	none	none	none	none
Cambodia	7	none	none	none	none	none	none
China	20	13: 1820-1940	none	none	3: 1820-1840	3: 1820-1840	none
Fiji	7	none	none	1: 1950	none	none	none
India	20	6: 1820-1870	none	none	3: 1820-1840	3: 1820-1840	none
Indonesia	20	13: 1820-1940	none	none	none	none	none
Laos	7	none	none	none	none	none	none
Malaysia	20	11: 1820-1920	none	none	none	4: 1820-1850	none
Mongolia	7	none	none	none	none	none	none
Myanmar	20	6: 1820-1870	none	none	none	none	none
Nepal	20	13: 1820-1940	none	none	9: 1820-1900	9: 1820-1900	9: 1820-1900
N. Korea	20	13: 1820-1940	none	none	none	none	none
Pakistan	7	none	none	none	none	none	none
P. N. Guinea	6	1: 2010	none	none	none	none	none
Philippines	21	10: 1820-1913	none	none	none	none	none
Sri Lanka	21	9: 1820-1900	none	none	1: 1820	1: 1820	1:1820
Thailand	20	9: 1820-1900	none	none	none	none	none
Vietnam	20	13: 1820-1940	none	none	8: 1820-1890	8: 1820-1890	8: 1820-1890
Sub-Saharan Africa							
Angola	7	1: 2010	none	none	none	none	none
Benin	7	none	none	none	none	none	none
Botswana	7	none	none	none	none	none	none
Burkina Faso	7	none	none	none	none	none	none
Burundi	7	none	none	none	none	none	1: 1950
Cameroon	7	none	none	none	none	none	none
Cape Verde	7	none	none	2: 1950-1960	none	none	1: 1950
C. Afr. Rep	7	1: 2010	none	none	none	none	none

Country	obs.	age dist.	labor force	invest rate	primary	secondary	higher
Chad	7	1: 2010	none	none	none	none	none
Comoros	7	1: 2010	none	2: 1950-1960	1: 1950	1: 1950	1: 1950
Congo	7	none	none	none	none	none	none
Cote d'Ivoire	7	1: 2010	none	none	none	none	none
Djibouti	7	2: 2000-2010	none	3: 1950-1970	none	none	1: 1950
Eq. Guinea	7	1: 2010	none	none	none	none	1: 1950
Eritrea	3	none	none	none	none	none	1: 1990
Ethiopia	7	none	none	none	none	none	none
Gabon	7	1: 2010	none	1: 1950	none	none	none
Gambia	7	1: 2010	none	2: 1950-1960	none	none	1: 1950
Ghana	15	8: 1870-1940	none	none	none	none	none
Guinea	7	none	none	none	none	none	1: 1950
G.-Bissau	7	1: 2010	none	2: 1950-1960	none	none	1: 1950
Kenya	7	none	none	none	none	none	none
Lesotho	7	none	none	1: 1950	none	none	none
Liberia	7	none	none	none	none	none	none
Madagascar	7	1: 2010	none	none	none	none	none
Malawi	7	none	none	none	none	none	none
Mali	7	none	none	none	none	none	none
Mauritania	7	none	none	none	none	none	none
Mauritius	7	none	none	none	none	none	none
Mozambique	7	none	none	none	none	none	none
Namibia	7	none	none	none	none	none	3: 1960-1980
Niger	7	none	none	none	none	none	none
Nigeria	7	none	none	none	none	none	none
Reunion	7	none	none	2: 1950-1960	none	none	none
Rwanda	7	none	none	none	none	none	1: 1950
Senegal	7	none	none	none	none	none	none
Seychelles	7	none	none	2: 1950-1960	none	none	none
Sierra Leone	7	none	none	none	none	none	none
Somalia	7	1: 2010	none	1: 1950	none	none	none
South Africa	23	11: 1800-1900	none	none	none	none	none
Sudan	7	1: 2010	none	none	none	none	none
Swaziland	7	none	none	2: 1950-1960	none	none	none
Tanzania	7	none	none	none	none	none	none
Togo	7	none	none	none	none	none	none
Uganda	7	none	none	none	none	none	none
Zaire	7	1: 2010	none	none	none	none	none
Zambia	7	none	none	none	none	none	none
Zimbabwe	7	1: 1950	none	none	none	none	none
Latin America							

Country	obs.	age dist.	labor force	invest rate	primary	secondary	higher
Argentina	22	9: 1800-1880	none	none	none	7: 1800-1860	none
Bahamas	7	1: 1950	none	none	none	none	none
Barbados	7	none	none	none	none	none	none
Belize	7	none	none	none	none	none	none
Bolivia	15	7: 1880-1940	none	none	none	none	none
Brazil	22	7: 1800-1860	none	none	none	none	none
Chile	22	10: 1800-1890	none	none	none	none	none
Colombia	22	3: 1890-1908	none	none	none	none	1: 1890
Costa Rica	10	none	none	none	none	none	none
Cuba	22	14: 1800-1930	none	none	none	none	none
Dom. Rep.	7	none	none	none	none	none	none
Ecuador	15	8: 1870-1940	none	none	none	none	none
El Salvador	10	1: 1920	none	none	none	none	none
Guatemala	10	none	none	none	none	none	none
Guyana	7	none	none	none	none	none	none
Haiti	8	1: 1940	none	none	none	none	1: 1940
Honduras	10	1: 1920	none	none	none	none	none
Jamaica	20	6: 1820-1870	none	none	none	none	none
Mexico	22	10: 1800-1890	none	none	none	none	none
Nicaragua	13	none	none	none	none	none	none
Panama	8	none	none	none	none	none	none
Paraguay	8	none	none	none	none	none	none
Peru	12	none	none	none	none	none	none
Puerto Rico	7	none	none	none	none	none	none
Suriname	7	none	none	1: 1950	none	none	none
Trinidad	7	none	none	none	none	none	none
Uruguay	23	9: 1800-1890	none	none	none	none	none
Venezuela	23	14: 1800-1928	none	none	7: 1820-1880	none	none
Middle East							
Bahrain	7	none	none	none	none	none	1: 1950
Iran	21	13: 1820-1940	none	none	none	none	none
Iraq	20	13: 1820-1940	none	none	none	none	none
Jordan	20	13: 1820-1940	none	none	8: 1820-1890	8: 1820-1890	10: 1820-1910
Kuwait	7	none	none	none	none	none	none
Lebanon	20	13: 1820-1940	none	none	8: 1820-1890	8: 1820-1890	12: 1820-1930
Oman	7	none	none	none	none	none	none
Qatar	7	none	none	none	none	1: 1950	1: 1950
Saudi Arabia	7	none	none	none	none	none	none
Syria	20	13: 1820-1940	none	none	none	none	none
UAE	7	none	none	none	2: 1950-1960	2: 1950-1960	2: 1950-1960
Yemen	7	none	none	none	none	none	none

Country	obs.	age dist.	labor force	invest rate	primary	secondary	higher
North Africa							
Algeria	22	9: 1820-1900	none	none	none	none	none
Egypt	21	9: 1820-1900	none	none	none	none	none
Libya	7	none	none	none	none	none	none
Morocco	20	13: 1820-1940	none	none	none	none	none
Tunisia	21	13: 1820-1940	none	none	none	none	none

Table 7: Extent of Data Imputation

Region	obs.	age	labor	invest rate	primary	schooling secondary	higher
Western Countries	355	56	0	0	1	34	3
Southern Europe	126	32	0	0	0	5	0
Central & Eastern Europe	206	36	0	0	0	10	18
N.I.C.'s	106	54	0	0	7	29	29
Asia	281	117	0	1	24	28	18
Sub-Saharan Africa	356	35	0	20	1	1	13
Latin America	373	101	0	1	7	7	2
Middle East	150	65	0	0	18	19	26
North Africa	91	44	0	0	0	0	0
World	2044	540	0	22	58	133	109

1.6 Initial Young Human Capital: 15 to 24 age group

In this section we characterize how the initial young human capital is chosen for each country, for the model with intergenerational human capital accumulation. We did not follow a strict rule, but nonetheless it can be reproduced reasonably easy with the following empirical specification. Of course all of the values of age specific human capital including initial conditions are available in the data set. Here we merely report on the general features of the initial conditions we used. Fairly simple rules are capable of reproducing the initial human capital of the youngest workers, those age 15 to 24, and all other age categories of human capital. These latter age categories are: 25 to 34, 35 to 44, 45 to 54 and 55 to 64. We present In Table 9 the empirical results, and in Table 8 the dummy variables. We present below in Table 9 two sets of results. The top panel presents the regressions of our initial log human capital of each category against fixed rule variables. The bottom panel presents the goodness of fit of the empirical model of our initial human capital conditions.

First we define some variables. Let our assumed initial, year t, human capital for 15 to 24 year olds in country i be given by h_{it}^{15-24}. Construct the log relative human capital of 15 to 24 year olds in country i in year t relative to their 15 to 24 year old counterparts in the US be:

$$\ln rhc_i^{15-24,t} = \ln[hc_{it}^{15-24}] - \ln[hc_{USt}^{15-24}]. \tag{17}$$

Compute the log of output per worker in country i in year t relative to output per worker in the US in year t as:

$$\ln r_{y_i t} = \ln[y_{it}] - \ln[y_{USt}]. \tag{18}$$

In both of these definitions we placed a country's initial observation year into its respective decade. For example if we observe a country for the first time in year 1876, we place the observation in the 1880 decade and normalize by using the US value for 1880. Let T be any decade year, a year ending in zero. Then a year t lies in decade T if $t \in [T-5, T+4]$.

We used Schoellman (2012) to construct quality adjusted human capital measures from schooling.[4] We form the relative age category human capital from schooling as:

$$\ln \Gamma_{it}^{j} = .2[E_{it}^{j} - E_{US,t}^{j}]. \tag{19}$$

where E is the years of schooling for country i, in cohort t, and in age categories $j = 15-24, 25-34, 35-44, 45-54$ and $55-64$. We use some *country* dummy variables listed in Table 8. Table 9 contains the results of the regressions of our initial relative human capital against the relative income of the country, the difference in schooling of the country's age category with the schooling of the comparable age cohort in the US. The regressions include region dummies, and the aforementioned *country dummies*. This model does a good job of capturing our choices of initial relative human capital for all age cohorts. In all five categories the R^2 exceeds .96.[5] In the second panel of Table 9 we report the regression of our initial log relative human capital against the model's predicted log relative human capital. We report the fit of these regressions separately by region, and age category to show that the model does a good job of capturing our initial relative human capital distribution. The first term is the slope coefficient on the predicted value, and the second term is the R^2 of the region regression. For each age category there are 9 region results, for a total of 45 goodness of fit measures. The lowest R^2 value is .918, and only 3 other times is the R^2 below .950. The slope coefficient is greater than 1.02 only 4 times. It is lower than .98 only 2 times.

1.7 Special experience return values

For the countries of the *Central & Eastern Europe* region, we used lower experience returns during their communist period. These are detailed in Table 10. The special value arises from the expression:

$$\text{experience return} = \alpha * .0495. \tag{20}$$

We report the value of α in Table 10. In addition during the transition from the communist period, we removed all returns to experience during the communist period. Thus the 2000 and 2010 values of human capital are computed using only the schooling acquired during the communist period, and experience after the fall of communism in 1990. Thus in 2000, the most experience any cohort could have would be 10 years. This generally reduced human capital, but for some countries with exceptionally low returns to experience during communism, the human capital could actually rise.

[4] These measures come from Schoellman (2012) and personal correspondence.

[5] We used the same specification for all relative human capital values, with the exception of including our initial relative human capital values for younger ages, that is we did not search for different country dummies to use in each age cohort regression. Obviously we could have produced a better fit of the initial relative human capital used for each country by allowing for different specifications, but what we are presenting is essentially a primer as to what our assumed initial relative

human capital values.

Table 8: Country or Region Group Dummy Variables

Variable	Country or Countries	Variable	Country
australia	Australia & New Zealand	austria	Austria & Switzerland
belgium	Belgium & Canada	denmark	Denmark
finland	Finland	france	France & Sweden
iceland	Iceland	ireland	Ireland & Norway
lux	Luxembourg & Netherlands	cyprus	Cyprus
greece	Greece	israel	Israel
malta	Malta & Turkey	albania	Albania
armenia	Armenia, Azerbaijan & Belarus	baltics	Estonia, Latvia & Lithuania
georgia	Georgia	hungary	Hungary
moldova	Moldova & Tajikistan	stans	Kyrgyzstan, Moldova, Turkmenistan,
yugo	Yugoslavia		& Ukraine
afghanistan	Afghanistan, Bangladesh, Malaysia,	japan	Japan
	Myanmar & Papua New Guinea	bhutan	Bhutan, Cambodia & Pakistan
fiji	Fiji	china	China & India
nk	North Korea	mongolia	Mongolia
thailand	Thailand & Vietnam	philippines	Philippines
benin	Benin, Burundi, Central African	angola	Angola, Djibouti, Mali, Mauritius &
	Republic, Chad, Comoros, Congo,		Mozambique
	Ivory Coast, Kenya, Liberia,	botswana	Botswana, Burkina Faso, Eritrea,
	Madagascar, Malawi, Namibia,		Ethiopia & Niger
	Sierra Leone, Somalia, Togo,	cameroon	Cameroon, Ghana
	Zaire, Zambia & Zimbabwe	equatorial	Equatorial Guinea
soafrica	South Africa	nigeria	Nigeria & Sudan
bahamas	Bahamas, Guyana, Honduras	argentina	Argentina & Chile
	& Panama	belize	Belize, El Salvador, Jamaica,
boliva	Bolivia		Paraguay & Peru
colombia	Colombia	brazil	Brazil
haiti	Haiti	cuba	Cuba
nicaragua	Nicaragua	mexico	Mexico
trinidad	Trinidad	venezuela	Venezuela
bahrain	Bahrain	iraq	Iraq & Lebanon
kuwait	Kuwait	saudi	Saudi Arabia
uae	UAE	algeria	Algeria
libya	Libya		

Table 9: Log Relative Initial Young Human Capital Regressions

Variable	$\ln rhc^{1524}$	$\ln rhc^{2534}$	$\ln rhc^{3544}$	$\ln rhc^{4554}$	$\ln rhc^{5564}$
r_y	0.2927	0.0570	-0.0179	0.1714	0.0045
	(0.0293)	(0.0223	(0.0215)	(0.0594)	(0.0263)
$\ln \Gamma$	0.2223	-0.0993	0.0548	0.0561	-0.0363
	(0.0567)	(0.0336)	(0.0341)	(0.0970)	(0.0430)
$\ln rhc^{1524}$	-	1.0542	-0.6924	1.3722	-0.1771
		(0.0552)	(.1165)	(.3800)	(.1725)
$\ln rhc^{2534}$	-	-	1.7297	-4.1008	0.7495
	-	-	(0.0995)	(0.5783)	(.3079)
$\ln rhc^{3544}$	-	-	-	3.6943	-1.7015
	-	-	-	(0.2945)	(0.2096)
$\ln rhc^{4554}$	-	-	-	-	2.1574
	-	-	-	-	(0.0456)
constant	-0.5841	-0.4213	0.1536	0.5648	0.0263
	(0.1849)	(0.1017)	(0.1038)	(0.2893)	(0.1248)
R^2	.9746	.9936	.9957	.9829	.9984
Actual value regressed on predicted value					
Region	(β, R^2)	(β, R^2)	(β, R^2)	(β, R^2)	(β, R^2)
WC	(1.00, .985)	(0.99, .992)	(1.04, .921)	(0.99, .998)	(1.00, .999)
SE	(1.00, .998)	(1.00, .999)	(1.00, .999)	(1.01, .991)	(0.99, .999)
CEE	(1.00, .957)	(0.99, .983)	(1.01, .996)	(0.98, .991)	(0.97, .982)
NIC	(1.00, .999)	(1.00, .999)	(1.00, .925)	(0.98, .991)	(1.02, .995)
Asia	(1.00, .949)	(0.99, .997)	(0.99, .997)	(0.99, .975)	(0.99, .999)
SSA	(0.98, .959)	(1.00, .996)	(1.00, .999)	(1.03, .955)	(1.02, .998)
LA	(0.98, .955)	(0.97, .970)	(1.01, .995)	(1.00, .956)	(0.98, .998)
ME	(1.13, .962)	(1.09, .987)	(0.98, .997)	(0.96, .994)	(0.99, .998)
NA	(0.96, .964)	(1.00, .999)	(1.00, .999)	(1.00, .992)	(1.00, .999)

Table 10: Special experience returns: *Central & Eastern Europe*

Country	α	period
Albania	.0625	1945-2000
Armenia	.0625	1970-2000
Armenia	.2500	2010
Azerbaijan	.0625	1970-2000
Azerbaijan	.2500	2010
Belarus	.0625	1970-2000
Bulgaria	.0625	1945-2000
Czech Republic	.0625	1945-2000
Czech Republic	.2500	2010
Estonia	.0625	1970-2000
Georgia	.0625	1970-2020
Hungary	.0625	1945-2000
Kazakhstan	.0625	1970-2000
Kazakhstan	.2500	2010
Kyrgyzstan	.0625	1970-2000
Kyrgyzstan	.2500	2010
Latvia	.0625	1970-2000
Lithuania	.0625	1970-2000
Moldova	.0400	1970-2010
Poland	.0625	1945-2000
Romania	.0625	1945-2000
Russia	.0625	1945-1999
Slovak Republic	.0625	1990-2000
Slovak Repubic	.2500	2010
Tajikistan	.0400	1970-2010
Turkmenistan	.0400	1970-2000
Turkmenistan	.125	2010
Ukraine	.0625	1970-2000
Uzbekistan	.0625	1970-2000
Uzbekistan	.2500	2010
Yugoslavia	.0625	1945-2000
Yugoslavia	.2500	2010

2 Special Depreciation Rates

In this section we detail the special depreciation rates for some countries that attempt to take into account special circumstances, like the end of centrally planned communism, and exceptionally bad government institutions. We model this by introducing a special depreciation rate, $\hat{\delta}$. These occur in general for the year 2000, and for a few additional countries, the year 2010.

$$k_{2000} = i_{2000} y_{1990} \frac{\frac{g_y^9}{g_w}\left[1 - \frac{(1-\delta)^{10}}{g_y g_w^{11}}\right]}{1 - \frac{(1-\delta)}{g_y g_w}} + \frac{(1-\delta)^{10} k_{1990}}{g_w^{10}}(1-\hat{\delta}). \tag{21}$$

Table **??** below contains the special cases in depreciation rates, $\hat{\delta}$. All countries in region 3 *Central & Eastern Europe* were affected by this additional physical capital depreciation rate for 2000. In addition, Table **??** lists the countries that were affected in 2010. With the exception of Azerbaijan, Kyrgyzstan and Lebanon, these countries are at the very bottom of 179 countries ranked for economic freedom, from Miller, Holmes, and Feulner. Obviously other countries presumably have insecure property rights, which would require adjustments, but that is beyond the scope of this paper.

Table 11: Special Physical Capital Depreciation Rates

country	$\hat{\delta}$	year	note (rank out of 179 countries)
All Central & East European Countries	$\frac{1}{2}$	2000	end of communism
Azerbaijan	$\frac{1}{7}$	2010	91
Belarus	$\frac{1}{7}$	2010	153
Bolivia	$\frac{1}{4}$	2000 & 2010	146
Congo	$\frac{1}{2}$	2000 & 2010	167
Cuba	$\frac{1}{2}$	2000 & 2010	177
Eritrea	$\frac{1}{2}$	2000 & 2010	175
East Germany	$\frac{1}{4}$	1990	end of communism
Iran	$\frac{1}{2}$	2000 & 2010	171
Iraq	$\frac{1}{2}$	2000 & 2010	N/A
Kyrgyzstan	$\frac{1}{7}$	2010	88
Lebanon	$\frac{1}{2}$	2000	90
Libya	$\frac{7}{20}$	2000 & 2010	176
Moldova	$\frac{1}{7}$	2010	124
Myanmar	$\frac{1}{2}$	2000 & 2010	173
North Korea	$\frac{1}{2}$	2000 & 2010	least economic freedom, 179
Sierra Leone	$\frac{1}{2}$	2000 & 2010	152
Russia	$\frac{1}{7}$	2010	144
Tajikistan	$\frac{3}{5}$	2010	129
Turkmenistan	$\frac{1}{7}$	2010	168
Ukraine	$\frac{1}{7}$	2010	163
Uzbekistan	$\frac{1}{7}$	2010	164
Venezuela	$\frac{1}{2}$	2000 & 2010	174
Vietnam	$\frac{1}{4}$	2000 & 2010	136
Zaire	$\frac{1}{2}$	2000 & 2010	172
Zimbabwe	$\frac{1}{2}$	2000 & 2010	second lowest economic freedom, 178

Notes: Table reports our exceptions in depreciation rates. Rank refers to the 2012 *Index of Economic Freedom* World Rankings, where 1 means most free country, and 179 is lowest economic freedom. Iraq is unrated due to lack of reliable data, in of itself a damning statistic.

3 Western Countries

3.1 Australia (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1861, 1871, 1881, 1891, 1901, 1911, 1921, 1933, 1947, 1954, 1961, 1971, 1980, 1990 come from Maa Table A5 p. 64, 65 and 66. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1861, 1871, 1881, 1891, 1901, 1911, 1921, 1933, 1947, 1954, 1961, 1971, 1981 and 1990 come from Maa Table A2 p. 28. The age distribution for Australia in 1820, 1830, 1840, 1850 are assumed to be that of 1861. The age distribution for Australia for 1990 is interpolated from 1981 and 1992 values. The age distribution for Australia for 2000 and 2010 come from *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force data for 1820 and 1830 come from Table 3.9 from *Forming a Colonial Economy: Australia 1810-1850* by N. G. Butlin. Labor force data for 1840, 1850, 1861, 1871, 1881, 1891, 1901, 1911, 1921, 1931, 1951, 1961, 1971, 1981, 1991, 2000, and 2010 come from Table A2 of *The Cambridge Economic History of Australia*, edited by Simon Ville and Glenn Withers. The Table was produced by Matthew Butlin, Robert Dixon and Peter Lloyd. Male and female labor force values for 1820 and 1830 come from Butlin. We applied Butlin's male and female shares for 1840 and 1850 to our labor force data from Butlin, Dixon and Lloyd values. For years 1861, 1871, 1881, 1891, we used the average male share of labor force for years 1840, 1850, and 1901. To compute male and female labor force values, we used information for 1901, 1911, 1921, 1931, 1941, 1951, 1961, 1971, 1980 and 1990 come from Maa Table B1 p. 102. Male and female shares of labor force for 2000 come from WDR, and male and female shares of labor force for 2010 come from *WDI*.

Real GNPs come from Maddison. We convert his estimate of output per capita by multiplying by population and dividing by our estimate of labor force. For 1820, 1830, 1840, and 1850 we used the sectoral output shares from Table A1 of *The Cambridge Economic History of Australia*, edited by Simon Ville and Glenn Withers. The Table was produced by Matthew Butlin, Robert Dixon and Peter Lloyd. We then applied the sectoral capital-output ratio for the US for those years, except that we used the 1850 US values for 1820, 1830 and 1840. These formed the initial conditions for our perpetual inventory calculations beyond 1850. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1861-1998 from Maa Table J1, pp. 1039, 1040 and 1041 and *WDR* (various years). For years 1971, 1981, 1991, 2000 we used the capital - output ratio from Pitketty and Zucman (2013). For 2010 we used the average investment rates from S & H online for 2000-2009, and perpetual inventory.

We assumed 0 enrollments in higher education in 1820, 1830, 1840 and 1850. We have .01 percent higher education enrollment rate and .1 percent secondary school enrollment rate for 1860. We used an estimate of 25 percent enrollment in primary school based on Steckel and Floud (1997) literacy estimate in 1820 and 1830 Australia. For 1820 and 1830 secondary school enrollment rates, we assumed 1.5 percent, compared with the 1840 datum of 1.89 percent. Enrollments in school from 1840-1991, come from Maa Table I1 pp. 992 and 993. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Our 1840 primary school enrollment rate is 29.5 percent. However since primary and secondary school enrollments are not reported separately until 1970, we used the following apportionment: 95 percent of students were in primary school in 1840, 1850, 1861, 1871. Eighty percent

of students in 1881 were primary students. Seventy-five percent of students in 1891, 1901, 1911, 1921 were primary students. Seventy percent of students in 1931 were primary students. Two thirds of students in 1947, 1954 were primary students. Fifty-five percent of students in 1961 were primary students. Higher education enrollments for 1906-1998 from Maa Table I2 p. 1006. For 2010 we used *HDR*. We interpolated for all enrollment rates in 2000. For years prior 1820 to 1901 we used enrollment rates of 0, 0, 0, 0, .0001, .001, .001, .001, .002. With our assumptions in the early years, we are able to match the Morrisson & Murtin education shares from 1870-1900. Our time series of years of schooling in the labor force for 1871-1921 is: 2.61 (1871), 4.13 (1881), 5.02 (1891), 5.77 (1901), 6.30 (1911) and 6.67 (1921). The Morrisson & Murtin time series of years of schooling is: 3.06 (1870), 4.15 (1880), 5.28 (1890), 6.25 (1900), 7.06 (1910) and 7.71 (1920).

3.2 Austria (1820-2010)

Population figures are for the Austrian Provinces of the Hapsburg Empire, 1820, 1830, 1842, 1850, 1860, 1869, 1880, 1890, 1900 and 1910, from Meu Table A2 p. 13. Populations for Republic of Austria 1923, 1934, 1951, 1961, 1971, 1981, 1991 come from Meu Table A2 p. 13. Population for Austria 2000 comes from *Time Almanac 2001*. Population for Austria 2010 comes from *Wikipedia*.

The age distributions of the population for 1869, 1880, 1890, 1900, 1910, 1923, 1934, 1951, 1961, 1971, 1981, 1991 come from M (1980) Table A2 p. 13. We assumed the same age distribution in 1820, 1830, 1842, 1850, 1860 as in 1869. Age distribution for Austria 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1869, 1880, 1890, 1900, 1910, 1920, 1934, 1939, 1951, 1961, 1971, 1981, 1991 are from Meu Table B1 p. 145. The 1869 labor force only contains information on all workers. We assumed the female labor force in 1869 was the same proportion of the male labor force as it was in 1880. Labor force data for 2000 come *WDR*. Labor force data for 2010 come from *WDI*. For 1820, 1830, 1842, 1850, 1860 we used the following procedure.[6] We used Clio infra for urbanization rates for 1800, 1850, 1900 and 1950, and *WDI* for 1960-2010. We interpolated for 1820-1840, 1860-1890, 1910-1940. We used rural 15-64 labor force participation rates of 95% for 1820-1869, 91.55% for 1880, 95% for 1890-1910, 78.62% for 1923, 52.32% for 1934, 65.1% for 1941, 84.25% for 1951, 95% for 1961-2010. We assumed urban 15-64 labor force participation rates of 80.8% for 1820-1869, 50% for 1880, 74.57% for 1890, 61.6% for 1900, 79.3% for 1910, 50% for 1923-1951, 61.14% for 1961, 53.05% for 1971, 55.7% for 1980, 51.95% for 1990, 56.13% for 2000 and 65% for 2010. We constructed the ratio of this labor force series with that from Meu, and *WDR* and *WDI*. The root mean ratio of the overlapping years 1869-2010 is 1.000. The 1869, 1880 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real GNPs come from Maddison. Sectoral output shares come from Meu Table J2 p. 929. They cover years 1910, 1923, 1934, 1951, 1961, 1971, 1980, 1990. For 2000 and 2010 we used *WDI*. For these years we produce capital by applying the US sectoral capital output ratios. These are used to check our calculations below. For the 1910-2000 period we used the US sectoral capital output ratios to produce estimates of aggregate physical capital. However these are not used, but serve as a check on our estimates

[6] Our 1820 value is nearly identical with R.C. Allen for 1820, with the same population as our 1820 population, and the assumption of more than 95% of the population living in rural areas. Our 1850 value is consistent with R.C. Allen adjusting for 1850 population.

given below. For 1941 we used the 1951 sectoral shares and the sectoral growth rates for agriculture and manufacturing relative to aggregates growth rates in Sabillon p. 179. He provides growth rates for the 1950s, and separately for 1900-1949 period. For 1820, 1830, 1842, 1850, 1860, 1869, 1880, 1890, 1900 and 1910, we used the agriculture and manufacturing sectoral growth rates relative to aggregate growth rates in Sabillon p. 179 applied to the 1910 base year. He provides separate growth rates for the 19th century. We then used the sectoral capital output ratios from the US for years 1850-1860 to produce capital stock measures for these years. For 1820, 1830 and 1840 we applied the 1850 US sectoral capital output ratios. For 1869-1910 we used capital stock estimates from Max-Stephan Schulze (2007). These for the initial conditions for our perpetual inventory calculations for 1923 onward. Physical capital investment rates comes from intraperiod averages of real gross capital formation and real income for 1913-1998 from Meu Table J1 pp. 908, 914 922. For 2010 we used the average investment rate from 2000-2009 from S & H.

For 1830 we used Lindert for primary school enrollments, which produces a 9.45 percent primary enrollment rate, which is indistinguishable from the 1842 and 1850 primary enrollment rates from Meu. We assumed a .5 percent enrollment rate for secondary school, and a .35 percent enrollment rate in higher education. These are in line with the .9 percent secondary enrollment rate and .4 percent higher education enrollment rate for 1842 from Meu. We assumed identical enrollment rates in 1820 as from 1830. This is consistent with the essentially constant literacy rate of birth cohorts 1801-1810, 1822-1820, 1821-1830, 1831-1840, Roser. Enrollments in primary and secondary school from 1842-1998 come from Meu Table I1 pp. 870, 873, 880 887. To calculate enrollment rates prior to 1971, we assumed 6-11 are primary school age and 12-17 are secondary school age. In 1971 we assumed that primary school lasts 8 years and secondary school lasts 4 years. Therefore 6-13 are primary school age and 14-17 are secondary school age. This switch occurred to fit with the enrollment rate data in WDR for 1980, 1990. Values for 2010 came from *HDR*. Higher education enrollments are from Meu Table I2 pp. 894, 895, 897, and 899. Our time series of years of schooling in the labor force for 1869-1923 is: .52 (1869), 1.47 (1880), 2.33 (1890), 3.08 (1900), 3.57 (1910) and 3.87 (1923). The Morrisson & Murtin time series of years of schooling is: 3.20 (1870), 3.44 (1880), 3.94 (1890), 4.63 (1900), 5.31 (1910) and 5.92 (1920).

3.3 Belgium (1820-2010)

Populations for 1820, 1830, 1840 1846, 1856, 1866, 1880, 1890, 1900, 1910, 1920, 1930, 1947, 1961, 1970, 1981, 1991 come from Meu Table A2 p. 14. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1846, 1856, 1866, 1880, 1890, 1900, 1910, 1920, 1930, 1947, 1961, 1970, 1981, 1991 come from Meu Table A2 p. 14. The age distribution for 1820, 1830 and 1840 are assumed to be identical to the age distribution in 1846. Age distribution for Belgium 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in 2011.

Labor force figures for 1846, 1856, 1866, 1880, 1890, 1900, 1910, 1920, 1930, 1947, 1961, 1970, 1981 and 1990 are from Meu Table B1 p. 146. Labor force data for 2000 come from *WDR*. Labor force data for 2010 comes from *WDI*. For 1820, 1830 and 1840 we used the following procedure. We used Clio infra for the urban population shares for 1800, 1850, 1900 and 1950, and *WDI* for 1960-2010. We assumed a rural 15-64 labor force participation rate of 88.115% for 1820-1840.[7] We assumed a rural 15-64 labor force

[7]This is our average 15-64 labor force participation rate for 1846-1880.

participation rate of 83.22% for 1846, 89.24% for 1856, 90% for 1866-1880, 86.65% for 1890, 86.3% for 1900, 83.3% for 1910, 65.43% for 1920, 72.5% for 1930, 65.91% for 1940, 61.24% for 1947, 95% for 1961-1970, 65% for 1980-1990, 85% for 2000-2010. We assumed an urban 15-64 labor force participation rate of 50% for 1820-1856, 55% for 1866, 55.55% for 1880, 50% for 1890-1947, 55.33% for 1961, 55.59% for 1970, 50.12% for 1981, 58.32% for 1990, 64.31% for 2000 and 66.58% for 2010.[8] We constructed the ratio of these labor force values and those from Meu, *WDR* and *WDI*. The root mean ratio for overlapping years 1846-2010 is 1.000. The 1846 and 1856 values are 1.000 and 1.000, respectively. The range of the ratio is 1.000 to 1.000.

Real GNP come from Maddison. Sectoral output shares for years 1947, 1961, 1970, 1980, 1990 come from Meu Table J2 p. 929. Sectoral output shares for 2000 and 2010 come from *WDI*. For years prior to 1947; 1820, 1830, 1840, 1846, 1856, 1866, 1880, 1890, 1900, 1910, 1920, 1930 and 1940, we used agriculture and manufacturing growth rates relative to aggregate growth rates in Sabillon p. 179. He provides these growth rates for the 1900-1949 period, and separately for the 19th century. We then applied the US sectoral capital output ratios for 1850-1940 to compute aggregate capital. For years prior to 1850, we used the 1850 US sectoral capital output ratios. Physical capital investment rates come from the intraperiod averages of real gross capital formation and real income for 1920-1998 from Meu Table J1 pp. 908, 914 922. Investment rate for 2010 is the average investment rate from 2000-2009 from S & H. For 1846-1970 we used Goldsmith (1985) to produce our capital stock estimates. From 1980-2010 we used perpetual inventory.

Enrollments in primary and secondary school from 1830-1993 come from Meu Table I1 pp. 870, 873, 880 887. We used the 1830 enrollment rates for 1820: 56.6 percent primary enrollment rate, 2.9 percent secondary enrollment rate and .1 percent higher education enrollment rate. This is consistent with the literacy rates in Belgium attained in 1800, Roser. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments are from Meu Table I2 pp. 894, 895, 897 and 899. The 2000 values from *WDR*. The 2010 enrollment rates come from *HDR*. Our time series of years of schooling in the labor force for 1866-1920 is: 4.68 (1866), 5.18 (1880), 4.40 (1890), 4.62 (1900), 5.18 (1910) and 5.62 (1920). The Morrisson & Murtin time series of years of schooling is: 4.27 (1870), 4.79 (1880), 5.17 (1890), 5.17 (1900), 5.33 (1910) and 5.57 (1920).

3.4 Canada (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1871, 1881, 1891, 1901, 1911, 1921, 1931, 1941, 1951, 1961, 1971, 1981 and 1991 come from Mam Table A2 p. 11. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distribution of the population for 1850, 1860, 1871, 1881, 1891, 1901, 1911, 1921, 1931, 1941, 1951, 1961, 1971, 1981 and 1991 come from Meu Table A2 p. 11. We used the age distribution in 1850 for 1820, 1830 and 1840. Age distribution for Canada 2000 and 2010 come from the *Demographic Yearbook*.

Labor force for 1871 comes from Maddison (2001) Table E-1, p.345. Labor force for 1881 comes from *Statistics Canada*, and assuming a 5% unemployment rate. Labor force figures for 1891, 1901, 1911, 1921, 1931, 1951, 1961, 1971, 1981 and 1991 come from Mam Table B1 p. 102. Labor force data for 2000 come from *WDR*. Labor force data for 2010 comes from *WDI*. Labor force data for 1820, 1830, 1840, 1850, 1860 is produced via the following procedure. We used Bairoch and Goertz for urban-rural population shares

[8] Our 1820 value is consistent with population adjustment of R.C. Allen in 1800, and assuming a 62% labor force participation rate of rural agricultural population, and a 38% labor force participation rate of non agricultural population and urban populations.

for 1820-1911, and Banks for 1921-1961, and *WDR* and *WDI* for 1971-2010. We assumed that the 15-64 rural labor force participation rate is 62.63% for 1820-1871, 59.3% for 1881, 58.4% for 1891, 57.5% for 1901, 62.5% for 1911, 61.91% for 1921, 62.5% for 1931-1961, 62.62% for 1971, 71.12% for 1980, 79.5% for 1990, 78.5% for 2000, for 78.5% for 2010. We assumed that the 15-64 urban labor force participation rate is 62.63% for 1820-1871, 50% for 1881-1901, 57.88% for 1911, 50% for 1921, 54.1% for 1931, 52.9% for 1941, 58.7% for 1951, 59.91% for 1961, 62.62% for 1971, 71.12% for 1980, 79.5% for 1990, 74.45% for 2000, 78.03% for 2010. We constructed the ratio of this labor force measure with that from Maddison, *Statistics Canada*, Mam, *WDR* and *WDI*. The root mean ratio for overlapping years 1871-2010 is 1.000. The ratios for 1871 and 1881 are 1.000, 1.000, respectively. The range of ratios is 1.000 to 1.000.

Real GNP for 1820, 1830, 1840, 1850, 1860, 1871, 1881, 1891, 1901, 1911, 1921, 1931, 1941, 1951, 1961, 1971, 1980, 1990, 2000, 2010 come from Maddison. Sectoral output shares for years 1921, 1931, 1941, 1951, 1961, 1971, 1980, 1990 come from come from Mam Table J2 pp. 788-789. Sectoral output shares for 2000 and 2010 come from *WDI*. We used the 1925-1929 value from Mam for 1921. For all years prior to 1921, 1820, 1830, 1840, 1850, 1860, 1871, 1881, 1891, 1901 and 1911 we used the agriculture and manufacturing growth rates relative to the aggregate growth rates from Sabillon p. 147. He reports these growth rates for 1900-1949, and the 19th century separately. We then used the US sectoral capital output ratios for 1850-2000. Physical capital investment rates come from the gross real capital formation and real income for 1870, 1890, 1900, 1910 and 1920 come from Mam Table J1, pp. 762, 763. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1926-1998 from Mam Table J1, pp. 763 and 767 and *WDR* (various years). We used the Piketty and Zucman (2014) estimates for 1971, 1980, 1990, 2000. For 2010 we used the average over the 2000-2009 period from S & H. Thus we used the 1860 physical capital stock from our sectoral calculations as the initial condition for our perpetual inventory calculations for years 1871-1961, inclusive, and perpetual inventory from 2000 to 2010 in order to create our 2010 value. The regression of log capital from perpetual inventory against the log capital from the sectoral calculation produces an R^2 of .8853, with a slope coefficient on $ln(k_{1985})$ of .897 and a standard error of (.114). The intercept term is not significant. The sectoral capital estimates are very close to our perpetual inventory calculations.

Enrollments for 1850 come from literacy rates, from Morris and Adelman (1988), of 65 percent (we assumed a 75 percent enrollment rate), and for 1860 the geometric average of 1850 and 1871 enrollment rates from Mitchell. We assumed 60 percent primary school enrollment rates for 1820-1840. Enrollments in school from 1868-1993 come from Mam Table I1 pp. 718, 721, 725 and 730. Since school enrollments are not broken down into primary and secondary categories, we tried to fit the enrollment rates to those from the WDR, mainly 1961 and 1971. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. We assumed 75 percent of total enrollments were in primary school in 1871, 1881, 1891, 1901. We assumed 70 percent of total enrollments were in primary school in 1911, 1921. We assumed that 67 percent of total enrollments were in primary school in 1931, 1941, 1951. This produces a 96 percent enrollment rate in primary school in 1871. We have 77 percent exposure rate to schooling in 1871, with 76 percent exposed to primary school, which exactly matches the 76 percent primary school exposure rate in Morrison & Murtin. Higher education enrollment rates are for 1920-1998 from Mam Table I2 pp. 751, 752 and 754. For prior years, 1871-1911 we used .002, .005, .005, .01 and .01. For 2010 we used data from *WDI*. Our 2010 enrollment rates come from *HDR*. Our time series share of labor force exposed to higher education is consistent with Morrisson & Murtin 1870-2010, although a tad low over the

1940-1970 period. Our time series of years of schooling in the labor force for 1871-1921 is: 4.67 (1871), 5.32 (1881), 5.67 (1891), 6.04 (1901), 6.24 (1911), 6.70 (1921). The Morrisson & Murtin times series of years of schooling is: 5.71 (1870), 6.20 (1880), 6.68 (1890), 7.10 (1900), 7.63 (1910), 8.04 (1920).

3.5 Denmark (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1901, 1911, 1921, 1930, 1940, 1950, 1960, 1970, 1981 and 1990 come from Meu Table A2 p.17. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1840, 1850, 1860, 1870, 1880, 1890, 1901, 1911, 1921, 1930, 1940, 1950, 1960, 1970, 1981 and 1990 come from Meu Table A2 p. 17. We assumed the 1820 and 1830 age distributions were identical to the 1840 age distribution. Age distribution for Denmark 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in 2013.

Labor force figures for 1870, 1880, 1890, 1901, 1911, 1921, 1930, 1940, 1950, 1960, 1970, 1981 and 1991 come from Meu Table B1 p. 147.[9] Federico (2010) provides agricultural labor force for 1820, 1910, 1938, 2000, we used the Mitchell value for 1880.[10] We interpolated these values to construct agricultural labor force for 1820-2000. We constructed the share of labor force that is agricultural for 1870-2000, and regressed this against log year. We predicted this share back to 1820, and applied it to produce our labor force values for 1820-1860. Labor force data for 2000 come from *WDR*, and the labor force data for 2010 comes from *WDI*.

Real GNP from 1820-2010 comes from Maddison. Sectoral output shares for Physical capital investment rates come from the intraperiod averages of gross real capital formation and real income for 1850-1998 from Meu Table J1, pp. 906, 909, 915 and 922 and *WDR* (various years). The 1820, 1830 and 1840 capital stock estimates come from our sectoral output shares. Mitchell provides sectoral output shares from 1820-1990, inclusive. We used *WDI* for sectoral shares in 2000 and 2010. We applied the US sectoral capital output ratios in 1850 to the sectoral output shares in 1820, 1830 and 1840 to create physical capital stock estimates. From 1850 onward we used perpetual inventory methods to compute physical capital. The 2010 value is the average investment rate over 2000-2009 from S & H.

We used a 70 percent enrollment rate in primary school in 1820 and 1830, taken from the introduction of compulsory primary schooling in 1814 and approximately half were enrolled by that date, Gold (1996). We kept this value for 1840, 1850, 1860 and 1870 so as to fit the 76 percent enrollment rate in 1880 from Lindert. These rates are lower than the estimated literacy rates of 95 percent in 1850 from Morris and Adelman, and 95 percent from Morris and Adelman in 1870, and 81 percent from Crafts. This produces a primary exposure rate of 68 percent, matching Morrison & Murtin. For 1820-1870 we used 1.5 percent for secondary enrollment rates, which matches the 1880 Lindert value. Enrollments in primary and secondary school from 1893-1993 come from Meu Table I1 pp. 874, 881 and 887. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments from 1893-1993 come from Meu Table I2 pp. 895 and 897. We used .05 percent for 1850-1880 and .01 percent for 1820-1840. For 2000 enrollment rates we used *WDR*. For 2010 enrollment rates we used *HDR*..

[9] These numbers are similar, but typically lower than the values contained in Henriksen. We chose to use Mitchell for consistency in source.

[10] Our 1820 value is interpolated using Federico (2010) 1800 value and our 1880 value from Mitchell.

Our time series share of labor force exposed to higher education is consistent with Morrisson & Murtin 1870-2000, but we are low over the 1950-1970 period. Our time series for years of schooling in the labor force for 1870-1921 is: 4.26 (1870), 4.27 (1880), 4.46 (1890), 5.16 (1901), 5.57 (1911) and 6.00 (1921). The Morrisson & Murtin time series of years of schooling is: 4.69 (1870), 4.94 (1880), 5.25 (1890), 5.61 (1900), 5.96 (1910) and 6.33 (1920).

3.6 Finland (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1865, 1880, 1890, 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970, 1980 and 1990 come from Meu Table A2 p. 18. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distribution of the population for 1850, 1865, 1880, 1890, 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970, 1980 and 1990 come from Meu Table A2 p. 18. We assumed the age distributions for 1820, 1830 and 1840 were identical to the age distribution in 1850. Age distribution for Finland 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1820, 1830, 1840, 1850, 1865, 1880, 1890, 1900, 1910 and 1920 use data from Grigg (1987) on labor force in agriculture, adjusting for share of labor in agriculture. For this agricultural share we use the 1880, 1900, 1910 and 1920 agriculture share from Meu. We adjusted the 1900 figure labor force for female workers in agriculture that is 203 instead of 103 in order to better fit the time series. We interpolated for 1890. Labor force for 1930, 1940, 1950, 1960, 1970, 1980, 1990 come from Meu Table B1 p. 148. Labor force data for 2000 come from *WDR*. Labor force data for 2010 comes from *WDI*.

Real GNP come from Maddison. Physical capital investment rates come from the intraperiod averages of gross real capital formation and real income for 1865-1998 from Meu Table J1, pp. 906, 909, 915 and 922 and *WDR* (various years). For 1820, 1830, 1840, 1850 we used sectoral capital estimates. Mitchell provides sectoral output shares for 1865-1990, and *WDI* provides shares for 2000 and 2010. We used Sabillon (2005) to produce sectoral output shares for 1820-1850. Sabillon provides agriculture, and manufacturing growth rates along with aggregate growth rates in order to produce these estimates. We applied the US 1850 capital - sectoral output ratios for 1820, 1830, 1840 and 1850. For years beyond 1850 we used perpetual inventory methods. The 2010 value was the average investment rate from 2000-2009 from S & H.

Enrollments in primary and secondary school from 1875-1993 come from Meu Table I1 pp. 874, 881 and 887. For 1820, 1830, 1840, 1850 and 1865 we used 10.9 percent for enrollment rates in primary school, respectively, and 3.5 percent for enrollment rates in secondary school (1820-1865). The secondary school enrollment rate series matches the share of the population in 1870 exposed to secondary school from Morrison & Murtin. These are also consistent with the literacy rate in Norway in 1800 from Roser. The 1865 primary enrollment rate is consistent with the12 percent literacy rate in 1870 in Crafts. These match with the data for 1880 (1890) values of 11 percent (17 percent) for primary school enrollment rates, and 3.6 percent (3.6 percent) for secondary school enrollment rate rates. For 1820 we used .22 percent for higher education enrollment rates, the average value from 1830-1865. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments from 1830-1993 are from Meu Table I2 pp. 894, 895, 897 and 899. We used *WDR* for 2000 enrollment rates, and we used *HDR* for 2010 enrollment rates. Our time series of years of schooling in the labor force for 1865-1920 is: .94 (1865), .96 (1880), .91 (1890), 1.13 (1900), 1.64 (1913) and 2.36 (1920). The Morrisson & Murtin

years of schooling is: 1.45 (1870), 1.53 (1880), 1.62 (1890), 1.73 (1900), 2.00 (1910) and 2.71 (1920).

3.7 France (1800-2010)

Populations for 1800, 1810, 1820, 1830, 1840, 1850, 1861, 1872, 1881, 1891, 1901, 1911, 1921, 1931, 1946, 1954, 1962, 1968, 1975, 1980 and 1991 come from Meu Table A2, pp. 19 and 20. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distribution of the population for 1850, 1861, 1872, 1881, 1891, 1901, 1911, 1921, 1931, 1946, 1954, 1962, 1968, 1975, 1980 and 1991 come from Meu Table A2, pp. 19, 20 and 21. The age distribution for 1800, 1810, 1820, 1830, 1840 are assumed to be identical with the age distribution from 1850. Age distribution for France 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force data for 1861, 1872, 1881 and 1891 come from interpolations of data for 1856, 1866, 1886, 1896 from Meu Table B1, p. 149. Labor force figures for 1901, 1911, 1921, 1931, 1946, 1954, 1962, 1968, 1975, 1981 and 1991 come from Meu Table B1 p. 149. We interpolated for our 1940 value. The 1981 value comes from interpolations using the 1975, 1982 values. Labor force data for 2000 come from *WDR*. The 2010 labor force comes from *WDI*. For 1800-1850, we used the following procedure. We used Clio infra for urban population shares for 1800, 1850, 1900, 1950, and *WDI* for 1962-2010. We interpolated for urban shares 1810-1840, 1860-1890, 1910-1946. WE assumed rural 15-64 labor force participation rate of 80.495% for 1800-1850, 67.9% for 1861, 74.805% for 1872, 79.315% for 1881, 81.94% for 1891, 85% for 1901-1946, 81.28% for 1954, 85% for 1962-2010.[11] We assumed an urban 15-64 labor force participation rate of 50% for 1800-1891, 61.2% for 1901, 70.3% for 1911, 74.23% or 1921, 58.65% for 1931, 59.32% for 1940, 58.6% for 1946, 50% for 1954, 57.84% for 1962, 55.16% for 1968, 55.4% for 1975, 51.36% for 1980, 60.8% for 1991, 63.09% for 2000 and 67.5% for 2010.[12] We constructed the ratio of this labor force with that from Meu, *WDR* and *WDI*. The root mean ratio for the overlapping years 1861-2010 is 1.000. The 1861 and 1872 values are 1.000 and 1.000. The range of the ratio is 1.000 to 1.000.

Real GNP come from Maddison. For all years prior to 2010, we used Piketty and Zucman (2014) capital output ratios to produce our estimates of physical capital. For 2010 we used perpetual inventory using the investment average rate over the 2000-2009 period from S & H.

The 1800 enrollment rate for primary school comes from literacy rates from Steckel and Floud. We assumed a .5 percent enrollment rate for secondary school, and a .05 percent enrollment rate for higher education. The 1830 and 1840 enrollment rates for primary school and secondary school come from Lindert. The 1810 and 1820 enrollment rates are interpolated from the 1800 and 1830 rates. Enrollments in primary and secondary school from 1850-1998 come from Meu Table I1, pp. 870, 874, 882 and 888. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1889-1993 are from Meu Table I2, pp. 895 897 and 899. We used .1 percent for 1800-1830, .2 percent for 1840-1861, .3 percent for 1869 and 1881 for higher education enrollment rates. The 2000 enrollment rates come from *WDI*. The enrollment rates for 2010 come from *HDR*.The higher education share fits the 1870-1920 shares in Morrisson & Murtin. Our time series for years of schooling in the labor force for 1869-1921 is: 4.20 (1869), 4.51 (1881), 4.73 (1891), 4.88 (1901), 4.99 (1911) and 5.10

[11] Our 1800-1850 value is the average of our rural 15-64 labor force participation rates for 1861-1931).

[12] Our 1800-1850 labor values are almost identical with those in Roser's "Agricultural Employment" from Wrigley.

(1921). The Morrisson & Murtin time series for years of schooling is: 4.12 (1870), 4.94 (1880). 5.79 (1890), 6.31 (1900), 6.99 (1910) and 7.53 (1920).

3.8 Germany (1800-2010)

Populations for 1800, 1810, 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1910, 1925, 1933 and 1939 come from Meu Table A1, p. 4. Populations for West Germany 1950, 1961, 1970 and 1980 come from Meu Table A1, p. 4. Population for West Germany in 1990 comes from S & H online. Population for Germany 2000 comes from *Time Almanac 2001*. Population for Germany 2010 comes from *Wikipedia*.

The age distributions of the population for 1871, 1880, 1890, 1900, 1910, 1925, 1933 and 1939 come from Meu Table A2 pp. 21 and 22. The age distributions for years prior to 1871 are assumed to be identical to the age distibution in 1871. The age distributions for West Germany 1950, 1961, 1970 and 1980 come from Meu Table A2 p. 22. Age distribution for West Germany for 1990 comes from DK (1994). Age distribution for Germany 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1882, 1895, 1907, 1925, 1933, 1939 come from Meu Table B1 p. 150. West German labor force data for 1946, 1950, 1961, 1970 and 1980 and 1990 are from Meu Table B1 p. 150. Labor force data for 2000 come from *WDR*. Labor force data for 2010 comes from *WDI*. Labor force figures for years prior to 1880 are derived from the following procedure. We used Clio infra for urbanization rates for 1800, 1850, 1900 and 1950, and *WDI* for urbanization rates for 1961-2010. We interpolated for 1810-1840, 1860-1890, 1910-1940. We assumed a rural 15-64 labor force participation rate of 68.65% for 1800-1880, 75.8% for 1890, 76.52% for 1900, 74.355% for 1910, 80% for 1918-1939, 75.56% for 1950, 80% for 1961-2010. We assumed uran 15-64 labor force participation rates of 50% for 1800-1910, 50.65% for 1918, 63.23% for 1925, 55.4% for 1933, 58.14% for 1939, 50% for 1950, 67.24% for 1961, 64.63% for 1970, 57.7% for 1980, 65.13% for 1990, 68.61% for 2000 and 76.59% for 2010.[13] We constructed the ratio of this labor force with that from Meu, *WDR* and *WDI*. The root mean ratio is 1.000. The ratio values for 1882, and 1895 are 1.000 and 1.000, respectively. The range of the ratio is 1.000 to 1.000.

Real GNP data for all years come from Maddison. For 1800-1840 we used sectoral output shares from *Germany: The Cambridge Economic History of Europe VII The Industrial Economies: Capital, Labour, Enterprise Part I*, where we used 1820 sectoral shares for 1800 and 1810. For 1850, 1860, 1870 and 1918 we used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares. Mitchell provides sectoral output shares for 1880-1910 and 1925-2000. For years prior to 1850 we used the sectoral output shares from Mitchell and the 1850 US capital - sectoral output ratios. For years 1850 and 1860 we used Goldsmith (1985). For 1870-2000 we used Piketty and Zucman (2014). For 2010 we used the average investment rate from 2000-2009 from S & H.

Enrollment rates in 1850-1880 are Easterlin enrollment data. The 1890 value is an interpolation of the 1880 data with the 1900 data. The 1830 and 1840 primary enrollment data are from Lindert. The 1800, 1810 and 1820 data are based on Steckel and Floud literacy rates for 1800, and 1830 Lindert data, and interpolations between these base years. Enrollments in primary, 1900-1993, and secondary school, 1910-1993, come from Meu Table I1, pp. 875, 882, and 888. Secondary enrollment rates from 1830-1910 are

[13] Our figures for 1800-1870 are consistent with information in Pierenkemper and Tilly in Graph 6. In particular, our 1800 figure is matched with their figure.

from Lindert. We assumed a 1 percent enrollment rate in years prior to 1830. For higher education prior to 1870, we assumed enrollment rates to be the average of higher education enrollment rates in France and England. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1872-1998 comes from Meu Table I2, pp. 895, 897 and 899. Our enrollment rates for 2000 come from *WDR*. Enrollment rates for 2010 come from *HDR*. Our time series share of labor force exposed to higher education is consistent with Morrisson & Murtin 1870-2010, but we are on the low side from 1950-1970. Our time series for years of schooling in the labor force for 1870-1925 is: 3.59 (1870), 3.83 (1880), 3.99 (1890), 4.25 (1900), 4.52 (1910), 4.70 (1918) and 4.89 (1925). The Morrisson & Murtin time series of years of schooling is: 5.44 (1870), 5.70 (1880), 6.10 (1890), 6.47 (1900), 6.99 (1910) and 7.45 (1920).

3.9 Iceland (1950-2010)

Population for 1950, 1960, 1970, 1980, 1990 come from S & H. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Iceland 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force for 1950, 1960, 1970, 1980 and 1990 come from *WDI*. Labor force for 2000 comes from internet government website. The labor force for 2010 comes from *WDI*.

Real GNP for 1950, 1960 and 1970 come from Penn World Tables 6.1. The 2010 comes from the World Bank rebenchmarking project, The 1980, 1990 and 2000 values come from the growth rates in *WDI* and the 2006 base. The 2000 value is 109 percent of PWT 6.1, The 1990 value is 100 percent of PWT 6.1. The 1980 value is 101 percent of PWT 6.1. Physical capital investment rates come from the intraperiod averages of gross real capital formation and real income for 1950-2000 PWT 6.1. For 1950 we used the sectoral output shares and the 1950 US capital - sectoral output ratios to produce our estimate. The 2000 and 2010 sectoral output shares come from *WDI*. We used Sabillon (2005) manufacturing growth rates and aggregate growth rates to produce industry shares for years 1950-1990. We assumed that farming was 25% of the non manufacturing sector, and that services were the remaining 75% of the non manufacturing sector for years 1950-1990. These are their respective shares for 2000 and 2010. The 2010 value is the average investment rate for 2000-2009 in S & H.

The 1960 primary school, secondary school and higher education enrollments are from 1954 *UN Statistical Yearbook*. Our 1960 enrollments are from the 1961 *UN Statistical Yearbook*. Our 1970 enrollments come from the 1973 *UN Statistical Yearbook*. Our 1980 enrollments are from the 1981 *UN Statistical Yearbook*. For 1990 we used *WDI*. For 2000 we used *WDR*. For 2010 we used *HDR*. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-18 are secondary school age.

3.10 Ireland (1820-2010)

Population for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1910, 1926, 1936, 1951, 1961, 1971, 1981 and 1991 come from Meu Table A2 p. 27. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1860, 1870, 1880, 1890, 1900, 1910, 1926, 1936, 1951, 1961, 1971, 1981 and 1991 come from Meu Table A2 p. 27. We assumed the 1820-1850 age distribution was the same as the 1860 age distribution. Age distribution for Ireland 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1910, 1926, 1936, 1951, 1961, 1971, 1981 and 1991 come from Meu Table B1 p. 152. We interpolated for 1918. Labor force data for 2000 come from *WDR*. The labor force for 2010 comes from *WDI*. For 1820 and 1830 we used the following procedure. We used Clio infra for urban-rural population shares for Ireland for 1800-1950 and *WDI* for 1961-2010. We interpolated for our 1820-1840, 1860-1890, 1910-1936 values. We assumed a rural 15-64 labor force participation rate of 73.84% for 1820-1840, 72.75% for 1850, 78.97% for 1860, 86.05% for 1870, 83.74% for 1880, 81.38% for 1890, 75.93% for 1900, 74.38% for 1910, 77.99% for 1918, 81.63% for 1926, 84.1% for 1936, 85.1% for 1951, 83.42% for 1961, 81.7% for 1971, 75.75% for 1980, 75.25% for 1990, 90% for 2000 and 2010. We assumed an urban 15-64 labor force participation rate of 50% for 1820-1990, 52.45% for 2000, 58.5% for 2010. We constructed the ratio of this labor force with that from Meu, *WDR* and *WDI*. The root mean ratio for the overlapping years of 1840-2010 is 1.000. The values of this ratio for 1840 and 1850 are 1.000, 1.000, respectively. The range of the ratios is 1.000 to 1.001.

Real GNP come from Maddison. Physical capital investment rates come from the intraperiod averages of gross real capital formation and real income for 1947-2000 from Meu Table J1, pp. 917 and 924 and *WDR* (various years). The 2010 value is the average investment rate for 2000-2009 in S & H. For years 1820-1936, inclusive, we use sectoral output shares and US capital - sectoral output ratios from 1850-1940 to produce our estimates. We applied the US 1850 capital sectoral output ratios for years 1820, 1830 and 1840. For 1951-2010 we used perpetual inventory to compute physical capital. Sectoral output shares for 1860-1910 come from Geary and Stark (2002). Mitchell provides sectoral output shares for years 1926-1990, and *WDI* provides these for 2000 and 2010. We used Sabillon (2005) farming and manufacturing sectoral growth rates relative to aggregate growth rates to compute sectoral shares from 1820-1918. The service sector share is the residual for these years. For years 1951-2000, the regression of log capital from perpetual inventory against the log capital from the sectoral calculation produces an R^2 of .9019, with a slope coefficient on $ln(k_{1985})$ of .826 and a standard error of (.136). The intercept term is not significant.

Primary school enrollments from 1850-1910 come from Lindert, except for 1880, which we interpolated. Secondary school enrollments from 1870-1910 come from Lindert. We assumed higher education enrollment rates of .01 percent for 1850, 1860, 1870, .02 percent for 1880, .1 percent for 1890. For years prior to 1850, we assumed all had the 1850 enrollment rates in primary school equal to the 1850 value. Thus our 1820-1850 primary school enrollment rate is constant at 37.6%. Our secondary school enrollment rate series for 1820-1860 are constant at their 1870 enrollment rate. These produce 1870 primary education exposure rate of 38%, about equal to the 43% primary exposure rate of Morrison & Murtin. It produces a 2% exposure rate to secondary school, which is below the 10% value reported in Morrison & Murtin. Enrollments in primary and secondary school for 1920-1993 come from Meu Table I1 p. 883 and 888. To calculate enrollment rates prior to 1971, we assumed 6-11 are primary school age and 12-17 are secondary school age. In 1971 we switched to 6-13 primary and 14-17 secondary in order to match enrollment rate data from 1980 and beyond. Prior to 1920 we assumed that the 1850-1870 values were constant at .01 percent, 1890 was .02 percent, 1880 was interpolated, 1900 was .1 percent, and 1820-1840 were three quarters of the succeeding

decade enrollment rate. Higher education enrollments for 1920-1998 come from Meu Table I2 pp. 898 and 899. The 2010 values come from *HDR*. Our higher education share series fits Morrison & Murtin 1870-1950 values. Our time series of years of schooling in the labor force for 1870-1920 is: 2.87 (1870), 3.49 (1880), 4.08 (1890), 4.69 (1900), 5.09 (1910) and 5.54 (1918). The Morrisson & Murtin years of schooling time series is: 2.65 (1870), 3.31 (1880), 3.89 (1890), 4.53 (1900), 5.08 (1910) and 5.88 (1920).

3.11 Luxembourg (1950-2010)

Population for 1950, 1960, 1970, 1980, 1990 come from S & H. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Luxembourg 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force for 1960, 1970, 1980, 1990, 2000 and 2010 come from *WDI*. For 1950 we used the following procedure. We used worldometers for 1950 urbanization rates (we took the growth rate from 1955 to 1960 to extrapolate back to 1950), and *WDI*. We assumed rural 15-64 labor force participation rates of 86.3% for 1950-1960, 71.25% for 1970, 95% for 1980-2010. We assumed urban 15-64 labor force participation rates of 50% for 1950-1970, 51.65% for 1980, 51.9% for 1990, 58.22% for 2000, 64.8% for 2010. We constructed the ratio of this labor force to that from *WDR* and *CIA Factbook*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real GNP for 1950, 1960 and 1970 come from Penn World Tables 6.1. The 2010 value comes from the World Bank rebenchmarking project, The 1980, 1990 and 2000 values come from the growth rates in *WDI* and the 2010 base. The 2000 value is 112 percent of PWT 6.1, The 1990 value is 133 percent of PWT 6.1. The 1980 value is 130 percent of PWT 6.1. Physical capital investment rates come from the intraperiod averages of gross real capital formation and real income for 1950-2000 PWT 6.1. The 2010 value is the average investment rate for 2000-2009 in S & H. We used the sectoral output shares from *WDI* for 2010 and 2000. We applied the Sabillon (2005) sectoral growth rates for manufacturing and services relative to aggregate growth rates to construct sectoral shares for 1990 and 1980. We used the Belgian sectoral relative growth rates for 1950, 1960 and 1970 decades in order to produce the sectoral shares for 1950, 1960 and 1970. Farming is the residual sectoral share. We applied the US 1950-1990 capital - sectoral output ratios in order to produce the 1950-1990 values. We used perpetual inventory calculations for 2000 and 2010.

For 1950 and 1960 enrollment rates, we used the 1954 and 1961 *UN Statistical Yearbooks*. For 1970-1990 we used *WDI*. http://www.statistiques.public.lu/stat/ReportFolders/ReportFolder.aspx?IF_Language=eng&MainTheme=3&FldrName=6 For 2000 we used WDR. For 2010 we used *HDR*. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-18 are secondary school age.

3.12 Netherlands (1800-2010)

Populations for 1800, 1810, 1820, 1830, 1840, 1849, 1859, 1869, 1879, 1889, 1899, 1909, 1920, 1930, 1940 1947, 1960, 1970, 1980 and 1991 come from Meu Table A2 pp. 29 and 30. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distribution of the population for 1840, 1849, 1859, 1869, 1879, 1889, 1899, 1909, 1920, 1930, 1940 1947, 1960, 1970, 1980 and 1990 come from Meu Table A2 pp. 29 and 30. The age distributions for 1800-1830 are assumed to be the same as the age distribution in 1840. Age distribution for Netherlands 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in 2011.

Labor force figures for 1849, 1859, 1869, 1879, 1889, 1899, 1909, 1920, 1930, 1940, 1947, 1960, 1971, 1981 and 1991 come from Meu Table B1 p. 154. Labor force data for 2000 come from *WDR*. The labor force for 2010 comes from *WDI*. For 1800, 1810, 1820, 1830 and 1840 we used the following procedure. We used Clio infra for urban population shares for 1800, 1950, 1900 and 1950. We interpolated for 1810-1840, 1860-1890, 1910-1940. We used *WDI* for 1960-2010 urban population shares. We assumed a rural 15-64 labor force participation rate of 73.64% for 1800-1849, 66.35% for 1859, 69.11% for 1869, 70.99% for 1879, 70.45% for 1889, 74.9% for 1899, 76.8% or 1909, 75.1% for 1920, 73.1% for 1930, 68.2% for 1940, 71.7% for 1947, 72.95% for 1960, 70.12% for 1970, 63.82% for 1980, 80% for 1990-2010. We assumed an urban 15-64 labor force participation rate of 50% for 1800-1980, 59.03% for 1990, 74.27% for 2000, 78.29% for 2010.[14] We constructed the ratio of this labor force with that from Meu, *WDR* and *WDI*. The root mean ratio for overlapping years 1849-2010 is 1.000. The ratios for 1849 and 1859 are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real GNP come from Maddison. For physical capital investment rates 1807-1920 we used Smits, Horlings and van Zanden. Physical capital investment rates come from the intraperiod averages of gross capital formation and gross income for 1921-1939, 1948-9 from Meu Table J1 p. 911. For the 1950-1998 period we used the intraperiod averages of real gross capital formation and real income from Meu Table J1 pp. 914 and 918 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. Mitchell provides sectoral output shares for 1930-1990, we interpolated the 1940 value. *WDI* provides sectoral output shares for 2000 and 2010. We used Sabillon (2005) sectoral growth rates for farming and manufacturing relative to aggregate growth rates to produce sectoral output shares from 1800-1920. We applied the US capital - sectoral output ratios from 1850 for 1800. For 1810 to 2010, we used perpetual inventory to compute physical capital stocks.

Enrollment rates for primary schooling 1800 are set at 67 percent, to reflect an estimate of 75 percent literacy in 1800 as estimated by Steckel and Floud. The same source claims a 75 percent literacy in 1830. We interpolated the 1810 and 1820 values using the 67 percent and 75 percent values in 1800 and 1830, respectively. We assumed a .2 percent enrollment rate in secondary school in 1800. Our 1840 value from Meu is .4 percent. We interpolated our 1810 and 1820 values from these boundary conditions. Enrollments in primary and secondary school from 1850-1998 come from Meu Table I1 pp. 871, 876, 883 and 889. To calculate enrollment rates, we assumed 6-13 are primary school age and 14-17 are secondary school age. Higher education enrollments for 1849-1993 are from Meu Table I2 pp. 894, 896, 898 and 899. We assumed .1 percent enrollment rates for higher education in 1800 and 1810. We assumed .2 percent enrollment rates for higher education in 1820 and 1830. Our 2010 values comes from *HDR*. Our time series of years of schooling in the labor force for 1869-1920 is: 5.82 (1869), 6.15 (1879), 6.27 (1889), 6.46 (1899), 6.94 (1909). 7.01 (1920). The Morrisson & Murtin time series of years of schooling is: 5.09 (1870), 5.25 (1880), 5.42 (1890), 5.60 (1900). 5.82 (1910), 6.09 (1920).

[14] Our 1800 value is very similar to that from Roser's "Agricultural Employment" from Wrigley. Our 1810-1840 values are nearly identical to the interpolated values using 1800 and 1849.

3.13 New Zealand (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1874, 1878, 1881, 1886, 1891, 1896, 1901, 1906, 1911, 1921, 1926, 1936, 1945, 1951, 1961 and 1971 come from Maa Table A2 p. 30. The 1980 and 1990 are from Maa Table A5 p. 66. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1874, 1878, 1881, 1886, 1891, 1896, 1901, 1906, 1911, 1921, 1926, 1936, 1945, 1951, 1961, 1971, 1980 and 1990 come from Maa Table A2 p. 30. The age distribution for 1820, 1830, 1840, 1850 and 1860 are assumed to be the same as in 1874. The age distribution for New Zealand for 1980 and 1990 are interpolated from 1971, 1981 and 1991 values. Age distribution for New Zealand 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in 2013.

Labor force figures for 1896, 1901, 1906, 1911, 1921, 1926, 1936, 1945, 1951 is interpolated from 1945 and 1956, 1961, 1971, 1980 and 1990 come from Maa Table B1 p. 103. We used Federico (2010) for work force in agriculture for years 1850, 1880, 1910 and 1938. We combined this with information on labor force from 1896-1936 from Mitchell. We interpolated the agriculture work force and constructed the share of the over all labor force that is in agriculture. We computed the average share of the labor force that is in agriculture from 1896-1936, and applied this to our agriculture work force to compute labor force values for 1860, 1881, 1886, 1891. For 1874 we used the minimum value of .178 as this was most consistent with a value from constant LF/population ratio. We interpolated for 1878. Finally for 1850, we used a value of .41 for the share of the labor force that was agricultural as this is consistent with the average LF/population ratio for 1860-1936. We used Bairoch and Goertz for urban-rural population shares for 1820-1901, where we used the other developed countries category, which includes the US, Canada, Australia and New Zealand. Typically New Zealand is slightly more urban in Banks data for 1910, but we assume that this category is representative. We interpolated between Bairoch and Goertz years to fill in the missing years. For 1906-1961 we used Banks for urban-rural population shares. Finally we used *WDR* and *WDI* for urban-rural population shares for 1971-2010. We assumed a rural 15-64 labor force participation rate for 1820-1891 of 90%, and 80% for rural 15-64 labor force participation rate for 1896-2010. We assumed an urban 15-64 labor force participation rate of 50% for 1820-1945, 55% for 1951, 65% for 1961, 1971, 1981, 72.5% for 1990 and 2000 and 77.5% for 2010. We constructed the ratio of this labor force with the one from Federico, Mam, *WDR* and *WDI*. The root mean ratio for the overlapping years 1850-2010 is 1.004. The ratio values for 1850 and 1860 are 1.042, .945, respectively. Labor force data for New Zealand for 1980 and 1990 are interpolated from 1971, 1981, 1986 and 1991 values. Labor force data for 2000 come from *WDR*. Labor force data for 2010 are from *WDI*.

Real GNPs come from Maddison. Physical capital investment rates come from the intraperiod average gross capital formation and income for 1939-1978 come from Maa Table J1, pp. 1040 and 1042; intraperiod average gross real capital formation and real income for 1979-1998 come from Maa Table J1 p. 1042 and *WDR* (various years). The 2010 value is the average of investment rates from 2000-2009 from S & H. Mitchell provides sectoral output shares for years 1971-2000, and *WDI* provides sectoral output shares for 2010. We used Sabillon (2005) sectoral growth rates of farming and manufacturing relative to aggregate growth rates in order to produce sectoral shares from 1820-1936. We applied the US capital - sectoral output ratios for 1850-1940 to produce our estimates. For 1820, 1830 and 1840 we used the US 1850 capital sectoral output ratio. For years 1945-2010, we used perpetual inventory to compute physical capital stocks.

We used the estimates of literacy of Morris and Adelman for 1870, as well as education exposure shares from Morrison & Murtin to estimate primary school enrollment rates in 1820-1860. Morris and Adelman estimate 75 percent literacy in 1870. We used primary school enrollment rates of 73 percent, 75 percent, 77 percent, 77 percent and 77 percent for 1820, 1830, 1840, 1850 and 1860, respectively. We assumed 3.9 percent enrollment rates in secondary school for 1820, 1830, 1840, 1850, 1860 and 1874. These produce a 73 percent exposure rate to schooling, 67 percent primary exposure and 4 percent secondary exposure in 1870. Morrison & Murtin report 73 percent education exposure, 56 percent primary and 16 percent secondary exposure. Our 3.9 percent secondary school enrollment rate matches the rate reported from Mitchell. We assumed .3 percent enrollment rates for higher education in 1820, 1830, 1840, 1850 and 1860. Enrollments in school from 1874-1993 come from Maa Table I1 pp. 992 and 993. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-18 are secondary school age. Higher education enrollments for 1878-1998 come from Maa Table I2 p. 1006. We used *HDR* for enrollment rates in 2010. Our time series share of labor force exposed to higher education is consistent with Morrisson & Murtin 1870-2010. Our years of schooling in the labor force for 1874-1921 is: 453 (1874), 4.98 (1878), 5.18 (1881), 5.45 (1886), 5.73 (1891), 5.91 (1896), 6.01 (1901), 6.07 (1906), 6.20 (1911) and 6.39 (1921). The Morrisson & Murtin time series of years of schooling is: 3.91 (1870), 4.46 (1880), 5.23 (1890), 6.25 (1900), 7.00 (1910) and 7.52 (1920).

3.14 Norway (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1865, 1875, 1890, 1900, 1910, 1920, 1930, 1946, 1960, 1970, 1980 and 1990 come from Meu Table A2 p. 31. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1850, 1865, 1875, 1890, 1900, 1910, 1920, 1930, 1946, 1960, 1970, 1980 and 1990 come from Meu Table A2 p. 31. We assumed the age distribution for 1820, 1830 and 1840 were the same as in 1850. Age distribution for Norway 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in 2011.

Labor force figures for 1875, 1891, 1900, 1910, 1920, 1930, 1946, 1960, 1970, 1980 and 1990 come from Meu Table B1 p. 154. Labor force data for 2000 come from *WDR*. The labor force for 2010 comes from *WDI*. For values before 1875 we used the following procedure. For years before 1875 we used the following procedure. We used Clio infra for urban population share for 1800, 1850, 1900 and 1950 and *WDI* for 1960-2010. We interpolated for our 1820-1840, 1860-1890, 1910-1940 values. We assumed a rural 15-64 labor force participation rate of 78.35% for 1820-1875, 72.8% for 1880, 70.7% for 1890, 69.25% for 1900, 72.33% for 1910, 70.83% for 1920, 69.27% for 1930, 1946, 74.05% for 1960, 75% for 1970, 79.86% for 1980, 79.2% for 1990, 82.92% for 2000, 81.52% for 2010. We assumed an urban 15-64 labor force participation rate of 50% for 1820-1960, 51.97% for 1970, 78.96% for 1980, 79.2% for 1990, 82.92% for 2000 and 81.52% for 2010. We constructed the ratio of this labor force with that from Meu, *WDR* and *WDI*. The root mean ratio for overlapping years 1875-2010 is 1.000. The ratios for 1875 and 1880 are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real GNP comes from Maddison. Physical capital investment rates come from the intraperiod averages of gross real capital formation and real income for 1865-1939, 1946-1998 from Meu Table J1, pp. 907, 912, 918 and 924 and *WDR* (various years). The 2010 value is the average investment rate from 2000-2009

from S & H. Mitchell provides sectoral output shares for years 1875-1990, except for 1880 and 1920, which we interpolated. For 2000 and 2010 we used *WDI*. Sabillon (2005) provides farming and manufacturing growth rates relative to aggregate growth rates. We used these to produce sectoral output shares from 1820-1865. We applied the US capital - sectoral output ratios from 1850-1870 to these observations. We used the 1850 US capital - sectoral output ratios for 1820, 1830 and 1840. We used perpetual inventory to compute physical capital stocks from 1865 - 2010.

The 1830 and 1840 enrollments for primary school come from Lindert, which produce 121 percent and 98 percent primary school enrollment rates, respectively. For 1820 we assumed an enrollment rate of 80 percent. We consider this reasonable in relation to the 1800 literacy rate from Roser. Enrollments in primary and secondary school from 1853-1993 come from Meu Table I1 pp. 872, 877, 884 and 889. For 1890 Lindert provides secondary school enrollments. For years prior to 1890, (1820, 1830, 1840, 1850, 1858, 1865, 1875, 1880) we extrapolated by assuming a 50 percent growth rate per decade in secondary school enrollments. This produces .10 percent, .26 percent, .31 percent, .43 percent, .48 percent, .52 percent, .70 percent and 1.0 percent secondary school enrollment rates. The 1890 Lindert secondary enrollment rate is 1.5 percent. From 1875-1993 to calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1850-1998 are from Meu Table I2, pp. 894, 896, 898 and 899. For 1820-1840 we assumed a higher education enrollment rate of .1 percent. For 2010 we used the *HDR*. Our time series share of labor force exposed to higher education is consistent with Morrisson & Murtin 1870-1940. Our years of schooling in the labor force time series for 1865-1920 is: 5.87 (1865), 5.95 (1875), 6.00 (1880), 5.97 (1890), 5.95 (1900), 6.10 (1910), and 6.20 (1920). The Morrisson & Murtin time series of years of schooling is: 5.68 (1870), 5.66 (1880), 5.71 (1890), 5.77 (1900) 5.98 (1910) and 6.26 (1920).

3.15 Sweden (1800-2010)

Populations for 1800, 1810, 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1910, 1920, 1930, 1940 1950 1960, 1970, 1980 and 1990 come from Meu Table A2 pp. 37-39. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1800, 1810, 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1910, 1920, 1930, 1940 1950 1960, 1970, 1980 and 1990 come from Meu Table A2 pp. 37-39. Age distribution for Sweden 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1850, 1860, 1870, 1880, 1890, 1900, 1910, 1920, 1930, 1945, 1950, 1960, 1970, 1980 and 1990 come from *Swedish Historical statistics*, http://www.historicalstatistics.org, accessed July 2, 2016. Labor force data for 2000 come from *WDR*. Our 2010 labor force comes from *WDI*. For 1800-1840, we followed the following procedure. Grigg (1980) provides the agricultural population of Sweden for years 1800, 1850-1930, decadal. We interpolated to produce estimates for 1810-1840. We then regressed log labor force against, year, log ag population share, and log ag population. This regression produced an R^2 of .9926. We used the predicted values from this regression for 1800-1840. For the overlapping years, 1850-1930, the ratio of the predicted values to the actual labor force values range from .9760 to 1.041 of the actual labor force data for 1850-1930. If we used the average LF/population ratios for 1860-1950, the correlation between these values and the predicted regression labor force values is .9988.

Real GNP comes from Maddison.[15] Physical capital investment rates come from the intraperiod averages of gross real capital formation and real income for 1871-1998 from Meu Table J1, pp. 907, 913, 920 and 926 and *WDR* (various years). For 2010 we used the average investment rate of 2000-2009 from S & H. For 1800-1870 we used historical physical investment from Swedish historical statistics, http://www.historicalstatistics.org, accessed July 2, 2016. We used perpetual inventory.

Enrollments in primary, 1850, 1865, 1870, 1890, 1900-1993 and secondary school, 1888-1998 come from Meu Table I1 pp. 872, 878, 885 and 889. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. For years prior to 1850 for primary school, prior to 1888 for secondary school and prior to 1910 for higher education we used Malmberg (2007). For 1910-1993 higher education enrollment rates are from Meu Table I2 pp. 894, 896, 898 and 899. For 2000 and 2010 we used *WDI*. These fit 1870-1910 education shares from Morrisson & Murtin. Our time series of years of schooling in the labor force for 1870-1920 is: 5.17 (1870), 5.44 (1880), 5.73 (1890), 5.92 (1900), 6.05 (1910) and 6.15 (1920). The Morrisson & Murtin time series of years of schooling is: 4.23 (1870), 4.61 (1880), 5.02 (1890), 5.49 (1900), 5.89 (1910) and 6.24 (1920).

3.16 Switzerland (1820-2010)

Population for 1820, 1830, 1840, 1850, 1860, 1870, 1888, 1900, 1910, 1920, 1930, 1941, 1950, 1960, 1970, 1980 and 1990 come from Meu Table A2 p. 40. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1820, 1830, 1840, 1850, 1860 are assumed to be the same as 1870. The age distributions of the population for 1870, 1888, 1900, 1910, 1920, 1930, 1941, 1950, 1960, 1970, 1980 and 1990 come from Meu Table A2 p. 40. Age distribution for Switzerland 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1888, 1900, 1910, 1920, 1930, 1941, 1950, 1960, 1970, 1980 and 1990 come from Meu Table B1 p. 159. Labor force for 1870 comes from Maddison (2001), Table E-1. We interpolated for 1880. Labor force data for 2000 come from *WDR*. The labor force for 2010 comes from *WDI*. For 1820, 1830, 1840, 1850, 1860, 1870 we used the following procedure. We used Clio infra for urban-rural population shares for 1800, 1850, 1900 and 1950. We interpolated for 1820-1840, 1860-1890, 1910-1940. We used a rural 15-64 labor force participation rate of 80% for 1820-1870, 79.87% for 1880, 78.4% for 1888, 71.74% for 1900, 82.8% for 1910, 79.3% for 1920, 76.2% for 1930, 73.6% for 1941, 75.45% for 1950, 84.1% for 1960, 95% for 1970-2010. We assumed an urban 15-64 labor force participation rate of 59% for 1820-1870, 50% for 1880-1960, 56.27% for 1970, 55.9% for 1980, 68.37% for 1990, 71.85% for 2000, 76.61% for 2010. We computed the labor force using the urban-rural population shares and produced the ratio of these values with Meu, Maddison, *WDR* and *WDI* data. The root mean ratio of this for years 1870-2010 is 1.000. The 1870 and 1880 values are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real GNPs are from Maddison. Physical capital investment rates come from the intraperiod averages of gross real capital formation and real income for 1950-1998 from Meu Table J1, pp. 920 and 926 and *WDR* (various years). For 2010 we used the average investment rates for 2000-2009 from S & H. For years 1880-1990 we used Goldsmith (1985). For 2000 and 2010 we used perpetual inventory. *WDI* provides

[15] This data produces similar per capita real income growth, for the three earliest periods: 1800-1840, 1840-1870, 1870-1910, as in Schön.

sectoral output shares for 1990, 2000 and 2010. We used Sabillon (2005) growth rates of manufacturing for 1820-1980s, and the growth rates of services for 1900-1980s, and aggregate growth rates for 1920-1980s in order to compute manufacturing and service shares from 1820-1980, and 1900-1980, respectively. For 1900-1980, farming is the residual share. For 1820-1890, we assumed farming and services are each half of the non manufacturing share in the economy. Since the capital - sectoral output ratios of these two sectors are very similar, it makes little difference. We used the US 1850-1990 capital - sectoral output ratios for years 1850-1990, and we used the US 1850 capital - sectoral output ratios for 1820, 1830 and 1840.

Primary enrollment rates for 1830, 1840, 1850, 1860 come from Lindert. Secondary enrollment rates for 1830, 1840, 1850, 1860, 1870, 1888 come from Lindert. For 1820 we used half the primary enrollment rates for 1830, and the same secondary enrollment rate for 1830. This produces a primary school exposure rate of 86% in 1870 compared with Morrisson & Murtin estimate of 83%. Our secondary exposure rate of 6% is below the Morrison & Murtin estimate of 15%. Enrollments in primary and secondary school from 1870-1961 come from Meu Table I1 pp. 879, 885. To calculate enrollment rates, we assumed 6-12 are primary school age and 13-18 are secondary school age. We used the WDR of various years for the enrollment rates in 1970, 1980 and 1990. Higher education enrollments for 1886-1998 are from Meu Table I2 pp. 896, 898 and 899. For 2010 we used *HDT*. For years 1820-1870 we assumed higher education enrollment rates of .2%. Our time series of years of schooling in the labor force for 1870-1920 is: 5.86 (1870), 6.14 (1880), 6.31 (1888), 6.47 (1900), 6.62 (1910), 6.84 (1920). The Morrisson & Murtin time series of years of schooling is: 6.07 (1870), 6.59 (1880), 6.91 (1890), 7.29 (1900), 7.55 (1910 and 7.71 (1920).

3.17 United Kingdom (1801-2010)

Populations for England, Scotland, Wales and Ireland for 1801, 1811, 1821, 1831, 1841, 1851, 1861, 1871, 1881, 1891, 1901, 1911, 1921, and England, Scotland, Northern Ireland and Wales for 1931, 1951, 1961, 1971, 1981 and 1991 come from Meu Table A1 p. 8. Population for UK for 2000 comes from *Time Almanac 2001*. Population for UK for 2010 comes from *Wikipedia*.

The age distributions of the population for England and Wales for 1841, 1851, 1861, 1871, 1881, 1891, 1901, 1911, 1921, 1931, 1951, 1961, 1971, 1981 and 1991 come from Meu Table A2 pp. 41 and 42. The age distribution for 1801, 1811, 1821 and 1831 are assumed to be the same as in 1841. Age distribution for UK 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in 2012.

Labor force figures for 1851, 1861, 1871, 1881, 1891, 1901, 1911, 1921, 1931, 1951, 1961, 1971, 1981 and 1991 come from Meu Table B1 p. 160. Labor force for 1821 comes from Maddison (1995). We interpolated for 1831, 1841. Labor force data for 2000 come from *WDR*. Labor force in 2010 is from *WDI*. Labor force data for 1801 are derived from Table 13B Broadberry, Campbell and van Leeuwen (2011) adjusted for population. Our 1801 value is also consistent with R.C. Allen. We interpolated for our 1811 value.

GDP are from Maddison. Physical capital investment rates come from the intraperiod averages of gross real capital formation and real income for 1830-1998 from Meu Table J1, pp. 905, 907, 913, and 926 and *WDR* (various years). The 2010 value comes from the average of 2000-2009 from S & H. However with the exception of the 2010 value, computed using perpetual inventory, all other values come from Piketty and Zucman (2014).

The primary school enrollment rate for 1821 comes from Lindert. The 1891 and 1901 secondary school enrollment rates come from Lindert. The 1801 and 1811 primary school enrollment rates come from

Lindert and comparable to the literacy rates from Steckel and Floud. Enrollments in primary, 1850-1993, and secondary school from 1904-1993 come from Meu Table I1 pp. 872, 879, 886 and 890. Secondary school enrollment rates from 1821-1881 are interpolated from our assumed .5 percent rate in 1801 and 1811, and Linderts value for 1891 of .91 percent. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. For primary enrollment rates in 1831 and 1841 we interpolated using 1821 and 1850 data. Higher education enrollments for 1922-1998 are from Meu Table I2 pp. 898 and 899. The 1921 enrollment rate in higher education is .9 percent. Prior to 1922 (1801, 1811, 1821, 1831, 1841, 1851, 1861, 1871, 1881, 1891, 1901, 1911) we used enrollment rates of 0, 0, .005%, .005%, .02%, .1%, .1%, .1%, .2%, .2%, .2% and .2%, respectively. For 2010 we use *HDR*. These enrollment rates produce education shares similar to Morrisson & Murtin. Overall in 1870 we have 70% exposed to any education level, and Morrisson & Murtin report a value of 72%. Our 1870 primary exposure share is 69& compared with their 63% value. Secondary exposure share in our data is 1% compared with their 8%. Our years of schooling in the labor force for 1871-1921 is: 4.27 (1871), 4.94 (1881), 5.23 (1891), 5.57 (1901), 5.78 (1911) and 6.00 (1921). The Morrisson & Murtin time series of years of schooling is: 3.59 (1870), 4.37 (1880), 4.84 (1890), 5.33 (1900), 5.86 (1910) and 6.31 (1920).

3.18 United States (1790-2010)

Populations for 1790, 1800, 1810, 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970, 1980 and 1990 come from Mam Table A1 p. 6. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970, 1980 and 1990 comes from Mam Table A2 pp. 21-23. We assumed the age distribution for 1790, 1800 and 1810 were the same as in 1820. The age distribution for 2000 comes from *The Statistical Abstract of the United States 2000*. The age distribution for 2010 is from *Statistical Abstract of the United States 2012*.

Labor force figures for 1870, 1880, 1890, 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970, 1980 and 1990 come from Mam Table B1 p. 107. The labor force for 1800, 1810, 1820, 1830, 1840, 1850, 1860 comes from *Historical Statistics of the United States, Millennial Edition*. We assumed the same LF/population ratio for 1790 as the average over 1800-2010. The labor force for 2000 comes from the *Statistical Abstract of the United States 2007*. The labor force for 2010 comes from the *Statistical Abstract of the United States 2012*.

Real GDP comes from Maddison from 1790-2010. For 2010 we used the average investment rate of 2000-2009 from S & H. For years 1790-2000 we used Piketty and Zucman (2014) capital output ratios to produce our capital stock values. For 2010 we used perpetual inventory.

For enrollment rates from 1840-1870 we use Turner, Tamura, Mulholland and Baier (2007). For 1830 we used Easterlin (1981). For years prior to 1830 we assumed primary enrollment rates of 60 percent for primary school for 1790, 1800, 1810 and 1820, respectively. This is broadly consistent with the high literacy rates reported in Grubb (1990). These are based on estimates of roughly 90 percent literacy for free white males in New England, and about 50 percent literacy for white women in New England, see Lynch (2011). Furthermore Gill (2011) reports that for white men in mid eighteenth century Virginia between 60 percent and 94 percent were literate in Middlesex county and York county. Women in York county at this time appeared to have a 50 percent literacy rate. For secondary school enrollment rates we assumed 2 percent for all years before 1840. For higher education we assumed .1 percent for 1790-1820 and .2 percent for

1830. Enrollments in primary and secondary school from 1871-1993 come from Mam Table I1 pp. 718, 720, 724, 729 and 734. To calculate enrollment rates, we assumed 6-13 are primary school age and 14-17 are secondary school age. For 2010 we used *WDI*. Our education shares from 1870-1900 are similar to those in Morrisson & Murtin. Our time series of years of schooling in the labor force for 1870-1930 is: 6.01 (1870), 6.92 (1880), 7.54 (1890), 7.92 (1900), 8.16 (1910), 8.49 (1920) and 9.02 (1930). The Morrisson & Murtin time series of years of schooling is: 5.57 (1870), 6.02 (1880), 6.60 (1890), 7.09 (1900), 7.46 (1910), 7.84 (1920) and 8.46 (1930).

4 Southern Europe

4.1 Cyprus (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Cyprus 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1990 come from *HDR*. Labor force figures for 1980 and 2000 come from *WDR* (various years). The labor force for 1950 uses the 1953 value from the 1961 *UN Statistical Yearbook*, and the 1960 value uses the 1959 value from the same issue. The 1970 labor force comes the 1973 *UN Statistical Yearbook*. Labor force for 2010 comes from *WDI*.

Real GNPs for 1950, 1960, 1970, 1980 and 1990 come from S & H online. The 2000 value comes from *WDR*. The 2010 value comes from *WDR*. The 1960-2009 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The 1980, 1990, 2000 and 2010 sectoral output shares are from the *WDI*. We used Sabillon (2005) growth rates of farming and manufacturing for Greece in the 1950s, 1960s and 1970s to construct estimates of sectoral output shares in Cyprus for 1950, 1960 and 1970. We used US 1950-1990 capital - sectoral output ratios for our capital estimates for years 1950-1990. We used perpetual inventory for years 2000 and 2010.

Enrollments in primary and secondary schools for 1950-2000 come from Maa Table I1 p. 986. We assumed the primary school and secondary school ages are 6-11 and 12-17. The tertiary school enrollments are from Maa Table I2 p. 1003. For 2010 we used enrollment rates from *HDR*.

4.2 Greece (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1910, 1920, 1928, 1950, 1961, 1971, 1981 and 1991 come from Meu Table A2 pp. 24 and 25. The 1910 value is an interpolation from the 1907 and 1920 values. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1870, 1880, 1890, 1900, 1910, 1920, 1928, 1950 1961, 1971, 1981 and 1991 come from Meu Table A2 pp. 24 and 25. The age distribution for 1910 is an interpolation from the 1907 and 1920 values. The age distribution prior to 1870 assumes the same age distribution as

in 1870. Age distribution for Greece 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in 2011.

Labor force figures for 1928, 1951, 1961 come from Meu Table B1 p. 151. Labor force data for 1971, 1981, 1990, 2000 come from *WDR*. We used the 1971, 1981 and 1990 values from the *WDR* instead of Meu because they better fit the time series of labor force to population ratios, both earlier in Meu as well as later with *WDI*. Labor force data for 2010 comes from *WDI*. For labor force figures for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1910, 1920 comes from the following procedure. We used Clio infra for the urban-rural population shares, 1800-1950. We used *WDI* for urban-rural shares for 1961-2010. We assumed the rural 15-64 labor participation rate is 72.8% for 1820-1928, 65% for 1940, 59.76% for 1951, 84.06% for 1961, 82.6% for 1971, 87.8% for 1980, 80.2% for 1990, 85% for 2000, and 2010. We assumed the urban 15-64 labor participation rate is 67% for 1820-1928, 58.4% for 1940, 50% for 1951-1990, 55.8% for 2000, 65.02% for 2010. We constructed the labor force using this measure relative to the labor force from Meu and *WDR*, and *WDI*. The root mean ratio for overlapping years 1928-2010 is 1.000. The 1928 and 1940 ratios are 1.000 and 1.000, respectively.[16] The range of the ratios is 1.000 to 1.000.

Real GNP are from Maddison. Physical capital investment rates come from the intraperiod averages of gross real capital formation and real income for 1949-1998 from Meu Table J1, pp. 916 and 923 and *WDR* (various years). For 2010 we used the average investment rates from 2000-2009 from S & H. Mitchell provides sectoral output shares for 1928-1990, except for 1940, which we interpoiated. The *WDI* provides sectoral shares for 2000 and 2010. We used Sabillon (2005) growth rates of farming and manufacturing together with the aggregate growth rates to produce sectoral output shares from 1820-1920.[17] We used US capital - sectoral output ratios from 1850-1940 to produce our physical capital estimates for 1820-1940. For 1820, 1830 and 1840 we used the 1850 US capital - sectoral output ratios. For years 1950-2010, we used perpetual inventory.

Enrollments in primary, 1901, 1910, 1926-1937, 1951-1993, and secondary school, 1926-1935, 1951-1993, come from Meu Table I1 pp. 875, 882 and 888. For 1870, 1880, 1890 we used Lindert for primary school enrollment rates. For years prior to 1870 we used three-quarters of the succeeding primary school enrollment rate. Our 1820, 1830, 1840, 1850, 1860 and 1870 primary school enrollment rate series is: 9.7%, 12.9%, 17.2%, 22.9%, 30.5% and 40% (datum). For secondary school enrollment rates 1820-1860 we used 1.6% compared with 1.7% the 1870 value. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. For 1920 and 1928 we used Lindert for secondary school enrollment rates. For 1870 - 1910 we interpolated from our 1860 value to the Lindert value for 1920. These primary and secondary enrollment rates produce an 1870 education exposed rate of 23 percent, which exactly matches Morrison & Murtin. Furthermore our exposed share consists of 21% primary exposed, and 2% secondary exposed share compares with Morrison & Murtin estimates of 18% and 4%, respectively. Our estimated years of schooling is 1.46 compared with their estimate of 1.45. Higher education enrollments for 1912-1993 are from Meu Table I2 pp. 895, 897 and 899. Our 2000 values come from *WDI*. All of our 2010 enrollment rates come from *HDR*. We used .00005 and .0001 for enrollment rates in higher education for 1820-1900

[16] Our labor force figures for 1850-1920 are nearly identical to those that come from data in Petmezas (2003). If we used his proportion of the population rural and applied this to the age distribution of the population in these years. For populations aged 15-64, we applied a 72.8% participation rate in the rural sector and a 67% participation rate in the urban sector, the correlation with the above is 1.000, a regression of the two series produces an ${}^2 of 1, a slope coefficient of 1.015, with a standard error of .0006, and an intercept of -2.04$.

[17] Our estimates of agriculture share of GDP are consistent with Petmezas, although a tad bit low in 1860-1930. In 1860 we estimate 64% of GDP in agriculture, while Petmezas reports 70%. Using our method we produce 85% GDP share in agriculture in 1820, which we deem to be a reasonable upper bound.

and 1910, respectively. Our higher education exposure shares in 1870-1910 also fits the Morrison & Murtin values. Our time series of years of schooling in the labor force for 1870-1920 is: 2.52 (1870), 2.97 (1880), 3.18 (1890), 3.32 (1900), 2.75 (1910) and 4.94 (1920). The Morrisson & Murtin time series is: 1.41 (1870), 1.76 (1880), 2.07 (1890), 2.37 (1900), 2.71 (1910) and 3.07 (1920).

4.3 Israel (1948-2010)

Populations for 1948, 1961 and 1972 comes from Maa Table A1 p. 8. Populations for 1956, 1980 and 1990 come from S & H (1991). Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distribution for 1948, 1956, 1961, 1972 comes from KF. The age distribution for Israel for 1980, 1990 come from DK. Age distribution for Israel 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1948, 1956, 1961, 1972 and 1983 come from Maa Table B1 p. 96. Labor force data for Israel 1990, 2000 and 2010 come from *WDR* (various years).

Real GNPs come from Maddison. Physical capital investment rates come from the intraperiod average gross capital formation and income for 1950-1954 and the intraperiod average gross real capital formation and real income for 1955-1998 from M (1992) Table J1, p. 1030 and *WDR* (various years). We used the average investment rate over 2000-2009 from S & H for 2010. For 1948-1990 we used the capital output ratios from Nehru & Dhareshwar (1993). For years 2000 and 2010 we used perpetual inventory.

Enrollments in primary and secondary schools from 1950-1993 come from Maa Table I1 p. 987. We used the 1950 enrollment rate for the 1948 observation. To calculate enrollment rates, we assumed 6-13 are primary school age and 14-17 are secondary school age. Higher education enrollments for 1945-1993 come from Maa Table I2 p. 1004. For 2010 we used *HDR*. We interpolated for all our 2000 enrollment rates.

4.4 Italy (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1861, 1871, 1881, 1901, 1911, 1921, 1931, 1951, 1961, 1971, 1980 and 1990 come from Meu Table A2 pp. 28 and 29. The population for 1940 is the geometric interpolation from 1931 and 1951 years in Meu. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distribution of the population for 1861, 1871, 1881, 1901, 1911, 1921, 1931, 1951, 1961, 1971, 1980 and 1990 come from Meu Table A2 pp. 28 and 29. For years 1820, 1830, 1840 and 1850 we assumed the same age distribution as in 1861. The age distribution for 1940 uses geometric interpolation from 1931 and 1951 years in Meu. Age distribution for Italy 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2001 age distribution, and for 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1871, 1881, 1901, 1911, 1921, 1931, 1941, 1951, 1961, 1971, 1980 and 1990 come from Meu Table B1 p. 153.[18] We interpolated for our 1890 value. Labor force data for 2000 come from *WDR*. The labor force data for 2010 comes from *WDI*. For 1820, 1830, 1840, 1850 and 1861 we used the following procedure. We used Clio infra for urban population share for 1800, 1850, 1900 and 1950. We

[18] Our workforce numbers are consistent with those contained in Federico.

interpolated for 1820-1840, 1860-1890, 1910-1940. We used *WDI* for 1960-2010. We assumed a rural 15-64 labor force participation rate of 85% for 1820-1911, 94.13% for 1921, 76.76% for 1931-1940, 71.58% for 1951, 71.1% for 1961, 69.72% for 1971, 74.72% for 1980, 77.13% for 1990, 78.59% for 2000 and 90% for 2010. We assumed an urban 15-64 labor force participation rate of 73.675% (which is the average urban 15-64 labor force participation rate we used for 1871, 1881, 1890, 1901) for 1820-1861, 60.53% for 1871, 91.13% for 1881, 77.44% for 1890, 65.6% for 1901, 51.54% for 1911, 50% for 1921-2000 and 50.75% for 2010.[19] We constructed the ratio of this labor force with that from Meu, *WDR* and *WDI*. The root mean ratio of the overlapping years 1871-2010 is 1.000. The 1871 and 1881 values are 1.000 and 1.000, respectively. The range of ratios is 1.000 to 1.000.

Real GNP come from Maddison. Physical capital investment rates come from the intraperiod averages of gross real capital formation and real income for 1861-1998 from Meu Table J1, pp. 907, 911, 917 and 924 and *WDR* (various years). The 2010 value is the average investment rate from 2000-2009 from S & H. Mitchell provides secular output shares for 1871-1990, and *WDI* provides sectoral output shares for 2000 and 2010. These shares are consistent with Cohen and Federico. We used Sabillon (2005) growth rates for farming, manufacturing and aggregate growth to produce sectoral output shares for 1820-1861. We used US capital - sectoral output ratios for 1850 to produce our capital estimates for 1820-1840. Our 1850 value comes from Goldsmith (1985). We used perpetual inventory to produce our estimates for 1861-1961. Piketty and Zucman (2014) provides our 1971-2000 values. We used perpetual inventory to produce our 2010 value.

Enrollments in primary and secondary school from 1861-1993 come from Meu Table I1 pp. 871, 876, 883 and 888. To calculate enrollment rates, we assumed 6-10 are primary school age and 11-18 are secondary school age. Primary school enrollments for 1830 and 1850 come from Lindert. The 1840 value is interpolated between 1830 and 1850. For 1820 we used 5 percent for primary enrollment rate, just slightly less than the 1830 value of 5.6 percent. This is consistent with the change in literacy between 1801-1810 birth cohorts and 1821-1830 birth cohorts from Roser. Secondary enrollment rates for 1820, 1830, 1840 and 1850 were extrapolated from 1861 (.5%), and are .10 percent, .15 percent, .20 percent and .26 percent, respectively. These fit the less than 1 percent 1870 secondary exposed share of Morrison & Murtin. Our education exposure in 1871 is 23% compared with Morrison & Murtin estimate of 17%. Our estimated years of schooling is 1.2 compared with their value of .8. Higher education enrollments for 1861-1993 are from Meu Table I2 pp. 894, 896, 898 and 899. We assume enrollment rates of .01 percent, .05 percent, .1 percent and .1 percent for 1820, 1830, 1840 and 1850, respectively. For 2000 we used *WDR*. For 2010 we used *HDR*. Our time series of the share of the labor force exposed to higher education is consistent with Morrisson & Murtin from 1870-1970. Our time series of years of schooling in the labor force for 1871-1921 is: 1.19 (1871), 1.80 (1881), 2.35 (1890), 2.71 (1901), 3.22 (1911) and 3.77 (1921). The Morrisson & Murtin time series of years of schooling is: .84 (1870), 1.33 (1880), 1.86 (1890), 2.38 (1900), 2.84 (1910) and 3.36 (1920).

4.5 Malta (1960-2010)

Population for 1960, 1970, 1980, 1990 come from *WDI*. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

[19] Our 1820 value is consistent with R. C. Allen's figure, adjusted for population in 1820, and assuming a 62% labor force participation rate in rural agricultural population and a 44% labor force participation rate for rural non agricultural and urban populations.

The age distributions of the population for 1960, 1970, 1980, 1990 come from KF. Age distribution for Malta 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force for 1960, 1970, 1980, 1990, 2000 and 2010 come from *WDI*.

Real GNP for 1960 and 1970 come from *WDI*. The 2010 value comes from the World Bank rebenchmarking project, The 1980, 1990 and 2000 values come from the growth rates in *WDI* and the 2010 base. Physical capital investment rates come from the intraperiod averages of gross real capital formation and real income for 1960-2000 *WDI*. The 2010 value is the average investment rate for 2000-2009 in S & H. *WDI* provides sectoral output shares for 1970-2010. We used Sabillon (2005) using 1960 Italy's farming and manufacturing relative to Malta's growth rate to produce our estimate of Malta's sectoral output share in 1960. We applied US capital - sectoral output ratios for 1960 - 1990 to produce capital estimates. We used perpetual inventory to compute 2000 and 2010 values.

For all years we used the enrollment rates from WDI, with the exception of 2010 which we used *HDR*. We assumed that primary schooling covers ages 6-11, and secondary schooling covers 12-18.

4.6 Portugal (1800-2010)

Our 1800 population comes from McEvedy and Jones. The 1910 observation on population is interpolated between our 1800 observation and our 1920 observation. Population for 1820, 1830, 1840, 1849, 1864, 1875, 1890, 1900, 1911, 1920, 1930, 1940, 1950, 1960, 1970, 1981, 1991 come from Meu Table A1 p. 7. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distribution of the population for 1864, 1875, 1890, 1900, 1911, 1920, 1930, 1940, 1950, 1960, 1970, 1981, 1990 come from Meu Table A2 p. 33 and 34. The 1800, 1810, 1820, 1830, 1840, 1849 data come from the age distribution for 1864. Age distribution for Portugal 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1890, 1900, 1911, 1920, 1930, 1940, 1950, 1960, 1970, 1981 and 1990 come from Meu Table B1 p. 155. Labor force data for 2000 come from *WDR*. Labor force data for 2010 come from *WDI*. For 1800, 1810, 1820, 1830, 1840, 1849, 1864, 1875 labor force we used the following procedure.[20] We used Clio infra for urban population shares for 1800-1950, and *WDI* for 1960-2010. We assumed a rural 15-64 rural labor force participation rate of 90% for 1800-1890, 83% for 1900, 83.35% for 1911, 90% for 1920, 95% for 1930, 72.9% for 1940, 68.1% for 1950, 72.3% for 1960, 69.81% for 1970, 71.25% for 1980, 80% for 1990-2010. We assumed an urban 15-64 labor force participation rate of 55.3% for 1800-1890, 50% for 1900, 1911, 84.7% for 1920, 91.75% for 1930, 50% for 1940-1980, 60.4% for 1990, 67.66% for 2000, 72.4% for 2010. We constructed the ratio of this labor force to that from Meu, *WDR* and *WDI*. The root mean ratio for overlapping years 1890-2010 is 1.000. The ratio values for 1890 and 1900 are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real GNPs are from Maddison. Physical capital investment rates come from the intraperiod averages of gross real capital formation and real income for 1953-1998 from Meu Table J1, p. 919 and 925 and *WDR* (various years). For 2010 we used the average investment rate from 2000-2009 from S & H. Mitchell provides sectoral output shares for 1950-1990, and *WDI* provides sectoral output shares for 2000 and

[20] Our data is similar to information in Lains (2009), although he only reports male labor force.

2010. We used Sabillon (2005) to compute sectoral output shares for 1800-1940; we used the farming and manufacturing growth rates relative to the aggregate growth rate. For years 1950-1990 we used the capital output ratios from Nehru & Dhareshwar (1993). For years 1864-1940 we used the differential capital growth rates compared with the aggregate growth rates from Pedro Lains (2003) paper. For 1800-1850 we used the US capital - sectoral output ratios for 1850, applied to years 1800, 1810, 1820, 1830 and 1840. For our 2000 and 2010 values we used perpetual inventory.

Enrollments in primary and secondary school from 1849-1993 come from Meu Table I1 pp. 872, 877, 884 and 889. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1849-1985 are from Meu Table I2 pp. 894, 896, 898 and 899. Enrollment rates for primary school for 1800, 1810, 1820, 1830, 1840 are 1.25%, 1.25%, 1.25%, 2.5% and 5%, respectively, reaching the 1849 figures of 10%. This is consistent with the slow increase in literacy from 1451-1500 (3%) to 1701-1800 (8%) in Spain from Roser. For secondary school enrollment rates for 1800, 1810, 1820, 1830 and 1840 we assumed, .01%, .01%, .01%, .03% .05% reaching the .1% secondary enrollment rate for 1849. Our 1870 no education exposure share of 91 percent exactly matches that of Morrison & Murtin. Our 9 percent primary share exactly fits Morrison Murtin as well. Our years of schooling value is .6 years compared with their value of .5 years. Finally for higher education enrollment rates for 1820, 1830 and 1840, we assumed 0, 0 and .01% rates, reaching the .01% 1849 value. Obviously our 0 percent higher education share fits the Morrison & Murtin value as well. For years 1800 & 1810 we used our 1820 values. For 2000 we used *WDR*, and for 2010 we used *HDR*. Our time series of years of schooling in the labor force for 1870-1920 is: .58 (1870), .77 (1875), 1.04 (1890), 1.51 (1900), 1.70 (1911) and 2.05 (1920). The Morrisson & Murtin time series of years of schooling is: .46 (1870), .72 (1880), 1.03 (1890), 1.38 (1900), 1.60 (1910) and 1.80 (1920).

4.7 Spain (1800-2010)

For 1800 we used McEvedy and Jones for population. For 1810 population we interpolated our 1800 and 1820 observations. Population for 1820, 1830, 1840, 1850, 1870, 1890, 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970, 1980 and 1990 come from Meu Table A2 p. 36. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distribution of the population for 1850, 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970, 1980 and 1990 come from Meu Table A2 p. 36. The age distribution for 1870 and 1890 are interpolated. The age distribution for 1800, 1810, 1820, 1830 and 1840 is identical to the 1850 age distribution. The age distribution for Spain for 1990 is interpolated from 1981 and 1992 values. Age distribution for Spain 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2001 age distribution, and for 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1850, 1900, 1910, 1920, 1940, 1950, 1964, 1970, 1980 and 1990 come from Meu Table B1 p. 157. The 1860, 1870, 1880, 1890 and 1930 values are interpolated. Labor force data for 2000 come from *WDR*. The labor force data for 2010 comes from *WDI*. For 1800-1840 labor data we followed the following procedure. We used Clio infra for urban population shares for 1800, 1850, 1900 and 1950, and *WDI* for 1960-2010. We interpolated for urban shares in 1810-1840, 1860-1890, 1910-1940. We assumed a rural 15-64 labor force participation rate of 85% for 1800-1860, 84.5% for 1870, 79% for 1880, 71.16% for 1890, 73.75% for 1900, 68.01% for 1910, 66.77% for 1920, 63.9% for 1930, 60.02% for 1940, 63.72% for

1950, 67.27% for 1960, 65.94% for 1970, 51.8% for 1980-2010. We assumed an urban 15-64 labor force participation rate of 63.1% for 1800-1850, 56.55% for 1860, 50% for 1870-1980, 63.02% for 1990, 70.13% for 2000 and 81.55% for 2010.[21] We constructed the ratio of this labor force with that from Meu, *WDR* and *WDI*. The root mean ratio of the overlapping years 1850-2010 is 1.000. The ratio values for 1850 and 1860 are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real GNPs come from Maddison. Physical capital investment rates come from the intraperiod averages of gross real capital formation and real income for 1954-1998 from Meu Table J1, pp. 919 and 925 and *WDR* (various years). For 2010 we used the average investment rate from 2000-2009 from S & H. Mitchell provides sectoral output shares for 1950-1990, and *WDI* provides sectoral output shares for 2000 and 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral shares for 1800-1940. We used the US 1850 capital - sectoral output ratios for years 1800-1850 to produce capital estimates. For years 1850-2000 we used Leandro Prados de la Escosura and Rosés (2009). We used perpetual inventory for 2010.

Enrollments in primary, 1850, 1870, 1885, 1909, 1914, 1926, 1932-1993 and secondary school, 1914-1993 come from Meu Table I1 pp. 878, 885 and 889. For 1830 and 1860 we used Easterlin (1981) Appendix Table 1. To calculate enrollment rates, we assumed 6-10 are primary school age and 11-17 are secondary school age. We assumed a 1 percent enrollment rate for secondary school in 1850. We interpolated from 1850 to 1914 (1.5% secondary school enrollment rate) for years 1860-1910. Higher education enrollments for 1884, 1914-1993 are from Meu Table I2 pp. 896, 898 and 899. We assumed higher education enrollment rates of .3 percent, .4 percent, .5 percent, .6 percent for 1850, 1870, 1890 and 1900, respectively. For 1800, 1810, 1820 primary school enrollment rates we assumed 10%, 10%, 18.1%. The 1800 and 1810 values are consistent with the 8% literacy of adults of the 1701-1800 cohort from Roser. For 1800, 1810, 1820, 1830 and 1840 secondary school enrollment rates we assumed .1%, .1%, .13%, .25%, and .5%. For higher education enrollment rates from 1820-1880 are .04%, .08%, .15%, .3%, .35%, .4% and .45%, respectively. The datum for 1890 is .5%. For 1800 and 1810 we assumed .02% higher education enrollment rate. For 2000 we used WDR. For 2010 we used *HDR*. Our time series share of the labor force exposed to higher education is consistent with Morrisson & Murtin from 1870-1970. Our 1870-1900 no education exposure shares are: 59%, 45%, 45% and 30%. This compares with Morrisson & Murtin estimates of 75%, 67%, 56% and 49%. Our years of schooling in the labor force for 1870-1920 are: 2.11 (1870), 2.81 (1880), 2.84 (1890), 3.61 (1900), 3.77 (1910) and 3.51 (1920). The Morrisson & Murtin time series of years of schooling is: 1.51 (1870), 2.04 (1880), 2.63 (1890), 3.12 (1900), 3.47 (1910) and 3.47 (1920).

4.8 Turkey (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1913, 1927, 1935, 1945, 1950, 1960, 1970, 1980 and 1990 come from Maa Table A2 p. 27. The 1945 value is interpolated from the 1935 and 1950 values. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1935, 1945, 1950, 1960, 1970, 1980 and 1990 come from Maa Table A2 p. 27. The age distribution for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1913 and 1927 are assumed to be the same as in 1935. The age distribution for Turkey for 1945 is interpolated from the 1935 and

[21] Our 1800 value is consistent with R.C. Allen adjusted for population in 1800, assuming a 50% labor force - population rate in rural agricultural population, and 34% labor force - population rate in rural non agricultural population and urban population. These labor force participation rates produce our 1850 labor force value for Spain with 65% rural agricultural population, and 35% non agricultural, both rural and urban.

1950 values. Age distribution for Turkey 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1950, 1960, 1970, 1980 and 1990 come from Maa Table B1 p. 101. For 1820, 1830, 1840, 1850, 1860 we used the following procedure. We used Issawi (1980) estimate of urbanization increased in the 19th century from .039 to .098. We used Banks for urban-rural population shares for 1921-1960, and *WDR* and *WDI* for 1971-2010. We assumed the rural 15-64 labor force participation rate was 80% from 1820-1890, 95% in 1900 and 1913, 105% in 1920, 115% in 1927 and 1935, 110% in 1940, 104% in 1950, 85% for 1960, 1970 and 1980, 60% in 1990 and 55% in 2000 and 2010. We assumed the urban 15-64 labor force participation rate was 50% for 1820-1900, 95% in 1913, 105% in 1920, 115% in 1927 and 1935, 110% in 1940, 104% in 1950, 75% in 1960, 70% in 1970, 65% in 1980, 55% in 1990, and 50% in 2000 and 2010. We constructed the ratio of this labor force with that from Maa, Pamuk, *WDR* and *WDI*. The root mean ratio for the overlapping years is .991. The ratios for 1870 and 1880 are 1.049, .990, respectively. We used Pamuk (2008) for agricultural labor force in 1870, 1880, and Mitchell for 1950. We interpolated between 1880 and 1950 for labor force agricultural. We used Pamuk (2008) for share of the aggregate labor force that is agricultural for 1880, 1900-1950. Labor force figures for Turkey 2000 come from *WDR*. Labor force figures for 2010 come from *WDI*.

Real GNP come from Maddison. Physical capital investment rates come from the intraperiod average gross capital formation and income for 1950-1959 and the intraperiod average gross real capital formation and real income for 1960-1998 from Maa Table J1, p. 1038 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. For 1880-2000 we used Altuğ, Filiztekin, Pamuk (2008). For 2010 we used perpetual inventory. Mitchell provides sectoral output shares for 1945-2000, and *WDI* provides sectoral output shares for 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1820-1935.[22] We applied the US capital - sectoral output shares for 1850-1940 to produce out capital estimates. We applied the US 1850 capital - sectoral output ratios for 1820, 1830 and 1840.

Enrollments in primary and secondary schools from 1925-1993 come from Maa Table I1 pp. 985 and 991. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1923-1993 come from Maa Table I2 pp. 1002 and 1005. For 1913 and 1920 we used Benavot and Riddle (1988) Appendix Table. For primary enrollment rates over 1820-1900 we assumed that each decade going back from 1913 had 85 percent of the enrollment rate in the succeeding decade. This produces a time series that has an initial enrollment rate of 2.8 percent in 1820, and a 1900 enrollment rate of 10.2 percent. For secondary school enrollment rates over 1820-1900 we assumed .5 percent for 1860-1900 for secondary enrollment rates. The 1870 value of exposed education share is 5 percent, which matches Morrison & Murtin. Furthermore the secondary exposed worker share of 1 percent matches Morrison & Murtin. Our years of schooling series for 1870-1900 is: .4, .4, .5 and .6 years. Their years of schooling series is: .3, .3, .4 and .5. For 1820-1900 we assumed a .01 percent higher education enrollment rate. This matches the 0 percent exposed higher education worker share of Morrison $ Murtin. For 2010 we used *HDR*. For 2000 we interpolated all of our enrollment rates. Our time series of years of schooling in the labor force for 1870-1920 is: .41 (1870), .46 (1880), .51 (1890), .56 (1900), .63 (1913) and

[22] Our shares are similar to Pamuk in 1880 and 1950, but with slightly lower agriculture shares for these years. We have 56% agriculture share in 1880 compared with 64%. Our 1950 value, from Mitchell, of 47% is lower than Pamuk's reported 53%. Out method produces agriculture share of 67% in 1820.

.69 (1920). The Morrisson & Murtin time series of years of schooling is: .26 (1870), .33 (1880), .39 (1890), .47 (1900), .53 (1910) and .59 (1920).

5 Central and Eastern Europe

5.1 Albania (1870-2010)

Population for 1870, 1890, 1900, 1910, 1920, 1929, 1940, 1950, 1960, 1970, 1980, 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950 come from Mitchell. Prior to 1950, we used the same age distribution as for 1950. Age distributions for 1960, 1970, 1980 are interpolated from 1950 and 1990 age distributions. The age distribution for 1990 comes from KF. Age distribution for Albania 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in 2012.

The labor force data for 1960, 1970 and 1980 come from the *WDI*. For values before 1960 we used the following procedure. We used Banks for urban-rural population shares from 1913-1960, *WDR* for 1970, 1980, 1990, and *WDI* for 2000 and 2010. Banks reports 0 urban share for years 1913-1939, and we assumed that Albania was no more urban prior to 1913. We assumed an urban 15-64 labor force participation rate of 50% for 1870-1990, 55% for 2000 and 65% for 2010. We assumed a rural 15-64 labor force participation rate of 89% for 1870-1960, 80% for 1970, 75% for 1980 and 1990, and 65% for 2000 and 70% for 2010. We constructed the labor force from this method and took the ratio relative to the data from *WDI* and *WDR*. The root mean ratio in the overlapping years 1960-2010 is 1.010. The 1960 and 1970 ratios are .999 and 1.007, respectively. Labor force figures for 1990 come from *HDR*. Labor force figures for 2000 come from *WDR*. The 2010 labor force figures come from *WDI*.

Real GNPs come from Maddison. The 2000 investment rate is the intraperiod average investment rate taken from *WDR* (various years). The 2010 value was the average of 2000-2009 investment rates from S & H. *WDI* provides sectoral output shares for 1980, 1990, 2000 and 2010. The 1950 and 1960 value comes from the *UN Statistical Yearbook* 1961. We interpolated the 1970 value. For years prior to 1950, we used Sabillon (2005) information on farming, manufacturing and aggregate growth rates. We used the US capital - sectoral output ratios for 1870-1990 to produce our estimates. We used perpetual inventory to calculated 2000 and 2010 values.

For 1870-1890, we assumed a 5%, 2% and .1% enrollment rate for primary, secondary and higher education. For 1900 and 1910 we assumed an 11% primary enrollment rate compared with the 1920 observation of 11.9%, from Benavot and Riddle. We assumed secondary enrollment rates of these years to be 40% of the concurrent primary enrollment rate. The higher education enrollment rates for these years are .1%. The 1920, 1929 and 1940 primary enrollment rates come from Benavot and Riddle. The secondary enrollment rates for these years are assumed to be 40% of the concurrent primary enrollment rate. The higher education enrollment rates are assumed to be .5%. The 1950 and 1960 enrollment rates were benchmarked to the literacy rate estimates of Banks, et. al. (1967). The 1970, 1980 figures come from *WDI*. Enrollments in primary and secondary schools for 1990 come from the 1996 *UN Statistical Yearbook*, Table 9. We adjusted the enrollments in primary school to include 50 percent of the reported secondary enrollments, and we reduced the secondary enrollments by 50 percent. This was done to fit

the 2000 enrollment rates reported in the WDR. The 2010 enrollment rates come from *HDR*. We assumed the primary school and secondary school ages are 6-13 and 14-17. We assumed that the initial stocks of primary, secondary and tertiary human capital are proportional to the Yugoslavian values in 1990, where the proportional constants are the different enrollment rates in 1990 between Albania and Yugoslavia. The tertiary school enrollments are from the 1996 *UN Statistical Yearbook*, Table 9.

5.2 Armenia (1970-2010)

Population for 1970 and 1980 come from *WDI*. Population for 1990 comes from *WDR*. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1970, 1980 and 1990 comes from KF. Age distribution for Armenia 2000 and 2010 come from the *Demographic Yearbook*. The 2010 age distribution is adjusted from 2011 by assuming that the 2010 age categories have the same share as in 2011.

Labor force data for 1970, 1980 and 1990 come from Easterly and Fischer. Labor force figures for 2000 come from *WDR*. The 2010 labor force data comes from *WDI*.

For real output we used Maddison. The 2000 investment rate is the intraperiod average investment rate taken from S & H years 1993-1999. The 2010 investment rate was the average of 2000-2009 investment rate from S & H. The *WDI* provides sectoral output shares for 2000 and 2010. We used Easterly and Fischer (1995) capital output ratios for 1970, 1980 and 1990, We used perpetual inventory for 2000 and 2010.

Enrollments in primary and secondary schools for 1990 come from the 1996 *UN Statistical Yearbook 1996*. We adjusted the enrollments in primary school to include 50 percent of the reported secondary enrollments, and we reduced the secondary enrollments by 50 percent. This was done to fit the 2000 enrollment rates reported in the *WDR*. We assumed the primary school and secondary school ages are 6-13 and 14-17. The 2010 rates are from *HDR*. For 1970 and 1980 we used *WDI*.

5.3 Azerbaijan (1970-2010)

Population for 1970 and 1980 come from Maddison. Population for 1990 comes from *WDR*. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1970, 1980 and 1990 comes from KF. Age distribution for Azerbaijan 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in 2012.

Labor force figures for 1970, 1980 and 1990 come from Easterly and Fischer. Labor force figures for 2000 come from *WDR*. The labor force data for 2010 comes from *WDI*.

Real GNPs for 1970, 1990, 2000 and 2010 come from Maddison. The 1980 observation is an interpolation of the 1970 and 1990 values. The 2000 investment rate is the intraperiod average investment rate taken from S & H for years 1993-1999. The 2010 investment rate was the 2000-2009 average investment rate from S & H. We used Easterly and Fischer (1995) for 1970, 1980 and 1990 capital output ratios. We used perpetual inventory for 2000 and 2010.

For 1970 and 1980 we used historical information from Didenko, Foldvari and Leeuwen (2013). They provide historical years of schooling for the former Soviet republics in 1939 and 1989, relative to Russia in those years. We interpolated the relative years of schooling for 1970 and 1980, and applied this to Russian enrollment rates in primary, secondary and higher education. Enrollments in primary and secondary

schools for 1990 come from the 1996 *UN Statistical Yearbook*, Table 9. We adjusted the enrollments in primary school to include 50 percent of the reported secondary enrollments, and we reduced the secondary enrollments by 50 percent. The 2010 data come from *HDR*. The 1990 tertiary school enrollments are from the 1996 *UN Statistical Yearbook*, Table 9. The 2000 enrollment rates are interpolated. We assumed the primary school and secondary school ages are 6-13 and 14-17.

5.4 Belarus (1970-2010)

Population for 1970, and 1980 come from *WDI*. Population for 1990 comes from *WDR*. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1970, 1980 and 1990 comes from KF. Age distribution for Belarus 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force for 1970, 1980 and 1990 come from Easterly and Fischer. Labor force figures for 2000 come from *WDR*. The labor force data for 2010 comes from *WDI*.

Real output comes from Maddison. The 2000 investment rate is the intraperiod average investment rate taken from S & H for years 1994-1999. The 2010 value is the average investment rate from 2000-2009 from S & H. For 1970-1990 we used Easterly and Fischer capital output ratios. For 2000 and 2010 we used perpetual inventory.

Enrollment rates in 1970 and 1980 come from *WDI*. Enrollments in primary and secondary schools for 1990 come from the 1996 *UN Statistical Yearbook*, Table 9. We adjusted the enrollments in primary school to include 50 percent of the reported secondary enrollments, and we reduced the secondary enrollments by 50 percent. We assumed the primary school and secondary school ages are 6-13 and 14-17. We assumed that the initial stocks of primary, secondary and tertiary human capital are proportional to the Soviet Union values in 1989, where the proportional constants are the different enrollment rates in 1989/1990 between Belarus and the Soviet Union. The tertiary school enrollments are from the 1996 *UN Statistical Yearbook*, Table 9. The 2010 values come from *HDR*. Out 2000 enrollment rates are interpolated.

5.5 Bulgaria (1870-2010)

Populations for 1870, 1880, 1890, 1900, 1910, 1920, 1934, 1946, 1956, 1965 come from Meu Table A2 p. 15. Populations for 1970, 1980 and 1990 come from interpolations of 1965, 1975, 1985 and 1992 figures from Meu Table A2 p. 15. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1890, 1900, 1910, 1920, 1934, 1946, 1956, 1965 come from Meu Table A2 p. 15. Age distribution for 1870 and 1880 is assumed to be identical with the 1890 age distribution. Age distributions for 1970, 1980, 1990 and come from interpolations of 1965, 1975, 1985, 1992 figures from Meu Table A2 p. 15. Age distribution for Bulgaria 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1910, 1920, 1934, 1946, 1956, 1965, 1975, 1985 and 1992 come from Meu Table B1 p. 146. The 1900 value comes from Damianova. For 1870, 1880, 1890 labor force data we used Banks for

urban-rural population shares. Banks begins in 1880, and goes through 1965. We used *WDR* and *WDI* for the 1975-2010 period. We assumed that the 1870 urban share was .024, compared with Banks data of .034 and .046 for 1880 and 1890. We assumed the urban 15-64 labor force participation rate was always 75%, and that the rural 15-64 labor force participation rate was 90% for 1880-1956 and 85% from 1965-2010. We construct the ratio of this labor force to that in Meu and *WDR* and *WDI*. The root mean ratio is 1.002, and the 1880 and 1890 ratios are .955 and .955, respectively. Labor force data for 2000 come from *WDR*. Labor force data for 2010 come from *WDI*.

Real GNP comes from Maddison. Physical capital investment rate from 1980-1998 comes from S & H online and *WDR* (various years). For 2010 we used the average investment rate from 2000-2009 from S & H. Mitchell provides sectoral output shares for 1934-2000, sans 1940, which we interpolated. *WDI* provides sectoral output shares for 2010. For years 1920-2000 we used capital output ratios from Bas van Leeuwen and Peter Foldvari. Sabillon (2005) provides farming, manufacturing and aggregate growth rates in order to produce sectoral output shares from 1870-1920. We used the US capital - sectoral output ratios for 1870-1920 to produce our estimates. We used perpetual inventory to produce our 2010 estimate.

Enrollments in primary and secondary school from 1890-1993 come from Meu Table I1 pp. 873, 880 and 887. For 1870 and 1880 we assumed that the primary schools enrollment rates are 99% of their succeeding decade value, 37.2%, 37.6% reaching the 1890 datum of 38%. The 37.2% enrollment rate in 1870 is consistent with a 19 percent education exposure rate, comparable to the 21 percent exposure rate from Morrison & Murtin. For secondary school enrollment rates in 1870 and 1880, we assumed 1% and 2%, reaching the 1890 datum of 3.5%. To calculate enrollment rates, in 1956 we assumed 6-11 are primary school age and 12-17 are secondary school age. For 1965 we assumed 6-13 are primary school age and 14-17 are secondary school age, this was to maintain consistency with the 1956 numbers and the change in secondary school enrollment rates since 1970. Higher education enrollments for 1895-1993 are from Meu Table I2 pp. 895, 897 and 899. We assumed .1% (the 1900 datum) enrollment rates in higher education for 1870, 1880 and 1890. The 2010 values come from *HDR*, and we interpolated all the enrollment rates for 2000. Our time series of higher education exposed labor force share is consistent with Morrisson & Murtin from 1870-1980. Our time series of years of schooling in the labor force for 1870-1920 is: .56 (1870), .82 (1880), 1.50 (1890), 2.39 (1900), 3.60 (1910) and 4.80 (1920). The Morrisson & Murtin time series for years of schooling is: 1.65 (1870), 1.91 (1880), 2.22 (1890), 2.72 (1900), 3.04 (1910) and 4.16 (1920).

5.6 Czechoslavakia (Czech Republic) (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1870, 1890, 1900, 1913, 1921, 1930, 1950, 1961, 1970, 1980 and 1990 come from Meu Table A2 p. 16. Population for Czech Republic 2000 comes from *Time Almanac 2001*. Population for Czech Republic 2010 comes from *Wikipedia*.

The age distributions of the population for 1921, 1930, 1950, 1961, 1970, 1980 and 1990 come from Meu Table A2 p. 16. For 1820, 1830, 1840, 1850, 1870, 1890, 1900, 1913 we assumed the same age distribution as in 1921. Age distribution for Czech Republic 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1921, 1930, 1947, 1961, 1970, 1980, 1991 come from Meu Table B1 p. 147. For 1913 we used Maddison (2001). For years prior to 1820-1900, we used two sources, Agnew for 1870, 1900 and 1913. We used the agricultural share of the active population, aged 15-64, with an 80% labor force

participation rate, and a 70% labor force participation rate for non agricultural population share. We interpolated for 1880 and 1890. For 1820-1860, we used Good. He provides the share of the labor force that is agricultural for 1756, 1846. We interpolated for 1820, 1830 and 1840 shares of population that are agricultural, and assumed that both agricultural and non agricultural labor force participation rates are 80%. We interpolated for 1860. Labor force data for 2000 come from *WDR*. Labor force data for 2010 come from *WDI*.

Real GNPs come from Maddison. Real physical capital investment rates come from the intraperiod averages of gross real capital formation and real income for 1920-1993 from Meu Table J1, pp. 908, 914 and 922 and *WDR* (various years). For 2010 we used the average of 2000-2009 investment rates from S & H. Mitchell provides sectoral output shares for 1921-2000, except for 1940, which we interpolated. The *WDI* provides sectoral output shares for 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1820-1913. We applied the US 1850-1910 capital - sectoral output ratios to produce our capital estimates for 1820-1913. We used the 1850 capital - sectoral output ratios of the US for 1820, 1830 and 1840. We used perpetual inventory to compute capital stocks for years 1921-2010.

Enrollments in primary and secondary school from 1921-1993 come from Meu Table I1 pp. 881 and 887. To calculate enrollment rates, we assumed 6-13 are primary school age and 14-17 are secondary school age. Higher education enrollments are for 1921-1993 are from Meu Table I2 pp. 897 and 899. We assumed that primary enrollment rates in 1820, 1830, 1840, 1850, 1860, 1870, 1890 and 1900 were 80% of German enrollment rates. These are consistent with Koralka (1990). We assumed that secondary enrollment rates in 1820, 1830, 1840, 1850, 1860, 1870, 1890 and 1900 were 40% of German enrollment rates. For 1913 interpolated between our 1900 value and our 1921 value. For 2010 we used *HDR* for all enrollment rates. We interpolated for all enrollment rates for 2000.

5.7 East Germany (1950-1990)

Populations for 1950, 1964, 1971, 1981 come from Meu Table A1 p. 4. Population for 1990 come from S & H (1991).

The age distributions of the population for 1950, 1964, 1971, 1981 come from Meu Table A2 p. 23. Age distribution for 1990 comes from KF.

Labor force figures for 1950, 1961, 1971 come from Meu Table B1 p. 150. Labor force data for 1990 are from *HDR*. Labor force data for 1981 is interpolated from 1971 and 1990 values.

Real GNP come from Maddison Physical capital investment rates and real income come from the intraperiod averages of gross capital formation and income for 1950-1993 from Meu Table J1, pp. 916 and 923. We used nominal values converted into our real international 1985 dollars because there was little evidence of sustained inflation in the entire East German series. We used UN Statistical Yearbooks for 1961, 1966, 1973, and 1981 for sectoral output shares for years 1950, 1964, 1971 and 1980, where the 1964 values are interpolated. Our 1990 value comes from Sabillon (2005) applying his 1980 growth rates of farming and manufacturing relative to aggregate growth rates. We used the 1950 US capital - sectoral output ratios to construct our 1950 value. We used perpetual inventory for all other years.

Enrollments in primary and secondary school for 1950-1989 are from Meu Table I1, p. 882 888 . To calculate enrollment rates, we assumed 6-13 are primary school age and 14-17 are secondary school age. Higher education enrollments for 1950-1988 are from Meu Table I2, pp. 897 and 899.

5.8 Estonia (1970-2010)

Population for 1970, 1980 come from Maddison. Population for 1990 comes from *WDR*. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1970, 1980 and 1990 comes from KF. Age distribution for Estonia 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force data for 1970, 1980 and 1990 are from Easterly and Fischer. Labor force figures for 2000 come from *WDR*. Labor force data for 2010 come from *WDI*.

Real output comes from Maddison. The 2000 investment rate is the intraperiod average investment rate taken from S & H for 1990-1999. The 2010 investment rate is the average of investment rates from 2000-2009 from S & H. For capital values in years 1970, 1980 and 1990 we used Easterly and Fischer (1995). For 2000 and 2010 we used perpetual inventory.

Enrollments in primary, secondary and tertiary schools for 1980-2010 come from the *WDI*. The 1970 rates are interpolated using the 1935-40 observation from Benavot & Riddle (1988). We assumed the primary school and secondary school ages are 6-13 and 14-17.

5.9 Georgia (1970-2010)

Population for 1970 and 1980 come from Maddison. Population for 1990 comes from *WDR*. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1970, 1980 and 1990 comes from KF. Age distribution for Georgia 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in 2012.

Labor force data for 1970, 1980 and 1990 come from Easterly and Fischer. Labor force figures for 2000 come from *WDR*. The labor force data for 2010 comes from *WDI*.

Real output comes from Maddison. The 2000 investment rate is the intraperiod average investment rate taken from S & H for years 1993-1999. The 2010 investment rate is the average investment rate for 2000-2009 from S & H. We used Easterly and Fischer (1995) capital output ratios to produce our estimates for 1970, 1980 and 1990. We used perpetual inventory for our 2000 and 2010 values.

For 1970 and 1980 we used historical information from Didenko, Foldvari and Leeuwen (2013). They provide historical years of schooling for the former Soviet republics in 1939 and 1989, relative to Russia in those years. We interpolated the relative years of schooling for 1970 and 1980, and applied this to Russian enrollment rates in primary, secondary and higher education. Enrollments in primary, secondary and tertiary schools for 1990 come from the *UN Statistical Yearbook 1996*, Table 9. We adjusted the enrollments in primary school to include 50 percent of the reported secondary enrollments, and we reduced the secondary enrollments by 50 percent. This was done to fit the 2000 enrollment rates reported in the WDR. The 2010 enrollment rates from the *HDR*.

5.10 Hungary (1869-2010)

Populations for 1869, 1880, 1890, 1900, 1910, 1920, 1930, 1941, 1949, 1960, 1970, 1980 and 1991 come from Meu Table A2 pp. 25 and 26. Population for 2000 comes from *Time Almanac 2001*. Population for 2010

comes from *Wikipedia*.

The age distributions of the population for 1869, 1880, 1890, 1900, 1910, 1920, 1930, 1941, 1949, 1960, 1971, 1980 and 1991 come from Meu Table A2 pp. 25 and 26. Age distribution for Hungary 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1869, 1880, 1890, 1900, 1910, 1920, 1930, 1949, 1960, 1970, 1980 and 1990 come from Meu Table B1 p. 151. Labor force data for 2000 come from *WDR*. The labor force data for 2010 comes from *WDI*.

Real GNP comes from Maddison. Physical capital investment rates come from the intraperiod averages of gross real capital formation and real income for 1925-1998 from Meu Table J1, pp. 910, 917 and 924 and *WDR* (various years). The 2010 value is the average investment rate from 2000-2009 from S & H. We used Max-Stephan Schulze (2007) capital output ratios for 1869-1910 for our estimates. Mitchell provides sectoral output shares for 1900-2000, except for 1940, which we interpolated. We applied the US 1920 capital - sectoral output ratios to construct our 1920 estimate. For years 1930-2000 we used perpetual inventory.

Enrollments in primary and secondary school from 1869-1993 come from Meu Table I1 pp. 875, 883 and 888. To calculate enrollment rates, we assumed 6-13 are primary school age and 14-17 are secondary school age. Higher education enrollments are from 1875-1993 are from Meu Table I2 pp. 895, 897 and 899. We used the higher education enrollment rate in 1880, .10%, for 1869. The 2010 values come from *HDR*. We interpolated for all enrollment rates in 2000. Our time series of years of schooling in the labor force for 1869-1920 is: 2.67 (1869), 3.08 (1880), 3.82 (1890), 4.62 (1900), 5.07 (1910) and 5.17 (1920). The Morrisson & Murtin time series of years of schooling is: 2.58 (1870), 2.97 (1880), 3.59 (1890), 4.05 (1900), 4.59 (1910) and 5.11 (1920).

5.11 Kazakhstan (1970-2010)

Population for 1970 and 1980 comes from Maddison. Population for 1990 comes from *WDR*. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1970, 1980 and 1990 comes from KF. Age distribution for Kazakhstan 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 1999 age distribution, and for 2010 we adjusted the 2008 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force data for 1970, 1980 and 1990 come from Easterly and Fischer. Labor force figures for 2000 come from *WDR*. The labor force data for 2010 come from *WDI*.

Real output comes from Maddison. The 2000 investment rate is the intraperiod average investment rate taken from S & H for years 1993-1999. The 2010 investment rate is the average investment rate for 2000-2009 from S & H. For 1970-1990 we used the capital output ratios from Easterly and Fischer (1995). For years 2000 and 2010 we used perpetual inventory.

For 1970 and 1980 we used historical information from Didenko, Foldvari and Leeuwen (2013). They provide historical years of schooling for the former Soviet republics in 1939 and 1989, relative to Russia in those years. We interpolated the relative years of schooling for 1970 and 1980, and applied this to Russian enrollment rates in primary, secondary and higher education. Enrollments in primary and secondary

schools for 1990 come from the 2996 *UN Statistical Yearbook*, Table 9. We adjusted the enrollments in primary school to include 50 percent of the reported secondary enrollments, and we reduced the secondary enrollments by 50 percent. This was done to fit the 2000 enrollment rates reported in the WDR, which we interpolated. We used *HDR* for enrollment rates in 2010. We assumed the primary school and secondary school ages are 6-13 and 14-17. The 1990 tertiary school enrollments are from the 1996 *UN Statistical Yearbook*, Table 9.

5.12 Kyrgyzstan (1970-2010)

Population for 1970 and 1980 comes from Maddison. Population for 1990 comes from *WDR*. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1970,1980 and 1990 comes from KF. Age distribution for Kyrgyzstan 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1970, 1980 and 1990 come from Easterly and Fischer. Labor force figures for 2000 come from *WDR*. The labor force data for 2010 come from *WDI*.

Real output comes from Maddison. The 2000 investment rate is the intraperiod average investment rate taken from S & H for years 1993-1999. The 2010 investment rate is the average investment rate for 2000-2009 from S & H. For 1970-1990 we used the capital output ratios from Easterly and Fischer (1995) to produce our estimates. For 2000 and 2010 we used perpetual inventory.

For 1970 and 1980 we used historical information from Didenko, Foldvari and Leeuwen (2013). They provide historical years of schooling for the former Soviet republics in 1939 and 1989, relative to Russia in those years. We interpolated the relative years of schooling for 1970 and 1980, and applied this to Russian enrollment rates in primary, secondary and higher education. Enrollments in primary and secondary schools for 1990 come from the *UN Statistical Yearbook 1996*, Table 9. We adjusted the enrollments in primary school to include 50 percent of the reported secondary enrollments, and we reduced the secondary enrollments by 50 percent. This was done to fit the 2000 enrollment rates, which we interpolated. We used *HDR* for enrollment rates in 2010. We assumed the primary school and secondary school ages are 6-13 and 14-17. The tertiary school enrollments are from the 1996 *UN Statistical Yearbook*, Table 9.

5.13 Latvia (1970-2010)

Population for 1970 and 1980 comes from Maddison. Population for 1990 comes from *WDR*. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1970, 1980 and 1990 comes from KF. Age distribution for Latvia 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1970, 1980 and 1990 come from Easterly and Fischer. Labor force figures for 2000 come from *WDR*. The labor force data for 2010 come from *WDI*.

Real output comes from Maddison. The 2000 investment rate is the intraperiod average investment rate taken from S & H for 1993-1999. The 2010 investment rate is the average investment rate for 2000-2009

from S & H. For 1970-1990 we used the capital output ratios from Easterly and Fischer (1995) to produce our estimates. We used perpetual inventory for 2000 and 2010.

For 1970 and 1980 enrollment rates we used *WDI*. Enrollments in primary, secondary and tertiary schools for 1990 come from the *UN Statistical Yearbook 1996*, Table 9. We adjusted the enrollments in primary school to include 50 percent of the reported secondary enrollments, and we reduced the secondary enrollments by 50 percent. This was done to fit the 2000 enrollment rates, which we interpolated. We used *HDR* for enrollment rates in 2010. We assumed the primary school and secondary school ages are 6-13 and 14-17.

5.14 Lithuania (1970-2010)

Population for 1970 and 1980 comes from Maddison. Population for 1990 comes from *WDR*. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1970, 1980 and 1990 come from KF. Age distribution for Lithuania 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1970, 1980 and 1990 come from Easterly and Fischer. Labor force figures for 2000 come from *WDR*. The labor force data for 2010 come from *WDI*.

Real output comes from Maddison. The 2000 investment rate is the intraperiod average investment rate taken from S & H for years 1993-1999. The 2010 investment rate is the average investment rate for 2000-2009 from S & H. For 1970-1990 we used the capital output ratios from Easterly and Fischer (1995) to produce our estimates. We used perpetual inventory for 2000 and 2010.

For 1970 and 1980 enrollment rates we used *WDI*. Enrollments in primary and secondary schools for 1990 come from the 1996 *UN Statistical Yearbook*, Table 9. We adjusted the enrollments in primary school to include 50 percent of the reported secondary enrollments, and we reduced the secondary enrollments by 50 percent. This was done to fit the 2000 enrollment rates, which we interpolated. We used *HDR* for enrollment rates in 2010. We assumed the primary school and secondary school ages are 6-13 and 14-17. The 1990 tertiary school enrollments are from the 1996 *UN Statistical Yearbook*, Table 9.

5.15 Moldova (1970-2010)

Population for 1970 and 1980 comes from Maddison. Population for 1990 comes from *WDR*. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributionsfor 1970, 1980 and 1990 comes from KF. Age distribution for Moldova 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force data for 1970, 1980 and 1990 are from Easterly and Fischer. Labor force figures for 2000 come from *WDR*. The labor force data for 2010 come from *WDI*.

Real output comes from Maddison. The 2000 investment rate is the intraperiod average investment rate taken from S & H for years 1992-1999. The 2010 investment rate is the average investment rate for

2000-2009 from S & H. For 1970-1990 we used the capital output ratios from Easterly and Fischer (1995) to produce our estimates. We used perpetual inventory for 2000 and 2010.

For 1970 and 1980 enrollment rates we used *WDI*. Enrollments in primary, secondary and higher education schools for 1990 come from the 1996 *UN Statistical Yearbook*, Table 9. We adjusted the enrollments in primary school to include 50 percent of the reported secondary enrollments, and we reduced the secondary enrollments by 50 percent. The 2000 enrollment rates are interpolated from our 1990 and 2010 data. We used *HDR* for enrollment rates in 2010. We assumed the primary school and secondary school ages are 6-13 and 14-17.

5.16 Poland (1870-2010)

Population for 1870, 1890, 1900, 1910, 1921, 1950, 1960, 1970, 1980, 1991 come from Meu Table A2 p. 32. Population for 1931 comes from the *Wikipedia*, https://en.wikipedia.org/wiki/Demographic_history_of_Poland, accessed July 24, 2016. The 1980 and population is interpolated using 1978, 1991 values. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1921, 1931, 1950, 1960, 1970, 1980 and 1991 come from Meu Table A2 p. 32. We assumed the age distributions for 1870, 1890, 1900, and 1910 are the same as in 1921. The 1980 age distribution is interpolated using 1978 and 1991 values. Age distribution for Poland 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we interpolated using the 1998 and 2002 age distributions, and for 2010 we adjusted the 2012 age distribution. We adjust the 2010 age distribution by assuming the same share by age category as in 2012.

Labor force figures for 1921, 1931, 1950, 1960, 1970, 1980 come from Meu Table B1 p. 155. We used Wolf (2006) for labor force data from 1870-1910. Wolf provides the share of Poland that is urbanized for 1900, 1910 and for regions in 1860. We assumed that 6% of Poland is urbanized in 1860, which is the average of the regions in Table 3. We interpolated the urban share from 1860 to 1900. We assumed that the labor force for 1870, 1890, 1900 and 1910 come from the 15-64 population, where rural and urban populations have the same demographic age structure. We assumed that the urban population has an 85% labor force participation rate and the rural population has a 100% labor force participation rate. These assumptions match well the Meu data for 1921 and 1931. The 1990 and 2000 values are from *WDR* (various years). The 2010 value comes from *WDI*.

Real GNP comes from Maddison. Physical capital investment rates come from the intraperiod averages of gross real capital formation and real income for 1947-1998 from Meu Table J1, pp. 918 and 925 and *WDR* (various years). For 2010 we used the average investment rate from 2000-2009 from S & H. Mitchell provides sectoral output shares for 1950-2000, and *WDI* provides sectoral shares for 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral shares for 1870-1940. We applied the US 1870-1940 capital - sectoral output ratios to produce our capital estimates. For years 1950-2010 we used perpetual inventory.

Enrollments in primary and secondary school from 1922-1937, 1945-1993 come from Meu Table I1 pp. 884 and 889. To calculate enrollment rates, we assumed 6-13 are primary school age and 14-17 are secondary school age. Higher education enrollments for 1920-1993 are from Meu Table I2 pp. 898 and 899. For 1870, 1890, 1900, 1910 we assumed primary enrollment rates of 10 percent, 25 percent, 30 percent and 44 percent, respectively. The 10 percent primary enrollment rate in 1870 is similar to the 5 percent adult literacy rate of the 1701-1800 cohort of adults. The 1921 primary school enrollment rate datum is 60 percent. For

secondary school enrollment rates we assumed 1 percent, 1 percent, 2 percent and 4 percent, respectively. The 1921 secondary school enrollment rate datum is 8.8 percent. For higher education enrollment rates we assumed .1 percent, .2 percent, .2 percent and .4 percent, respectively. The 1921 higher education enrollment rate datum is .8 percent. For 2000 we interpolated, and for 2010 we used *HDR*.

5.17 Romania (1870-2010)

Populations for 1870, 1880, 1890, 1899, 1912, 1920, 1948, 1956, 1966, 1970, 1980 come from Meu Table A2 p. 34. Population for 1930 and 1938 come from *Wikipedia*, https://en.wikipedia.org/wiki/Demographics_of_Romania#Between_WWI_and_WWII, accessed July 24, 2016. The 1990 value is from S & H online. Population for 2000 comes from *Time Alamanc 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1899, 1912, 1920, 1930, 1956, 1966, 1980 and 1990 come from Meu Table A2 p. 34. The age distribution for 1870, 1880 and 1890 are assumed to be the same as in 1899. We interpolated for the age distributions in 1938, 1948 and 1970. The age distribution for 1980 and 1990 are interpolations of the 1977 and 1993 values. Age distribution for Romania 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1912, 1930, 1956, 1966, 1980 and 1990 come from Meu Table B1 p. 156, we interpolated for 1920. For 1870, 1890, 1899 we used Banks for urban-rural population shares for 1870-1966, and *WDR* and *WDI* for urban-rural shares for 1980-2010. We assumed an urban 15-64 labor force participation rate of 80%, and a rural 15-64 labor force participation rate of 95% for 1870-1956 For 1966-2010, we assumed an urban 15-64 labor force participation rate of 65% and a rural 15-64 labor force participation rate of 80%. We constructed the ratio of this labor force with that from Meu and *WDR* and *WDI*. The root mean ratio is .990, and the 1880 and 1890 ratios are .992, and .968, respectively. Labor force data for 1980, 1990 are interpolations using 1966 and 1992 values. Labor force data for 2000 come from *WDR*. Labor force data for 2010 comes from *WDI*.

Real incomes come from Maddison. Physical capital investment rates for 1960-1998 are from S & H online and *WDR* (various years). For 2010 we used the average investment rate from 2000-2009 from S & H. Mitchell provides sectoral output shares for 1956-2000, except for 1970 and 1980. *WDI* provides sectoral output shares for 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral shares for 1870-1948, and 1970 and 1980. For years 1870-1956 we used the US 1870-1960 capital - sectoral output ratios to produce our capital estimates. We used perpetual inventory for years 1966-2010.

Enrollments in primary and secondary school from 1870, 1890, 1899, 1912, 1930-1993 come from Meu Table I1 pp. 877, 884 and 889. For 1920 we used Lindert. To calculate enrollment rates, we assumed 6-13 are primary school age and 14-17 are secondary school age. Higher education enrollments for 1900-1993 are from Meu Table I2 pp. 896, 898 and 899. For 1870, 1880 and 1890 we assumed higher education enrollment rates of .20%, .24% and .30%, respectively. The 1900 datum is .60%. For 2000 we interpolated, and for 2010 we used *HDR*.

5.18 Soviet Union/Russia (1820-2010)

Population for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1897 come from A.G. Rashin (1956) via email correspondence with Paul Dower, Ocotber 31, 2016. We used the population of the 50 European provinces. Rashin provides population data for 1811, 1863, 1885 and 1897. We interpolated for 1820, 1830, 1840, 1850, 1860, 1870, 1880. For 1926, 1939, 1949, 1959, 1970, 1989 come from Meu Table A1 p. 7. We interpolated for 1907 and 1917. Population for Soviet Union in 1980 comes from S & H online at NBER. Population for Russia in 2000 comes from *Time Almanac 2001*. Population for Russia in 2010 comes from *Wikipedia*.

The age distribution of the population for 1897, 1917, 1926, 1939, 1959, 1970, 1987 comes from Meu Table A2 p. 35. For 1820, 1830 1840, 1850, 1860, 1870, 1880, 1890 we assumed the age distribution of 1897 was identical. The 1907, 1917 values are interpolations of the 1897 and 1926 values. Age distribution for Russia 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in 2012.

Labor force figures for 1926, 1959 and 1970 come from Meu Table B1 p. 156. Labor force data for 1939 and 1949 are interpolations of 1926, and 1959 values. The 1980 labor force is interpolated from the 1970 and 1987 labor force. Labor force data for 1987 comes from the HDR (1990). Labor force data for 2000 come from *Time Almanac 2001*. The 2010 labor force data comes from *WDI*. In email correspondence, Dower argued that the best estimate of labor force prior to 1926 would be an estimate of the 16 and older population, roughly 70% of the Russian population in 1811-1865 period, and 65% of Russian population in 1865-1913 period. For 1917 we used ithe population 15 and older, which is 69.4% of the Russian population. We used the 15 and older population in 1907, which is 69.1% of the Russian population. For 1897 we used the 16 and older population, which is 70.8% of the Russian population.[23] The 1890 labor force given by the employment growth rate in Meliantsev, is within 1.6% of the population 16 and older. The 1880 labor force, again given by the employment growth rate in Meliantsev, relative to our population estimate of 16 and older, it is within 2.5%. For 1820, 1830, 1840, 1850, 1860 and 1870 we used the population 16 and older for each year. These are 70.8% of the Russian population, consistent with Dower's suggestion.

Real GNPs are from Maddison. Physical capital investment rates come from the intraperiod averages of gross real capital formation and real income for 1938-1998 from Meu Table J1, pp. 912, 919 and 925 and *WDR* (various years). For 2010 we used the average investment rate over 2000-2009 from S & H. We used Goldsmith (1985) for capital values 1917-1970. Mitchell provides sectoral output shares for 1959-1980, and *WDI* provides sectoral output shares for 1989, 2000 and 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1820-1949. We applied US 1850-1950 capital - sectoral output ratios to produce our estimates of physical capital for 1820-1949, where we used the 1850 ratios for 1820, 1830 and 1840. For physical capital stock estimates of 1980-2010 we used perpetual inventory.

Enrollments for 1850 and 1860 are based on literacy rates of Morris and Adelman. They report 5-10% literacy rates in 1850 and 1870. For 1820, 1830, 1840 we assumed the same primary and higher education enrollment rates as in 1850, and for secondary enrollment rates, we assumed half the 1850 secondary enrollment rate. Our 1820-1840 primary enrollment rates of 7 percent are consistent with estimates of literacy for Russia, 8% in 1820 from Roser and also Meliantsev, and the aforementioned 5-10% rates by

[23] Our labor force is within 1% of the estimate from Kessler and Valetov. They used 56.4% of the population, applied to the Russian population of 116,237,800, which includes non European Russian population.

Morris and Adelman. For 1870, 1890, 1897 we used Cameron. For 1917, 1926, 1939, 1959, and 1970 we used Clarke (1972). Enrollment rates in primary, secondary and higher education for 1980, 1987 and 1996 are from *WDR* (various years). For 2000 we used WDR, and for 2010 we used *HDR*. Our time series of the share of the labor force exposed to higher education is consistent with Morrisson & Murtin from 1870-1980. Our 1870 years of schooling estimate is .6, compared with Morrisson & Murtin of .9. Our time series of years of schooling in the labor force for 1870-1917 is: .57 (1870), .57 (1880), .69 (1890), 1.03 (1897), 1.57 (1907), 1.94 (1917). The Morrisson & Murtin time series of years of schooling is: .90 (1870), 1.00 (1880), 1.10 (1890), 1.20 (1900), 1.30 (1910), 1.91 (1920).

5.19 Slovak Republic (1990-2010)

Population for 1990 comes from *WDR*. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distribution for 1990 comes from KF. Age distribution for Slovak Republic 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1990 come from *HDR*. Labor force figures for 2000 come from *WDR*. The labor force data for 2010 comes from *WDI*.

Real output comes from Maddison. The 1990 investment rate was the average physical capital investment rate from S & H for 1987-1989. The 2000 investment rate is the intraperiod average investment rate taken from S & H for years 1990-1999 The 2010 investment rate is the average investment rate for 2000-2009 from S & H. *WDI* provides sectoral output shares for 2000 and 2010. We used Sabillon (2005) information on 1990 Czech farming, manufacturing and aggregate growth rates to construct 1990 sectoral output shares. We used the US 1990 capital - sectoral output ratios to construct our 1990 capital estimate. We used perpetual inventory to compute our 2000 and 2010 values.

Enrollments in primary and secondary schools for 1990 come from the 1996 *UB Statistical Yearbook*, Table 9. We adjusted the enrollments in primary school to include 50 percent of the reported secondary enrollments, and we reduced the secondary enrollments by 50 percent. This was done to fit the 2000 enrollment rates reported in the *WDR*. We assumed the primary school and secondary school ages are 6-13 and 14-17. We assumed that the initial stocks of primary, secondary and tertiary human capital are proportional to the Soviet Union values in 1989, where the proportional constants are the different enrollment rates in 1989/1990 between the Slovak Republic and the Czechoslavakia. The tertiary school enrollments are from the 1996 *UN Statistical Yearbook*, Table 9. The 2010 values are from *HDR*. We interpolated for all 2000 enrollment rates.

5.20 Tajikistan (1970-2010)

Population for 1970 and 1980 come from Maddison. Population for 1990 comes from *WDR*. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1970, 1980 and 1990 come from KF. Age distribution for Tajikistan 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in 2012.

The labor force data for 1970, 1980 and 1990 come from Easterly and Fischer. Labor force figures for 2000 come from *WDR*. The labor force data for 2010 comes from *WDI*.

Real output comes from Maddison. The 2000 investment rate is the intraperiod average investment rate taken from S & H for years 1993-1999. The 2010 investment rate is the average investment rate for 2000-2009 from S & H. For 1970-1990 we used Easterly and Fischer (1995) to produce our capital estimates. We used perpetual inventory for 2000 and 2010.

Enrollment rates for 1970 and 1980 come from *WDI*. Enrollments in primary and secondary schools for 1990 come from the 1996 *UN Statistical Yearbook*, Table 9. We adjusted the enrollments in primary school to include 50 percent of the reported secondary enrollments, and we reduced the secondary enrollments by 50 percent. This was done to fit the 2000 enrollment rates, which we interpolated. We assumed the primary school and secondary school ages are 6-13 and 14-17. We assumed that the initial stocks of primary, secondary and tertiary human capital are proportional to the Soviet Union values in 1989, where the proportional constants are the different enrollment rates in 1989/1990 between Tajikstan and the Soviet Union. The tertiary school enrollments are from the 1996 *UN Statistical Yearbook*, Table 9. The 2010 data come from *HDR*.

5.21 Turkmenistan (1970-2010)

Population for 1970 and 1980 come from Maddison. Population for 1990 comes from *WDR*. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1970, 1980 and 1990 come from KF. Age distribution for Turkmenistan 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in the reference year.

The labor force data for 1970, 1980 and 1990 come from Easterly and Fischer. Labor force figures for 2000 come from *WDR*. The labor force data for 2010 comes from *WDI*.

Real output comes from Maddison. The 2000 investment rate is the intraperiod average investment rate taken from S & H for years 1993-1999. The 2010 investment rate is the average investment rate for 2000-2009 from S & H. For 1970-1990 we used Easterly and Fischer (1995) to produce our capital estimates. We used perpetual inventory to produce our 2000 and 2010 values.

For 1970, 1980 and 1990 we used historical information from Didenko, Foldvari and Leeuwen (2013). They provide historical years of schooling for the former Soviet republics in 1939 and 1989, relative to Russia in those years. We interpolated the relative years of schooling for 1970, 1980 and 1990, and applied this to Russian enrollment rates in primary, secondary and higher education. For 2010 secondary school enrollment rate we used Tajikistan's value. For primary and tertiary school enrollment rates we used *WDI*. We interpolated for 2000 enrollment rates.

5.22 Ukraine (1970-2010)

Population for 1970 and 1980 come from Maddison. Population for 1990 comes from *WDR*. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 19709, 1980 and 1990 come from KF. Age distribution for Ukraine 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2013 age distribution. We adjust by

assuming the same share by age category as in 2013.

The labor force data for 1970, 1980 and 1990 come from Easterly and Fischer. Labor force figures for 2000 come from *WDR*. The labor force data for 2010 comes from *WDI*.

Real output comes from Maddison. The 2000 investment rate is the intraperiod average investment rate taken from *WDR* (various years). The 2010 investment rate is the average investment rate for 2000-2009 from S & H. For 1970-1990 we used Easterly and Fischer (1995) to produce our capital estimates. We used perpetual inventory to produce our 2000 and 2010 values.

Enrollment rates for 1970 and 1980 come from *WDI*. Enrollments in primary and secondary schools for 1990 come from the 1996 *UN Statistical Yearbook*, Table 9. We adjusted the enrollments in primary school to include 50 percent of the reported secondary enrollments, and we reduced the secondary enrollments by 50 percent. This was done to fit the 2000 enrollment rates, which we interpolated. We assumed the primary school and secondary school ages are 6-13 and 14-17. We assumed that the initial stocks of primary, secondary and tertiary human capital are proportional to the Soviet Union values in 1989, where the proportional constants are the different enrollment rates in 1989/1990 between Ukraine and the Soviet Union. The tertiary school enrollments are from the 1996 *UN Statistical Yearbook*, Table 9. The 2010 data come from *HDR*.

5.23 Uzbekistan (1970-2010)

Population for 1970 and 1980 come from Maddison. Population for 1990 comes from *WDR*. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1970,1980 and 1990 come from KF. Age distribution for Uzbekistan 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2001 age distribution. We adjust by assuming the same share by age category as in 2001.

The labor force data for 1970, 1980 and 1990 come from Easterly and Fischer. Labor force figures for 2000 come from *WDR*. The labor force data for 2010 comes from *WDI*.

Real output comes from Maddison. The 2000 investment rate is the intraperiod average investment rate taken from S & H for years 1993-1999. The 2010 investment rate is the average investment rate for 2000-2009 from S & H. For 1970-1990 we used Easterly and Fischer (1995) to produce our capital estimates. We used perpetual inventory to produce our 2000 and 2010 values.

Enrollment rates for 1970 and 1980 come from *WDI*. Enrollments in primary and secondary schools for 1990 come from the 1996 *UN Statistical Yearbook*, Table 9. We adjusted the enrollments in primary school to include 50 percent of the reported secondary enrollments, and we reduced the secondary enrollments by 50 percent. This was done to fit the 2000 enrollment rates, which we interpolated. We assumed the primary school and secondary school ages are 6-13 and 14-17. We assumed that the initial stocks of primary, secondary and tertiary human capital are proportional to the Soviet Union values in 1989, where the proportional constants are the different enrollment rates in 1989/1990 between Tajikistan and the Soviet Union. The tertiary school enrollments are from the 1996 *UN Statistical Yearbook*, Table 9. The 2010 data come from *HDR*.

5.24 Yugoslavia (1910-2010)

Populations for 1910, 1920, 1930 and 1940 come from Maddison. Populations for 1950, 1960, 1970, 1980 and 1990 come from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1921, 1931, 1961, 1971, 1981, 1990 come from Meu Table A2 p. 44. The age distribution for 1910 is assumed to be identical to the age distribution in 1921. We interpolated using the 1931, 1948 and 1953 values in Meu Table A2 p. 44 to generate the 1941 and 1951 age distribution. Age distribution for Yugoslavia 2000 and 2010 come from the *Demographic Yearbook*. For 2000 and 2010 we used the age distribution for Serbia and Montenegra.

Labor force figures for 1921, 1931, 1961, 1971, 1981 come from Meu Table B1 p. 160. The labor force figures for 1910 comes from the same labor force participation rate, LF/population, as 1921 from Lampe and Jackson. We interpolated using the 1948 and 1953 values to generate the 1941 and 1951 labor force observations. The 1990 observation comes from HDR 1994. The labor force figures for 2000 come from *WDR*. The 2010 labor force figures come from assuming the same labor force participation rate as in 2000.

Real GNPs for 1910, 1921, 1931, 1941, 1950, 1960, 1970, 1980, 1990, 2000 and 2010 come from Maddison. The 1920-1952 investment rates are for Italy over the same period and the intraperiod average investment rate taken from Meu Table J1 pp. 921 and 926, for 1953-1993 and *WDR* (various years). We used the 1920 value for 1910. For 2010 we used the average investment rate for Serbia from 2000-2009 from S & H. The *WDI* provides sectoral output shares for 1990, 2000 and 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1910-1980. We applied the US 1910-1950 capital - sectoral output ratios to construct our capital estimates for 1910-1950. We used perpetual inventory for years 1960-2010.

Enrollments in primary and secondary schools for 1910-2000 come from Meu Table I1 p. 886 and 890. We assumed the primary school and secondary school ages are 6-9 and 10-17. The tertiary school enrollments are from Meu Table I2 pp. 898 and 899. For 2010 we used *HDR*.

6 Newly Industrialized Countries

6.1 Hong Kong (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1910, 1920, 1930, 1940, 1950, 1960 and 1970 come from Maa Table A1 p. 7. Populations for 1980 and 1990 come from S & H (1991). Population for 2000 comes from extrapolating DK population figure for 1993, using annual population growth rate from 1988-1993. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1961, 1970, 1980, 1990 come from personal internet correspondence with Ms. Vivian Chan, for the Commissioner for Census and Statistics. We used the 1961 age distribution for years 1820-1950. Age distribution for Hong Kong 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2013 age distribution. We adjusted by assuming that the age distribution was the same as in the reference year.

The labor force data for 1960 and 1970 come from *WDR* (various years). The labor force for 1980, 1990 and 2000 from personal correspondence with Ms. Vivian Chan, for the Commissioner for Census and Statistics. The labor force for 2010 comes from *WDI*. The labor force for 1820, 1830, 1840, 1850, 1860,

1870, 1880, 1890, 1900, 1910, 1920, 1930, 1940, 1950 come from the following procedure. We used *UN 1969* for urbanization rates for 1920, 1930, 1940, 1950, and *worldometers* for urbanization rates for 1960, 1970, 1980, 1990, and *WDI* for 2000 and 2010. We used urbanization rates of 51.9%, 53.8% and 55.8% in 1820, 1830 and 1840, respectively. These come from a regression of log urban share for years 1955-2010 on year, and forecasts of urbanization rate for 1820-1840. The stationary values of urbanization from 1920-1960 are ignored in this regression. The British take control of Hong Kong as a port for the opium trade in 1841. We then took an average of the predicted value of the regression, without the stationary values between 1920-1960, and the geometric interpolations from 1840 to 1920. We assumed a 50% rural 15-64 labor force participation rate for 1820-2010. We assumed an urban 15-64 labor force participation rate of 85.13% for 1820-1960, 78.35% for 1970, 75.1% for 1980, 71.74% for 1990, 65.33% for 2000 and 69.32% for 2010. We constructed the ratio of this labor force with the overlapping years from *WDR*, *WDI* and Ms. Vivian Chan, for 1960-2010. The root mean ratio is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000. The range of ratios is 1.000 to 1.000.

Real GNPs come from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1963-1975 from Maa Table J1, p. 1028 and from 1976-1998 from S & H online and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. The *WDI* provides sectoral output shares for 2000 and 2010. We used Sabillon (2005) information on manufacturing, services and aggregate growth rates to construct sectoral output shares for 1820-1990. We applied the US 1840-1960 capital - sectoral output ratios, using 1850 values for 1820, 1830 and 1840, to construct our capital estimates for 1820-1960. We used perpetual inventory for 1970-2010 values.

For 1950 we used data from UNESCO. Enrollment rates for 1960, 1970, 1980, 1990 and 2000 come from personal internet correspondence with Education Department Rm 1420 Wu Chung, 213 Queens Road East, Hong Kong. For 1890, 1900, 1910, 1920, 1930 and 1940 we used Benavot Riddle (1988). For years prior to 1890, we assumed that each enrollment rate was 90 percent of the succeeding decade's enrollment rate. For primary school enrollment rates, this produces the following time series: 8% (1820), 9% (1830), 10% (1840), 11% (1850), 12% (1860), 13% (1870), 15% (1880), 16% (1890 datum). For secondary school enrollment rates we used the same procedure but with 85%, this produces the following time series: 2% (1820), 3% (1830), 3% (1840), 4% (1850), 5% (1860), 5% (1870), 6% (1880), 8% (1890), 9% (1900), 10% (1910), 12% (1920), 14% (1930), 17% (1940) and 20% (1950 datum). For higher education enrollment rates, this produces the following time series: .06% (1820), .07% (1830), .08% (1840), .10% (1850), .12% (1860), .14% (1870), .16% (1880), .19% (1890), .22% (1900), .26% (1910), .31% (1920), .36% (1930), .42% (1940) and .5% (1950 datum). For 2010 we used government data on the internet.

6.2 Japan (1800-2010)

Populations for 1872, 1880, 1890, 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970, 1980 and 1990 come from Maa Table A1 p. 8. Population for 1800 and 1850 comes from McEvedy and Jones. The 1820, 1830 and 1840 values are interpolated between these two McEvedy and Jones data. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1890, 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970, 1980 and 1990 come from Maa Table A2 p. 22. The 1890 age distribution is interpolated from 1884 and 1893 values. The age distributions prior to 1890 are assumed to be identical to the age distribution in 1890. Age distribution for

Japan 2000 and 2010 come from the *Demographic Yearbook*.

Labor force figures for 1950, 1960, 1970, 1980 and 1990 come from Maa Table B1 p. 97. Labor force for 1880, 1890, 1900, 1910, 1920 come from Nakamura and data for 1930 and 1940 come from the Bank of Japan. The 1820 and 1872 values come from Maddison (1995). We interpolate for 1830, 1840, 1850 and 1860. For 1800 and 1810 we used the following procedure. We used Rozman for urban-rural share of population in 1800 through 1872, and Banks for 1880-1960, with adjustments to fit the share of the population in cities over 10,000. We then used *WDR* and *WDI* for 1970, 1980, 1990, 2000 and 2010. We assumed a rural 15-64 labor force participation rate of 90% for 1800-2010, and urban 15-64 labor force participation rate of 70% for 1800-1903, and 68% for 1913-2010. We constructed the ratio of this labor force to that from Maa, Nakamura, Bank of Japan and Maddison in the overlapping years of 1820-2010. The root mean ratio is .998, and the 1820 and 1830 values are .950 and .987, respectively. No ratio is below .95, and only one ratio is as large as 1.082. Labor force figure for Japan 2000 come from *WDR*. The labor force data for 2010 comes from *WDI*.

Real GNPs come from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1885-1998 from Maa Table J1, pp. 1025, 1026 and 1031 and *WDR* (various years). The 2010 value is the average investment rate from 2000-2009 from S & H. Mitchell provides sectoral output shares for 1880-2000. Saito and Takashima (2014) provide sectoral output shares for 1800-1872. We used Goldsmith (1985) for years 1872-1960. Piketty and Zucman (2014) provide our estimates for 1970-2000. We used perpetual inventory for 2010. For 1800-1860 we used US 1850-1860 capital - sectoral capital ratios, applying the 1850 value for 1800, 1810, 1820, 1830 and 1840, to produce our capital estimates.

For 1830, 1840 and 1850 we used Steckel and Floud. For 1800-1820 we assumed the same primary and secondary enrollment rates as in 1830. Enrollments in primary and secondary schools from 1873-1993 come from Maa Table I1 pp. 981, 983 and 987. These produce schooling exposure rates of 25 percent in 1870, similar to the 20 percent in Morrisson & Murtin. All of the exposure is to primary school in both our calculations and in Morrisson &Murtin. Our years of schooling estimate for 1870 is 1.5 compared with their value of 1.0. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1873-1993 come from Maa Table I2 pp. 1001 and 1004. For years prior to 1872 we assumed a rate of .05. For 2000 we used *HDR*. For 2010 we used *HDR*. Our time series of years of schooling in the labor force for 1872-1920 is: 1.51 (1872), 1.62 (1880), 2.21 (1890), 2.86 (1903), 3.97 (1913), 5.01 (1920). The Morrisson & Murtin time series of years of schooling is: .97 (1870), 1.13 (1880), 1.54 (1890), 2.18 (1900), 2.88 (1910), 3.86 (1920).

6.3 Singapore (1820-2010)

Population for 1820-1950 comes from Maddison. Population for 1963 is geometrically interpolated from 1957 and 1970 from Maa Table A1 p. 9. Populations for 1970 and 1980 come from Maa Table A1 p. 9. Population for 1990 comes from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1963, 1970, 1980, 1990 come from KF. We assumed the age distribution for 1820-1950 is the same as the age distribution in 1963. Age distribution for Singapore 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in the reference year.

The labor force data for 1950, for 1963 is interpolated from 1957, 1970, 1980 and 1989 come from Maa Table B1 p. 99. The 1947 labor force value comes the 1956 *UN Statistical Yearbook*. The labor force for Singapore 2000 comes from *WDR*. Labor force data for 2010 comes from *WDI*. For 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1913, 1920, 1930 we used the following procedure. We used Clio infra for urbanization rates for 1800, 1850, 1900 and 1950. We interpolated for 1820-1840, 1860-1890, 1910-1940. Singapore became 100% urbanized by 1900, but in 1820 it was only 7% urbanized. We assumed the rural and urban 15-64 labor force participation rates were identical through all years. For 1820-1940 we used the average 15-64 labor force participation rate for 1947-2010, which is 67.68%. We used 15-64 labor force participation rate of 61.15% for 1947, 73.2% for 1950, 52.1% for 1963, 53.36% for 1970, 71.25% for 1980, 89.63% for 1990, 67.67% for 2000 and 73.08% for 2010. We constructed the ratio of this labor force relative to that from Maa, and *UN Statistical Yearbook*. The root mean ratio for the overlapping years 1947-2010 is 1.000. The 1947 and 1950 ratios are 1.000, 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real output comes from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1960-1998 from Maa Table J1, p. 1036 and *WDR* (various years). For 2010 we used the average investment rates for 2000-2009 from S & H. Mitchell provides sectoral output shares for 1950-2000, and *WDI* provides sectoral output shares for 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1820-1940. We applied the US 1850-1950 capital - sectoral output ratios, applying the 1850 values for 1820, 1830 and 1840, in order to produce our capital estimates for 1820-1950. We used perpetual inventory for 1960-2010.

Enrollments in primary and secondary schools from 1947-1993 come from Maa Table I1 pp. 985 and 990. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1951-1993 come from Maa Table I2 p. 1005. For 1947 enrollment rates, we assumed them to be 90 percent of the 1950 enrollment rates. For years prior to 1947, we assumed that enrollment rates were 80 percent of the succeeding decade's enrollment rates. For primary school enrollment rates, this produces the following time series (by comparison we also list the primary enrollment rate for [Malaysia, Hong Kong & China] in the same year): 3% ([2%, 8%, 9%], 1820), 4% ([2%, 9%, 9%], 1830), 5% ([2%, 10%, 9%], 1840), 7% ([2%, 11%, 9%], 1850), 8% ([2%, 12%, 9%], 1860), 10% ([2%, 13%, 9%], 1870), 13% ([2%,15%, 9%], 1880), 16% ([5%, 16%, 9%], 1890), 20% ([4%, 12%, 9%], 1900), 25% ([9%, 11%, 9%], 1913), 49% ([10%, 17%, 11%], 1920), 31% ([21%, 26%, 14%], 1930), 48% ([30%, 23%, 23%], 1940) and 67% ([78%, 80%, 40%], 1950 datum). Our measure for primary school enrollments is consistent with the evidence from Turnbull (1989), *History of Singapore: 1819-1988*, which indicates the founding of the first higher education school in 1823. It is also similar to the primary enrollment rates in Malaysia. For secondary school enrollment rates, this produces the following time series: 1% (1820), 1% (1830), 1% (1840), 1% (1850), 1% (1860), 2% (1870), 2% (1880), 2% (1890), 2% (1900), 2% (1913), 3% (1920), 3% (1930), 3% (1940) and 4% (1950 datum). For higher education enrollment rates, this produces the following time series: .12% (1820), .13% (1830), .14% (1840), .16% (1850), .18% (1860), .20% (1870), .22% (1880), .24% (1890), .27% (1900), .30% (1910), .33% (1920), .37% (1930), .41% (1940) and .51% (1950 datum). For 2000 we used *WDR*. For 2010 we used data from the *UIS Global Database*.

6.4 (South) Korea (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1910 and 1920 come from Maddison. Populations for 1940 and 1950 are geometrically interpolated from 1930, 1944 and 1960 from Maa Table A2 p. 23. Populations for 1960, 1967, 1975, 1980 and 1990 come from Maa Table A2 p. 23. The 1990 value is interpolated from the 1980 and 1994 value. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions in 1820-1920 are extrapolated from the age distribution in 1930. The age distributions for 1940, 1950 are geometrically interpolated from 1930 and 1944 and 1960 from Maa Table A2 p. 23. The age distributions for South Korea for 1960, 1980 and 1990 come from Maa Table A2 p. 23. The 1967 and 1975 age distributions are interpolated. The age distribution for 1990 is interpolated from the 1980 and 1994 values. Age distribution for South Korea 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution. We adjust by assuming the same share by age category as in 2002.

Labor force figures for 1960, 1970, 1980 come from Maa Table B1 p. 97. We interpolated for labor force data in 1967 and 1975. Labor force figures for the South Korea 1990 and 2000 come from *WDR* (various years). Labor force data for 2010 come from *WDI*. For 1820-1950 we used the following procedure. We used Clio Infra for urbanization rates for 1820-1950, and *WDR* and *WDI* for 1960-2010. We interpolated between 1800 and 1850 for 1820-1840, and we interpolated between 1850 and 1900 for 1860-1890, and we interpolated between 1900 and 1950 for 1913-1940. We assumed a rural 15-64 labor force participation rate of 68% for 1820-2010. We assumed an urban 15-64 labor force participation rate of 30% for 1820-1975, 48% in 1980, 56% in 1990, 61% in 2000 and 2010. We constructed the ratio of this labor force to that from *WDI* and *WDR*. The root mean ratio for the overlapping years 1960-2010 is 1.013. The 1960 and 1967 values are 1.067 and 1.072. The range of the ratio is .983 to 1.072.

Real GNPs come from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1911-1998 from Maa Table J1, pp. 1025, 1026 and 1032 and *WDR* (various years). For 2010 we used the average investment rate from 2000-2009 from S & H. Mitchell provides sectoral output shares for 1950-2000, except for 1967. Our 2010 sectoral shares come from *WDI*. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1820-1940 and 1967. We applied the US 1840-1910 capital - sectoral output ratios, applying the 1850 values for 1820, 1830 and 1840, to produce our capital estimates for 1820-1910. We used perpetual inventory for 1920-2010.

Enrollments in primary and secondary schools from 1910-1938 for Korea Maa Table I1 p. 983. The primary school enrollment rate in 1910 is 2%. The secondary school enrollment rate in 1910 is .08%. In 1930 the higher education enrollment rate is .029%. For years prior to 1910, we assume a 1 percent enrollment rate in primary school, .05 percent enrollment rate in secondary school and .001 percent enrollment rate in higher education. This is consistent with Easterlin (1991). In 1910, Easterlin (1981) reports primary school enrollments of 27 per 10,000 population in Korea, compared with 1240 per 10,000 population in Japan. Our primary school enrollment rate for Japan in 1910 is 93%, so a proportionate Korean estimate would be 2%, which is what we use. If this proportion held in 1850, the 25% primary school enrollment rate in Japan would imply a .5% enrollment rate for Korea, which we round up to 1%. This is also consistent with Steckel and Floud's estimate of 25% literacy in Japan in 1830, which if the proportion held, implies a Korean literacy rate of .5% in 1830, which we rounded up to 1%. For South Korea enrollments in primary

and secondary schools from 1946-1993 come from Maa Table I1 p. 988. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1925-1993 come from Maa Table I2 pp. 1001 and 1004. We assumed 1 percent enrollment rate in 1900 for primary school, .005 percent enrollment rate for secondary school and .001 percent enrollment rate for higher education. For 2010 we used *HDR* for all enrollment rates. Our time series of years of schooling in the labor force for 1870-1950 is: .06 (1870-1910), .11 (1920), .38 (1930), 1.21 (1940), 2.46 (1950). The Morrisson & Murtin time series of years of schooling is: 1.11 (1870), 1.16 (1880), 1.21 (1890), 1.26 (1900), 1.31 (1910), 1.46 (1920), 1.62 (1930), 2.31 (1930), 3.10 (1950).

6.5 Taiwan (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1905, 1915, 1920, 1930, 1940, 1950, 1956, 1964, 1970 and 1980 come from Maa Table A2 p. 26. Population for 1990 comes from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1905, 1915, 1920, 1930, 1940, 1956, 1966, 1970 and 1980 come from Maa Table A2 p. 26. We interpolated for the age distributions in 1950, 1964. We assume that the age distribution prior to 1905 is identical to the age distribution in 1905. The age distribution for Taiwan for 1990 comes from DK. The age distributions for 2000 and 2010 come from *WDI*.

Labor force figures for 1905, 1915, 1920, 1930 1940 1954, 1964, 1970, 1980 come from Maa Table B1 p. 100. Labor force figures for Taiwan 1990 and 2000 come from *WDR* (various years). The labor force figures for Taiwan for 2010 comes from NationMaster, http://www.nationmaster.com/country-info/profiles/Taiwan/Labor. For years prior to 1905, we used the following procedure. Liao provides urbanization rates in Taiwan for 1900, 1920, 1940, 1958, 1972, 1980. We used these to interpolate for 1905, 1915, 1950, 1956, 1964, 1970. From WorldOMeters, we got urbanization rates for 1990, 2000 and 2010. From WorldOMeters, we also got the urbanization rates for China for 1950-2010. From Clio Infra we got urbanization rates for China in 1800, 1850 and 1900. The ratio of urbanization rates of Taiwan to China is about 2 for all overlapping years. This is just about the value of the 1900 overlapping observation. Thus we used the same value for the 19th century as for 1900. We assumed a rural 15-64 labor force participation rate of 76.5% for 1820-1930, 87.5% for 1940, 76.5% for 1950, 66.5% for 1956, 60% for 1964, 76.5% for 1970, 70% for 1980, 64% for 1990, 66% for 2000 and 70% for 2010. We assumed an urban 15-64 labor force participation rate of 50% for 1820-1930, 67% for 1940, 50% for 1950, 1956, 1964, 58.5% for 1970, 55.5% for 1980, 1990, 63.3% for 2000 and 69.3% for 2010. We constructed the ratio of this labor force with that from Maa, *WDR* and NationMaster. The root mean ratio in the overlapping years 1905-2010 is 1.004. The range of this ratio is .998 to 1.027.

Real GNPs come from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1904-1998 from Maa Table J1, pp. 1025, 1026 and 1037 and *WDR* (various years). We used the average investment rate for 2000-2009 from S & H for 2010. The WDR provides sectoral output shares for 1960 and 1978. We used Sabillon (2005) to construct our 1956-2000 sectoral output shares, and 1820-1950 sectoral output shares. We applied US 1850-1900 capital - sectoral output ratios, applying the 1850 values for 1820, 1830 and 1840, to construct our capital estimates for 1820-1900. We used perpetual inventory for years 1905-2010.

Enrollments in schools from 1905-1937 and for primary and secondary schools from 1940-1993 come from Maa Table I1 pp. 985 and 991. To calculate enrollment rates, we assumed 6-11 are primary school age

and 12-17 are secondary school age. For 1905, 1915, 1920 and 1930 we assumed that primary enrollments were the bulk of school enrollments. For years prior to 1905 we assumed that primary enrollment rates were 85 percent of the Chinese primary enrollment rate, which is constant at 9 percent. This ratio is the 1910 Taiwan/China enrollment ratio. For 1820-1860 we assumed a .01 percent secondary enrollment rate, and for 1870-1900, we assumed a .05 percent secondary enrollment rate. The 1905 observation for secondary school enrollment rate is .08%. For 1820-1860 higher education enrollment rates, we assume 0. For 1870-1900 we assumed a rate of .01 percent. Higher education enrollments for 1920-1993 come from Maa Table I2 pp. 1002 and 1005. The 1920 higher education enrollment rate observation is .18%. For 2000 we used WDR. For 2010 we used data from government internet site.

7 Asia

7.1 Afghanistan (1950-2010)

Population for 1950, 1960, 1970, 1980 and 1990 come from KF. Population for 2000 comes from *WDR*. Population for 2010 comes from *Wikipedia*.

The age distribution for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Afghanistan 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in 2012.

Labor force for 1960, 1970 come from *WDI*. Labor force for 1980, 1990, 2000 come from *nationmaster*. Labor force for 2010 come from CIA Factbook. For 1950 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for 1960-2010. We assumed rural 15-64 labor force participation rates of 73.48% for 1950-1960, 82.15% for 1970, 79.1% for 1980, 79.9% for 1990, 73.05% for 2000, 56.66% for 2010. We assumed urban 15-64 labor force participation rates of 50% for 1950-2010. We constructed the ratio of this labor force to that from *WDR*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

The 1950, 1960, 1970, 1980 1990 and 2000 real output data come from Maddison. The 2010 value comes from the 2008 value from Maddison and then using the real growth rate per capita from 2008-2010 using WDR 2011 and 2012. Physical capital investment rates from *WDI* for 1960-1979 are averaged to produce 1970 and 1980 investment rates. We used S & H for 1980-2009 for the 1990 and 2000 and 2010 observations. The *WDI* provides sectoral output shares for 2000 and 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1950-1990. We applied the US 1950 and 1960 capital - sectoral output ratios to construct our 1950 and 1960 estimates. We used perpetual inventory for 1970-2010.

For 1950 enrollment rates we used the 1954 *UN Statistical Yearbook*. For 1960 enrollment rates we used the 1963 *UN Statistical Yearbook*. We used enrollment rates from *WDR* for 1980, 1990. For 2010 we used *HDR*. We interpolated the enrollment rates in 1970 and 2000. To calculate enrollment rates, we assumed 6-13 are primary school age and 14-17 are secondary school age. Higher education enrollments for were essentially 0.

7.2 Bangladesh (1950-2010)

Population for 1950, 1960, 1970, 1980, 1990 come from KF. Population for 2000 comes from *WDR*. Population for 2010 comes from *Wikipedia*.

The age distribution for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Bangladesh 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in 2011.

Labor force figures for Bangladesh 1961, 1974, 1981 and 1991 come from Maa Table B1 p. 95. The 2000 and 2010 labor force data comes from *WDI*. For 1950 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for 1960-2010. We assumed rural 15-64 labor force participation rates of 93.93% for 1950-1961, 92.38% for 1970, 95% for 1980, 84.4% for 1990, 92.42% for 2000, 83.57% for 2010. We assumed urban 15-64 labor force participation rates of 50% for 1950, 1961 and 1970, 70.45% for 1980, 1990, 2000 and 2010. We constructed the ratio of this labor force to that from *WDR*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

The real output data come from Maddison. Physical capital investment rates for 2000 and 2010 come from the intraperiod average gross investment rates from S & H online and *WDR* (various years). Mitchell provides sectoral output shares for 1961-2000. We applied Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct 1950 sectoral output shares. We applied the US 1950 capital - sectoral capital output ratios to construct our 1950 estimate. We used Nehru and Dhareshwar (1993) capital output ratios to construct our 1961-1990 values. We used perpetual inventory to compute our 2000 and 2010 values.

Enrollments in primary and secondary schools from 1947-1992 come from Maa Table I1 pp. 982 and 986. To calculate enrollment rates, we assumed 6-10 are primary school age and 11-17 are secondary school age. Higher education enrollments for 1947-1993 come from Maa Table I2 p. 1003. For 2010 we used *HDR*. We interpolated for our 2000 observations. Our time series of years of schooling in the labor force for 1950-2010 is: 1.64 (1950), 2.18 (1961), 3.07 (1970), 3.22 (1980), 4.05 (1990), 4.69 (2000) and 5.60 (2010). The Morrisson & Murtin time series of years of schooling is: 1.12 (1950), 1.23 (1960), 1.70 (1970), 2.26 (1980), 3.16 (1990), 4.42 (2000) and 4.69 (2010).

7.3 Bhutan (1980-2010)

Population for 1980, 1990 come from KF. Population for 2000 comes from *WDR*. Population for 2010 comes from *Wikipedia*.

The age distribution for 1980, 1990 come from KF. Age distribution for Bhutan 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2001 age distribution, and for 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1980, 1990, 2000 and 2010 come from *WDI*.

We used the 2007 value comes from the World Bank rebenchmarking project. We then applied the real per capita growth rate using *WDR 2010, 2011, 2012*. We used the *WDI* growth rates and the 2006 base level to produce estimates for 2000, 1990, and 1980. We assumed that the most accurate estimates come from 2006, and that the growth rates from *WDI* are more accurate than original *WDI* or WDR levels. These values are quite similar, nonetheless. The 1980 value is 116 percent of the original *WDI*. The 1990

value is 133 percent of the original *WDI*. The 2000 value is 184 percent of the WDR. Physical capital investment rates for 1990, 2000 and 2010 come from the intraperiod average gross investment rates from S & H online and *WDI* (various years). The *WDI* provides sectoral output shares for all years. We applied the US 1980 capital - sectoral output ratios to produce our 1980 estimate. We used perpetual inventory for 1990, 2000 and 2010.

Enrollments in primary and secondary schools from 1980 and 1990 come from *WDI*. The 2000 and 2010 data come from UNESCO's *UIS Global Database*. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-15 are secondary school age

7.4 Cambodia (1950-2010)

Population for 1950, 1960, 1970, 1980 come from *WDR* (various years). Population for 1990 comes from *Population Pyramid*, https://populationpyramid.net/cambodia/1990/, accessed July 24, 2016. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Cambodia 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we interpolated the 1998 and 2002 age distributions, and for 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force data for 1960 come from *WDI*. The labor force data for 1970 is an interpolation of data for 1960 and 1980. Labor force figures for 1990 come from *HDR*. Labor force figures for 1980 and 2000 come from *WDR*. The labor force data for 2010 come from *WDI*. For 1950 we used Banks for urban-rural population shares for 1950 and 1960, assuming his 1953 value was the same as in 1950. The 1953-1956 values are all constant. We used *WDR* and *WDI* for 1970, 1990, 2000 and 2010. We interpolated for the 1980 value, even though the Khmer Rouge imposed a sweeping de urbanization of the country. The 1970 and 1990 values are both essentially 12%. We assumed a rural 15-64 labor force participation rate of 95% for 1950, 1960 and 1970, 77% for 1980 and 1990, 87.5% for 2000 and 2010. we assumed an urban 15-64 labor force participation rate of 65% for 1950-2010. We constructed the ratio of this labor force with that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is .993. The 1960 and 1970 ratios are .951 and 1.007, respectively. All ratios are between .951 and 1.011.

Real output come from Maddison. The 1970, 1980, 1990 and 2000 investment rates are from *WDI*. The 2010 value is the average investment rate for 2000-2009 from S & H. The 1961 UN Statistical Yearbook provides sectoral output shares for 1950 and 1960. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1970, 1980, 1990, 2000 and 2010. We applied the US 1950 and 1960 capital - sectoral output ratios to produce our 1950 and 1960 estimates. We used perpetual inventory for 1970-2010.

Enrollment rates in 1950 come from Mitchell. Enrollment rates in primary and secondary schools for 1960-1990 come from *WDI* (various years). The 2000 values come from UNESCO's *UIS Global Database*, and the 2010 values come from *HDR*. We assume primary school and secondary school ages are 6-11, 12-18. Tertiary enrollment rates are from *WDR* (various years).

7.5 China (1820-2010)

for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1913, 1920, 1930, 1940, 1950, 1960, 1970, 1982, 1990 come from Maddison (1998). Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1953, 1982, and 1990 comes from Maa Table A2 p. 19. The age distribution for all years prior to 1953 are assumed to be identical to the age distribution in 1953. The age distributions for 1960 and 1970 come from KF. Age distribution for China 2000 and 2010 come from the *Demographic Yearbook*.

The labor force data for 1933 (Table 3.13) and 1953, 1960, 1970, 1980, 1990 (Table 3.17) come from Maddison (1998). The labor force data for 1970 and 1982 come from *WDR* (various years). The labor force for 2000 comes from the WDR. The 2010 labor force data comes from the *WDI*. For 1820 we used the 1800 value from Moll-Murata, adjusted for population, but reducing the agriculture labor force participation rate from 100% of population 15-60 to 70%. We do this in order to make the data more consistent with the data from Maddison. For years 1830-1850 we increased the labor force by the annualized growth rate of population from 1820-1850. From 1850 to 1870 the population of China declines from 412 million to 358 million. For 1860 and 1870 we reduced the labor force by the annualized decline rate over this period. For 1880-1920, we increased the labor force by the annualized growth rate of population over this period.

Real GNPs come from Maddison. Physical capital investment rates from 1960-1998 come from S & H online at NBER web site and *WDR* (various years). Physical capital investment rates for 1890-1960 come from Mn2 Table 3.9, p. 64. We interpolated between our 1930 and 1950 values for 1940. For 2010 we used the average investment rate for 2000-2009 from S & H. Broadberry, Guan, and Li (2014) provide estimates of agriculture's share of GDP in 1820 and 1830. Zhihong, Yuping , van Leeuwen (2015) provide data to estimate agriculture, industry and service share of GDP in 1840, 1850, 1860, 1870, 1880. Their 1880 value is dearly identical to that in Richardson. Maddison (1998) (Table 2.5) provides sectoral output shares for 1890-1952. The *WDI* provides sectoral output shares for 1990, 2000 and 2010. We applied the US capital - sectoral output ratios for 1840-1950, using 1840 values for 1820 and 1830, to produce our capital estimates. We used perpetual inventory for 1960-2010.

Enrollment rates for the 1930s, 1953, 1982 and 1990 come from Mn2 Table 3.7, p. 63 and the age distribution of the population. Enrollment rates for 1960 and 1970 are from the *WDR* various issues. For 1850-1913 we used Morris and Adelman. For years prior to 1850 we assumed the same enrollment rate as in 1850. Higher education enrollments for 1946-1993 come from Maa Table I2 p. 1003. Prior to 1946 enrollments come from Mn2 Table 3.7, p. 63. For 2000 we used *WDR*, and for 2010 we used *WDI*. Our time series for labor force share exposed to education is consistent with Morrisson & Murtin from 1870-2010. We produce an estimated 1.04 years of schooling in the labor force in 1870, compared with Morrisson & Murtin's estimate of 1.01. For 1870-1920, we produce a time series of years of schooling in the labor force of 1.04 (1870), 1.05 (1880), 1.05 (1890), 1.05 (1900), 1.06 (1913), 1.05 (1920). The Morrisson & Murtin time series for these years is 1.01 (1870), 1.06 (1880), 1.11 (1890), 1.16 (1900), 1.21 (1910), 1.26 (1920).

7.6 Fiji (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1960, 1970, 1980, 1990 come from KF. The age distribution for 1960 was used for 1950. Age distribution for Fiji 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2008 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1990 come from *HDR*. Labor force figures for 1960, 1970, 1980 and 2000 come from *WDR* (various years). For 1950 we used the average labor force/population ratio for 1960-2010. The labor force figures for 2010 comes from *WDI*.

Real GNPs for all years come from Maddison except for 2010. We used the World Bank PPP real growth rate of per capita income from 2000 to 2010 on our 2000 base to produce our 2010 estimate. The 1960-2009 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1960-2010. We used the Oceania (Australia, New Zealand and Papua New Guinea) relative growth rates of farming and manufacturing compared with aggregate growth rates to produce the 1950 sectoral shares. We applied the US 1950 capital - sectoral output ratios to produce our 1950 capital estimate. We used perpetual inventory for 1960-2010 for the rest of our capital estimates.

Enrollments in primary and secondary schools for 1960-2000 come from Maa Table I1 p. 993. For 1950 we used *UN Statistical Yearbook 1953*. We assumed the primary school and secondary school ages are 6-11 and 12-17. The tertiary school enrollments are from *WDR* (various years). The 2010 primary school and tertiary school enrollment rate data come from *HDR*. The secondary school enrollment rate comes from UNESCO's *UIS Global Database*.

7.7 India (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1890, 1901, 1911, 1921, 1931, 1951, 1961, 1971 and 1981 come from Maa Table A2 pp. 19 and 20. Population for 1990 comes from S & H online. We interpolated for 1880 and 1941. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1890, 1901, 1911, 1921, 1931, 1951, 1961, 1971, 1981 and 1990 come from Maa Table A2 pp. 19 and 20. We assumed the age distribution for 1820, 1830, 1840, 1850, 1860, 1870 and 1880 to be identical to the age distribution in 1890. We interpolated for 1941. The age distribution for India for 1990 is interpolated from 1981 and 1993 values. Age distribution for India 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2001 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1901, 1911, 1921, 1931, 1951, 1961, 1971, 1981 and 1990 come from Maa Table B1 p. 95. Labor force for 1860, 1880, and 1890 come from *Cambridge Economic History of India*. The 1870 value is interpolated. We used Federico (2010) for agricultural labor force for 1800, 1850, 1880, 1910. We interpolated between these years to construct a time series for agricultural labor force for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900 and 1910. We used the time series on labor force for 1860-1910 to construct a share of labor force that is agricultural. The average for these years is 62.8%, with almost no variation. The range of shares is between 61.1% to 64.0%. We then applied this average to construct labor force for 1820, 1830, 1840 and 1850. We assumed the labor force for 1820, 1830, 1840, 1850 come from the average labor force participation rate, LF/population, for 1860, 1880-1911. Labor force figures for India for 1990 is interpolated from 1981 and 1991 values. Labor force figures for 2000 comes from *WDR*. The

labor force figures for 2010 comes from *WDI*.

Real output comes from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1950-1993 from Maa Table J1, p. 1028 and *WDR* (various years). For 2000 and 2010 we used intraperiod averages from S & H. Mitchell provides sectoral output shares for 1951-2000. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1850-1940. We used Broadberry, Custodis and Gupta (2015) for sectoral output shares in 1820, 1830 and 1840. The only adjustment we made was to farming. Our estimated share of manufacturing in these years agrees with Broadberry, Custodis and Gupta. However, instead of a share of farming to 65% in 1820 we assumed that it was 70%. This compares with 67.5% in 1870. We chose this assuming that farming is the least subject to measurement error, and that there was quite a bit of change in export industry activity between 1600-1870. We applied the US capital - sectoral output ratios for 1850-1940, applying 1850 to 1820, 1830 and 1840. We used perpetual inventory for 1951-2010.

Enrollments in primary and secondary schools for 1850, 1870, 1890, 1901 are from Lindert. Enrollments in primary and secondary schools from 1911-1993 come from Maa Table I1 pp. 980, 982 and 986. For years prior to 1850 we used the 1850 primary enrollment rate in primary school (.36%), and the 1850 secondary enrollment rate (.16%) in secondary school, and 0, 0 and .005 percent for higher education enrollment rates in 1820, 1830 and 1840, respectively. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1873-1986 come from Maa Table I2 pp. 1001 and 1003. For 2000 we interpolated all of our enrollment rates, and for 2010 we used *HDR*. Our time series for labor force share exposed to education is consistent with Morrisson & Murtin from 1870-2010. Our time series of years of schooling for 1870-1931 is .05 (1870), .09 (1880), .17 (1890), .34 (1901), .46 (1911), .62 (1921), .78 (1931). The Morrisson & Murtin time series for these years is: .08 (1870), .11 (1880), .15 (1890), .29 (1900), .37 (1910), .51 (1920), .70 (1930).

7.8 Indonesia (1820-2010)

Population for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1910, 1920 come from Maddison. Populations for 1930, 1940, 1951, 1961, 1971 and 1980 come from Maa Table A1 p. 7. The population for 1951 is interpolated from the 1930 and 1961 values. Population for 1990 from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1961, 1971 and 1980 come from Maa Table A2 p.20. The age distribution for all years prior to 1961 are assumed to be identical with the age distribution of 1961. The age distribution for Indonesia for 1990 comes from DK. Age distribution for Indonesia 2000 and 2010 come from the *Demographic Yearbook*.

Labor force for 1820-1920 come through the following procedure. We use van Zanden (2003) for agricultural labor force share in 1880. We used *WDI* and *WDR* for labor force share for 1960, 1970, 1980, 1990, 2000 and 2010. The labor force share in 1960 is 75%, so we assumed 76% labor force share for 1890-1950. We used De Zwalt, Luiten and van Zanden (2015) for labor force share in 1820. We interpolated between 1820 and 1880. We then took the population between 15-64 and assumed that 65% is the labor force participation rate in agriculture and 50% is the labor force participation rate for non agricultural population. This method almost exactly fits the data for 1930-1960. Labor force figures for 1930, 1940, 1951 come from Christensen (1966). Labor force figures for 1961, 1971 and 1980 comes from Maa Table B1 p. 96. Labor

force figures for Indonesia for 1990 and 2000 come from *WDR* (various years). Labor force figures for 2010 come from *WDI*.

Real output comes from Maddison. Physical capital investment rates come from the intraperiod average gross capital formation and income for 1951-1957 and the intraperiod average gross real capital formation and real income for 1958-1998 from Maa Table J1, p. 1029 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. Mitchell provides sectoral output shares for 1971-2000. The *WDI* provides sectoral output shares for 1961. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1820-1951. Our data is consistent with van Zanden. In 1880 we have sectoral output shares of 66%, 7% and 27% in agriculture, manufacturing and services. The same year from van Zanden has sectoral output shares of 64%, 9% and 27%, respectively. In 1927, van Zanden's data are 52%, 18%, 30% for agriculture, industry and services. Interpolating our data produces sectoral output shares of 57%, 8% and 35%. So the major difference is in industry. However the *WDI* 1961 data is much closer to our estimate of industry, 9% compared with our 9% 1950 value and 8% 1930 value compared with van Zanden's 18% for 1927. Thus we choose to stay with Sabillon. We used the 1850-1870 US capital - sectoral output ratios, applying the 1850 US values to 1820, 1830 and 1840, to produce capital estimates. We used Pierre van der Eng (2010) to produce estimates of capital for 1880-2000. Pierre van der Eng (2010) provides capital without residential capital. We applied the US 1880-2000 capital - sectoral output ratios to the sectoral shares and averaged these with Pierre van der Eng's estimates to produce our 1880-2000 capital estimates. We used perpetual inventory for 2010.

Enrollments in primary and secondary schools from 1866-1938, 1954-1992 come from Maa Table I1 pp. 980, 982 and 987. We used the 1870 enrollment rates for all year prior to 1870. This is .54% for primary school enrollment rates, and .003% for secondary school enrollment rates. The 1870 exposure share is 1 percent, which matches Morrisson & Murtin. We assumed 0 (1920 datum) higher education enrollment rates prior to 1920. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1920-1993 come from Maa Table I2 pp. 1001 and 1003. For 2010 we used *HDR*. For 2000 we interpolated. Our time series share of labor force years of schooling for 1870-1930 is consistent with Morrisson & Murtin. Our years of schooling time series for these years is: .03 (1870), .04 (1880), .05 (1890), .10 (1900), .14 (1910), .24 (1920), .50 (1930). The time series produced by Morrisson & Murtin is: .05 (1870), .04 (1880), .04 (1890), .05 (1900), .08 (1910), .13 (1920), .25 (1930).

7.9 Laos (1950-2010)

Populations for 1950, 1960 come from *WDI*. Population for 1970 comes from Maddison. Population for 1980 comes from *WDR*. Population for 1990 comes from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Laos 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in 2013.

The labor force data for 1960 comes from the *WDI*. Labor force figures for 1970, 1980 1990 come from *HDR*. Labor force figures for 2000 come from *WDR*. Labor force data for 2010 comes from *WDI*. For 1950 we used Clio infa for urban-rural population shares for 1950 and *WDR* for 1960-2010. We assumed a rural 15-64 labor force participation rate of 84.83% for 1950, 1960, 86.6% for 1970, 83.65% for 1980, 86.94%

for 1990, 95% for 2000 and 2010. We assumed an urban 15-64 labor force participation rate of 50% for 1950-1990, 58% for 2000 and 60.75% for 2010. We constructed the ratio of this labor force with that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively. All ratios are between 1.000 and 1.001.

Real output comes from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. The 1990 investment rate comes from the intraperiod average investment rate from S & H online. The 2000 investment rate is the intraperiod average investment rates taken from *WDR* (various years). The 2010 investment rate is the 2000-2009 average investment rate from S & H. The *WDI* provides sectoral output shares for 1990, 2000 and 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1950, 1960, 1970 and 1980. We applied the US 1950-1980 capital - sectoral output ratios to produce our capital estimates. We used perpetual inventory for 1990, 2000 and 2010.

The 1950 enrollment rates are from Mitchell. Enrollment rates in primary and secondary schools for 1960-1990 come from *WDI* (various years). The 2000 enrollment rates come from UNESCO's *UIS Global database*. The 2010 primary and tertiary enrollment rates come from *HDR*. The secondary school enrollment rate for 2010 comes from *WDR*. We assumed that primary school and secondary school ages are 6-10 and 11-16. Tertiary school enrollment rates are from *WDR* (various years).

7.10 Malaysia (1820-2010)

Population for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1913, 1920, 1930, 1940, 1950 comes from Mitchell. Populations for 1960, 1970, 1980 and 1990 come from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. We assumed the age distributions for all years prior to 1950 were the same as in 1950. Age distribution for Malaysia 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1990 come from *HDR*. Labor force figures for 1960, 1970, 1980 and 2000 come from *WDR* (various years). The 1950 labor force figures come from Mitchell. The 2010 labor force comes from *WDI*. For 1820-1940 we used the following procedure. We used Clio Infra for urbanization rates for 1820-1950, and *WDR* and *WDI* for 1960-2010. We interpolated between 1800 and 1850 for 1820-1840, and we interpolated between 1850 and 1900 for 1860-1890, and we interpolated between 1900 and 1950 for 1910-1940. We assumed a rural 15-64 labor force participation rate of 77% for 1820-2010. We assumed an urban 15-64 labor force participation rate of 53% for 1820-2010. We constructed the ratio of this labor force to that from *WDI* and *WDR*. The root mean ratio for the overlapping years 1950-2010 is 1.019. The 1950 and 1960 values are 0.986 and 1.062. The range of the ratio is .962 to 1.073.

Real output come from Maddison. The 1950-2000 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The 2010 investment rate is the average 2000-2009 investment rates from S & H. The *WDI* provides sectoral output shares for 1960-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate output growth to produce sectoral output shares for 1820-1950. We applied the US 1850-1940 capital - sectoral output ratios, applying the 1850 US values for 1820, 1830 and 1840, to produce our capital estimates for 1820-1940. We used perpetual

inventory for 1950-2010.

Enrollments in primary and secondary schools for 1864-1990 come from Maa Table I1 pp. 980, 981 and 988. We used the enrollments for 1864 for the 1860 observation. For all years prior to 1860 we assumed that each decade's enrollment rates were equal to the average enrollment rate for 1860, 1870 and 1880 (2.17%), because they were all essentially 2%. This produces a 2 percent exposure rate to education, which matches Morrisson & Murtin. It is mostly primary school exposure, 1.8 percent and only .3 percent secondary school exposure, as opposed to Morrisson & Murtin values of 1 percent for each. For secondary school enrollment rates, we assumed that for years prior to 1860 they were equal to the average secondary enrollment rate for 1860, 1870 and 1880 (0.28%), because they were all essentially 0.3%. For higher education enrollment rates, this produces a time series of: .0002% (1820), .0003% (1830), .0004% (1840), .0006% (1850), .0007% (1860), .0010% (1870), .0013% (1880), .0018% (1890), .0023% (1900), .0031% (1913), .0044% (1920), .0071% (1930), .011% (1940), .019% (1950), .077% (1960 datum). We assumed the primary school and secondary school ages are 6-11 and 12-18. The tertiary school enrollments are from Maa Table I2 p. 1004. The 2010 rates for all three categories come from *HDR*. We interpolated for our 2000 enrollment rates. Our time series share of labor force exposed to higher education is consistent with Morrisson & Murtin 1870-2010. Our time series for years of schooling in the labor force for 1870-1920 is: .14 (1870), .14 (1880), .15 (1890), .22 (1900), .24 (1913), .41 (1920). The Morrison & Murtin schooling time series for these years is: .11 (1870), .12 (1880), .13 (1890), .20 (1900), .24 (1910), .37 (1920).

7.11 Mongolia (1950-2010)

Population for 1950, 1960, 1970, 1980 come from KF. Population for 1990 comes from *Population Pyramid*, https://populationpyramid.net/mongolia/1990/, accessed July 24, 2016. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Mongolia 2000 and 2010 come from the *Demographic Yearbook*.

Labor force data for 1960, 1970 and 2010 come from *WDI*. The labor force data for 1980 comes from *WDR*. The 2000 labor force comes from CIA Factbook. Labor force figures for 1990 is interpolated from the 1980 and 2000 data. For 1950 we used the following procedure. We used Banks for urban-rural population shares for 1950, 1960, and *WDI* for 1970, 1980, 1990, 2000 and 2010. We assumed a rural 15-64 labor force participation rate of 93.5% for 1950-1980, and 76.7% for 1990, 80% for 2000, 78% for 2010. We assumed an urban 15-64 labor force participation rate of 75% for 1950-1980, and 50% for 1990 and 54% for 2000-2010. We constructed the ratio of this labor force to that from WDR and *WDI*. The root mean ratio is 1.000 for the overlapping years 1960-2010. The 1960 and 1970 ratios are 1.043, and .986, respectively. This is also the range of values.

Real output come from Maddison. We took the Maddison 2008 value and then applied the real PPP per capita income growth rate from *WDR 2011 & 2012*. The 1980, 1990 and 2000 investment rates are from *WDI*. The 2010 value is the average investment rate for 2000-2009 from S & H. The *WDI* provides sectoral output shares for 1981, 1990, 2000 and 2010. We used Sabillon (2005) to produce the 1980 value. In addition we used information on farming, manufacturing and aggregate output growth rates to produce sectoral output shares for 1950, 1960, and 1970. We applied the US 1950-1970 capital - sectoral output ratios to produce our 1950-1970 estimates. We used perpetual inventory to produce our 1980, 1990, 2000 and 2010 values.

Enrollment rates in 1950 come from Banks. Enrollment rates in primary and secondary schools and higher education for 1960-1990 come from *WDI* (various years). Our 2010 data come from *HDR*. We interpolated for our 2000 enrollment rate observations. We assume primary school and secondary school ages are 6-10, 11-17.

7.12 Myanmar (1820-2010)

Population for 1820, 1830, 1840, 1850 come from McEvedy and Jones. Populations for 1870, 1881, 1891, 1901, 1911, 1921 and 1931 come from Mitchell. Populations for 1941, 1951, 1961 come from Maddison. Population for 1973 and 1983 come from Maa Table A1 p. 9. Population for 1990 comes from Maa Table A5 p. 63. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1881, 1891, 1901, 1911, 1921, 1931, 1941, 1951, 1961, 1973 and 1983 come from Maa Table A2 p. 24. The 1941, 1951 values are interpolated from the 1931 age distribution and the 1961 age distribution. The age distributions for 1820, 1830, 1840, 1850, 1860, 1870 are assumed to be the same as the age distribution in 1881. Age distribution for Myanmar 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in 2012.

Labor force figures for 1983 come from Maa Table B1 p. 98. The labor force for 1961 and 1973 and 2000 are from *WDR* (various years). The 2010 labor force data come from *WDI*. For the labor force series 1820-1951 we used the following procedure. We used Clio Infra for urbanization rates for 1820-1950, and *WDR* and *WDI* for 1960-2010. We interpolated between 1800 and 1850 for 1820-1840, and we interpolated between 1850 and 1900 for 1860-1890, and we interpolated between 1900 and 1950 for 1910-1940. We assumed a rural 15-64 labor force participation rate of 82.5% for 1820-1961, 90% for 1973, 92% for 1983, 100% for 1990, 85% for 2000 and 72.5% for 2010. We assumed an urban 15-64 labor force participation rate of 50% for 1820-1961, 76% for 1973, 90% for 1983, 100% for 1990, 75% for 2000 and 63.5% for 2010. We constructed the ratio of this labor force to that from Maa, *WDI* and *WDR*. The root mean ratio for the overlapping years 1961-2010 is .997. The 1961 and 1973 values are 1.030 and .997, respectively. The range of the ratio is .984 to 1.001.[24]

Real output come from Maddison. Physical capital investment rates come from the intraperiod average gross capital formation and income for 1947-1962 and the intraperiod average gross real capital formation and real income for 1963-1998 from Maa Table J1, pp. 1033 and *WDR* (various years). For 2010 we used the average investment rate from 2000-2009 from S & H. We used *WDI* for sectoral output shares in 1973-2000. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce 1820-1961 sectoral output shares. We used the US 1850-1941 capital - sectoral output ratios, applying the 1850 US values on 1820, 1830 and 1840. This produces our capital estimates for 1820-1941. We used perpetual inventory for 1951-2010.

Enrollments in primary and secondary schools from 1860-1993 come from Maa Table I1 pp. 981, 984 and 988. For 1820-1850 we used the average primary, secondary and higher education enrollment rates in 1860 and 1870. These are 1.24% primary school enrollment rate, .08% secondary school enrollment rate and 0% higher education enrollment rate. This produces a 1 percent exposure to schooling in 1870, identical

[24]For 1881, 1891, 1901, 1911, 1921, 1931 we combined the index values of land/labor ratios from Williamson (2000), with the land data from Mitchell to produce our labor force. Williamson indexes 1911 as 1. Our data also produces landl/labor ratios and we also index 1911 at 1. His index value series is 1.019, .924, .976, 1, .983, .992. Our index value series is 1.126, .937, .972, 1, .969, .980. Thus our series and his are very correlated.

with Morrisson & Murtin. To calculate enrollment rates, we assumed 6-10 are primary school age and 11-16 are secondary school age. Higher education enrollments for 1894-1993 come from Maa Table I2 pp. 1001, 1002 and 1004. For 2010 we used *HDR*. WE interpolated for our enrollment rates in 2000. Our time series share of labor force exposed to higher education is consistent with Morrisson & Murtin 1870-2010. Our years of schooling in the labor force time series for 1870-1921 is: .48 (1870), .49 (1881), .56 (1891), .72 (1901), .93 (1911) and 1.18 (1921). The Morrisson & Murtin time series of schooling in the labor force for these years is: .03 (1870), .04 (1880), .20 (1890), .44 (1900), .67 (1910), .89 (1920).

7.13 Nepal (1820-2010)

Populations for 1850 and 1900 come from McEvedy and Jones. The 1820, 1830, 1840 values come from interpolation of McEvedy and Jones 1800 and 1850 values. Our 1860-1890 population values are interpolations of McEvedy and Jones 1850 and 1900 values. For years between 1900 and 1950 we interpolated using McEvedy and Jones datum for 1900 and Maddison's population for 1950. Populations for 1960, 1970, 1980, 1986 come from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1986 come from KF. Age distribution for Nepal 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2001 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year. For all years prior to 1950, we assumed the same age distribution as 1950.

Labor force figures for 1990 come from *HDR*. Labor force figures for 1960, 1970, 1980, 1986 and 2000 come from *WDR* (various years). Labor force data for 1950 and all years prior assumes the same labor force participation rate as in 1960, but this 100% of the 15-64 population in rural and urban populations. This 100% rate is true not just for 1960, but for 1970, and only drops to 96% in 1980 and 1990! The 2010 labor force data come from *WDI*.

Real output come from Maddison, except for 2010. Maddison reports the 2008 value, and we used the PPP real per capita income growth rate from *WDR 2011* and *WDR 2012* to produce the 2010 observation. The 1970-2000 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The 2010 investment rate is the average investment rate for 2000-2009 from S & H. The *WDI* provides sectoral output shares for 1960-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1820-1950. We applied the US 1840-1960 capital - sectoral output ratios, applying the 1850 values for 1820, 1830 and 1840. This produces our estimates for 1820-1960 capital stocks. We used perpetual inventory to produce 1970-2010 capital stocks.

Enrollments in primary and secondary schools for 1950-1990 come from Maa Table I1 p. 989. We assumed the primary school and secondary school ages are 6-10 and 11-15. The tertiary school enrollments are from *WDR* (various years). The 1950 values for primary, secondary and higher education were: .88%, .18% and .1%, respectively. For years 1910-1940 we used the 1950 enrollment rates. For all years 1820-1900 we assumed primary, secondary and higher education enrollment rates of: .725%, .18% and .025%, respectively. We make these assumptions on enrollment rates over the 1820-1940 period in order to match the 95-99% illiteracy rates of adults in Nepal for 1950, see Table 7 of UNESCO *World Illiteracy at Mid-Century*. The 2010 rates come from *HDR*. We interpolated for our 2000 enrollment rates.

7.14 North Korea (1820-2010)

Population for 1950, 1960, 1970, 1980 and 1990 come from KF. Population prior to 1950 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. We assumed that the age distribution prior to 1950 were identical to the 1950 age distribution. Age distribution for North Korea 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2008 age distribution. We adjust by assuming the same share by age category as in 2008.

Labor force data for 1960, 1970, 1980, 1990 and 2010 come from *WDI*. The 2000 labor force comes from *CIA Factbook*. Labor force figures for 1820-1950 come from the following procedure. We used Clio Infra for urbanization rates for 1820-1950, and *WDR* and *WDI* for 1960-2010. We interpolated between 1800 and 1950 for 1820-1940. We assumed a rural 15-64 labor force participation rate of 85% for 1820-1980, 94.5% for 1990, 87.5% for 2000, 84.25% for 2010. We assumed an urban 15-64 labor force participation rate of 71.5% for 1820-1960, 62% for 1970, 68% for 1980, 94.5% for 1990, 87.5% for 2000 and 84.25% for 2010. We constructed the ratio of this labor force to that from *WDI* and *WDR*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 values are 1.000 and .999, respectively. The range of the ratio is .999 to 1.001.

Real output come from Maddison, except for 2010. We used Maddison 2008 value updated to include the real PPP per capita income growth rate for 2008-2010 to produce our observation. We used *WDR 2011* and *WDR 2012* to get the real PPP per capita income growth rate. For 1920, 1930 and 1940 we used Mitchell. For 2010, when we used the average from *WDI*. We used the 1820 sectoral output shares from South Korea in 1820, but with 5% point lower farming share, and 5% point higher manufacturing share. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produces sectoral output shares for 1830-2000. We applied the US 1850-2000 capital - sectoral output ratios, applying the 1850 values for 1820, 1830 and 1840, to produce capital estimates. We used perpetual inventory to produce our 2010 value.

Enrollment rates in 1950 and 1960 come from Mitchell. WDR was used to provide enrollment rates in 1970, 1980 and 1990. We assumed 90 percent enrollment rates in primary school in 2000 and 2010, and 90 percent enrollment rates in secondary school for the same years. We assumed 7 percent higher education enrollment rates for these years as well. For 1910, 1920, 1930 and 1935-40 primary enrollment rates we used Benavot Riddle (1988). For 1910 they provide 1.1% primary enrollment rate. We assumed that the 1820-1900 primary enrollment rate values are constant at 1%. Again this is consistent with Table 5.1 of Easterlin (1998). Prior to 1950 we assumed that each decade's secondary enrollment rates were 60 percent of the succeeding decade's enrollment rate. This produced the secondary school enrollment rate time series is: .04% (1820), .07% (1830), .11% (1840), .18% (1850), .31% (1860), .51% (1870), .85% (1880), 1.4% (1890), 2.4% (1900), 3.9% (1913), 6.5% (1920), 11% (1930), 18% (1940) and 30% (1950 datum). For higher education enrollment rates, the time series is: .0010% (1820), .0016% (1830), .0027% (1840), .0045% (1850), .0075% (1860), .012% (1870), .021% (1880), .035% (1890), .058% (1900), .096% (1913), .16% (1920), .27% (1930), .45% (1940) and .74% (1950 datum). We assume primary school and secondary school ages are 6-9, 10-15.

7.15 Pakistan (1950-2010)

Populations for 1950, 1961, 1972 and 1981 come from Maa Table A2 p. 24. Population for 1990 comes from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1961, 1972 and 1981 comes from Maa Table A2 p. 24. The age distribution for Pakistan 1990 comes from DK. Age distribution for Pakistan 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 1998 age distribution, and for 2010 we adjusted the 2007 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1950, 1961, 1972, 1981 come from Maa Table B1 p. 98. Labor force data for Pakistan 1990 and 2000 come from *WDR* (various years). The labor force data for 2010 comes from *WDI*.

Real output come from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1959-1998 from S & H online. Mitchell provides sectoral output shares for 1950-2000. We used the US 1950 capital - sectoral output ratios to produce our 1950 capital stock. For 2010 we used the average investment rate for 2000-2009 from S & H. We used Nehru & Dhareshwar (1993) to produce capital stock estimates for 1961-1991. We used perpetual inventory to compute our 2000 and 2010 values.

Enrollments in primary and secondary schools from 1947-1993 come from Maa Table I1 pp. 984 and 989. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1947-1993 come from Maa Table I2 p. 1004. The 2010 enrollment rates come from *HDR*. We interpolated for our 2000 enrollment rates. Our time series of years of schooling in the labor force for 1950-2010 is: .67 (1950), 1.09 (1961), 1.38 (1972), 2.60 (1981), 2.94 (1991), 3.55 (2000) and 4.46 (2010). The Morrisson & Murtin time series of years of schooling is: 1.17 (1950), 1.43 (1960), 1.72 (1970), 2.38 (1980), 3.25 (1990), 4.22 (2000) and 5.97 (2010).

7.16 Papua New Guinea (1960-2010)

Populations for 1960, 1970, 1980 and 1990 come from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1960, 1970, 1980, 1990 and 2000 come from KF. The age distribution for 2010 is assumed to be identical to the age distribution in 2000.

Labor force figures for 1990 come from *HDR*. Labor force figures for 1960, 1970, 1980 and 2000 come from *WDR* (various years). Labor force data for 2010 comes from *WDI*.

Real output come from Maddison, except for 2010. For 2010 we used the 2000 Maddison value and applied the growth rate in real PPP per capita income from 2000 to 2010 using *WDI*. The 1961-2000 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The 2010 investment rate is the 2000-2009 average investment rate from S & H. The *WDI* provides sectoral output shares for 1961, 1970, 1980, 1990 and 2000. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce our sectoral output shares for 1960. We applied the US capital - sectoral output ratios for 1960 to produce our 1960 value. We used perpetual inventory to produce our capital estimates for 1970-2010.

Enrollments in primary and secondary schools for 1960-1990 come from Maa Table I1 p. 993. We assumed the primary school and secondary school ages are 6-11 and 12-17. The tertiary school enrollments

are from *WDR* (various years). The 2010 enrollment rates come from *HDR*. Our 2000 enrollment rates are interpolated.

7.17 Philippines (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1913, 1939, 1948, 1960, 1970, 1980 and 1990 come from Maa Table A2 p. 25. For 1920 and 1930 we interpolated. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1939, 1948, 1960, 1970, 1980 and 1990 come from Maa Table A2 p. 25. The age distributions in 1820-1900 and 1913 are assumed to be identical to the age distribution in 1939. The age distribution for the Philippines 1990 is interpolated from 1980 and 1991 values. Age distribution for Philippines 2000 and 2010 come from the *Demographic Yearbook*.

Labor force figures for 1939, 1948, 1960, 1970, 1975, 1980 and 1990 come from Maa Table B1 p. 98. Labor force figures for the Philippines 1980 and 1990 are interpolated from 1975 and 1993 values. Labor force figure for 2000 comes from *WDR*. Labor force figure for 2010 comes from *WDI*. Labor force figures for 1820-1900 and 1913, 1920 and 1930 come from the following procedure. We used Clio Infra for urbanization rates for 1820-1950, and *WDR* and *WDI* for 1960-2010. We interpolated between 1800 and 1850 for 1820-1840, and we interpolated between 1850 and 1900 for 1860-1890, and we interpolated between 1900 and 1950 for 1910-1940. We assumed a rural 15-64 labor force participation rate of 72.5% for 1820-1948, 60% for 1960, 1970, 67.5% for 1976, 1980, 1990, 2000 and 2010. We assumed an urban 15-64 labor force participation rate of 65% for 1820-1939, 50% for 1948, 1960, 57.5% for 1970, 62.5% for 1976, 1980, 60% for 1990, 63.5% for 2000 and 60% for 2010. We constructed the ratio of this labor force to that from *WDI* and *WDR*. The root mean ratio for the overlapping years 1939-2010 is .998. The 1939 and 1948 values are 1.000 and 1.016, respectively. The range of the ratio is .983 to 1.016.

Real output come from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1946-1998 from Maa Table J1, p. 1035 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2006 from S & H. Sectoral output shares for 1900-1990 come from Hooley (2005). We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1820-1890. We the US 1840-1940 capital - sectoral output ratios, applying the 1850 values for 1820, 1830 and 1840, to produce our capital estimates for 1820-1940. We used perpetual inventory for 1948-2010.

For enrollments in primary and secondary schools for 1900 and 1913 we used Lindert. The primary school enrollment rate in 1900 is 11.6%. The secondary school enrollment rate in 1900 is .21%. Enrollments in primary and secondary schools from 1920-1993 come from Maa Table I1 pp. 984 and 989. For 1820-1890 primary enrollment rates, we assumed each preceding decade is 75 percent of the current decade primary enrollment rate. This produces an 1820 initial primary enrollment rate of 1.2 percent and an 1890 primary enrollment rate of 8.7 percent. It produces a 3 percent school exposure rate in 1870, consistent with Morrison and Murtin. It also matches the 3 percent primary exposure and 0 percent secondary exposure. For 1820-1890 secondary enrollment rates we assume .2 percent. For 1820-1900 higher education enrollment rates we assumed .005 percent. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1910-1993 come from Maa Table I2 pp. 1002 and 1005. For 2010 enrollment rates we used *HDR*. We interpolated for our 2000 enrollment rates. Our time series of schooling in the labor force for 1870-1920 is: .30 (1870), .31 (1880), .50 (1890), .60 (1900),

.67 (1913) and 1.68 (1920). The corresponding time series of schooling from Morrisson & Murtin is: .14 (1870), .17 (1880), .19 (1890), .23 (1900), .29 (1910), .85 (1920).

7.18 Sri Lanka (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1913, 1946, 1953, 1963, 1971, 1981 and 1991 from Maa Table A1 p. 9. Population for 2000 comes from *Time Almanac 2001*. For 1920, 1930 and 1940, we interpolated. Population for 2010 comes from *Wikipedia*.

The age distributions for 1946, 1953, 1963, 1971, 1981 and 1991 come from Maa Table A2 p. 25. The age distributions for all years prior to 1946 are assumed to be identical to the age distribution in 1946. Age distribution for Sri Lanka 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 1998 age distribution, and for 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1946, 1953, 1963, 1971, 1981 come from Maa Table B1 p. 99. Labor force figures for Sri Lanka 1991 and 2000 come from *WDR* (various years). The labor force figures for 2010 comes from *WDI*. Labor force figures for 1820-1940 come from the following procedure. We used Clio Infra for urbanization rates for 1820-1950, and *WDR* and *WDI* for 1960-2010. We interpolated between 1800 and 1850 for 1820-1840, and we interpolated between 1850 and 1900 for 1860-1890, and we interpolated between 1900 and 1950 for 1910-1940. We assumed a rural 15-64 labor force participation rate of 68.5% for 1820-1953, 57% for 1963, 51% for 1971, 46% for 1981, 63% for 1991, 70% for 2000 and 62.5% for 2010. We assumed an urban 15-64 labor force participation rate of 50% for 1820-1971, 46% for 1981, 50% for 1991, 2000 and 2010. We constructed the ratio of this labor force to that from *WDI* and *WDR*. The root mean ratio for the overlapping years 1946-2010 is 1.001. The 1946 and 1953 values are 1.000 and .996, respectively. The range of the ratio is .996 to 1.006.

Real output come from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1950-1990 and 1994-1998 come from Maa Table J1, p. 1036 and *WDR* (various years). For 2010 we used the average investment rate from 2000-2009 from S & H. The *WDI* provides sectoral output shares for 1963-2000. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1820-1953. We applied the US 1850-1950 capital - sectoral output ratios, applying the 1850 values for 1820, 1830 and 1840, to produce capital estimates from 1820-1946. We used perpetual inventory to compute capital stocks from 1953-2010.

Enrollment rates for primary school and higher education in 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900 and 1913 come from Lindert. The 1830 primary school enrollment rate is 10%. For 1820 we assumed primary, secondary and higher education enrollment rates of: 10 percent, .5 percent and .001 percent, respectively. For secondary rates over this period, except for 1913, we used one tenth of the enrollment rate in primary school for secondary school enrollment rates. Enrollments in schools from 1946-1993 come from Maa Table I1 p. 985 and 990. We interpolated for years 1920-1940. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1921-1990 come from Maa Table I2 pp. 1002 and 1005. For 2010 we used the online UNESCO *UIS Global Database* for secondary school and higher education enrollment rates. We used *HDR* for primary enrollment rates in 2010. We interpolated for our 2000 enrollment rates.

7.19 Thailand (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1917, 1927, 1937, 1947, 1960, 1970, 1980 and 1990 come from Maa Table A2 p. 27. We interpolated the 1910 value. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1937, 1947, 1960, 1970, 1980 and 1990 come from Maa Table A2 p. 27. The age distributions for 1820-1910, 1917 and 1927 are assumed to be identical to the age distribution in 1937. The age distribution for Thailand for 1990 is interpolated from 1980 and 1991 values. Age distribution for Thailand 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1937 1947, 1960, 1970, 1980 and 1990 come from Maa Table B1 p. 101. Labor force for 1927 comes from Williamson (2000) land labor ratio, and Mitchell's land data. Labor force for 1820-1917 are from our calculations. Warr (1996) presents the labor force share in agriculture from 1960-1990. In 1960, 90% of workers are in agriculture. We ran a time series regression on this share and it implied 100% of the work force in agriculture for all years prior to 1950. We assumed that agriculture's share was top coded at 95% from 1820-1860, and then declines at 1% until reaching 90% in 1910 and remains there until 1960. We computed the population aged 15-64, and assumed that agriculture population has a 100% labor force participation rate, and non agriculture population has a 70% labor force participation rate. Comparing this method with Maa data over 1947-1990 shows that in three years we are within .1% of the Maa data, and in the other two observations, within 5%. Labor force figures for 1820-1900 are assumed to have the same labor force participation rate, LF/population, as the average labor force participation rate for years 1910-2010. Labor force figure for Thailand 2000 comes from *WDR*. Labor force figures for 2010 come from *WDI*.

Real GNPs come from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1957-1998 from Maa Table J1, p. 1038 and *WDR* (various years). We used the intraperiod average gross capital formation and income for 1952-1956. For 2010 we used the average investment rate for 2000-2009 from S & H. Mitchell provides sectoral output shares for 1947-2000. The UN Statistical Yearbook 1954 provides sectoral output shares for 1937. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1820-1927. We applied US 1850-1950 capital - sectoral output ratios, applying US 1850 values for 1820, 1830 and 1840, to produce our capital estimates for 1820-1947. We used perpetual inventory for 1960-2010.

For primary school enrollment rates in 1890 and 1910 we used Benavot and Riddle. For 1900 we used Lindert for enrollment rates in primary, secondary and higher education. The 1900 primary school enrollment rate is 6.8%. The secondary school enrollment rate in 1900 is .05%. The higher education enrollment rate for 1900 is .05%. For primary enrollment rates prior to 1890, we assumed that each decade's enrollment rate was 90.5 percent of the succeeding decade's primary enrollment rate. This produces a primary school enrollment rate time series of: .35% (1820), .37% (1830), .39% (1840), .41% (1850), .43% (1860), .45% (1870), .475% (1880). For secondary enrollment rates prior to 1900 we assumed .01 percent. For higher education enrollment rates prior to 1900, we assumed .001 percent. Enrollments in primary and secondary schools from 1913-1993 come from Maa Table I1 pp. 985 and 991. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1915-1993 come from Maa Table I2 pp. 1002 and 1005. For 2010 we used *HDR*. We interpolated for

our enrollment rates in 2000. Our time series share of labor force exposed to higher education is consistent with Morrisson & Murtin 1870-2010. Our time series for schooling in the labor force for 1870-1920 is: .03 (1870), .03 (1880), .03 (1890), .03 (1900), .23 (1910) and .28 (1917). The Morrisson & Murtin time series of schooling for these years is: .17 (1870), .22 (1880), .26 (1890), .31 (1900), .36 (1910) and .40 (1920).

7.20 Vietnam (1820-2010)

McEvedy and Jones provide population for 1800, 1850, 1900. For years 1820, 1830, 1840, 1860, 1870, 1880, 1890, 1913, 1920, 1930 and 1940 we interpolated. Populations for 1950, 1960 and 1970 come from *WDI*. Populations for 1980 and 1990 come from *WDR* (various years). Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1960, 1970, 1980, 1990 come from KF. We assumed the age distributions for years prior to 1960 were identical with the age distribution in 1960. Age distribution for Vietnam 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2009 age distribution. We adjust by assuming the same share by age category as in 2009.

Labor force figures for 1960, 1970, 1980 come from *WDI*. Labor force figures for 1990 come from *HDR*. Labor force data for 2000 comes from WDR, and labor force data for 2010 comes from *WDI*. Labor force figures for 1820-1950 come from the following procedure. We used Clio Infra for urbanization rates for 1820-1950, and *WDR* and *WDI* for 1960-2010. We interpolated between 1800 and 1850 for 1820-1840, and we interpolated between 1850 and 1900 for 1860-1890, and we interpolated between 1900 and 1950 for 1910-1940. We assumed a rural 15-64 labor force participation rate of 103% for 1820-1960, 92% for 1970, 91% for 1980, 110% for 1990, 90% for 2000 and 85% for 2010. We assumed an urban 15-64 labor force participation rate of 103% for 1820-1960, 92% for 1970, 91% for 1980, 110% for 1990, 90% for 2000, and 80% for 2010. We constructed the ratio of this labor force to that from *WDI* and *WDR*. The root mean ratio for the overlapping years 1960-2010 is .996. The 1960 and 1970 values are .999 and .996, respectively. The range of the ratio is .983 to 1.005.

Real output come from Maddison. The 1990 and 2000 investment rates are from the intraperiod average investment rate taken from *WDR* (various years). The 2010 investment rate is the average investment rate from S & H for years 2000-2009. The *WDI* provides sectoral output shares for 1960, 1970, 1990-2010, and we interpolated for 1980. We applied Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1820-1950. We used US 1850-1980 capital - sectoral output ratios, applying the US 1850 values for 1820, 1830 and 1840, to produce our capital estimates for 1820-1980. We used perpetual inventory for 1990-2010.

Enrollment rates in primary school for 1900, 1913, 1920, 1930 and 1940 come from Benavot & Riddle. For secondary school enrollment rates, we assumed they are one twentieth of primary school enrollment rates, which is the ratio in 1950. For tertiary enrollment rates we assumed a constant .02% value for years 1920-1940, and .005% for 1900 and 1913. This compares with the 1950 observation of .04%. The 1900 primary enrollment rate is .5% primary school enrollment rates. For years 1820-1890 we assumed schooling enrollment rates were the same as their 1900 values. Enrollment rates in primary, secondary and tertiary schools for 1950-1970 come from Mitchell. come from *WDR* (various years). The 1980, 1990 rates come from *WDI*. The 2000 rates come from UNESCO *UIS Global Database*. The 2010 rates come from *HDR*. We assumed that primary school and secondary school ages are 6-10 and 11-17.

8 Sub-Saharan Africa

8.1 Angola (1950-2010)

Populations for 1950, 1960, 1970 come from Maa Table A1 p. 3. Populations for 1980 and 1990 come from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distribution for 1950, 1960 come from Maa Table A2 p. 14. The age distributions for Angola for 1970, 1980, 1990 and 2000 comes from KF. We assumed the age distribution in 2010 was the same as the age distribution in 2000.

Labor force figures for Angola for 1980, 1990 and 2000 come from *WDR* (various years). Labor force figures for 2010 come from *WDI*. For 1950, 1960, 1970 we followed the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 91% for 1950-1980, 82.25% for 1990, 100% for 2000 and 83.85% for 2010. We assumed urban 15-64 labor force participation rates of 57.5% for 1950-1980, 50% for 1990, 91% for 2000 and 60% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1980-2010 is 1.000. The 1980 and 1990 ratios are 1.000 and 1.001, respectively. The range of the ratios is .999 to 1.001.

Real output come from Maddison. Physical capital investment rates for 1980,1990, 2000, 2010 come from the intraperiod average investment rate from S & H online and *WDR* (various years). The 1978 and 1982 WDR provide sectoral output shares for 1960, 1976 and 1980. We interpolated our 1970 value. The *WDI* provides sectoral output shares for 1990 and 2000. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates for Sub-Saharan African countries to produce our 1950 sectoral output shares. We applied the US 1950-1970 capital - sectoral output ratios to produce our capital estimates. We used perpetual inventory to compute our 1980-2010 values. We did not use Nehru and Dhareshwar (1993) values for 1950-1970 as their estimates have capital output ratios between .85 and 1.22, which we viewed as implausibly too low.

Enrollments in primary and secondary schools from 1940-1960 come from Maa Table I1 p. 973. To calculate enrollment rates, we assumed 6-9 are primary school age and 10-15 are secondary school age. Higher education enrollments for 1978-1992 come from Maa Table I2 p. 997. Prior to 1978 we assume an enrollment rate of 0 for higher education. For years 1970-1990, inclusive, we used *WDI* for primary and secondary schooling enrollment rates, but Mitchell for higher education enrollment rates. For 2000 we used UNESCO *UIS Global Database*. For 2010 primary school enrollment rate we used WDR. For 2010 tertiary and secondary school enrollment rates we used *HDR*. Our time series for years of schooling in the labor force for 1940-2010 is: .02 (1940), .04 (1950), .09 (1960), .41 (1970), 2.42 (1980), 5.00 (1990), 4.80 (2000), 4.55 (2010). The Morrisson & Murtin time series for years of schooling is: .03 (1940), .04 (1950), .12 (1960), .30 (1970), .96 (1980), 1.91 (1990), 2.39 (2000), 2.98 (2010).

8.2 Benin (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Benin 2000

and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in 2011.

Labor force figures for 1990 come from *HDR*. Labor force figures for 1960, 1970, 1980 and 2000 come from *WDR* (various years). For 2010 we used *WDI*. For 1950 we followed the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 92.5% for 1950-1970, 89% for 1980, 83.75% for 1990, 90.5% for 2000 and 75% for 2010. We assumed urban 15-64 labor force participation rates of 52% for 1950-1980, 50% for 1990, 75% for 2000 and 68% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are .990 and 1.001, respectively. The range of the ratios is .990 to 1.002.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rate from *WDR 2011* and *WDR 2012*. The 1960-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1960-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates for Sub-Saharan African countries to produce our 1950 values. We applied the US 1950 capital - sectoral output ratios to produce our 1950 value. We used perpetual inventory to compute our 1960-2010 capital values.

Enrollments in primary and secondary schools for 1950-1990 come from Maa Table I1 p. 973. We assumed the primary school and secondary school ages are 6-11 and 12-18. The tertiary school enrollments are from Maa Table I2 p. 997. Our 2000 primary and secondary school enrollment rates are averages of 1998 and 2002 rates in *WDR*. Our 2000 higher education enrollment rate is from *WDR*. For 2010 we used *WDR* for tertiary and secondary school enrollment rates. We used *HDR* for primary school enrollment rates in 2010. Our time series for years of schooling in the labor force for 1950-2010 is: .50 (1950), .74 (1960), 1.31 (1970), 2.26 (1980), 4.08 (1990), 4.41 (2000) and 6.83 (2010). The Morrisson & Murtin time series of years of schooling is: .33 (1950), .42 (1960), .56 (1970), .98 (1980), 1.82 (1990), 2.33 (2000) and 2.89 (2010).

8.3 Botswana (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Population Pyramid* at https://populationpyramid.net/botswana/2000/, accessed July 24, 2016. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Botswana 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2001 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1990 come from *HDR*. Labor force figures for 1980 and 2000 come from *WDR* (various years). For 2010 we used *WDI*. For 1950, 1960, 1970 we followed the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 105% for 1950-1980, 95% for 1990, 85% for 2000 and 90% for 2010. We assumed urban 15-64 labor force participation rates of 105% for 1950-1980, 87.5% for 1990, 72% for 2000 and 74.25% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1980-2010 is 1.001. The 1980 and 1990 ratios are 1.001 and

1.001, respectively. The range of the ratios is 1.000 to 1.001.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. The 1960-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1960, 1980-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct our 1950 and 1970 values. We applied the US 1950 capital - sectoral output ratios to produce our 1950 capital estimates. We used perpetual inventory to produce our 1960-2010 values.

Enrollment rates for primary and secondary schools for 1970-2010 come from *WDR* (various years), except for 2010 primary school enrollment rates which comes from *HDR*. We assumed the primary school and secondary school ages are 6-12 and 13-17. The tertiary school enrollments are from *WDR* (various years). For 1950 enrollment rates come from 1954 *UN Statistical Yearbook*. Our 1960 enrollment rates come from 1963 *UN Statistical Yearook*.

8.4 Burkina Faso (1950-2010)

Populations for 1950, 1960, 1970, 2000 and 2010 come from *Population Pyramid*, https://populationpyramid.net/burkina-faso/1960/, accessed July 24, 2016. Populations for 1980 and 1990 come from Maddison.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Burkina Faso 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 1996 age distribution, and for 2010 we adjusted the 2009 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1990 come from *HDR*. Labor force figures for 1970, 1980 and 2000 come from *WDR* (various years). Labor force for 2010 comes from *WDI*. For 1950, 1960 we followed the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 80% for 1950-2000, 97.5% for 2010. We assumed urban 15-64 labor force participation rates of 55% for 1950-1970, 50% for 1980, 1990, 62.25% for 2000 and 93% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1970-2010 is 1.001. The 1970 and 1980 ratios are 1.000 and 1.001, respectively. The range of the ratios is 1.000 to 1.002.

Real output come from Maddison. The 1971-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1960-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct 1950 sectoral output shares. We applied the US 1950 and 1960 capital - sectoral output ratios to construct our 1950 and 1960 capital values. We used perpetual inventory for 1970-2010.

Enrollments in primary and secondary schools for 1960-2000 come from Maa Table I1 p. 973. We assumed the primary school and secondary school ages are 6-11 and 12-18. The tertiary school enrollments are from Maa Table I2 p. 997. For 1950 we used 1954 *UN Statistical Yearbook* for all enrollment rates. For 2010 we used *HDR* for enrollment rates.

8.5 Burundi (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Burundi 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2008 age distribution. We adjust by assuming the same share by age category as in 2008.

Labor force figures for 1990 come from *HDR*. Labor force figures for 1950, 1960, 1970, 1980 and 2000 come from *WDR* (various years). Labor force data for 2010 comes from *WDI*.

Real output come from Maddison, except for 2010. For 2010 we used the Maddison 2008 value and applied the real PPP per capita income growth rate from *WDR 2011* and *WDR 2012*. The 1970-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1970-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct 1950 and 1960 sectoral shares. We used the US 1950 and 1960 capital - sectoral output ratios to construct our 1950 and 1960 capital values. We used perpetual inventory for 1970-2010.

Enrollments in primary and secondary schools for 1960-2000 come from Maa Table I1 p. 973. We assumed the primary school and secondary school ages are 6-11 and 12-18. The tertiary school enrollments are from Maa Table I2 p. 997. For 1950 we used the 1954 *UN Statistical Yearbook* for all enrollment rates. For 2010 we used *HDR* for all enrollment rates.

8.6 Cameroon (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Cameroon 2000 and 2010 come from the *Demographic Yearbook*.

Labor force figures for Cameroon for 1980 and 1990 come from Maa Table B1 p. 90. The labor force for 1960 and 1970 come from the *WDI*. The labor force for 2000 and 2010 come the WDR. We used Clio infra to compute the 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 88.15% for 1950-1960, 83.71% for 1970, 88% for 1980, 85.23% for 1990, 94.5% for 2000 and 95% for 2010. We assumed urban 15-64 labor force participation rates of 50% for 1950-2000, 54.6% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real output come from Maddison. Physical capital investment rates for 1970-2010 come from the intraperiod average investment rate from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1970-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1950 and 1960. We applied the US 1950 and 1960 capital - sectoral output ratios to construct our 1950 and 1960 capital estimates. We used perpetual inventory for 1970-2010.

Enrollments in primary and secondary schools from 1950-1993 come from Maa Table I1 p. 973. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-18 are secondary school age.

Higher education enrollments for 1967-1990 come from Maa Table I2 p. 997. Prior to 1967 we assume an enrollment rate of .001 for higher education. For 2010 we used *HDR*. We interpolated for our 2000 enrollment rates. Our time series for years of schooling in the labor force for 1950-2010 is: .55 (1950), .81 (1960), 2.45 (1970), 4.20 (1980), 5.99 (1990), 6.97 (2000) and 7.61 (2010). The Morrisson & Murtin time series of years of schooling is: .61 (1950), 1.37 (1960), 1.88 (1970), 3.05 (1980), 4.10 (1990), 4.66 (2000) and 5.10 (2010).

8.7 Cape Verde (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Cape Verde 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1960, 1970, 1980, 1990, 2000 and 2010 come from *WDI* (various years). For 1950 we used the following procedure. We used Clio infra to compute the 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 60.725% for 1950-1960, 71.4% for 1970, 70% for 1980, 69.2% in 1990, 95% for 2000 and 2010. We assumed urban 15-64 labor force participation rates of 50% for 1950-1990, 51.25% for 2000, 78.5% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1980 and 1990 ratios are .999 and 1.001, respectively. The range of the ratios is .999 to 1.001.

Real output come from Maddison, except 2010. The 2010 value comes from the Maddison 2008 observation that we applied the real PPP per capita income growth rate from *WDR 2011* and *WDR 2012*. The 1970-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1970-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1950 and 1960. We applied the US 1950 and 1960 capital - sectoral output ratios to construct our 1950 and 1960 capital estimates. We used perpetual inventory for 1970-2010.

For 1950 enrollment rates we used 1954 *UN Statistical Yearbook*. For 1960 we used UNESCO *UIS Global Database*. Enrollments in primary and secondary schools for 1970-1990 come *WDI*. For 2000 we used UNESCO *UIS Global Database*. For 2010 enrollment rates we used *HDR*. We assumed the primary school and secondary school ages are 6-11 and 12-16.

8.8 Central African Republic (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. We assumed the same age distribution for 2010 as in 2000.

Labor force figures for 1990 come from *HDR*. Labor force figures for 1960, 1970, 1980, 2000, and 2010 come from *WDR* (various years). For 1950 we followed the following procedure. We used Clio infra for 1950

urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 1.0075% for 1950-1960, 104.5% for 1970, 100% for 1980, 95% for 1990, 97% for 2000, 80% for 2010. We assumed urban 15-64 labor force participation rates of 100.75% for 1950-1960, 104.5% for 1970, 93.5% for 1980, 91.5% for 1990, 97% for 2000, and 77.25% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.001, respectively. The range of the ratios is 1.000 to 1.001.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDT 2011 and 2012*. The 1960-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1970-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1950 and 1960. We applied the US 1950 capital - sectoral output ratios to construct our 1950 estimates of capital. We used perpetual inventory for 1960-2010.

Enrollments in primary and secondary schools for 1950-2000 come from Maa Table I1 p. 974. We assumed the primary school and secondary school ages are 6-11 and 12-18. The tertiary school enrollments are from Maa Table I2 p. 997. For 2010 we used *HDR*.

8.9 Chad (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. We assumed the same age distribution in 2010 as in 2000.

Labor force figures for 1980, 1990 come from *WDI*. Labor force figures for 2000 and 2010 come from *WDR* (various years). For 1950, 1960 and 1970 we followed the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 60% for 1950-1980, 75% for 1990, 73.05% for 2000, 75% for 2010. We assumed urban 15-64 labor force participation rates of 50% for 1950-1980, 65% for 1990, 50% for 2000, 67.75% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1980-2010 is 1.000. The 1980 and 1990 ratios are .999 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDT 2011 and 2012*. The 1960-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1960-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1950. We applied the US 1950 capital - sectoral output ratios to construct our 1950 estimates of capital. We used perpetual inventory for 1960-2010.

Enrollments in primary and secondary schools for 1950-2000 come from Maa Table I1 p. 974. We assumed the primary school and secondary school ages are 6-11 and 12-18. The tertiary school enrollments are from Maa Table I2 p. 997. For 2010 we used *HDR*.

8.10 Comoros (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. We assumed the same age distribution in 2010 as in 2000.

Labor force figures for 1980-2010 come from *WDI*. For 1950, 1960, 1970 we used the following procedure. We used Clio infra to compute the 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 85.6% for 1950-1960, 91% for 1970, 94.9% for 1980, 52.4% in 1990, 58.1% for 2000, 60% for 2010. We assumed urban 15-64 labor force participation rates of 50% for 1950-2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are .999 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. The 1970-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1980-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates on Sub-Saharan African countries to construct sectoral output shares for 1950,-1970. We applied the US 1950-1960 capital - sectoral output ratios to construct our 1950 and 1960 estimates of capital. We used perpetual inventory for 1970-2010.

Enrollments in primary and secondary schools for 1960-2000 come from *WDR* various issues. For 2010 we used *HDR*. For 1950 we assumed 5 percent elementary school enrollment rate, and .5 percent secondary school enrollment rate, and 0 higher education. We assumed the primary school and secondary school ages are 6-11 and 12-18.

8.11 Congo (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Congo 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2009 age distribution. We adjust by assuming the same share by age category as in 2009.

Labor force figures for 1960-1990 come from *WDI*. Labor force figures for 2000 and 2010 come from *WDR* (various years). For 1950 we followed the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 90% for 1950-1980, 80% for 1990, 90% for 2000, 80% for 2010. We assumed urban 15-64 labor force participation rates of 71.5% for 1950-1970, 70.3% for 1980, 62.22% for 1990, 73.25% for 2000, 70.07% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1980 and 1990 ratios are 1.000 and 1.001, respectively. The range of the ratios is 1.000 to 1.001.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. The 1970-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various

years). The *WDI* provides sectoral output shares for 1960-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates for Sub-Saharan African countries to construct sectoral output shares for 1950. We applied the US 1950 and 1960 capital - sectoral output ratios to construct our 1950 and 1960 estimates of capital. We used perpetual inventory for 1970-2010.

Enrollments in primary and secondary schools for 1950-2000 come from Maa Table I1 p. 974. We assumed the primary school and secondary school ages are 6-11 and 12-18. The tertiary school enrollments are from Maa Table I2 p. 997. For 2010 we used *HDR*.

8.12 Cote de Ivoire (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. We assumed the same age distribution in 2010 as in 2000.

Labor force figures for 1960-1990 come from *WDI*. Labor force figures for 2000 and 2010 come from *WDR* (various years). For 1950 we followed the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 87% for 1950-1980, 90% for 1970-2000, 80% for 2010. We assumed urban 15-64 labor force participation rates of 64.4% for 1950-1960, 66.5% for 1970, 65.5% for 1980, 57.1% for 1990, 66.2% for 2000, 57.1% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1980 and 1990 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real output come from Maddison. The 1970-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1960-2010. We used Sabillon (2005) information on farming, manufacturing, and aggregate growth rates for Sub-Saharan African countries to construct our 1950 sectoral output shares. We applied the US 1950 capital - sectoral ratios to construct our capital estimates for 1950. We used Nehru and Dhareshwar (1993) for 1960-1990 capital values. We used perpetual inventory for 2000 and 2010.

Enrollments in primary and secondary schools for 1950-2000 come from Maa Table I1 p. 975. We assumed the primary school and secondary school ages are 6-11 and 12-18. The tertiary school enrollments are from Maa Table I2 p. 998. For 2010 we used *HDR*. Our time series of years of schooling in the labor force for 1950-2010 is: .24 (1950), .34 (1960), 1.35 (1970), 2.48 (1980), 3.98 (1990) 4.82 (2000) and 5.92 (2010). The Morrisson & Murtin time series of years of schooling is: .16 (1950), .28 (1960), .59 (1970), 1.51 (1980), 2.52 (1990), 3.22 (2000) and 3.78 (2010).

8.13 Djibouti (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. We assumed the same age distribution in 2000 and 2010 as in 1990.

Labor force figures for 1990 come from *Time Almanac 2003*. For 2000 we used *WDR*. Labor force figures for 2010 we used *WDI*. For 1950, 1960, 1970 and 1980 we used the following procedure. We used

Clio infra to compute the 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 52% for 1950-1980 (which is the average of the rural LF participation rates used below), 44% in 1990, 48.1% for 2000, 63.5% for 2010. We assumed urban 15-64 labor force participation rates of 47% for 1950-1980 (which is the average or the urban LF participation rates used below), 44% for 1990, 48.1% for 2000, 50% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1990-2010 is 1.001. The 1990 and 2000 ratios are .999 and 1.002, respectively. The range of the ratios is .999 to 1.002.

Real output come from Maddison, except for 2010. We used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. The 2000-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1990 and 2000. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates of Sub-Saharan African countries to construct our 1950-1990 sectoral output shares. We applied the US 1950-1990 capital - sectoral output ratios to construct our capital estimates. We used perpetual inventory to compute our 2000 and 2010 capital estimates.

For 1950 we used the 1954 *UN Statistical Yearbook* for primary and secondary school enrollments, and we assumed 0 higher education enrollments. For 1960 enrollments we used the 1963 *UN Statistical Yearbook*. Enrollments in primary and secondary schools for 1970-1990 come from *WDI*. For 2000 we used UNESCO *UIS Global Database*. We interpolated higher education enrollment rates for 1980 and 1990. For 2010 we used *HDR*. We assumed the primary school and secondary school ages are 6-11 and 12-18.

8.14 Equatorial Guinea (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. We assumed the same age distribution in 2010 as in 2000.

Labor force figures for 1960-2000 come from *WDI*. Labor force figures for 2010 assumes the same labor force participation rate, LF/population, as in 2000. For 1950 we used the following procedure. We used Clio infra to compute the 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 95% for 1950-1990, 101.35% in 2000, 88.45% for 2010, 63.5%. We assumed urban 15-64 labor force participation rates of 75% for 1950-1960, 50.1% for 1970, 66% for 1980, 90.5% for 1990, 101.35% for 2000, 88.45% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.001, respectively. The range of the ratios is 1.000 to 1.001.

Real output come from Maddison, except for 2010. We used the 2008 Maddison value and applied the real PPP per capita income growth rate from *WDR 2011* and *WDR 2012*. The 1970-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The Geohive site *GDP World 2000 by sector*, at http://www.geohive.com/charts/ec_sector.aspx, provides sectoral output shares for 2000. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct our 1950-1990 sectoral output shares. We applied the US 1950-1990 capital - sectoral output ratios to produce our capital estimates for these years. We used perpetual inventory to estimate capital for 2000 and 2010.

Our 1950 enrollments come from the 1954 *UN Statistical Yearbook*. Our 1960 enrollments from from

the 1963 *UN Statistical Yearbook*. Enrollments in primary, secondary schools and higher education for 1970-2000 come from *WDI*. For 2010 we used *HDR*. We assumed the primary school and secondary school ages are 6-11 and 12-17.

8.15 Eritrea (1990-2010)

Populations for 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1990 comes from KF, we used the Ethiopian age distribution. We assumed the same age distribution in 2010 and 2000 as in Ethiopia in 2010 and 2000.

Labor force figures for 1990 come from *WDI*. Labor force figures for 2000 come from *CIA World Factbook*. Labor force figures for 2010 come from *WDI*.

Output comes from Maddison. The 1990-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The 2002 print edition of the *World Development Indicators* provides sectoral output shares for 1990 and 2000. We applied the US 1990 capital - sectoral output ratios to construct our 1990 value. We used perpetual inventory to produce our 2000 and 2010 estimates.

Enrollment rates in primary, secondary schools for 1990 come from UNICEF, Division of Policy and Practice, Statistics and Monitoring Section. For higher education enrollment rate in 1990, we assumed the same 1990 value as in Ethiopia, .9%. For 2010 we used *WDR*. We interpolated enrollment rates for 2000. We assumed the primary school and secondary school ages are 6-11 and 12-17.

8.16 Ethiopia (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from KF. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Ethiopia 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2008 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1960, 1980 and 2000, 2010 come from *WDR* (various years). The labor force for 1970 and 1990 come from interpolation of the 1960, 1980 and 2000 values. For 1950 we used the following procedure. We used Clio infra to compute the 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 102.1% for 1950-1960, 94% for 1970, 93% for 1980, 90.5% for 1990, 89.1% in 2000, 93.75% for 2010, 63.5%. We assumed urban 15-64 labor force participation rates of 102.1% for 1950-1960, 53% for 1970, 50% for 1980, 1990, 2000, 50.25% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real output come from Maddison. Physical capital investment rates for 1960-2006 come from the intraperiod average investment rate from S & H online and *WDR* (various years). The 1982 *World Development Report* provides sectoral output shares for 1960 and 1980, and we interpolated for 1970. The 2002 *World Development Indicators* provides sectoral output shares for 1990 and 2000. We used Sabillon (2005)

information on farming, manufacturing and aggregate growth rates for Sub-Saharan African countries to produce our 1950 sectoral output shares. We used the US 1950 - 1990 capital - sectoral output ratios to construct capital stock values for 1950 - 1990. We used perpetual inventory for 2000 and 2010 estimates.

Enrollments in primary and secondary schools from 1948-1993 come from Maa Table I1 pp. 968 and 974. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1955-1991 come from Maa Table I2 p. 998. Prior to 1955 we assume an enrollment rate of 0 for higher education. For 2010 we used *HDR*. We interpolated enrollment rates for 2000. Our years of schooling in the labor force time series for 1950-2010 is: .06 (1950), .09 (1960), .24 (1970), .78 (1980), 1.83 (1990), 2.18 (2000) and 3.42 (2010). The Morrisson & Murtin time series of years of schooling is: .05 (1950), .12 (1960), .22 (1970), .53 (1980), 1.29 (1990), 1.93 (2000) and 2.51 (2010).

8.17 Gabon (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. We assumed the same age distribution in 2010 as in 2000.

Labor force figures for 1990 come from *WDI*. Labor force figures for 2000 and 2010 come from *WDR* (various years). For 1950, 1960, 1970 and 1980 we followed the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 50% for 1950-1980, 45.1% for 1990, 70% for 2000, 96% for 2010. We assumed urban 15-64 labor force participation rates of 50% for 1950-1980, 45.1% for 1990, 61.75% for 2000, 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1990-2010 is 1.000. The 1980 and 1990 ratios are 1.000 and 1.001, respectively.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rate from *WDR 2011* and *WDR 2012*. The 1970-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1990, 2000 and 2010. The 1973 *UN Statistical Yearbook* provides sectoral output shares for 1960. We interpolated to produce 1970 and 1980 values. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates for Sub-Saharan African countries to produce our 1950 sectoral output shares. We used the US 1950 and 1960 capital - sectoral output ratios to produce our 1950 and 1960 capital estimates. We used perpetual inventory for 1970-2010.

Enrollments in primary and secondary schools for 1950-2000 come from Maa Table I1 p. 974. We assumed the primary school and secondary school ages are 6-11 and 12-18. The tertiary school enrollments are from Maa Table I2 p. 998. For 2010 we used *HDR*.

8.18 Gambia, The (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. We assumed the same age distribution in 2010 as in 2000.

Labor force figures for 1960-1990 come from *WDI*. Labor force figures for 2000 and 2010 come from *WDR* (various years). For 1950 we followed the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 95% for 1950-1960, 96% for 1970, 1980, 90% for 1990, 75% for 2000, 90% for 2010. We assumed urban 15-64 labor force participation rates of 62.75% for 1950-1960, 95% for 1970, 1980, 70% for 1990, 68.7% for 2000, 71.25% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1980 and 1990 ratios are 1.000 and 1.002, respectively.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rate from *WDR 2011* and *WDR 2012*. The 1970-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1990, 2000 and 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates for Sub-Saharan African countries to produce our sectoral output shares for 1950-1980. We used the 1950 and 1960 US capital - sectoral output ratios to construct our physical capital stock values for these years. We used perpetual inventory for 1970-2010 values.

For the 1950 primary school and secondary school enrollments we used the 1954 *UN Statistical Yearbook*. For 1950 we assumed a 0 percent higher education enrollment rate. Enrollment rates in primary and secondary schools for 1960-2000 come from *WDR* (various years). We interpolated the 1960 secondary school enrollment rate. We assumed the primary school and secondary school ages are 6-11 and 12-18. The tertiary school enrollment rates are from *WDR* (various years). For 2010 we used *HDR*.

8.19 Ghana (1870-2010)

Populations for 1870, 1880, 1890, 1900, 1913, 1920, 1830, 1940, 1950, 1960, 1970, 1980, 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1960, 1970, 1980 and 1990 come from Maa Table A2 p. 15. The age distributions for 1870, 1880, 1890, 1900, 1913, 1920, 1830, 1940, and 1950 are assumed to be identical to the age distribution in 1960. Age distribution for Ghana 2000 and 2010 come from the *Demographic Yearbook*.

Labor force figures for 1960, 1970, 1980 and 1990 come from *WDI* (various years). The 2000 and 2010 labor force figures come from *WDR*. For 1870-1950 we used the following procedure. We used Clio Infra for urbanization rates for 1850-1950, and *WDR* and *WDI* for 1960-2010. We interpolated between 1850 and 1900 for 1870-1890, and we interpolated between, and we interpolated between 1900 and 1950 for 1910-1940. We assumed a rural 15-64 labor force participation rate of 95% for 1870-1970, 94% for 1980, 75% for 1990, 84.5% for 2000, and 80% for 2010. We assumed an urban 15-64 labor force participation rate of 85% for 1870-1960, 94.5% for 1970, 92.2% for 1980, 65% for 1990, 2000 and 63% for 2010. We constructed the ratio of this labor force to that from *WDI* and *WDR*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 values are 1.001 and 1.000, respectively. The range of the ratio is .995 to 1.001.

Real output come from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1959-1986 and 1994-1998 and average gross capital formation and income for 1987-1993 from Maa Table J1, p. 1014 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. The 1960 farming share, and 1970-2010 sectoral shares

are from the *WDI*. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce our 1870-1950 sectoral shares, and manufacturing and service shares for 1960. We used the US 1870-1990 capital - sectoral output ratios to produce our capital stock estimates for these years. We used perpetual inventory for 2000 and 2010.

Enrollments in primary and secondary schools from 1903-1993 come from Maa (2003) Table I1 pp. 969 and 975. The 1900 observation for primary and secondary school enrollment rates are 4% and .27%, respectively. The 1890 observation comes from Benavot and Riddle (1988). We used the 1890 value, 1.1% for primary enrollment rates in 1870 and 1880. For secondary enrollment rates for these years we assumed .1, .1 percent and .1 percent, respectively. For higher education enrollment rates, for these years, we assumed 0. All of these fit exactly the schooling shares of primary, secondary and higher education from Morrisson & Murtin for 1870, 1880, 1890. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1949-1990 come from Maa Table I2 p. 998. We used *HDR* for 2010. We interpolated our 2000 enrollment rates for all schooling categories. Our time series for years of schooling in the labor force for 1870-1930 is: .03 (1870), .05 (1880), .06 (1890), .07 (1900), .16 (1913), .60 (1920), 1.02 (1930). The Morrisson & Murtin time series of years of schooling is: .04 (1870), .05 (1880), .06 (1890), .07 (1900), .12 (1910), .18 (1920), .30 (1930).

8.20 Guinea (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Guinea 2010 comes from the *Demographic Yearbook*. For 2010 we adjusted the 2009 age distribution. We adjust by assuming the same share by age category as in 2009.

Labor force figures for 1980, 1990, 2000 and 2010 come from *WDR* (various years). For 1970 we used the average labor force participation rate, LF/population, for years 1980-2010. For 1960 we used *WDR*. For 1950 we used the 1960 labor force population ratio.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rate from *WDR 2011* and *WDR 2012*. The 1970-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1990-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1950-1980. We applied the US 1950 and 1960 capital - sectoral output ratios to produce our capital estimates for these years. We used perpetual inventory for 1970-2010.

Enrollments in primary and secondary schools for 1960-2000 come from Maa Table I1 p. 975. We assumed the primary school and secondary school ages are 6-11 and 12-18. The tertiary school enrollments are from Maa Table I2 p. 998 and *WDR* (various years). For 2010 we used *HDR* for all enrollment rates. For 1950 we assumed the same enrollment rates as in 1950 Senegal.

8.21 Guinea-Bissau (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. We assumed the same age distribution in 2010 as in 2000.

Labor force figures for 1970, 1980, 1990, 2000 and 2010 come from *WDR* (various years). For 1950 and 1960 we followed the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 77% for 1950-2010. We assumed urban 15-64 labor force participation rates of 61% for 1950-1970, 55.5% for 1980, 56% for 1990, 54.5% for 2000, 58.75% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1970-2010 is 1.000. The 1970 and 1980 ratios are .999 and 1.001, respectively.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rate from *WDR 2011* and *WDR 2012*. The 1960-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1970-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates of Sub-Saharan African countries to produce sectoral output shares for 1950 and 1960. We applied the US 1950 capital - sectoral output ratio to produce our 1950 capital value. We used perpetual inventory for 1960-2010.

For 1950 we used the 1954 *UN Statistical Yearbook* for all enrollment rates. For higher education enrollment rate in 1950 we assumed 0, the same value as 1960-1990. Enrollment rates for primary and secondary schools for 1960-2000 come from *WDR* (various years). We assumed the primary school and secondary school ages are 6-11 and 12-16. The tertiary school enrollments are from *WDR* (various years). Our 2010 enrollment rates for all three categories are from *HDR*.

8.22 Kenya (1950-2010)

Population for 1950 and 1990 come from Maddison. Populations for 1962, 1969 and 1979 come from Maa Table A2 p. 15. Population for 2000 comes from *Population Pyramid*, https://populationpyramid.net/kenya/2000/, accessed July 24, 2016. Population for 2010 comes from *Wikipedia*.

The age distributions for 1962, 1969, 1979 and 1990 come from Maa Table A2 p. 15. The age distribution in 1950 and 2000 come from KF. Age distribution for Kenya 2010 comes from the *Demographic Yearbook*. For 2010 we adjusted the 2009 age distribution. We adjust by assuming the same share by age category as in 2009.

Labor force figures for 1962, 1969 and 1979 come from *WDI* (various years). The labor force figures for 1990, 2000 and 2010 come from *WDR*. For 1950 we used the following procedure. We used Clio infra to compute the 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 104.6% for 1950-1962, 103.3% for 1969, 101.7% for 1980, 75.2% for 1990, 71.45% in 2000, 73.55% for 2010. We assumed urban 15-64 labor force participation rates of 104.6% for 1950-1962, 103.3% for 1969, 101.7% for 1980, 50% for 1990, 2000, 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real output come from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1964-1988 from Maa Table J1, p. 1015 and for 1994-1998 from *WDR* (various years). For the 1962 observation we used the average gross capital formation and income for 1957-1961 from Maa Table J1 p. 1015. For 2010 we used the average investment rate for 2000-

2009 from S & H. Mitchell provides sectoral output shares for 1950-2000. We used Nehru and Dhareshwar (1993) for capital values 1950-1990. We used perpetual inventory for years 2000 and 2010.

Enrollments in primary and secondary schools from 1935-1993 come from Maa Table I1 pp. 969 and 975. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1960-1992 come from Maa Table I2 p. 998. For 2010 we used *HDR* for all enrollment rates. We interpolated for our 2000 values. Our years of schooling in the labor force time series for 1950-2010 is: 1.20 (1950), 1.76 (1960), 3.88 (1970), 5.77 (1980), 8.11 (1990), 8.41 (2000) and 8.41 (2010). The Morrisson & Murtin time series of years of schooling is: 1.19 (1950), 1.87 (1960), 2.82 (1970), 4.01 (1980), 5.25 (1990), 6.08 (2000) and 6.46 (2010).

8.23 Lesotho (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Lesotho 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2001 age distribution, and for 2010 we adjusted the 2006 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1990 come from *HDR*. Labor force figures for 1950, 1960, 1970, 1980, 2000 and 2010 come from *WDR* (various years).

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rate from *WDR 2011* and *WDR 2012*. The 1970-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1960-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates of Sub-Saharan African countries to construct our 1950 sectoral shares. We applied the US 1950 and 1960 capital - sectoral output ratios to produce our 1950 and 1960 capital values. We used perpetual inventory for 1970-2010.

For all enrollment rates in 1950, we used the 1954 *UN Statistical Yearbook*. Enrollment rates in primary and secondary schools for 1960-2000 come from *WDR* (various years). We assumed the primary school and secondary school ages are 6-12 and 13-17. The tertiary school enrollments are from *WDR* (various years). Our 2010 enrollment rates come from *HDR*.

8.24 Liberia (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Liberia 2010 comes from the *Demographic Yearbook*. For 2010 we adjusted the 2008 age distribution. We adjust by assuming the same share by age category as in 2008.

Labor force figures for 1960, 1970, 1980, 1990 and 2000 come *WDI*. For 1950 we used Banks for urban-rural population shares for 1960, 1960. We used *WDR* and *WDI* for urban shares for 1970, 1980, 1990, 2000 and 2010. We assumed an 85% labor force participation rate for 15-64 for both urban and rural workers in years 1950-1980. For 1990-2010 we assumed a common 61.25% labor force participation rate.

We computed the ratio of this labor force with that from *WDR* and *WDI*. The root mean ratio is 1.010. The 1960 and 1970 ratios are .991 and 1.030, respectively. The 2010 labor force data come from the *WDR*.

Real GNP come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rate from *WDR 2011* and *WDR 2012*. Physical capital investment rates come from the intraperiod average real gross physical capital and real income for 1965-1989 come from Maa Table J1 1016 and for 1994-1998 from *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. The *WDI* provides sectoral output shares for 1980, 1989, which we used for 1990, and 2010. We interpolated between these for 2000. The 1973 *UN Statistical Yearbook* provides sectoral output shares for 1970. We used Sabillon (2005) information on farming, manufacturing and aggregate output growth rates to produce our 1960 and 1950 sectoral output shares. We used the 1950 and 1960 capital - sectoral output ratios of the US to produce our 1950 and 1960 capital estimates. We used perpetual inventory for 1970-2010.

Enrollments in primary and secondary schools from 1950-1986 come from Maa Table I1 p. 975. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1973-1993 come from Maa Table I2 p. 998. Prior to 1973 we assume an enrollment rates of 0 for higher education. For 1990 and 2000 higher education enrollment rates we used *WDR* various years. For 2010 we used *HDR* for all enrollment rates.

8.25 Madagascar (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. We assumed the same age distribution in 2010 as in 2000.

Labor force figures for 1960, 1970, 1980 and 1990 come from *WDI* (various years). For 2000 and 2010 we used *WDR*. For 1950 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 95% for 1950-2010. We assumed urban 15-64 labor force participation rates of 83% for 1950-1960, 92% for 1970, 88.5% for 1980, 52.75% for 1990, 67.2% for 2000, 77.9% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1980 and 1990 ratios are 1.000 and 1.000, respectively.

Real output come from Maddison. The 1961-1990 investment rates are the intraperiod average investment rate taken from WDI, and the 2000, 2010 investment rates are the intraperiod average investment rates taken from S & H online. The *WDI* provides sectoral output shares for 1970-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1950 and 1960. We used US 1950 and 1960 capital - sectoral output ratios to produce our capital estimates for 1950 and 1960. We used perpetual inventory for years 1970-2010.

Enrollment rates for primary and secondary schools for 1950-2000 come from Maa Table I1 p. 976. We assumed the primary school and secondary school ages are 6-10 and 11-17. The tertiary school enrollments are from Maa Table I2 p. 998. For 2010 we used *HDR* for all enrollment rates. Our time series for years of schooling in the labor force for 1950-2010 is: 1.11 (1950), 1.62 (1960), 2.04 (1970), 3.20 (1980), 4.42 (1990), 4.95 (2000) and 5.24 (2010). The Morrisson & Murtin time series for years of schooling is: .96 (1950), 1.43 (1960), 1.55 (1970), 2.12 (1980), 3.04 (1990), 3.74 (2000) and 4.13 (2010).

8.26 Malawi (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Malawi 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 1998 age distribution, and for 2010 we adjusted the 2008 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1990 come from Maa Table B1 p. 92. The 1980, 2000, 2010 labor force figures come from *WDR*. For 1950, 1960 and 1970 we used the following procedure. We used Clio infra to compute the 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 98.4% for 1950-1980, 104.7% for 1990, 86.45% for 2000, 94.17% for 2010. We assumed urban 15-64 labor force participation rates of 98.4% for 1950-1980, 104.7% for 1990, 86.45% for 2000, 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1980-2010 is 1.000. The 1980 and 1990 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.001.

Real output come from Maddison. Physical capital investment rates come from the intraperiod average real gross physical capital and real income for 1955-1998 from Maa Table J1 p. 1017 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. Mitchell provides sectoral output shares for 1950-2000. We used US 1950 and 1960 capital - sectoral output ratios to produce our capital estimates for 1950 and 1960. We used Nehru and Dhareshwar (1993) for capital stocks in 1970, 1980 and 1990. We used perpetual inventory for years 2000 and 2010.

Enrollments in primary and secondary schools from 1945-1993 come from Maa Table I1 pp. 969 and 976. To calculate enrollment rates, we assumed 6-13 are primary school age and 14-17 are secondary school age. Higher education enrollments for 1966-1990 come from Maa Table I2 p. 998. Prior to 1966 we assume an enrollment rates of .0001 and .0005 for higher education for 1950 and 1960, respectively. For 2010 we used *HDR* for primary and secondary enrollment rates. We assumed 0 for higher education enrollment rate in 2010. Our time series of years of schooling in the labor force for 1950-2010 is: 1.47 (1950), 2.54 (1960), 2.91 (1970), 2.99 (1980), 4.27 (1990), 5.08 (2000) and 6.62 (2010). The Morrisson & Murtin time series of years of schooling is: .40 (1950), .89 (1960), 1.14 (1970), 2.13 (1980), 3.07 (1990), 4.01 (2000) and 4.77 (2010).

8.27 Mali (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Mali 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2009 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1990 come from *WDI*. For 2000 and 2010 we used *WDR*. For 1950, 1960, 1970, and 1980 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 70% for 1950-2000, and 73% for 2010. We assumed urban 15-64 labor force participation rates of 52.1% for 1950-1990,

55.7% for 2000, 67.25% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1990-2010 is 1.000. The 1990 and 2000 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real output come from Maddison. The 1970-2000 investment rates are the intraperiod average investment rate taken from *WDI*. The 2010 investment rate is the average investment rate for 2000-2009 from S & H. The *WDI* provides sectoral output shares for 1970-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1950 and 1960. We used US 1950 and 1960 capital - sectoral output ratios to construct 1950 and 1960 capital values. We used perpetual inventory for 1970-2010.

Enrollments in primary and secondary schools for 1950-2000 come from Maa Table I1 p. 976. We assumed the primary school and secondary school ages are 6-11 and 12-17. The tertiary school enrollments are from Maa Table I2 p. 999. We used *HDR* for all 2010 enrollment rates. Our time series of years of schooling in the labor force for 1950-2010 is: .14 (1950), .26 (1960), .54 (1970), 1.50 (1980), 1.86 (1990), 1.71 (2000), 4.41 (2010). The Morrisson & Murtin time series of years of schooling is: .15 (1950), .22 (1960), .31 (1970), .71 (1980), .96 (1990), 1.18 (2000) and 1.51 (2010).

8.28 Mauritania (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Population Pyramid*, https://populationpyramid.net/mauritania/2000/, accessed July 24, 2016. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Mauritania 2010 comes from the *Demographic Yearbook*. For 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in 2013.

Labor force figures for 1990 come from *WDI*. For 2000 and 2010 we used *WDR*. For 1950, 1960, 1970, and 1980 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 52.25% for 1950-1990, 60% for 2000 and 78.6% for 2010. We assumed urban 15-64 labor force participation rates of 50% for 1950-1990, 51.5% for 2000, 60% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1990-2010 is 1.000. The 1990 and 2000 ratios are 1.000 and 1.000, respectively. The range of the ratios is .999 to 1.000.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. The 1970-2000 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The 2010 investment rate is the average investment rate for 2000-2009 from S & H. The *WDI* provides sectoral output shares for 1960-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1950. We used US 1950 and 1960 capital - sectoral output ratios to produce our 1950 and 1960 capital values. We used perpetual inventory to produce our 1970-2010 values.

Enrollments in primary and secondary schools for 1950-2000 come from Maa Table I1 p. 976. We assumed the primary school and secondary school ages are 6-11 and 12-17. The tertiary school enrollments are from Maa Table I2 p. 999. For all enrollment rates in 2010 we used *HDR*.

8.29 Mauritius (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Mauritius 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1960, 1970 and 1990 come from Maa Table B1 p. 92. Labor force figures for Mauritius for 1980 are interpolated from 1970 and 1990 values. The labor force for 2000 comes from the average labor force participation rate of 1980 and 1990. For 2010 labor force we used WDR. For 1950 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for 1960-2010. We assumed rural 15-64 labor force participation rates of 57.79% for 1950-1960, 47.85% for 1970, 51% for 1980, 71.15% for 1990, 67.8% for 2000, 75.6% for 2010. We assumed urban 15-64 labor force participation rates of 50% for all years except 1970, where we used 47.85%. We constructed the ratio of this labor force to that from *WDR*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.001.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. For 1960-2010 we used the average investment rate for 1951-2009 from S & H. Mitchell provides sectoral output shares for 1950-2000. We used Nehru and Dhareshwar (1993) for capital stock values for 1950-1990. We used perpetual inventory for 2000 and 2010.

For 1950 enrollment rates in primary, secondary and higher education, we used the 1954 *UN Statistical Yearbook*. For all enrollment rates in 1960, we used the 1963 *UN Statistical Yearbook*. Enrollments in primary and secondary schools from 1970-1993 come from Maa Table I1 p. 976. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-18 are secondary school age. Higher education enrollments for 1967-1991 come from Maa Table I2 p. 999. For all 2010 enrollment rates we used *HDR*. We interpolated for our 2000 enrollment rates.

8.30 Mozambique (1950-2010)

Population for 1950 and 1990 come from Maddsion. Populations for 1960, 1970 and 1980 come from Maa Table A1 p. 4. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1960, 1970, 1980 and 1990 come from Maa Table A2 p. 16. The age distribution in 1950 is assumed to be identical to the age distribution in 1960. The age distribution for 1990 is interpolated from 1980 and 1995 values. Age distributions for Mozambique 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2007 age distribution. We adjust by assuming the same share by age category as in 2007.

Labor force figures for 1980 come from Maa Table B1 p. 93. The labor force data for 1990, 2000 and 2010 come from *WDR*. For 1950-1970 we used the following procedure. We used Clio infra to compute the 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 105.95% for 1950-1980, 73% for 1990, 87.33% for 2000 and, 93.9% 2010. We

assumed urban 15-64 labor force participation rates of 105.95% for 1950-1980, 50% for 1990, 71.25% for 2000 and 91.3% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1980-2010 is 1.000. The 1980 and 1990 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real output come from Maddison. Physical capital investment rates for 1970-2010 come from the intraperiod average investment rate from S & H online and *WDR* (various years). The *WDR 1982* provides sectoral output shares for 1960 and 1980. The *WDI* provides sectoral output shares for 1990-2000. We interpolated for 1970. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates for Sub-Saharan African countries to produce 1950 sectoral output shares. We used Nehru and Dhareshwar (1993) for 1950-1990 capital values. We used perpetual inventory to construct 2000 and 2010 values.

Enrollments in primary and secondary schools from 1950-1992 come from Maa Table I1 p. 977. To calculate enrollment rates, we assumed 6-10 are primary school age and 11-17 are secondary school age. Higher education enrollments for 1976-1992 come from Maa Table I2 p. 999. Prior to 1976 we assume an enrollment rate of 0 for higher education. For 2010 enrollment rates we used *HDR*. We interpolated for our 2000 enrollment rates. Our time series for years of schooling in the labor force for 1940-2010 is: .26 (1940), .44 (1950), .59 (1960), 1.43 (1970), 1.85 (1980), 4.02 (1990), 3.80 (2000) and 4.70 (2010). The Morrisson & Murtin time series of years of schooling is: .19 (1940), .24 (1950), .45 (1960), .80 (1970), 1.05 (1980), 2.03 (1990), 2.39 (2000) and 2.68 (2010).

8.31 Namibia (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Namibia 2010 comes from the *Demographic Yearbook*. For 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in 2011.

Labor force figures for 1990 come from *WDI*. For 2000 and 2010 we used *WDR*. For 1950, 1960, 1970, and 1980 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 64.25% for 1950-1980 (which is the average 15-64 rural labor force participation for 1990-2010), 47% for 1990, 65.75% for 2000 and 80% for 2010. We assumed urban 15-64 labor force participation rates of 50% for 1950-1980, 45.75% for 1990, 50% for 2000, 55.9% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1990-2010 is 1.001. The 1990 and 2000 ratios are 1.000 and 1.001, respectively.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 value from Maddison and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. The 1970-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1980-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct 1950-1970 sectoral shares. We used US 1950 and 1960 capital - sectoral output ratios to construct our 1950 and 1960 values. We used perpetual inventory to compute our 1970-2010 capital stocks.

For all three enrollment rates in 1950, we used the 1954 *UN Statistical Yearbook*, using the Southwest

Africa aggregate. The 1960 secondary enrollment rate comes from the 1963 *UN Statistical Yearbook*, using the Southwest Africa aggregate and one-tenth of the primary and secondary school enrollment aggregate. The 1960-1980 primary enrollment rates, higher education enrollment rates and 1970-1980 secondary school enrollment rates are interpolated. We assumed the primary school and secondary school ages are 6-12 and 13-17. Our 1990 enrollment rates are from *WDR*. All 2010 enrollment rates are from *HDR*. We interpolated our 2000 enrollment rates.

8.32 Niger (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Niger 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2008 age distribution. We adjust by assuming the same share by age category as in 2008.

Labor force figures for 1960, 1970, 1980 and 1990 come from *WDI* (various years). For 2000 and 2010 we used *WDR*. For 1950 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 93.5% for 1950-1960, 100% for 1970, 1980, 61.5% for 1990 71% for 2000, and 72% for 2010. We assumed urban 15-64 labor force participation rates of 83.5% for 1950-1960, 92% for 1970, 100% for 1980, 50% for 1990, 52% for 2000, 53.25% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is .999. The 1960 and 1970 ratios are 1.000 and 1.000, respectively.

Real output come from Maddison. The 1970-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1960-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares in 1950. We applied 1950 and 1960 US capital - sectoral output ratios to produce our 1950 and 1960 estimates. We used perpetual inventory to produce capital stock values for 1970-2010.

Enrollments in primary and secondary schools for 1950-2000 come from Maa Table I1 p. 977. We assumed the primary school and secondary school ages are 6-11 and 12-18. The tertiary school enrollments are from Maa Table I2 p. 999. For 2010 we used *HDR* for all enrollment rates. Our years of schooling in the labor force for 1950-2010 is: .06 (1950), .08 (1960), .19 (1970), .53 (1980), 1.34 (1990), 1.95 (2000) and 2.54 (2010). The Morrisson & Murtin years of schooling time series is: .04 (1950), .07 (1960), .14 (1970), .39 (1980), .76 (1990), 1.03 (2000) and 1.31 (2010).

8.33 Nigeria (1950-2010)

Populations for 1950, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*. Population for 1963 comes from Maa Table A2 p. 17

The age distribution for 1963 come from Maa Table A2 p. 17. We assumed the 1950 age distribution is identical to the 1963 age distribution. The age distributions for Nigeria for 1970, 1980, 1990 come from DK. Age distribution for Nigeria 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted

the 2006 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for Nigeria for 1963, 1972, 1980, 1990, 2000 and 2010 come from *WDR* (various years). For 1950 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for 1960-2010. We assumed rural 15-64 labor force participation rates of 75.911% for 1950-1963, 88.35% for 1972, 93.06% for 1980, 70.98% for 1990, 60.98% for 2000 and 57.78% for 2010. We assumed urban 15-64 labor force participation rates of 50% for 1950-2010. We constructed the ratio of this labor force to that from *WDR*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real output come from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1951-1998 from Maa Table J1, p. 1019 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. Mitchell provides sectoral output shares for 1950-2000. We used US 1950 and 1960 capital - sectoral output ratios to produce capital estimates for 1950 and 1960. We used Nehru and Dhareshwar (1993) for capital stocks in 1972-1990. We used perpetual inventory for 2000 and 2010.

Enrollments in primary and secondary schools from 1944-1993 come from Maa Table I1 p. 970 and 977. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1964-1993 come from Maa Table I2 p. 999. We used .0005 for the higher education enrollment rate for 1952. These fit the education shares from Morrisson & Murtin. For 2010 we used *HDR*. All 2000 enrollment rates are interpolated. Our years of schooling in the labor force time series for 1950-2010 is: .67 (1950), .95 (1963), 1.65 (1972), 3.77 (1980), 5.36 (1990), 5.45 (2000) and 6.18 (2010). The Morrisson & Murtin years of schooling time series is: .38 (1950), 1.06 (1960), 1.30 (1970), 1.47 (1980), 2.72 (1990), 3.99 (2000) and 4.48 (2010).

8.34 Reunion (1950-2010)

Populations for 1950, 1960, 1980 and 1990 come from Maddison. Population for 1970 comes from *Population Pyramid*, https://populationpyramid.net/runion/1970/, accessed July 24,2016. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Reunion 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 1999 age distribution. We adjust by assuming the same share by age category as in 1999.

Labor force figures for 1970, 1980 come from, *World Tables of economic and social indicators, 1950-1981*. For 1990, 2000 and 2010 we used *WDI*. For 1950, 1960 we used the following procedure. We used WorldoMeters to compute the 1950-2010 urbanization rates. For 1950 we used the growth rate from 1955 to 1960 to extrapolate back to 1950. We assumed rural 15-64 labor force participation rates of 95% for 1950-1980, 88.5% for 1990, 2000, and 2010. We assumed urban 15-64 labor force participation rates of 70.875% for 1950-1970, 78% for 1980, 50% for 1990, 60.95% for 2000, 64.7% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1970-2010 is 1.000. The 1980 and 1990 ratios are .999 and 1.000, respectively. The range of the ratios is .999 to 1.000.

Real output prior to 2000 come from Maddison. The 2000 figure comes from *WDR*. The 2010 figure comes from the *CIA World Factbook 2011-2012*. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1970-1990 from S & H online and *WDR* (various

years). The 1981 and 1996 *UN Statistical Yearbooks* provide sectoral output shares for 1970 and 1990, we interpolated 1980. We used information from Sabillon (2005) on Sub-Saharan African countries farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1950, 1960 and 2000. We used US 1950-1990 US capital - sectoral output ratios to construct our estimates for 1950-2000. We used perpetual inventory for 2010.

For 1950 enrollment rates in all three categories we used the 1954 *UN Statistical Yearbook*. For 1960 enrollment rates in all three categories we used the 1963 *UN Statistical Yearbook*. Enrollments in secondary schools from 1970, 1980 and 1990 come from UNESCO *UIS Global Database*. Enrollments in primary schools from 1980 and 1990 come from UNESCO *UIS Global Database*. We interpolated our 1970 primary school enrollment rate. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollment rates for 1960-2000 are interpolated. Our 2000 and 2010 enrollment rates for primary and secondary schools come from the 2003 *Europa World Yearbook* and 2012/2013 *Europa World Yearbook*.

8.35 Rwanda (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Rwanda 2010 comes from the *Demographic Yearbook*. For 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in 2012.

Labor force figures for 1960, 1970, 1980 and 1990 come from *WDI* (various years). For 2000 and 2010 we used *WDR*. For 1950 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for 1960-2010. We assumed rural 15-64 labor force participation rates of 127.1% for 1950-1960, 128.6% for 1970, 138.5% for 1980, 129.7% for 1990, 150.1% for 2000, and 88.3% for 2010. We assumed urban 15-64 labor force participation rates of 127.1% for 1950-1960, 128.6% for 1970, 138.5% for 1980, 129.7% for 1990, 150.1% for 2000, 88.3% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 value from Maddison and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1970-2009 from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1970-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1950 and 1960. We applied the US 1950-1990 capital - sectoral output ratios to construct our 1950-1990 estimates. We used perpetual inventory for 2000 and 2010.

For 1950 we used the 1954 *UN Statistical Yearbook* for all enrollment rates. Enrollments in primary and secondary schools from 1960-1991 come from Maa Table I1 p. 977. The 1960 observations for primary and secondary school enrollment rates are 49% and 2%, respectively. To calculate enrollment rates, we assumed 6-13 are primary school age and 14-19 are secondary school age. Higher education enrollments for 1967-1990 come from Maa Table I2 p. 999. We used 0 for the higher education enrollment rate prior to 1967. For 2010 we used *HDR* for all enrollment rates. We interpolated for 2000 enrollment rates.

8.36 Senegal (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Senegal 2010 comes from the *Demographic Yearbook*. For 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in 2013.

Labor force figures for 1960, 1970, 1980 and 1990 come from *WDI* (various years). For 2010 we used WDR. For 1950 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for 1960-2010. We assumed rural 15-64 labor force participation rates of 95% for 1950-2010. We assumed urban 15-64 labor force participation rates of 56% for 1950-1960, 85.9% for 1970, 83.7% for 1980, 58.2% for 1990, 56.35% for 2000, 51.8% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real output come from Maddison. Physical capital investment rates come from the intraperiod average gross capital formation and income for 1968-1969 and the intraperiod average gross real capital formation and real income for 1970-1998 from Maa Table J1, p. 1020 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. We used Nehru and Dhareshwar (1993) for capital estimates 1950-1990. We used perpetual inventory for 2000 and 2010.

Enrollments in primary and secondary schools from 1950-1992 come from Maa Table I1 p. 977. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-18 are secondary school age. Higher education enrollments for 1949-1993 come from Maa Table I2 p. 999. For 2010 we used *HDR* for all enrollment rates. We interpolated for 2000 enrollment rates. Our years of schooling in the labor force time series for 1950-2010 is: .73 (1950), .67 (1960), 1.22 (1970), 2.37 (1980), 3.13 (1990), 4.09 (2000) and 4.16 (2010). The Morrisson & Murtin years of schooling time series is: .22 (1950), .39 (1960), .61 (1970), 1.26 (1980), 1.94 (1990), 2.57 (2000), and 2.99 (2010).

8.37 Seychelles (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Seychelles 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 1998 age distribution, and for 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1960, 1970, 1980 come from *World Tables of Economic and Social Indicators, 1950-1981*. For 1990, 2000 and 2010 we used the indexmundi. For 1950 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for 1960-2010. We assumed rural 15-64 labor force participation rates of 81.75% for 1950-1960, 90% for 1970, 1980, 84.5% for 1990, 90% for 2000 and 81.1% for 2010. We assumed urban 15-64 labor force participation rates of 50% for 1950-1960, 50.8% for 1970, 60.7% for 1980, 50% for 1990, 54.6% for 2000, 50% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively. The range of the ratios is .999 to 1.000.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1970-2010 from S & H online and *WDI* (various years). The *WDI* provides sectoral output shares for 1980-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates for Sub-Saharan African countries on the *WDI* 1976 sectoral shares to produce sectoral output shares for 1950-1970. We used US 1950 and 1960 capital - sectoral output ratios for construct our 1950 and 1960 estimates. We used perpetual inventory for 1970-2010.

The 1950 enrollment rates come from the 1954 *UN Statistical Yearbook*. Enrollments in primary and secondary schools from 1960-1990 come from UNESCO. The 1970-2010 primary school and secondary school enrollment rates come from UNESCO *UIS Global Database*. The 2010 higher education enrollment rate comes from *HDR*. We interpolated higher education enrollment rates from 1960-2000. We interpolated our 1960 primary school and secondary school enrollment rates. To calculate enrollment rates, we assumed 6-13 are primary school age and 14-17 are secondary school age.

8.38 Sierra Leone (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Sierra Leone 2010 comes from the *Demographic Yearbook*.

Labor force figures for 1960, 1970, 1980 and 1990 come from *WDI* (various years). For 2000 and 2010 we used *WDR*. For 1950 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 77.8% for 1950-1961, 80.25% for 1970, 80.7% for 1980, 77.25% for 1990, 66.4% for 2000, and 84% for 2010. We assumed urban 15-64 labor force participation rates of 50% for 1950-2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are .999 and 1.001, respectively. The range of ratios is .999 to 1.001.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. The 1960-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1970-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce our 1950-1961 sectoral output shares. We used the US 1950 capital - sectoral output ratios to construct our 1950 value. We used Nehru and Dhareshwar (1993) for capital in 1961-1990. We used perpetual inventory for 2000 and 2010.

For all enrollment rates in 1950 we used the 1954 *UN Statistical Yearbook*. Enrollments in primary and secondary schools for 1960-2000 come from Maa Table I1 p. 977. We assumed the primary school and secondary school ages are 6-12 and 13-19. The tertiary school enrollments are from Maa Table I2 p. 999. For 2010 we used *HDR*. Our years of schooling in the labor force time series for 1950-2010 is: .46 (1950), .70 (1960), 1.36 (1970), 2.34 (1980), 4.23 (1990), 4.45 (2000) and 8.56 (2010). The Morrisson & Murtin years of schooling time series is:.38 (1950), .80 (1960), 1.15 (1970), 2.01 (1980), 2.92 (1990), 3.63 (2000) and 3.99 (2010).

8.39 Somalia (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Somalia 2000 comes from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution. We assumed the same age distribution in 2010 as in 2000.

Labor force figures for 1990 come from *WDI* . For 2000 and 2010 we used *WDR*. For 1950, 1960, 1970 and 1980 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 55% for 1950-1990, 65% for 2000 and 2010. We assumed urban 15-64 labor force participation rates of 50% for 1950-1990, 54.4% for 2000, 58.3% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1990-2010 is 1.000. The 1990 and 2000 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. The 1970-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). *UN Statistical Yearbooks* provides sectoral output shares for 1960 and 1970. The *WDI* provides sectoral output shares for 1980 and 1990. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates for Sub-Saharan African countries to construct 1950 sectoral output shares. We used US 1950 and 1960 capital - sectoral output ratios to construct our 1950 and 1960 estimates. We used perpetual inventory for 1970-2010.

The 1950 enrollment rates for all categories come from the 1954 *UN Statistical Yearbook*. Enrollments in primary and secondary schools for 1960-1990 come from Maa Table I1 p. 978. We assumed the primary school and secondary school ages are 6-11 and 12-18. The tertiary school enrollments are from Maa Table I2 p. 1000. For primary school enrollment rate in 2010 we used WDR. We interpolated for our 2000 primary school enrollment rate. We assumed the 2000 and 2010 higher education enrollment rates (both .25%) was essentially unchanged from the 1990 value, (.24%). The secondary school enrollment rates for 2000 and 2010 were computed from the average ratio of secondary to primary school enrollment rates for 1950-1990.

8.40 South Africa (1800-2010)

Population for 1800, 1810, 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1910, 1920, 1930, 1940 come from McEvedy and Jones. Population for 1951 is geometrically interpolated from 1946 and 1960 from Maa Table A2 p. 17. Populations for 1913, 1920, 1930, 1940, 1950, 1960, 1970 and 1980 comes from Maa Table A2 p. 17. Population for 1990 comes from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distribution for 1951 is geometrically interpolated from 1946 and 1960 from Maa Table A2 p. 17. The age distributions for South Africa for 1960, 1970, 1980 and 1990 come from Maa Table A2 p. 17. We used the age distribution in 1951 for all years prior to 1951. The age distribution for South Africa 2000 comes from KF. Age distribution for South Africa 2010 comes from the *Demographic Yearbook*. For 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in 2013.

Labor force figures for South Africa 1913, 1920, 1930, 1940, 1946 are from Maa Table B1 p. 93. The

1920, 1930 and 1940 observations are interpolations. Labor force figures for 1951, 1960, 1970, 1980 and 1990 are from Feinstein. Labor force figures for South Africa for 2000 and 2010 come from *WDR* (various years). For years prior to 1911, we used the following procedure. We used Clio Infra for urbanization rates for 1800-1950, and *WDR* and *WDI* for 1960-2010. We interpolated between 1700 and 1850 for 1800-1850, and between 1850 and 1900 for 1860-1890. We assumed a rural 15-64 labor force participation rate of 103.5% for 1820-1913, 96% for 1920, 90% for 1930, 85.5% for 1940, 81.5% for 1946, 75% for 1951, 1960, 70% for 1970, 71% for 1980, 70% for 1990, 68% for 2000, and 60% for 2010. We assumed an urban 15-64 labor force participation rate of 103.5% for 1820-1913, 96% for 1920, 79% for 1930, 60% for 1940, 58% for 1946, 54% for 1951, 65% for 1960, 64.5% for 1970, 70% for 1980, 67.75% for 1990, 64% for 2000, and 52% for 2010. We constructed the ratio of this labor force to that from *WDI* and *WDR*. The root mean ratio for the overlapping years 1913-2010 is 1.000. The 1913 and 1920 values are 1.003 and .994, respectively. The range of the ratio is .994 to 1.002.

Real output come from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1953-1992, 1994-1998 and the intraperiod average gross capital formation and income from 1950-1952 from Maa Table J1, p. 1021 and *WDR* (various years). We used the average investment rate for years 1950-2000 for all years prior to 1951. For 2010 we used the average investment rate for 2000-2009 from S & H. The *WDI* provides sectoral output shares for 1960-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1800-1951. We used US capital - sectoral output ratios for 1850-1940, applying 1850 for 1800-1840, to construct our capital estimates for 1800-1940. We used perpetual inventory for 1951-2010.

For years 1870-1940 we used Benavot Riddle (1988). For 1800-1820, we assumed that the primary enrollment rate was constant at 6%, and from 1830-1860 it was 75 percent of the succeeding decade's primary enrollment rate. This produces the following time series for primary school enrollment rates: 6% (1800-1820), 7.8% (1830), 10.4% (1840), 13.9% (1850), 18.5% (1860), 25% (1870 datum). For years 1800-1840 we assumed that the secondary enrollment rate was constant at .1%. For years 1850-1900 we assumed the enrollment rate in secondary school was half of the succeeding decade value. This produces the following time series for secondary school enrollment rates: .1% (1800-1840), .14% (1850), .29% (1860), .58% (1870), 1.2% (1880), 2.3% (1890), 4.6% (1900). Enrollments in secondary schools from 1914-1993 come from Maa Table I1 pp. 971 and 978. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Prior to 1960 we assumed that 75 percent of European students were primary students, 80 percent of Colored and Asian students were primary students and 95 percent of Native students were primary students. From 1960 onward we assumed that two thirds of European students were primary students, 75 percent of Colored and Asian students were primary students and 90 percent of Native students were primary students. Our 1950 enrollment rates for all categories are from the 1954 *UN Statistical Yearbook*. Our 1960 and 1970 enrollment rates for all categories come from Mitchell. Our 1920-1940 higher education and secondary school enrollment rates are interpolated. For years prior to 1913, 1800-1840 we assumed 0 enrollment rate in high education, and for 1850-1900 we assumed each decade's higher education enrollment rate was 75 percent of the succeeding decade's higher education enrollment rate. These match our data in 1970 and 1996. Higher education enrollments for 1934-1993 come from Maa Table I2 pp. 996 and 1000. For 1980 and 1990 enrollment rates we used *WDI*. For 2010 we used *HDR* for all enrollment rates. We interpolated our 2000 enrollment rates. Our years of schooling in the

labor force for 1870-1920 is: .91 (1870), 1.14 (1880), 1.77 (1890), 1.34 (1900), 1.29 (1913) and 1.51 (1920). The Morrisson & Murtin time series of years of schooling is: 1.10 (1870-1900), 1.12 (1910) and 1.44 (1920).

8.41 Sudan (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. We assumed the same age distribution in 2010 as in 2000.

Labor force figures for 1960, 1970, 1980 and 1990 come from *WDI* (various years). For 2000 and 2010 we used *WDR*. For 1950 we used the following procedure. We used Worldometers to compute the 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. Worldometers provides 1955 and 1960 values of urbanization, and the 1960 value fits with the *WDI* value. We assumed the same 5 year rate of change from 1950 to 1955 as from 1955 to 1960. We assumed rural 15-64 labor force participation rates of 84.75% for 1950-1960, 72.1% for 1970, 72.65% for 1980, 60.6% for 1990, 54.975% for 2000 and 61.25% for 2010. We assumed urban 15-64 labor force participation rates of 50% for 1950-2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively. The range of ratios is 1.000 to 1.000.

Real output come from Maddison. For 1960 and 1970 investment rates are the intraperiod average investment rate taken from Maa Table J1 p. 1020. The 1980-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1960-2010. We used information on farming, manufacturing and aggregate growth rates for Sub-Saharan African countries to construct our 1950 sectoral output shares. We used the US 1950-1990 capital - sectoral output ratios to construct our capital estimates for 1950-1990. We used perpetual inventory to create our 2000 and 2010 values.

Our 1950 enrollment rates for all categories come from the 1954 *UN Statistical Yearbook*. Enrollments in primary and secondary schools for 1960-2000 come from Maa Table I1 p. 978. We assumed the primary school and secondary school ages are 6-11 and 12-17. The tertiary school enrollments are from Maa Table I2 p. 1000. For 2010 we used *HDR*. Our time series of years of schooling in the labor force for 1950-2010 is: .36 (1950), .52 (1960), 1.06 (1970), 2.02 (1980), 3.06 (1990), 3.89 (2000) and 4.92 (2010). The Morrisson & Murtin years of schooling time series is: .52 (1950), 1.08 (1960), 1.42 (1970), 2.09 (1980), 2.51 (1990), 2.87 (2000) and 3.08 (2010).

8.42 Swaziland (1950-2010)

Population for 1950 comes from Maddison. Populations for 1960, 1970, 1980, 1990, 2000 and 2010 come from *Population Pyramid*, https://populationpyramid.net/swaziland/, accessed July 24, 2016.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Swaziland 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in 2013.

Labor force figures for 1960, 1970, 1980, 1990, 2000 come from *WDI* (various years). For 2010 we used *WDR*. For 1950 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI*

for 1960-2010. We assumed rural 15-64 labor force participation rates of 67.6% for 1950-1960, 69.8% for 1970, 68.35% for 1980, 60.7% for 1990, 58.65% for 2000 and 59.25% for 2010. We assumed urban 15-64 labor force participation rates of 50% for 1950-2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.001.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. The 1980-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides 1970-2010 sectoral output shares. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1950 and 1960. We used the US 1950 and 1960 capital - sectoral output ratios to create our 1950 and 1960 capital estimates. We used perpetual inventory for 1970-2010.

Our 1950 enrollment rates are from the 1954 *UN Statistical Yearbook*. Enrollments in primary and secondary schools for 1960-2000 we used *WDR* various years. Our 1970-2000 higher education enrollment rates are from *WDI*. We interpolated our 1960 higher education enrollment rate. Our 2010 enrollment rates are from *HDR*. We assumed the primary school and secondary school ages are 6-12 and 13-17.

8.43 Tanzania (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Tanzania 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in 2013.

Labor force figures for 1960, 1970, 1980 and 1990 come from *WDI* (various years). For 2000 and 2010 we used *WDR*. For 1950 we used the following procedure. We used Clio infra to compute the 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 109.75% for 1950-1960, 107.35% for 1970, 103.75% for 1980, 97.65% for 1990, 95% for 2000 and 2010. We assumed urban 15-64 labor force participation rates of 109.75% for 1950-1960, 107.35% for 1970, 103.75% for 1980, 97.65% for 1990, 63% for 2000 and 88.8% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively.

Real output come from Maddison. The 1970-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). We used Nehru and Dhareshwar (1993) for 1950-1990 capital estimates. We used perpetual inventory for 2000 and 2010.

Our 1950 enrollment rates come from the 1954 *UN Statistical Yearbook*. Enrollments in primary and secondary schools for 1960-1990 come from Maa Table I1 p. 978. We assumed the primary school and secondary school ages are 6-12 and 13-18. The tertiary school enrollments are from Maa Table I2 p. 1000. For 2010 we used *HDR*. We interpolated our enrollment rates for all categories in 2000.

8.44 Togo (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Togo 2010 comes from the *Demographic Yearbook*.

Labor force figures for 1960, 1970, 1980 and 1990 come from *WDI* (various years). For 2000 and 2010 we used *WDR*. For 1950 we used the following procedure. We used Clio infra to compute the 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 89.5% for 1950-1960, 93.5% for 1970, 94% for 1980, 87.25% for 1990, 95% for 2000 and 100% for 2010. We assumed urban 15-64 labor force participation rates of 50% for 1950-2000, 55.5% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. The 1970-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1960-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates of Sub-Saharan African countries to construct our 1950 sectoral output shares. We used the US 1950 and 1960 capital - sectoral output ratios to create our 1950 and 1960 capital values. We used perpetual inventory for 1970-2010.

Our 1950 enrollment rates are from 1954 *UN Statistical Yearbook*. Enrollments in primary and secondary schools for 1960-1990 come from Maa Table I1 p. 978. We assumed the primary school and secondary school ages are 6-11 and 12-18. The tertiary school enrollments are from Maa Table I2 p. 1000. For 2010 we used *HDR*. We interpolated our enrollment rates in 2000.

8.45 Uganda (1950-2010)

Population for 1950 and 1990 come from Maddison. Populations for 1959, 1969, 1980 come from Maa Table A1 p. 6. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1959, 1969 and 1990 come from Maa Table A2 p. 18. The 1990 value is interpolated from the 1969 and 1991 values. The age distribution in 1950 is assumed to be identical to the age distribution in 1959. The age distributions for Uganda for 1980 and 2000 come from KF. Age distribution for Uganda 2010 comes from the *Demographic Yearbook*. For 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in 2012.

Labor force figures for 1960, 1970, 1980 and 1990 come from *WDI* (various years). For 2000 and 2010 we used *WDR*. For 1950 we used the following procedure. We used Clio infra to compute the 1950 urbanization rates, and *WDI* for urbanization rates from 1960-2010. We assumed rural 15-64 labor force participation rates of 95% for 1950-1959, 105.9% for 1969, 115.45% for 1980, 90% in 1990, 85% for 2000 and 2010. We assumed urban 15-64 labor force participation rates of 64.5% for 1950-1959, 105.9% for 1969, 115.45% for 1980, 68% for 1990 and 65.25% for 2000, 65.25% for 2010. We constructed the ratio of this labor force to that from *WDR* and *WDI*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1980 and 1990 ratios are 1.000 and 1.000, respectively.

Real output come from Maddison. Physical capital investment rates come for 1959-2010 from the

intraperiod average investment rate from S & H online and *WDR* (various years). Mitchel provides sectoral output shares for 1959-2000. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates of Sub-Saharan African countries to construct sectoral output shares for 1950. We used US 1950, 1960 and 1970 capital - sectoral output ratios to construct our 1950, 1960 and 1969 capital estimates. We used Nehru and Dhareshwar (1993) for 1980 and 1990 capital values. We used perpetual inventory for 2000 and 2010.

Our 1950 enrollment rates are from the 1954 *UN Statistical Yearbook*. Enrollments in primary and secondary schools from 1960-1993 come from Maa Table I1 p. 979. To calculate enrollment rates, we assumed 6-12 are primary school age and 13-18 are secondary school age. Higher education enrollments for 1965-1992 come from Maa Table I2 p. 1000. For 1960 we interpolated for our enrollment rate of higher education. For 2010 we used *HDR*. We interpolated for our enrollment rates in 2000. Our time series of years of schooling in the labor force for 1950-2010 is: 1.10 (1950), 1.60 (1960), 2.26 (1970), 2.65 (1980), 3.65 (1990), 5.22 (2000) and 7.04 (2010). The Morrisson & Murtin time series of years of schooling is: .93 (1950), 1.20 (1960), 1.81 (1970), 2.18 (1980), 2.62 (1990), 3.33 (2000) and 4.54 (2010).

8.46 Zaire (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. We assumed the same age distribution in 2010 as in 2000.

Labor force figures for 1960, 1970, 1980, 1990, 2000 and 2010 come from *WDR* (various years). For 1950 we used UN (1969) for urban-rural population shares for 1950, 1960, and *WDR* and *WDI* for 1970, 1980, 1990, 2000 and 2010. We assumed a rural 15-64 labor force participation rate of 90% for 1950-1980, and 85% for 1990, and 75% for 2000-2010. We assumed an urban 15-64 labor force participation rate of 50% for 1950-2000, and 60% for 2010. We constructed the ratio of this labor force to that from WDR and *WDI*. The root mean ratio is .998 for the overlapping years 1960-2010. The 1960 and 1970 ratios are 1.002, and .948, respectively. The range of values is .948 to 1.015.

Real output come from Maddison. The 1960-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1970-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates of Sub-Saharan Africa countries applied to the *WDI* data on 1965 sectoral output shares to construct our 1950 and 1960 sectoral output shares. We used US 1950-1980 capital - sectoral output ratios to construct our 1950-1980 capital values. We used Nehru and Dhareshwar (1993) for 1990 capital. We used perpetual inventory for 2000 and 2010.

Our 1950 enrollment rates are from the 1954 *UN Statistical Yearbook*. Enrollments in primary and secondary schools for 1960-2000 come from Maa Table I1 p. 979. We assumed the primary school and secondary school ages are 6-11 and 12-17. The tertiary school enrollments are from Maa Table I2 p. 1000. For 2010 we used *HDR*.

8.47 Zambia (1950-2010)

Population for 1950 comes from Maa Table A1 p. 6. Populations for 1960 and 1990 come from S & H online. Populations for 1969 and 1980 comes from Maa Table A2 p. 18. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Malawi 2000 and 2010 come from the *Demographic Yearbook*.

Labor force figures for 1969 and 1980 come from Maa Table B1 p. 94. The 1960, 1990, 2000 and 2010 labor force data come from *WDR* (various years). For 1950 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for 1960-2010. We assumed rural 15-64 labor force participation rates of 95% for 1950-2010. We assumed urban 15-64 labor force participation rates of 63.45% for 1950-1960, 73.5% for 1969, 83.2% for 1980, 63.85% for 1990, 72.1% for 2000, 51.1% for 2010. We constructed the ratio of this labor force to that from *WDR*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real output come from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1954-1990 and average intraperiod gross capital formation and income for 1950-1953 and 1991-1998 from Maa Table K1, p. 1024 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. Mitchell provides sectoral output shares for 1960-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates of Sub-Saharan African countries to produce our 1950 sectoral output shares. We used Nehru and Dhareshwar (1993) for capital estimates for 1950-1990. We used perpetual inventory for 2000 and 2010.

Our 1950 enrollment rates are from the 1954 *UN Statistical Yearbook* for Northern Rhodesia. Our 1960 enrollment rates are from the 1963 *UN Statistical Yearbook* for Northern Rhodesia. Enrollments in primary and secondary schools from 1970-1993 come from Maa Table I1 pp. 972 and 979. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1967-1990 come from Maa Table I2 p. 1000. For 2010 we used *HDR*. We interpolated for our enrollment rates in 2000.

8.48 Zimbabwe (1950-2010)

Population for 1950 comes from Maa Table A1 p. 6. Populations for 1960 and 1990 come from S & H online. Populations for 1969 and 1980 comes from Maa Table A2 p. 18. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1969, 1980 and 1988 come from Maa Table A2 p. 18. The 1980 and 1990 age distributions are interpolated from the 1969, 1982 and 1994 values. The age distribution for 1960 is from DK. The age distribution in 1950 is assumed to be identical to the age distribution in 1960. Age distribution for Zimbabwe 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 1999 age distribution, and for 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in the reference year.

The labor force figures for 1960, 1969, 1980, 1990, 2000 and 2010 come from *WDR* (various years). For 1950 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for 1960-2010. We assumed rural 15-64 labor force participation rates of 83.6% for 1950-1960, 96.5% for 1969,

1980, 84% for 1990, 95% for 2000 and 2010. We assumed urban 15-64 labor force participation rates of 50% for 1950-1960, 51.5% for 1970, 64.8% for 1980, 50% for 1990, 75% for 2000, 81.4% for 2010. We constructed the ratio of this labor force to that from *WDR*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real output come from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1954-2000 from Maa Table J1 p 1024. For 2010 we the average investment rate for years 2000-2009 from S & H. Mitchell provides sectoral output shares for 1950-2000. We used Nehru and Dhareshwar (1993) for capital values for 1950-1990. We used perpetual inventory for years 2000 and 2010.

Our 1950 enrollment rates come from the 1954 *UN Statistical Yearbook* for Southern Rhodesia. Our 1960 enrollment rates come from the 1963 *UN Statistical Yearbook* for Southern Rhodesia. Enrollments in primary and secondary schools from 1970-1993 come from Maa Table I1 pp. 972 and 979. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1965-1992 come from Maa Table I2 p. 1000. For 2010 we used *HDR*. We interpolated for our 2000 enrollment rates. Our time series of years of schooling in the labor force for 1950-2010 is: 2.61 (1950), 3.50 (1960), 5.16 (1970), 5.50 (1980), 8.63 (1990), 9.40 (2000) and 9.71 (2010). The Morrisson & Murtin time series of years of schooling is: 2.00 (1950), 2.49 (1960), 3.32 (1970), 4.06 (1980), 4.96 (1990), 7.26 (2000) and 7.92 (2010).

9 Latin America

9.1 Argentina (1800-2010)

Populations for 1800, 1810, 1820, 1830, 1840, 1850, 1860 come from McEvedy and Jones. The 1810, 1820, 1830, 1840 values are interpolations. The 1860 value is an interpolation of McEvedy and Jones and Mitchell. Populations for 1870, 1880, 1895, 1914, 1947, 1960, 1970, 1980 and 1991 come from Mam Table A1 p. 7. Population for 1922 and 1935 are geometrically interpolated from 1914 and 1947. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1895, 1914, 1947, 1960, 1970, 1980 and 1990 come from Mam Table A2 p. 24. We assumed the age distribution in 1800-1880 were the same as the age distribution in 1895. Age distribution for Argentina 2000 and 2010 come from the *Demographic Yearbook*.

Labor force figures for 1947, 1960, 1970 and 1980 come from Mam Table B1 p. 108. For 1869, 1895 and 1914 we used Francis. We interpolated for 1880 and 1890. The 1914 Francis figure is within 2.5% of the Mitchell figure for the same year. The 1869 figure is also comparable to assuming a 100% labor force participation rate of 15-64 year olds in rural-agriculture, 65% of the population, and 77% labor force participation rate of 15-64 year olds in cities and towns, 35% of the population. The population distribution in 1869 is from Fajgelbaum and Redding. The 1869 figure is consistent with Harris as well. For 1820 and 1860 we used the urbanization rate from Salvatore and Newland, and the same rural-urban labor force participation rates. We assumed the labor force for years 1800-1810 arise from the urbanization rate indicated by the time trend between 1820 and 1860. Labor force data for Argentina for 1990, 2000 and 2010 are from *WDR*.

Real output come from Maddison. Physical capital investment rates come the intraperiod average gross real capital formation and real income for (1900-1998) from Mam Table J1, pp. 775, 776 and 782 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. The *WDI* provides sectoral output shares for 1970-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1800-1960. We used US 1850-1880 capital - sectoral output ratios, applying 1850 to 1800-1840, to produce capital estimates for 1800-1880. We used Tafunell Ducoing (2016) capital estimates for 1890, 1895 and 1914 averaged with the capital stock values produced by applying the US 1890-1910 capital - sectoral output ratios to our sectoral data. For years 1922-2010 we used perpetual inventory.

We used Lindert for enrollment rates in primary and secondary school for 1870. We assumed a .1 percent enrollment rate in higher education in 1870. For years prior to 1870 we used the 1870 enrollment rates. While this produces an 1870 primary share of 12.5 percent, Morrisson & Murtin report primary exposure share of 25 percent. However our more conservative value is consistent with the enrollments found in Lindert, Easterlin, and Mitchell. Enrollments in primary and secondary school from 1882, 1892-1993 come from Mam Table I1 pp. 735, 736, 738, 742 and 746. To calculate enrollment rates, we assumed 6-12 are primary school age and 13-17 are secondary school age. Prior to 1885 we assumed .1 percent higher education enrollment rates. Higher education enrollments for 1885-1992 are from Mam Table I2 pp. 755, 756 and 757. For 2010 we used *HDR*. Our labor force exposed to higher education share time series is consistent with Morrisson & Murtin until 2000. We interpolated enrollment rates in all three categories in the following years: 1880, 1890, 1929, and 2000. Our years of schooling in the labor force for 1870-1920 is: .91 (1870), .92 (1880), 1.05 (1890), .92 (1895), 1.26 (1914), 2.73 (1922). The Morrisson & Murtin time series is: 1.50 (1870), 1.64 (1880), 1.79 (1890), 2.12 (1900), 2.53 (1910), 3.11 (1920).

9.2 Bahamas (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1960, 1970, 1980 and 1990 come from KF. We assumed the 1950 age distribution was the same as the 1960 age distribution. Age distribution for Bahamas 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in 2013.

Labor force figures for 1960, 1970, 1980, 1990, 2000 come from *WDR*. Labor force data for 2010 come from *CIA Factbook*. For 1950 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for 1960-2010. We assumed rural 15-64 labor force participation rates of 90% for 1950-2010. We assumed urban 15-64 labor force participation rates of 66.79% for 1950-1960, 71.2% for 1970, 62.7% for 1980, 73.2% for 1990, 70% for 2000, 78.3% for 2010. We constructed the ratio of this labor force to that from *WDR* and *CIA Factbook*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.001 and 1.000, respectively. The range of the ratios is 1.000 to 1.001.

Real GNP comes from Maddison, except for 2010. For 2010 we used the 2000 Maddison value and applied the real PPP per capita income growth rate from 2000 to 2010 using *WDI*. Physical capital investment rates come the intraperiod average gross real capital formation and real income for 1980-2010 come from S & H. The *WDI* provides sectoral output shares for 1990-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1950-1980. We

used the US 1950-1970 capital - sectoral output ratios to construct our capital values for 1950-1970. We used perpetual inventory for 1980-2010.

For 1950 we used "Educational Reform in the Bahamas "by Keva Bethel. Bethel provides 1947 a primary and secondary total enrollment that we allocate two thirds to primary school, and one third to secondary school. We assumed that the higher education enrollment rate in 1950 was the same proportion of the secondary enrollment rate as the 1960 higher education enrollment rate was to 1960 secondary enrollment rate. Enrollment rates come from *WDR*, various years, for 1960, 1970, 1980, 1990 and 2000. For 2010 we used *HDR*. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. We assumed that the 2010 higher education enrollment rate was the same proportion of the 2010 secondary enrollment rate as the 2000 higher education enrollment rate was to the 2000 secondary enrollment rate.

9.3 Barbados (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1950, 1960, 1970, 1980 and 1990 come from KF. We interpolated between 1990 and 2010 for the age distribution in 2000. Age distribution for Barbados 2010 comes from the *Demographic Yearbook*.

Labor force figures for 1960, 1970, 1980, 1990, 2000 come from *WDR*. Labor force data for 2010 come from *CIA Factbook 2011-2012*. For 1950 we used the average labor force participation rate, LF/population, for years 1960-2010.

Real GNP come from Maddison, except for 2010. For 2010 we used the 2000 Maddison value and applied the real PPP per capita income growth rate from 2000 to 2010 using *WDI*. Physical capital investment rates come the intraperiod average gross real capital formation and real income for 1970-2010 come from S & H. The 1966, 1973 and 1981 *UN Statistical Yearbooks* provide sectoral output shares for 1960-1980. The *Human Development Report* 2000 provides the 1998 sectoral output shares, which we used for 2000. We used this to interpolate for the 1990 sectoral output shares. We used Sabillon (2005) information from farming, manufacturing and aggregate growth rates using Bahamas, to produce the 1950 sectoral output shares. We use the US 1950 and 1960 capital - sectoral output ratios to produce our 1950 and 1960 capital values. We used perpetual inventory for 1970-2010.

For 1950 we used Arthur Layne's estimate of 1900 primary enrollments and 1945 secondary enrollments to interpolate with our 1960 data. Enrollment rates come from *WDR*, various issues, for 1960, 1970, 1980, 1990 and 2000. For 2000 higher education enrollment rate, and primary and secondary school enrollment rates in 2010 we used UNESCO *UIS Global Database*. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. We assumed the 1950 and 1960 higher education enrollment rates were the same proportion of the contemporaneous secondary school enrollment rates as the 1970 higher education enrollment rate was to the 1970 secondary school enrollment rate. The 2010 higher education enrollment rate was the same proportion of the 2010 secondary school enrollment rate as the 2000 higher education enrollment rate was to the 2000 secondary school enrollment rate.

9.4 Belize (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1950, 1960, 1970, 1980 and 1990 come from KF. Age distributiosn for Belize 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 1998 age distribution, and for 2010 we interpolated using the 2006 and 2013 age distributions. We adjust by assuming the same share by age category as in 1998.

Labor force figures for 1960, 1970, 1980, 1990, 2000 come from *WDR*. Labor force data for 2010 come from *CIA Factbook 2011-2012*. For 1950 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for 1960-2010. We assumed rural 15-64 labor force participation rates of 73% for 1950-1960, 70.5% for 1970, 74.85% for 1980, 71.6% for 1990, 76.2% for 2000, 82.8% for 2010. We assumed urban 15-64 labor force participation rates of 50% for 1950-2010. We constructed the ratio of this labor force to that from *WDR* and *CIA Factbook*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.001 and 1.000, respectively. The range of the ratios is 1.000 to 1.001.

Real GNP come from Maddison, except for 2010. For 2010 we used the 2000 Maddison value and applied the real PPP per capita income growth rate from 2000 to 2010 from *WDI*. Physical capital investment rates come the intraperiod average gross real capital formation and real income for 1990-2010 come from S & H. The *WDI* provides sectoral output shares for 1980-2010. The 1973 and 1981 *UN Statistical Yearbooks* provide information to compute the 1960 and 1970 sectoral output shares. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates for the Caribbean countries to produce our 1950 estimates. We used the US 1950-1980 capital - sectoral output ratios to produce our 1950-1980 capital stocks. We used perpetual inventory to produce 1990-2010 capital values.

For 1950 we used Lewis (2000), who provides late 1940s enrollments in primary schools and secondary schools separately. Primary and secondary enrollment rates come from *WDR* , various issues, for 1960, 1980, 1990 and 2000. We interpolated for 1970 primary and secondary school enrollment rates. For 2010 we used *HDR* for primary and secondary enrollment rates. To calculate enrollment rates, we assumed 6-12 are primary school age and 13-17 are secondary school age. Our 2000 higher education enrollment rate comes from *WDR* and our 2010 higher education enrollment rate comes from *HDR*. For 1950-1990 we assumed that the 2000 higher education enrollment rate relative to the 2000 secondary school enrollment rate held for the 1950-1990 period.

9.5 Bolivia (1880-2010)

Populations for 1880, 1890, 1900, 1913, 1920, 1930, 1940, 1950 and 1976 come from Mam Table A1 p. 7. Population for 1962 is geometrically interpolated. Population for 1980 and 1990 comes from S & H (1991). Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1950, 1976 and 1990 come from Mam Table A2 p. 24. The age distribution for 1962 is geometrically interpolated. The age distribution for years prior to 1950 are assumed to be identical to 1950. The age distribution for 1980 comes from DK (1994). Age distribution for Bolivia 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in 2012.

Labor force figures for 1950, 1976 and 1991 come from Mam Table B1 p. 108. To compute labor force data for years prior to 1950 we used Banks for urbanization rate for 1880, 1888, 1900, 1913, 1924, 1930, 1940, 1950. We assumed that rural labor force participation rate of 15-64 year olds is 70% and urban labor force participation rate of 15-64 year olds is 66.5%. These rates exactly fit the 1950 and 2000 values of labor force. The 2000 urbanization rate comes from *WDR*. Labor force data for 1962 is geometrically interpolated. Labor force data for Bolivia for 1980 is from *WDR*. Labor force data for 2000 come from DK. Labor force data for 2010 come from *WDR*.

Real output come from Maddison. Physical capital investment rates come the intraperiod average gross real capital formation and real income for 1950-1998 from Mam Table J1, p. 776 and 782 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. The *WDI* provides sectoral output shares for 1975-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce 1880-1968 sectoral output shares. We used the US 1880-1950 US capital - sectoral output ratios to produce our capital estimates for 1880-1950. We used perpetual inventory to compute our 1962-2010 capital stocks.

For primary school and secondary school enrollment rates in 1880 and 1888 we used Lindert. For 1901 secondary school enrollment rate we used Lindert. Enrollments in primary and secondary school from 1901 (only primary), 1923, 1950-1993 come from Mam Table I1 pp. 736, 738, 742 and 746. To calculate enrollment rates, we assumed 6-13 are primary school age and 14-17 are secondary school age. Higher education enrollments for 1901-1990 are from Mam Table I2 pp. 755, 756 and 757. For 2010 we used *HDR*. We interpolated for our 1913, 1930, 1940, 1968 and 2000 enrollment rates.

9.6 Brazil (1800-2010)

Population for 1800 comes from McEvedy and Jones. We interpolated our 1800 and 1820 value for 1810. Population for 1820, 1830, 1840, 1850 and 1860 come from Maddison. Populations for 1872, 1889, 1900, 1920, 1940, 1950, 1960, 1970, 1980 and 1990 come from Mam Table A1 p. 7. Population for 1906, 1928 are geometrically interpolated. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1872, 1889, 1900, 1920, 1940, 1950, 1960, 1970, 1980 and 1990 come from Mam Table A2 p. 25. We assumed the age distributions in 1820, 1830, 1840, 1850 and 1860 are identical to the age distribution in 1872. Age distributions for Brazil 2000 and 2010 come from the *Demographic Yearbook*.

Labor force figures for 1872, 1900, 1920, 1940, 1950, 1960, 1970, 1980 and 1990 come from Mam Table B1 p. 108. For 1800, 1820, 1860, 1880 we used Lamounier. We used the information for the slave population for 1860, 1872 and 1880. For 1800 and 1820 we used Alexander estimates of the slave population. We then assumed 100% labor force participation of slaves, and we used a 50% labor force participation rate for the non slave population over the age of 4 and under the age of 65. We interpolated for 1810, 1830, 1840, 1850. Labor force data for 2000 and 2010 come from *WDR*.

Real output come from Maddison. Physical capital investment rates come the intraperiod average gross real capital formation and real income for 1948-1998 from Mam Table J1, pp. 777 and 782 and *WDR* (various years). For 2010 we used the average of 2000-2009 from S & H online. The *WDI* provides sectoral output shares for 1960-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to compute sectoral output shares for 1800-1950. We used US 1840-1940 capital - sectoral

output shares, applying the 1850 values for 1800-1840, to produce our capital estimates for 1800-1940. We used perpetual inventory for years 1950-2010.

Enrollment rates for 1850 and 1860 use Morris and Adelman. The 1850 enrollment rates (data) for primary, secondary and tertiary schools are: 7%, .13% and 0%, respectively. For years prior to 1850, we assumed the same enrollment rates as in 1850 for primary, secondary and higher education. These produce 1870 education shares of 7 percent exposure, and 7 percent primary exposure. Morrison & Murtin report 9 percent exposure, and 8 percent primary exposure and 1 percent secondary exposure. Enrollments in primary and secondary school from 1871, 1906, 1927-1993 come from Mam Table I1 pp. 735, 736, 738, 742 and 746. To calculate enrollment rates, we assumed 6-13 are primary school age and 14-16 are secondary school age. Higher education enrollments for 1907-1993 come from Mam Table I2 pp. 755, 756 and 757. Prior to 1907 we used higher education enrollment rates of 0 and .001. For 2010 we used *HDR*. We interpolated for our 1880 and 2000 enrollment rates. Our time series of years of schooling in the labor force for 1870-1920 is: .56 (1872), .58 (1880), .61 (1889), .64 (1898), .86 (1906) and 1.32 (1920). The Morrisson & Murtin time series of schooling year is: .46 (2870), .55 (1880), .64 (1890), .74 (1900), .87 (1910) and 1.04 (1920).

9.7 Chile (1800-2010)

Population for 1800, 1810 come from interpolating between the 1778 and 1813 population from Llorca-Jaña and Navarrete-Montalvo (2016). For 1820 and 1830 we interpolated between 1813 and 1835 from Llorca-Jaña and Navarrete-Montalvo (2016). Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1895, 1907, 1920, 1930, 1940, 1952, 1960, 1970, 1980 and 1990 come from Mam Table A1 p. 7. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1895, 1907, 1920, 1930, 1940, 1952, 1960, 1970, 1980 and 1990 come from Mam Table A2 p. 26. For years prior to 1895, we assumed the 1895 age distribution. The age distribution for 1980 and 1990 are interpolated using 1970 and 1982 and 1991 values. Age distributions for Chile 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1920, 1930, 1940, 1952, 1960, 1970, 1980 and 1990 come from Mam Table B1 p. 109. We used Federico (2010) for agricultural labor force for 1800, 1850, 1880, 1910, 1938 and 2000. We interpolated to fill in the missing years. We used the overlapping years from 1920-2000 to produce a time series of the share of total labor force in agriculture. We then fit a time series of this share on the log of year and predicted the share back to 1800. We capped the share at 80% for years 1800, 1810 and 1820. Labor force data for 1980 are interpolated from 1970 and 1982 data. Labor force data for 2000 and 2010 come from *WDR*.

Real output come from Maddison. Physical capital investment rates come the intraperiod average gross real capital formation and real income for 1940-1998 from Mam Table J1, pp. 777 and 782 and *WDR* (various years). For 2010 we used the average of 2000-2009 from S & H online. The *WDI* provides sectoral output shares for 1960-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1800-1952. We used US 1850-1940 capital - sectoral output ratios, applying 1850 values for 1800-1840, to produce capital estimates for 1800-1940. Our 1890 and 1895 estimates are modified by averaging these capital values with the non residential capital stock

estimates of Tafunell and Ducoing (2016). We used perpetual inventory for 1952-2010.

Enrollments in primary and secondary school from 1895-1993 come from Mam Table I1 pp. 735, 736, 739, 743 and 747. The 1870 primary school enrollment rate observation comes from Benavot Riddle (1988); it is 18.7%. The 1870 observations for primary and secondary school enrollment rates are 18.7% and 1.6%, respectively. The 1880 and 1890 primary and secondary enrollment rates come from Lindert. Enrollment rates in primary school from 1800 to 1870 are assumed to be 15 percent for all years. Secondary enrollment rates from 1800 to 1870 are assumed to be 1.5 percent for all years. Higher education enrollment rates from 1800 to 1860 are assumed to be .05 percent for all years. These produce 1870 education shares of 19 percent, 17 percent primary exposure and 1.5 percent secondary exposure. Morrison & Murtin report 20 percent exposure, 18 percent primary exposure and 2 percent secondary exposure. To calculate enrollment rates, we assumed 6-13 are primary school age and 14-17 are secondary school age. Higher education enrollments for 1886-1993 are from Mam Table I2 pp. 755, 756 and 757. For 2010 we used *HDR*. We interpolated for enrollment rates in 2000. Our 1870-1920 time series for years of schooling in the labor force is: 1.20 (1870), 1.41 (1880), 1.24 (1890), 1.51 (1895), 2.01 (1907), 2.75 (1920). The Morrisson & Murtin time series of schooling years is: .94 (1870), 1.05 (1880), 1.57 (1890), 1.73 (1900), 2.04 (1910), 2.89 (1920).

9.8 Colombia (1800-2010)

Population for 1800, 1850 come from McEvedy and Jones. For 1810, 1820, 1830, 1840, 1860, 1870, 1880 we interpolated. Populations for 1890, 1900, 1908, 1917, 1938, 1951, 1964 and 1973 come from Mam Table A1 p. 7. Population for 1980 and 1990 comes from S & H (1991). Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1917, 1938, 1951, 1964, 1973, 1980 and 1990 come from Mam Table A2 p. 27. We assumed the age distribution prior to 1917 were identical to the age distribution in 1917. The age distribution for Colombia for 1980 and 1990 are interpolations using 1973, 1985 and 1991 values. Age distributions for Columbia 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1938, 1951, 1964 and 1970 come from Mam Table B1 p. 109. Labor force data for Colombia for 1980, 1990, 2000 and 2010 are from *WDR* (various years). For 1800, 1810, 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1908, 1917, 1928 we used the following procedure. We used Clio infra for the urban-rural population shares 1800, 1850, 1900, 1951. WE interpolated for 1810-1840, 1860-1890 and 1910-1940. We assumed that the rural 15-64 labor force participation rate is 66% for 1800-1951, 61.8% for 1964, 71.6% for 1973, 61.98% for 1980, 73.58% for 1990, 95% for 2000 and 2010. We assumed the urban 15-64 labor force participation rate is 51% for 1800-1990, 57% for 2000 and 54.37% for 2010. We constructed the ratio of this labor force to the one from Mam, and *WDR*. The root mean ratio for the overlapping years of 1938-2010 is 1.000. The ratios for 1938 and 1951 are 1.000 and 1.000. The range of ratios is 1.000 to 1.000.

Real output come from Maddison. Physical capital investment rates come the intraperiod average gross real capital formation and real income for 1935-1998 from Mam Table J1, pp. 775, 778 and 783 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. The *WDI* provides sectoral output shares for 1973-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce our 1800-1964 sectoral output shares. We used the US 1850-1930 capital

- output ratios, applying 1850 values for 1800-1840, to produce our capital stocks for 1800-1928. We used perpetual inventory for 1938-2010.

Enrollments in primary and secondary school from 1900-1993 come from Mam Table I1 pp. 739, 743 and 747. For 1890 we used Lindert. For 1870 and 1880 primary school enrollment rates we used Benavot and Riddle. For secondary school enrollment rates in 1870 and 1880 we assumed they were the same proportion of primary school enrollment rates as the 1890 secondary school enrollment rate was to the 1890 primary school enrollment rate. For higher education enrollment rates in 1870 and 1880 we assumed they were .01%. For our 1870 data we have primary, secondary and tertiary school enrollment rates of 5.9%, .3% and .01%. For 1800-1860 we assumed primary, and tertiary school enrollment rates of 4%, and .01%, respectively. To calculate enrollment rates, we assumed 6-10 are primary school age and 11-16 are secondary school age. Higher education enrollments for 1897-1991 are from Mam Table I2 pp. 755, 756 and 757. For 1890 we assumed a higher education enrollment rate of .10%. For 2010 we used *HDR*. We interpolated enrollment rates in 2000. Our time series of schooling in the labor force for 1870-1920 is: .22 (1870), .29 (1880), .39 (1890), .70 (1900), .85 (1908), 1.27 (1917).

9.9 Costa Rica (1920-2010)

Population for 1920 comes from Maddison. Populations for 1950, 1963, 1973, 1980, 1990 come from Mam Table A2 p. 12. Population for 1980 is interpolated using 1973 and 1990 values. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1950, 1963, 1973, 1980 and 1990 come from Mam Table A2 p. 12. We assumed the age distributions for 1927, 1930 and 1940 were identical to the age distribution in 1950. The age distribution for Costa Rica for 1980 is interpolated from the 1973 and 1990 values. Age distributions for Costa Rica 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in 2013.

Labor force figures for 1951, 1963, 1973, 1980 and 1990 come from Mam Table B1 p. 102. Labor force data for 1980, 1990 are interpolations from 1973, 1984 and 1993. Labor force data for 2000 and 2010 come from *WDR*. For 1892, 1900, 1910, 1920, 1930 and 1940 we used the following procedure. We used Banks for the rural and urban shares of the population up to 1950. We used *WDI* for urban shares for 1960-2010. We assumed a rural 15-64 labor force participation rate of 64% for 1892-1963, 62.8% for 1973, 64% for 1980-2010. We assumed an urban 15-64 labor force participation rate of 59% for 1892-1951, 53% for 1963, 50% for 1973, 62% for 1980, 1990, 63.8% for 2000 and 58.45% for 2010. We constructed the ratio of this labor force to that from Mam and *WDR*. The root mean ratio for overlapping years 1951-2010 is 1.000. The 1941 and 1963 ratios are .997, 1.001, respectively. The range of the ratios is .997 to 1.001.

Real output come from Maddison. Physical capital investment rates come from the intraperiod average nominal gross capital formation and nominal income for 1950-1965, average gross real capital formation and real income for 1966-1998 from Mam Table J1, p. 767 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. The *WDI* provides sectoral output shares for 1990, 2000 and 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares to produce sectoral output shares for 1920-1980. We used the US 1920-1940 capital - sectoral output ratios to produce our 1920-1940 capital values. We used Nehru and Dhareshwar (1993) for capital values for 1951-1990. We used perpetual inventory for years 2000 and 2010.

Enrollments in primary and secondary school from 1890-1993 come from Mam Table J1 pp. 719, 721,

725 and 730. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1951-1993 are from Mam Table I2 pp. 752 and 754. For 1920 we assumed a higher education enrollment rate of 1% compared to the 1% rate in 1950. These produce 1920 education shares of 49 percent exposed, 47.5 percent primary exposed, 1 percent secondary exposed and 1 percent higher education exposed. Morrison & Murtin report 49 percent exposed, 47.4 percent primary exposed, 0 percent secondary exposed and 1.1 percent higher education exposed. We interpolated between these for our 1930 and 1940 values. For 2010 we used *HDR*. We interpolated enrollment rates for 1900, 1910, 1930, 1940 and 2000. Our 1890-1930 time series of schooling in the labor force is: 1.50 (1892), 1.54 (1900) 1.71 (1910), 1.55 (1920), 2.33 (1930). The Morrisson & Murtin time series of schooling for this period is: 1.52 (1890), 1.69 (1900), 2.01 (1910), 2.13 (1920) 2.38 (1930).

9.10 Cuba (1800-2010)

Populations for 1800 and 1850 come from McEvedy and Jones. For 1810, 1820, 1830, 1840, 1860, 1870, 1880 and 1890 we interpolated. Populations for 1900, 1910, 1920, 1930, 1940, 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1950, 1960, 1970, 1980 and 1990 come from KF. We assumed the age distributions for years prior to 1950 are identical to the age distribution in 1950. Age distributions for Cuba 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1920 and 1940 come from Mitchell. We interpolated between these years to get an estimate of labor force for 1930. The labor force figures for 1960, 1970, 1980, 1990 come from *WDI*. We interpolated between 1940 and 1960 labor force to get an estimate for 1950. For 1820, 1830, 1850 and 1860 we used the 1817, 1827, 1846 and 1862 population figures for slaves and non slave population of Cuba from Lamounier, Table 7. We assumed 100% labor force participation for slaves, and after deducting the Cuban population 65 and older, as well as the relative population of 0-4 males and females, we assumed a 50% labor force participation rate of this age cohort. We adjusted these numbers by the relative population figures between the years of Lamounier and our population figures. We interpolated for 1840, 1870, 1880, 1890, 1900, 1910. Labor force data for 2000 and 2010 come from *WDR*. For 1800 and 1810 we used the following procedure. We used Clio infra for 1800, 1850, 1900 and 1950 urbanization rates, and *WDI* for 1960-2010. We interpolated between 1800 and 1850 for 1820-1840 values, and interpolated between 1850 and 1900 for 1860-1890 values, and interpolated between 1900 and 1950 for 1910-1940 values. We assumed rural 15-64 labor force participation rates of 102.85% for 1800-1820, 105.85% for 1830, 102.2% for 1840, 98.5% for 1850, 93.9% for 1860, 93% for 1870, 92% for 1880, 91.1% for 1890, 90.25% for 1900, 73.35% for 1910, 64.55% for 1920, 64.31% for 1930, 63.55% for 1940, 59.95% for 1950, 69.85% for 1960, 64.6% for 1970, 84% for 1980, 83.65% for 1990, 88.3% for 2000, 90% for 2010. We assumed urban 15-64 labor force participation rates of 102.85% for 1800-1820, 105.85% for 1830, 102.2% for 1840, 98.5% for 1850, 93.9% for 1860, 93% for 1870, 92% for 1880, 91.1% for 1890, 90.25% for 1900, 50% for 1910-2000, 59.05% for 2010. We constructed the ratio of this labor force to that from *WDR*. The root mean ratio for the overlapping years 1820-2010 is 1.000. The 1820 and 1830 ratios are 1.000 and 1.001, respectively. The range of the ratios is 1.000 to 1.001.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real per capita income growth rates from *WDR 2011* and *WDR 2012*. Physical capital investment rates come the intraperiod average gross real capital formation and real income for 1960-2010 come from Mitchell (through 1979) and S & H (1980 onward). The *WDI* provides sectoral output shares for 1970-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1800-1960. We applied the US 1850-1950 capital - sectoral output ratios, applying the 1850 values for 1800-1840, to produce capital values for 1800-1950. We used perpetual inventory for 1960-2010.

For primary enrollment rates 1880-1940 Benavot and Riddle. For secondary enrollment rates in 1930 and 1940 we used Astorga, et al. The 1880 primary enrollment rate datum is 9.7%. The secondary enrollment rate datum for 1930 is 1%. The higher education enrollment rates for 1930 and 1940 are .5%. For all secondary enrollment rates 1800-1920 we assumed 1%. For all primary enrollment rates for 1800-1870 we assumed 9%. For all higher education enrollment rates 1800-1920 we assumed .1%. For 1950 we used Mitchell. Enrollment rates come from WDR for 1960, 1970, 1980, and 1990 come from *WDI*. For 2000 we used UNESCO *UIS Global Database*. For 2010 we used *WDR*. To calculate enrollment rates, we assumed 6-10 are primary school age and 11-16 are secondary school age. Our time series of schooling in the labor force for 1870-1920 is: .65 (1870), .93 (1880), 1.25 (1890), 1.65 (1900), 1.94 (1910) and 2.33 (1920). The Morrisson & Murtin time series for schooling is: .83 (1870), .94 (1880), 1.04 (1890), 1.22 (1900), 1.51 (1910) and 2.08 (1920).

9.11 Dominican Republic (1950-2010)

Populations for 1950, 1960, 1970, 1990 come from Mam Table A2 p. 13. Population for 1980 comes from *Population Pyramid*, https://populationpyramid.net/dominican-republic/1980/, accessed July 24, 2016. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1950, 1960, 1970, 1980 and 1990 come from Mam Table A2 p. 13. The age distribution for Dominican Republic for 1980 is interpolated using the 1970 and 1990 values. The age distribution for 2000 comes from KF. Age distribution for Dominican Republic 2010 comes from the *Demographic Yearbook*. For 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in 2013.

Labor force figures for 1950, 1960, 1970, 1981 come from Mam Table B1 p. 103. Labor force data for 1990, 2000 and 2010 are from *WDR*.

Real output come from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1950-1998 from Mam Table J1, p. 768 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. Mitchell provides sectoral output shares for 1950-2000. We used the 1950 US capital - sectoral output ratios to construct our 1950 capital stock estimates. We used Nehru and Dhareshwar (1993) for capital stocks 1960-1990. We used perpetual inventory for 2000 and 2010.

Enrollments in primary and secondary school from 1935-1993 come from Mam Table I1 pp. 721, 726 and 731. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1935-1985 are from Table I2 pp. 752 and 754. For 2010 we used enrollment rates from *HDR*. We interpolated for our 2000 enrollment rates. Our time series for schooling in the labor force for 1935-1960 is: .78 (1935), 1.09 (1950), and 2.75 (1960). The Morrisson &

Murtin 1930-1960 time series for years of schooling is: 1.26 (1930), 1.60 (1940), 1.99 (1950), 2.55 (1960).

9.12 Ecuador (1870-2010)

Populations for 1870-1940 come from Maddison. Populations for 1950, 1962, 1974 come from Mam Table A1 p. 7. Population for 1980 comes from *Population Pyramid*, https://populationpyramid.net/ecuador/1980/, accessed July 24, 2016. Population for 1990 is from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1950, 1962, 1974, 1980 and 1990 come from Mam Table A2 p. 27. We assumed the age distribution in 1940 was identical to the age distribution in 1950. The age distribution for Ecuador for 1980 and 1990 are interpolations using 1974, 1982 and 1991. Age distributions for Ecuador 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2001 age distribution, and for 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1950, 1962, 1974, 1980 and 1990 come from Mam Table B1 p. 109. For the 1870-1950 labor force values we used Clio infra for urban-rural population shares. We assumed the rural 15-64 labor force participation rate was 70%for 1870-1950, 65% for 1962, 57.5% for 1974, 50% for 1980, 70% for 1990-2000, and 79% for 2010. We assumed the urban 15-64 labor force participation rate was 57.5% for 1870-2010. For the years of 1962, 1974, 1980, 1990, 2000 and 2010 we used *WDI* for urban share. The root mean square ratio for the overlapping years, 1950-2010 is 1.004. The two closest years, 1950, 1962 have ratio values of .975 and 1.009, respectively. The range of the ratios is .975 to 1.015. Labor force data for Ecuador for 1980 are interpolated from 1974 and 1982 values. Labor force data for 2000 and 2010 come from *WDR*.

Real output come from Maddison. Physical capital investment rates come the intraperiod average gross real capital formation and real income for 1960-1998 from Mam Table J1, pp. 778 and 783 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. The *WDI* provides sectoral output shares for 1962-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral output shares for 1870-1950. We used US 1870-1940 capital - sectoral output ratios to produce capital stocks for 1870-1940. We used Nehru and Dhareshwar (1993) for capital stocks in 1950-1990. We used perpetual inventory for 2000 and 2010.

For 1870 primary school enrollment rate comes from Benavot and Riddle. For 1890-1930 primary and secondary school enrollment rates come from Lindert. Lindert supplies our 1900-1930 higher education enrollment rates. We interpolated for our 1880 primary and secondary school enrollment rates. For secondary enrollment rate in 1870 we assumed .95%, the 1880 value. For higher education enrollment rates 1870-1890, we assumed .05%, or half the 1900 value. Enrollments in primary and secondary school from 1928-1992 come from Mam Table I1 pp. 739, 743 and 747. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1928-1989 are from Mam Table I2 pp. 756 and 757. For 2010 we used *HDR*. We interpolated for our 2000 enrollment rates.

9.13 El Salvador (1920-2010)

Populations for 1920, 1930, 1940, 1950, 1961, 1971, 1981 and 1991 come from Mam Table A2 p. 13. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1930, 1950, 1961, 1971, 1981 and 1991 come from Mam Table A2 p. 13. The age distribution in 1920 is assumed to be identical to the 1930 age distribution. We interpolated for 1940. Age distributions for El Salvador 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1950, 1961, 1971 come from Mam Table B1 p. 103. Labor force data for El Salvador for 1980, 1991, 2000 and 2010 are from *WDR* (various years). For 1920, 1930 and 1940, 1950, 1961, 1971 we used the urban-rural population shares from Banks. For 1980-2010 we used *WDR* and *WDI*. We assumed the 15-64 rural labor force participation rate was 65% for 1920, 1930, 1940, 1950, 1961, 70.5% for 1971, 87% for 1980, 72.1% for 1990, 56% for 2000 and 57.35% for 2010. We assumed the urban 15-64 labor force participation rate was 57.5% for 1920-1990, 56% for 2000 and 57.35% for 2010. We constructed the ratio of this labor force with that from Mam and *WDR*. The geometric mean of the ratio for overlapping years 1950-2010 is 1.003. The values for 1950 and 1961 are .999 and 1.018, respectively. The range of ratios is .998 to 1.018.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. Physical capital investment rates come from the intraperiod average gross capital formation and income for 1951-1998 from Mam Table J1, p. 769 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. The *WDI* provides sectoral output shares for 1990-2010. We used Sabillon (2005) information on farming, manufacturing, and aggregate growth rates to construct sectoral output shares for 1920, 1930, 1950-1980. We used US 1920 and 1930 capital - sectoral output ratios to construct our capital estimates for those years. We used Nehru and Dhareshwar (1993) for capital stocks covering 1950-1990. We used perpetual inventory for 2000 and 2010.

For 1920 and 1930 we used information from Astorga, Berges and Fitzgerald. Enrollments in primary and secondary school from 1942-1993 come from Mam Table I1 pp. 722, 726 and 731. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1944-1991 come from Mam Table I2 pp. 752 and 754. For 2010 we used *HDR*. We interpolated for our 2000 enrollment rates. Our time series of years of schooling in the labor force for 1910-1940 is: .79 (1910), .83 (1920), .84 (1930), .89 (1940). The Morrisson & Murtin time series of schooling is: 1.27 (1910), 1.26 (1920), 1.35 (1930) and 1.33 (1940).

9.14 Guatemala (1921-2010)

Populations for 1921, 1930, 1940, 1950, 1964, 1973, 1980 and 1991 come from Mam Table A2 p. 14. Population for 1980 is interpolated using 1973, 1981 values. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1921, 1930, 1940, 1950, 1964, 1973, 1980 and 1991 come from Mam Table A2 p. 14. The age distribution for Guatemala for 1980 is interpolated using 1973, 1981 values. Age distributions for Guatemala 2000 and 2010 come from the *Demographic Yearbook*. For 2000

we adjusted the 2001 age distribution. We adjust by assuming the same share by age category as in 2001.

Labor force figures for 1950, 1964, 1973, 1981 come from Mam Table B1 p. 103. For 1973, 1980 and 1991 we used *WDR* For 2000 and 2010 we used *WDI*. To produce our 1921-1940 labor force data we used Banks for the urban-rural population share. We also used Banks for 1950 and 1964. We used *WDI* for years 1973-2010. We assumed a rural 15-64 labor force participation rate of 64.5% for 1890-1950, 61.8% for 1964, 62.1% for 1973, 56.6% for 1980, 73.66% for 1990, 69.6% for 2000 and 78.32% for 2010. We assumed an urban 15-64 labor force participation rate of 50% for 1890-2010. In the overlapping years, the root mean of the ratio of this labor force compared to those from Mam and *WDR* is 1.000. The 1950 and 1964 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real output come from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1946-1998 from Mam Table J1, p. 769 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. The *WDI* provides sectoral output shares for 2000 and 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1921-1990. We used US 1920-1940 capital - sectoral output ratios to construct our capital estimates for 1921-1940. We used Nehru and Dhareshwar (1993) for capital stocks in 1950-1990. We used perpetual inventory for 2000 and 2010.

Enrollments in primary and secondary school from 1920-1993 come from Mam Table I1 pp. 722, 726 and 731. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1950-1986 are from Mam Table I2 pp. 752 and 754. The 1950 value is .50%. We assumed a 1921-1940 values of .5%. We used data from Benavot & Riddle from 1890-1910 to help produce our initial 1921 exposure rates. We produce 1921 exposure rates of 21 percent, 20 percent primary, .5 percent each for secondary and higher education. Morrisson & Murtin report 25 percent exposure, 24 percent primary, 0 percent secondary and 1 percent higher education. Higher education enrollment rate for 1990 is from *WDR* All enrollment rates for 2010 are from *HDR*. We interpolated for our 1930 and 2000 enrollment rates for all categories. Our time series for schooling in the labor force for 1890-1930 is: .51 (1890), .72 (1900), .97 (1910), 1.16 (1921) and 1.44 (1930). The Morrisson & Murtin time series for schooling is: .83 (1890), .97 (1900), 1.09 (1910), 1.14 (1920) and 1.42 (1930).

9.15 Guyana (1946-2010)

Populations for 1946, 1960, 1970, 1980 and 1990 come from Mam Table A1 p. 7. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1946, 1960, 1970, 1980 and 1990 come from Mam Table A2 p. 28. Age distribution for Guyana 2010 comes from the *Demographic Yearbook*. For 2000 we interpolated.

Labor force figures for 1946, 1960 and 1980 come from Mam Table B1 p. 110. Labor force data for Guyana for 1970, 1990, 2000 and 2010 are from *WDR* (various years).

Real output come from Maddison, except 2010. For 2010 we used the 2000 Maddison value and applied the real PPP per capita income growth rate from 2000 to 2010 from *WDI*. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1960-1998 from Mam Table J1, pp. 779 and 783 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. The Mitchell provides sectoral output shares for 1946-2000. We used Nehru and Dhareshwar (1993) for capital stocks 1946-1990, we used the 1950 capital- output ratio for 1946. We used perpetual inventory for 2000 and 2010.

Enrollments in primary, 1911-1988, and secondary school, 1948-1986, come from Mam Table I1 pp. 740, 744 and 748. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1964-1993 are from Mam Table I2 p. 757. For 1946 and 1960 higher education enrollment rates we used the 1954 *UN Statistical Yearbook* and the 1963 UN Statistical Yearbook. For years prior to 1946 we used higher education enrollment rates of .001. For 2000 we used *WDI*. For 2010 enrollment rates we used *HDR*.

9.16 Haiti (1940-2010)

Populations for 1940, 1950, 1971, 1980 and 1991 come from Mam Table A2 p. 15. We used *WDR* for 1960. Population for 1980 is interpolated using 1972 and 1982 values. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1950, 1971, 1980 and 1991 come from Mam Table A2 p. 15. We assumed the age distribution in 1940 was identical to the age distribution in 1950. The 1960 age distribution is interpolated. The age distribution for Haiti for 1980 is interpolated using the 1972 and 1982 values. Age distributions for Haiti 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 1999 age distribution. We adjust by assuming the same share by age category as in 1999.

Labor force figures for 1950, 1971, 1980 and 1991 come from Mam Table B1 p. 103. For 1960 we interpolated. Labor force data for Haiti for 1980 are interpolated of 1971 and 1982 values. Labor force data for 2000 and 2010 come from *WDR*. For 1940 we used Banks for urban-rural population shares covering 1940, 1950, 1960, and *WDR* for 1971, 1980, 1991 and *WDI* for 2000. We assumed a rural 15-64 labor force participation rate of 95% for 1940, 1950, 89.14% for 1960, 99% for 1971, 95% for 1980, 79.5% for 1985, 80.2% for 2000 and 73.28% for 2010. We assumed an urban 15-64 labor force participation rate of 80.5% for 1940, 1950, 50% for 1960, 99% 1971, 74.8% for 1980, 50% for 1985, 2000 and 2010. Our root mean ratio of this labor force with those from Mam and *WDR* is 1.000 for overlapping years 1950-2010. Our 1950 and 1960 ratios are 1.000 and 1.000, respectively. The range of ratios is 1.000 to 1.001.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. Physical capital investment rates come from the intraperiod average gross capital formation and income for 1955-1962 from Mam Table J1, p. 770. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1963-1998 from Mam Table J1, p. 770 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2006 from S & H. Mitchell provides sectoral output shares for 1950, 1971-2000. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1940 and 1960. We used the US 1940 and 1950 capital - sectoral output ratios to produce capital estimates for these years. We used perpetual inventory for 1960-2010.

Enrollments in primary and secondary school from 1970-1992 come from Mam Table J1 pp. 722, 727 and 732. We have information on primary school enrollment rates from 1910-1930 from Astorga, Berges and Fitzgerald. Our 1950 enrollment rates for all categories come from the 1954 *UN Statistical Yearbook*. Our 1960 enrollment rates for all categories come from the 1963 *UN Statistical Yearbook*. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollment rates for Haiti from 1980 onward come from *WDR* (various years). We interpolated our 1971 higher education enrollment rate. We assumed for 1940 0%. For 2000 we used *WDR* for all

enrollment rates. For 2010 we used UNESCO *UIS Global Database* for higher education enrollment rates and primary school enrollment rates. For 2010 secondary school enrollment rates we assumed that the 2000 ratio of secondary school enrollment rates to primary school enrollment rates was constant.

9.17 Honduras (1920-2010)

Populations for 1920, 1930, 1940, 1950, 1961, 1974, 1980 and 1990 come from Mam Table A1 p. 4. The 1980 value is interpolated from the 1974 and 1990 value. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1930, 1940, 1950, 1961, 1974, 1980 and 1990 come from Mam Table A2 p. 15. We assumed the age distribution in 1920 is identical to the age distribution in 1930. The age distribution for Honduras for 1980 is interpolated using 1974 and 1990 values. Age distribution for Honduras 2010 come from the *Demographic Yearbook*. For 2000 we adjusted interpolated.

Labor force figures for 1950, 1961, 1974, 1980 and 1990 come from Mam Table B1 p. 104. For 1920, 1930 and 1940 we used Banks for the urban-rural population shares. We assumed 63% rural 15-64 labor force participation rate and 57% urban 15-64 labor force participation rate. We used Banks for urban shares in 1950, 1961, *WDR* for 1974, 1980, 1990 and *WDI* for urban shares in 2000 and 2010. We constructed the ratio of this labor force with the Mam and *WDR*. The root mean value of the 1950-2010 ratios is .984. The values for 1950 and 1961 are .945 and 1.016, respectively. Labor force data for Honduras for 1980 are interpolated using 1974 and 1990 values. Labor force data for 2000 and 2010 comes from *WDR*.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. Physical capital investment rates come the intraperiod average gross real capital formation and real income for 1925-1998 from Mam Table J1, pp. 764 and 770 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. Mitchell provides sectoral output shares for 1940-2000 and the *WDI* provides sectoral output shares for 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1920 and 1930. We used the US 1920 and 1930 capital - sectoral output ratios to produce our 1920 and 1930 capital stock estimates. We used perpetual inventory for 1940-2010.

Enrollments in primary, 1919-1993, and secondary school 1922-1993 come from Mam Table I1 pp. 722, 727 and 732. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. We also have information from 1890-1920 on primary school enrollment rates from Engerman and Sokolof. We assumed that the secondary enrollment rates during this period was the same ratio of primary school enrollment rates as in 1930. We assume that the higher education enrollment rates were .001 throughout. Higher education enrollments for 1922-1992 are from Mam Table I2 pp. 751, 752 and 754. For 2010 we used *HDR*. We interpolated for all enrollment rates in 2000. Our time series of labor force schooling years for 1890-1930 is: .81 (1890), 1.15 (1900), 1.33 (1910), 1.44 (1920) and 1.49 (1930). Morrisson & Murtin time series for years of schooling is: 1.05 (1890), 1.25 (1900), 1.48 (1910), 1.51 (1920) and 1.58 (1930).

9.18 Jamaica (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1910, 1921, 1943, 1953, 1960, 1970 and 1991 come from Mam Table A1 p. 5. Population for 1980 is from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1880, 1890, 1900, 1910, 1921, 1943, 1953, 1960, 1970, 1980 and 1991 come from Mam Table A2 p. 16. The age distribution for years prior to 1880 are assumed to be identical to the age distribution in 1880. The 1900 age distribution is interpolated using the 1890 and 1910 age distributions. The age distribution for Jamaica for 1980 is interpolated using 1970 and 1982 values. Age distributions for Jamaica 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2001 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1943, 1953, 1963, 1973, 1980 and 1990 come from Mam Table B1 p. 104. Labor force data for Jamaica for 1980 are interpolated using 1973 and 1982 values. Labor force data for 2000 and 2010 come from *WDR*. For 1820-1950 we used the following procedure. The 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1910, 1921, 1932 labor force comes from the following procedure. We used Clio Infra for urbanization rates for 1820-1950, and *WDR* and *WDI* for 1960-2010. We interpolated between 1800 and 1850 for 1820-1840, and we interpolated between 1850 and 1900 for 1860-1890, and we interpolated between 1900 and 1950 for 1910-1940. We assumed a rural 15-64 labor force participation rate of 72.5% for 1820-1960, 85% for 1970, 62.5% for 1980, 72.5% for 1990, 2000. We assumed an urban 15-64 labor force participation rate of 52% for 1820-1943, 68.5% for 1950, 263% for 1960, 85% for 1970 and 50% for 1980, 69% for 1990, 70% for 2000 and 61% for 2010. We constructed the ratio of this labor force to that from *WDI* and *WDR*. The root mean ratio for the overlapping years 1943-2010 is 1.001. The 1943 and 1950 values are 1.001 and .999, respectively. The range of the ratio is .999 to 1.004.

Real output come from Maddison. Physical capital investment rates from the intraperiod average gross real capital formation and real income for 1953-1988 and from 1989-1998 we used the average gross capital formation and income. Mam Table J1, p. 771 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. Mitchell provides sectoral output shares for 1950-2000. and the *WDI* provides sectoral output shares for 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1820-1943. We used US 1850-1940 capital - sectoral output ratios, applying 1850 to 1820, 1830 and 1840, to construct capital estimates for 1820-1940. We used Nehru Dhareshwar (1993) for capital stocks 1950-1990; we used 1953 capital - output ratio to estimate 1950 capital. We used perpetual inventory for years 2000 and 2010.

For 1870, 1880, 1890, 1900, 1910, we used Lindert for enrollment rates for primary school. The 1870 observation for primary school enrollment rate is 42.7%. For years prior to 1870, we assumed primary, secondary enrollment rates that were 45 percent of the succeeding decade's primary, secondary enrollment rates. For primary school enrollment rates, this produces the following time series: 0.8% (1820), 1.8% (1830), 3.9% (1840), 8.6% (1850), 19.2% (1860). Enrollments in primary and secondary school from 1918-1991 come from Mam Table I1 pp. 723, 727 and 732. The 1920 secondary school enrollment rate is 1%. For years 1870-1910, for secondary school enrollment rates we assumed .25 percent. For 1820-1860, we assumed the following time series for secondary school enrollment rates: .01% (1820), .02% (1830), .03% (1840), .06% (1850) and .13% (1860). These produce 1870 values of 13 percent education exposure and 13 percent primary exposure. Morrisson & Murtin report 1870 values of 12 percent education exposure

and 12 percent primary exposure. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1948-1992 are from Mam Table I2 pp. 753 and 754. We assume 0 higher education enrollment rates for 1820-1860. For 1870-1910, for higher education we assumed .005 percent. For 1921 we assumed .01%, for .02% (1932), .05% (1943) and .1% (1950 datum). These produce a 0 percent higher education exposure rate in 1950, which matches Morrsion & Murtin. For 2010 we used *HDR*. We interpolated for all enrollment rates for 2000. Our time series for labor force schooling for 1870-1920 is: .91 (1870), 1.80 (1880), 2.91 (1890), 4.08 (1900), 4.75 (1910), and 4.60 (1921). Morrisson & Murtin time series for years of schooling is: .51 (1870), .70 (1880), 1.19 (1890), 1.66 (1900), 2.24 (1910), 2.50 (1920).

9.19 Mexico (1800-2010)

Population for 1800 comes from McEvedy and Jones. We interpolated our 1800 and 1820 population values for 1810. Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1895, 1900, 1910, 1921, 1930, 1940, 1950, 1960, 1970, 1980 and 1990 come from Mam Table A2 p. 17. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1895, 1900, 1910, 1921, 1930, 1940, 1950, 1960, 1970, 1980 and 1990 come from Mam Table A2 p. 17. We assumed the age distributions in years prior to 1895 were identical to the age distribution in 1895. Age distribution for Mexico 2000 and 2010 come from the *Demographic Yearbook*.

Labor force figures for 1895, 1910, 1921, 1930, 1940, 1950, 1960, 1980 and 1990 come from Mam Table B1 p. 105. Labor force data for 2000 come from DK. Labor force for 2010 comes from *WDR*. We replaced 1970 with *WDR* as it is more consistent with the labor force to population ratios in both 1960 and 1980; the value from Mam is abnormally low. For 1800-1950 we used Clio infra for urban-rural population shares. We interpolated for 1810-1840, 1860-1890, 1910-1940. We assumed 93.06% 15-64 rural labor force participation rate for 1800-1895, 98.26% for 1910, 74.8% for 1921, 64% for 1930, 58.5% for 1940, 57.74% for 1950, 74.5% for 1960, 71.4% for 1970, 86.5% for 1980, 72% for 1990, 95% for 2000 and 92% for 2010. We assumed urban 15-64 labor force participation rate of 50% for 1800-1895, 98.26% for 1910, 50% for 1921-1990, 54.09% for 2000 and 50.36% for 2010. Constructing the ratio of the computed labor force with the data, the values for 1895 and 1910 are 1.000 and 1.000, respectively. The root mean value for overlapping years 1895-2010 is 1.000. The range in values is 1.000 to 1.000.

Real output come from Maddison. Physical capital investment rates come the intraperiod average gross capital formation and income for 1939-1959 from Mam Table J1, pp. 764 and 771. Physical capital investment rates come the intraperiod average gross real capital formation and real income for 1960-1998 from Mam Table J1, p. 771 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. Mitchell provides sectoral output shares for 1895-2000. The *WDI* provides sectoral output shares for 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1800-1890. We used the US 1840-1940 capital - sectoral output ratios, applying the 1840 values for 1800-1830, to produce our 1800-1940 capital estimates. We used perpetual inventory for 1950-2010.

Enrollments in primary and secondary school from 1927-1993 come from Mam Table I1 pp. 720, 723, 728 and 733. Primary school enrollment rates for 1800-1840 are assumed to be 4 percent. Primary school enrollment rates for 1850 are assumed to be 12 percent. Primary school enrollment rate for 1860 is assumed

to be 16 percent. The 1870-1920 values come from Benavot Riddle (1988). Secondary school enrollment rates for 1800-1870 are assumed to be .1 percent, half the .2 percent 1895 value. These produce a .1% share of the labor force exposed to secondary school in 1870, compared with Morrisson & Murtin reported 1 percent share. For 1880 and 1890 they are interpolated between the 1870 rate and the 1895 rate. For 1800-1870 we assumed the higher education enrollment rate was .01 percent. For 1880 and 1890 we interpolated the higher education rate between 1860 and 1895. For higher education enrollment rate in 1870 we assumed it to be one tenth of the 1895 rate. Our 1870 exposure shares are 12 percent, and 12 percent primary exposure. Morrisson & Murtin report 12 percent exposure share, 11 percent primary exposure share and 1 percent secondary exposure. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1923-1993 come from Mam Table I2 pp. 751, 753 and 754. For 1820-1870 we assumed .001% enrollment rate for higher education. For the 1880 & 1890 values we interpolated between the 1870 value and the datum for 1895, .1%. For 2010 we used *HDR*. We interpolated for all our 2000 enrollment rates. Our time series of the share of labor force exposed to higher education is consistent with Morrisson & Murtin from 1870 - present. Our 1870-1920 time series of years of schooling in the labor force is: .65 (1870), .80 (1880), .91 (1890), .94 (1895), 1.08 (1910), 1.33 (1921). The Morrisson & Murtin time series of years of schooling is: .56 (1870), .74 (1880), .91 (1890), 1.09 (1900), 1.26 (1910), 1.36 (1920).

9.20 Nicaragua (1920-2010)

Populations for 1920, 1938 come from Mam Table A2 p. 18. Population for 1950, 1960, 1971, 1980 and 1991 come from *Population Pyramid*, https://populationpyramid.net/nicaragua/1950/, accessed July 24, 2016. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1920, 1938, 1950, 1963, 1971, 1980 and 1991 come from Mam Table A2 p. 18. The age distribution for Nicaragua for 1980 is interpolated using 1971 and 1991 values. Age distribution for Nicaragua 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2009 age distribution. We adjust by assuming the same share by age category as in 2009.

Labor force figures for 1938, 1950, 1963 and 1971 come from Mam Table B1 p. 106. Labor force data for Nicaragua for 1980, 1990, 2000 and 2010 are from *WDR* (various years). For 1920 and 1930 we used Banks for the urban-rural population shares for 1890-1964 and *WDR* and *WDI* for 1971-2010. We assumed a rural 15-64 labor force participation rate of 68.2% for 1890-1938, 62.5% for 1950, 65% for 1963, 60.1% for 1971, 59.1% for 1980, 75% for 1990-2010. We assumed urban 15-64 labor force participation rate of 62.5% for 1890-1938, 51.5% for 1950, 60.8% for 1963, 50% for 1971, 1980, 62.25% for 1990, 56.55% for 2000, 57.65% for 2010. We constructed the ratio of this labor force to that from Mam and *WDR*. The root mean ratio for years 1938-2010 is 1.000, and the 1938, 1950 values are 1.000 and 1.001, respectively. The range of ratios is 1.000 to 1.001.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 201 and 2012*. Physical capital investment rates come the intraperiod average gross real capital formation and real income for 1945-1975 from Mam Table J1, pp. 765 and 772. Physical capital investment rates for 1976-1998 are from S & H online and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. Mitchell provides sectoral output shares for 1950-2000, and the *WDI* provides sectoral output shares for 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to produce sectoral

output shares for 1920-1938. We used the US capital - sectoral output ratios for 1920-1940 to produce our capital estimates for 1920, 1930 and 1938. We used perpetual inventory for 1950-2010.

For 1890-1920 we used Lindert to construct enrollment rates in primary and secondary school. We interpolated these rates with rates in 1938 to construct enrollment rates in 1930. Enrollments in primary and secondary school from 1938, 1950-1993 come from Mam Table I1 pp. 723, 728 and 733. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1944-1992 are from Mam Table I2 pp. 753 and 754. For years 1920 and 1938 we assumed .1%, the 1950 value. For 2010 we used *HDR* for all enrollment rates. For 2000 we extrapolated our enrollment rates. Our time series share of labor force exposed to higher education is consistent with Morrisson & Murtin 1920-2010. Our time series of years of schooling in the labor force for 1890-1920 is: .48 (1890), .75 (1900), 1.01 (1910), 1.00 (1920). The Morrisson & Murtin years of schooling time series is: .89 (1890), .99 (1900), 1.07 (1910), 1.11 (1920).

9.21 Panama (1940-2010)

Populations for 1940, 1950, 1960, 1970, 1980 and 1990 come from Mam Table A1 p. 5. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1940, 1950, 1960, 1970, 1980 and 1990 come from Mam Table A2 p. 18. Age distribution for Panama 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in 2013.

Labor force figures for 1940, 1950, 1960, 1970, 1980 and 1990 come from Mam Table B1 p. 106. Labor force data for 2000 come from *DK*. Labor force data for 2010 come from *WDR*.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. Physical capital investment rates come the intraperiod average gross real capital formation and real income for 1946-1998 from Mam Table J1, pp. 765 and 772 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. Mitchell provides sectoral output shares for 1940-2000. The *WDI* provides sectoral output shares for 2010. We used the US 1940 capital - sectoral output ratios to produce our 1940 estimate. We used perpetual inventory for 1950-2010.

Enrollments in primary and secondary school from 1910-1930 come from Astorga, Berges and Fitzgerald, and from 1940-1993 come from Mam Table I1 pp. 724, 728 and 733. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1945-1993 are from Mam Table I2 pp. 753 and 754. For 1910-1940 we assumed 0, .1%, .25%, and .5% compared with our 1.2% 1950 observation. These produce a 1 percent higher education exposure, the same as Morrisson & Murtin. For 2010 we used *HDR*. We interpolated for our 2000 observation. Our 1910-1940 time series of years of schooling in the labor force is: 1.16 (1910), 1.55 (1920), 2.32 (1930), and 2.92 (1940). The Morrisson & Murtin time series of years of schooling is: 1.79 (1910), 2.04 (1920), 2.36 (1930), and 2.91 (1940).

9.22 Paraguay (1939-2010)

Populations for 1939, 1950, 1962, 1972, 1980 and 1990 come from Mam Table A1 p. 8. Population for 1980 and 1990 are interpolated from 1972, 1982 and 1991 values. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1939, 1950, 1962, 1972, 1980 and 1990 come from Mam Table A2 p. 28. The age distribution for Paraguay for 1936 is extrapolated from 1950. The age distribution 1980 and 1990 is interpolated from 1972, 1982 and 1991 values. Age distribution for Paraguay 2010 come from the *Demographic Yearbook*. For 2000 we interpolated, and for 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in 2013.

Labor force figures for 1950, 1962, 1972, 1980 and 1990 come from Mam Table B1 p. 110. For 1939 we used Banks for urban-rural population shares for 1939-1962, and *WDR* and *WDI* for 1972-2010. We assumed for years 1939-1980 a 15-64 labor force participation rate of 63.5% for both rural and urban populations. For 1990-2010, we assumed a rural 15-64 labor force participation rate of 75% and an urban 15-64 labor force participation rate of 70%. We constructed the ratio of this labor force to that from Meu and *WDR*. For the 1950-2010 period, the root mean ratio is .995. The 1950 and 1962 ratios are 1.030, .927, respectively. Labor force data for Paraguay for 1980 and 1990 are interpolated from 1972, 1982 and 1994 values. Labor force data for 2000 and 2010 come from *WDR*.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1953-1998 from Mam Table J1, pp. 779 and 783 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. Mitchell provides sectoral output shares for 1939-2000. We used the *WDI* for 2010. We applied the US 1940 and 1950 capital - sectoral output ratios to produce our 1939 and 1950 estimates of capital. We used Nehru and Dhareshwar (1993) for capital stocks 1961-1990. We used perpetual inventory for 2000 and 2010.

Enrollments in primary 1890-1930, and secondary school, 1900-1930, come from Lindert. For 1890 secondary school enrollments we assumed the same proportion between secondary school enrollment rate and primary school enrollment rate in 1900 held in 1890. Primary and secondary enrollments for 1950-1993 come from Mam Table I1 pp. 740, 744 and 748. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1915-1993 are from Mam Table I2 pp. 755, 756 and 757. For 1890 and 1900 we assumed higher education enrollment rate of .10%. For 2010 we used *HDR*. For 2000 we interpolated all of our enrollment rates. Our 1890-1930 years of schooling in the labor force time series is: 1.01 (1890), 1.43 (1900), 1.42 (1910), 2.36 (1920), and 3.08 (1930). The Morrisson & Murtin time series of year of schooling is: .93 (1890), 1.10 (1900), 1.37 (1910), 1.81 (1920) and 2.44 (1930).

9.23 Peru (1870-2010)

Population for 1870 comes from McEvedy and Jones. Populations for 1880, 1890, 1900, 1908, 1919, 1930, 1960, 1972, 1981 and 1991 come from Mam Table A1 p. 8. The 1900, 1908, 1919 and 1930 populations are interpolations using 1876 and 1940 values. 1950 population comes from McEvedy and Jones. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1900, 1908, 1919, 1930, 1940, 1960, 1972, 1980 and 19901 come from Mam Table A2 p. 29. The age distribution for Peru for 1870, 1880, 1890, 1900, 1908, 1919 and 1930 are extrapolations of the age distribution in 1940. The 1950 age distribution is interpolated. The age distribution for Peru for 1980 and 1990 are interpolations of the 1972, 1981 and 1991 values. Age distribution for Peru 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 1998 age distribution, and for 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1940, 1961, 1972, 1980 and 1990 come from Mam Table B1 p. 110. We interpolated for 1950. For labor force for 1870, 1880, 1890, 1900, 1908, 1919 and 1930 we used Banks for urban-rural population shares for 1870-1960, and *WDR* and *WDI* for 1972-2010. We assumed a rural 15-64 labor force participation rate of 71.25% for all years. We assumed an urban 15-64 labor force participatiopn rate of 50% for 1870-1980, and 60% for 1990, 65% for 2000 and 75% for 2010. We constructed the ratio of this labor force with that from Mam and *WDR*. The root mean ratio for 1940-2010 is 1.022, and the 1940 and 1950 ratios are .938 and 1.028, respectively. Labor force data for Peru for 1980 and 1990 are interpolations of 1972, 1981 and 1991 values. Labor force data for 2000 and 2010 come from *WDR*.

Real GNPs are from Maddison. Physical capital investment rates come from the intraperiod average gross capital formation and income for 1942-1950, and physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1950-1998 from Mam Table J1, pp. 780 and 784 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. Mitchell provides sectoral output shares for 1940-2000, except for 1950 which we interpolated. We used the Salvucci chapter in *Cambridge Economic History of Latin America* for sectoral output shares for 1900-1930. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1870-1890. We used US 1870-1940 capital - sectoral output ratios to construct our capital estimates for 1870-1940. We used perpetual inventory for 1950-2010.

Enrollment rates in primary and secondary school from 1890-1930 come from Lindert. For 1870 and 1880 we assumed that the primary and secondary school enrollment rates of 1890 held. Enrollments in primary and secondary school from 1940-1993 come from Mam Table I1 pp. 737, 740, 744 and 748. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-16 are secondary school age. Higher education enrollment rates for 1900-1930 come from Lindert. For 1870-1890 we assumed higher education enrollment rates of .10%. Higher education enrollments for 1930-1993 are from Mam Table I2 pp. 755, 756 and 757. For years 1890, 1900, 1908, 1919, 1930 we used Lindert. For 2010 we used *HDR*. We interpolated for all of our 2000 enrollment rates. Our time series of years of schooling in the labor force for 1870-1920 is: .36 (1870), .50 (1880), .63 (1890), .67 (1900), .94 (1908), 1.22 (1919). The Morrisson & Murtin years of schooling time series is: .28 (1870), .35 (1880), .42 (1890), .48 (1900), .63 (1910) and 1.16 (1920).

9.24 Puerto Rico (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Puerto Rico 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in 2012.

Labor force figures for 1990 come from *HDR*. Labor force figures for 1950, 1960, 1970, 1980, 2000 and 2010 come from *Statistical Abstracts of the United States* (various years).

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. The 1960-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The 1954 *UN Statistical Yearbook* provides sectoral output shares for 1950. The *WDI* provides sectoral output shares for 1970-2010. We interpolated for 1960 sectoral output shares. We used the US 1950 capital - sectoral output ratios to construct our 1950 value. We used perpetual inventory for 1960-2010.

Enrollments in primary and secondary schools for 1910-2000 come from Mam Table I1 pp. 729 and 734. We assumed the primary school and secondary school ages are 6-13 and 14-17. The tertiary school enrollments are from Mam Table I2 pp. 753 and 754. For 2010 we used UNESCO *UIS Global Database*.

9.25 Suriname (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Suriname 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in 2011.

Labor force figures for 1960, 1970, 1980, 1990 and 2000 come from *WDI*. The 2010 labor force figures come from the *CIA Factbook 2011-2012*. For 1950 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for 1960-2010. We assumed rural 15-64 labor force participation rates of 64.2% for 1950-1960, 52.2% for 1970, 47.25% for 1980, 79.8% for 1990, 2000 and 71.9% for 2010. We assumed urban 15-64 labor force participation rates of 50% for 1950-1970, 47.25% for 1980, 50% for 1990, 52.7% for 2000, 50.5% for 2010. We constructed the ratio of this labor force to that from *WDR*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.001 and 1.000, respectively. The range of the ratios is 1.000 to 1.001.

Real output come from Maddison, except for 2010. For 2010 we used the 2000 Maddison value and applied the real PPP per capita income growth rate from 2000 to 2010 from *WDI*. The 1970-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDI* (various years). The *WDI* provides sectoral output shares for 1960-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates for Caribbean nations to produce our 1950 sectoral output shares. We used the US 1950-1960 capital - sectoral output ratios to construct our capital estimates for 1950 and 1960. We used perpetual inventory for 1970-2010.

Enrollment rates in primary, secondary and higher education schools for 1980-2000 come from *WDR*. We have primary enrollment rates from 1890 to 1940 from Benavot and Riddle (1988). The 1980 observations for these enrollment rates are 90%, 40% and 6%, respectively. For 2010 we used *HDR*. For 1950 primary, secondary and higher education enrollment rates we used the 1954 *UN Statistical Yearbook*. For 1960 primary, secondary and higher education enrollment rates we used the 1963 *UN Statistical Yearbook*. For 970 enrollment rates we interpolated. We assumed the primary school and secondary school ages are 6-11 and 12-17. For all enrollment rates in 2010 we used *HDR*.

9.26 Trinidad & Tobago (1946-2010)

Populations for 1946, 1960, 1970, 1980 and 1990 come from Mam Table A1 p. 6. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1946, 1960, 1970, 1980 and 1990 come from Mam Table A2 p. 20. Age distributions for Trinidad & Tobago 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 1997 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1946, 1960, 1971, 1980 and 1990 come from Mam Table B1 p. 107. Labor force data for Trinidad Tobago for 1990 is interpolated from the 1980 and 1993 values. Labor force data for 2000 come from DK. Labor force for 2010 comes from *WDR*.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. Physical capital investment rates come from the intraperiod average gross capital formation and income for 1951-1974 from Mam Table J1, p. 773. Physical capital investment rates for 1975-1998 come from the intraperiod average gross real capital formation and real income Mam Table J1, p. 773 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. Mitchell provides sectoral output shares for 1946-2000. We used the US 1950 capital - sectoral output ratios to construct our 1946 capital estimate. We used Nehru and Dhareshwar (1993) for capital stocks in 1960-1990. We used perpetual inventory for 2000 and 2010.

Primary school enrollment rates for 1860-1931 come from Lindert. Enrollments in primary school from 1946-1993 come from Mam Tale I1 pp. 724, 729 and 734. Enrollments in secondary school from 1910-1993 come from Mam Table I1 pp. 724, 729 and 734. For 1860-1900 we extrapolated assuming the same proportion between secondary school enrollment rates to primary school enrollment rates in 1910 held. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1964-1993 are from Mam Table I2 pp. 753 and 754. Prior to 1964 we used enrollment rates for higher education of .01%. For 2010 we used *HDR*. We interpolated for all of our 2000 enrollment rates.

9.27 Uruguay (1800-2010)

Population for 1800 and 1850 are from McEvedy and Jones. For 1810, 1820, 1830, 1840, 1860 we interpolated. Populations for 1870, 1880, 1890, 1900, 1908, 1939 and 1949 are interpolated from 1852 and 1963 come from Mam Table A1 p. 8. Populations for 1961 and 1975 come from Mam Table A1 p. 8. Populations for 1980 and 1990 comes from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1900, 1908, 1939, 1949 are interpolated from 1908 and 1961 from Mam Table A2 p. 30. The age distribution for 1961 and 1975 come from Mam Table A2 p. 30. The age distributions for 1800-1890 are assumed to be identical to the age distribution in 1900. The age distribution for Uruguay for 1980 and 1990 is interpolated from the 1975, 1985 and 1991 values. Age distribution for Uruguay 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in 2011.

Labor force figures for 1961, 1975 and 1990 come from Mam Table B1 p. 110. For 1980 we substituted

using *WDR* as this is more consistent with the surround years labor force to population ratios. For 1800-1949 we used Banks for urban-rural population shares. Banks goes back only to 1828. However from 1828-1853 the urban share is 0, so we used 0 for 1800-1820. We used a 75% rural 15-64 labor force participation rate and 50% urban 15-64 labor force participation rate for 1800-1968. We maintain a 75% rural 15-64 labor force participation rate for all years 1975-2010. For urban 15-64 labor force participation rates we used 60% for 1975 and 1980, 70% for 1990 and 77.5% for 2000 and 2010. We constructed the ratio of this labor force with Mam and *WDR*. The root mean ratio for 1961-2010 is 1.018. The 1961 and 1968 ratios are 1.055 and 1.023, respectively. Labor force data for Uruguay for 1980 and 1990 are interpolated from 1975, 1985 and 1992 values. Labor force data for 2000 and 2010 come from *WDR*.

Real output come from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1961-1998 from Mam Table J1, pp. 780 and 784 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. The Salvucci chapter in *Cambridge Economic History of Latin America* provides sectoral output shares for 1870, 1908 and 1929. We interpolated intervening years, 1880, 1890, 1900, 1919. Mitchell provides sectoral output shares for 1939-2000, except for 1968, which we interpolated. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1800-1870. We used the US 1850-1940 capital - sectoral output ratios, applying 1850 values for 1800-1840, to construct our capital estimates for 1800-1940. We used Nehru and Dhareshwar (1993) for 1949-1990 capital stocks, we used the 1950 capital - output ratio for 1949. We used perpetual inventory for 2000 and 2010.

Enrollments in primary and secondary school from 1869-1993 come from Mam Table I1 pp. 735, 737, 741, 745 and 749. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. For primary school enrollment rates prior to 1870 (26%), we assumed 25% for 1800-1860. For secondary school enrollment rates prior to 1870 (.1%) we assumed .1% for all years. Higher education enrollments for 1886-1992 are from Mam Table I2 pp. 755, 756 and 757. For years between 1870 and our first observation 1900, we interpolated our 1870 assumed value .05% and the 1900 datum, .20%. For years 1800-1860 we assumed higher education enrollment rates of .01%. These produce 1870 education shares of 25 percent, 24 percent primary and 1 percent secondary. These are close to the values in Morrison & Murtin, 34 percent, 29 percent primary and 5 percent secondary. For 2010 we used *HDR*. For 2000 all enrollment rates were interpolated. Our time series for labor force years of schooling for 1870-1920 is: 1.49 (1870), 1.52 (1880), 1.59 (1890), 1.73 (1900), 1.93 (1908) and 2.74 (1919). The Morrisson & Murtin time series of years of schooling is: 1.61 (1870), 1.77 (1880), 1.94 (1890), 2.23 (1900), 2.50 (1910) and 2.50 (1920).

9.28 Venezuela (1800-2010)

Population for 1800 comes from McEvedy and Jones. We interpolated our 1810 value. Population for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1913, 1920, 1936, 1941, 1950, 1961 and 1971 come from Mam Table A1 p. 8. Population for 1980 and 1990 comes from S & H (1991). Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions of the population for 1936, 1941, 1950, 1961, 1971, 1980 and 1990 come from Mam Table A2 p. 30. We assumed the age distributions for years prior to 1936 are identical to the age distribution in 1936. The age distribution for Venezuela for 1980 and 1990 are interpolations using the 1971, 1981 and 1991 values. Age distributions for Venezuela 2000 and 2010 come from the *Demographic Yearbook*.

For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1941, 1950, 1961, 1971, 1980 and 1990 come from Mam Table B1 p. 110. To construct labor force data for 1800, 1810, 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1913, 1920, 1936, we used Banks for urban-rural population shares for 1830-1961. There are no urban share information for 1800-1820, however the urban share for 1830-1863 is 0, so we used that value for 1800-1820. We assumed a rural 15-64 labor force participation rate of 65% and an urban 15-64 labor force participation rate of 50% for 1800-1950. For 1961-2010, we maintain the 65% rural 15-64 labor force participation rate. For urban 15-64 labor force participation rate, we assumed 55% for 1961, 1971, and 1980, and 60% for 1990 and 70% for 2000 and 2010. With these rates we constructed labor force for 1800-2010. In the 1941-2010 overlapping years, we constructed the ratio of these labor force numbers with Mam and *WDR*. The root mean ratio is 1.020, and the 1950 and 1941 values are 1.036 and 1.046, respectively. Labor force data for Venezuela for 1980 and 1990 are interpolated from the 1971, 1981 and 1991 values. Labor force data for 2000 and 2010 come from *WDR*.

Real output come from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1950-1998 from Mam Table J1, pp. 781 and 785 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. Mitchell provides sectoral output shares for 1936-2000, and the *WDI* provides sectoral output shares for 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1800-1928. We used the US 1840-1940 capital - sectoral output ratios, applying the 1840 values for 1800-1830, to produce out capital estimates for 1800-1941. We used Nehru and Dhareshwar (1993) for 1950-1990. We used perpetual inventory for 2000 and 2010.

For primary school enrollment rates in years 1890, 1900, 1913, 1920, 1930, 1936 come from Benavot & Riddle (1988). For 1800-1880 we assumed a primary school enrollment rate of 1.6 percent, the 1890 value. Our figure is quite low relative to the Morrisson & Murtin 1870 primary share of 23%. However the Benavot & Riddle values smoothly splice with the Mitchell data. Enrollments in primary and secondary school from 1941-1993 come from Mam Table I1 pp. 741, 745 and 749. The 1920 secondary school enrollment rate observations is .31%. We assumed .31 percent enrollment rates in secondary school for 1800-1913. In 1870 we have .3% share of the labor force exposed to secondary school compared with 1% reported in Morrisson & Murtin. We assumed no higher education enrollments in these years. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-16 are secondary school age. Higher education enrollments for 1926-1991 are from Mam Table I2 pp. 756 and 757. For 1820-1920 we assumed 0 enrollments in higher education. For 2010 we used *HDR*. For 2000 we interpolated for all our enrollment rates. Our time series for labor share exposed to higher education is consistent with Morrisson & Murtin from 1870-1990. Our time series for years of schooling in the labor force for 1870-1940 is: .11 (1870-1900), .19 (1913), .36 (1920), .43 (1928), .69 (1936). The Morrisson & Murtin time series for years of schooling is: 1.10 (1870), 1.17 (1880), 1.25 (1890), 1.38 (1900), 1.51 (1910), 1.58 (1920), 1.66 (1930) and 1.81 (1940).

10 Middle East

10.1 Bahrain (1950-2010)

Population for 1950, 1960, 1970 come from Maddison. Populations for 1980 and 1990 come from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 acome from KF. Age distributions for Bahrain 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1980, 1990, 2000 and 2010 come from *WDR* (various years). For 1950, 1960 and 1970 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for 1960-2010. We assumed rural 15-64 labor force participation rates of 50% for 1950-2010. We assumed urban 15-64 labor force participation rates of 66.9% for 1950-1970 (which is the average urban labor force participation rate for 15-64 for 1980-2010), 58.5% for 1980, 67.59% for 1990, 63.82% for 2000, 77.7% for 2010. We constructed the ratio of this labor force to that from *WDR*. The root mean ratio for the overlapping years 1980-2010 is 1.001. The 1980 and 1990 ratios are 1.002 and 1.000, respectively. The range of the ratios is 1.000 to 1.002.

Real GNP come from Maddison. The values prior to 1970 are consistent with the fact of 1932 as the peak year of oil discovery as well as applying 2000 terms of trade on the early years. The 1980-2000 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). For 2010 we used the average investment rate from 2000-2009 from S & H. The *WDI* provides sectoral output shares for 1980-2010. We used Sabillon (2005) information farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1950-1970. We used the US 1950-1970 capital - sectoral output ratios to construct our 1950-1970 capital estimates. We used perpetual inventory for 1980-2010.

Enrollment rates for 1950 for all categories come from the 1954 *UN Statistical Yearbook*. Enrollment rates in primary, secondary schools and higher education for 1960-2000 come from *WDR* (various years). The 1960 observations for primary school enrollment rate, secondary school enrollment rate, and higher education enrollment rate are 72%, 9.5% and .37%, respectively. The latter two enrollment rates are interpolated. For 2000 we used UNESCO *UIS Global Database*. For 2010 we used *HDR*. We assumed the primary school and secondary school ages are 6-11 and 12-17.

10.2 Iran (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1913, 1920, 1930, 1940, 1950, 1956, 1966, 1971 and 1981 come from Maa Table A2 p. 21. Population for 1990 comes from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1956, 1966, 1971, 1980 and 1990 come from Maa Table A2 p.21. We assumed the age distributions prior to 1956 are identical to the age distribution in 1956. The age distribution for Iran for 1980 and 1990 are interpolated from 1971, 1986 and 1991 values. Age distributions for Iran 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 1996 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1956, 1966, 1976 and 1986 comes from Maa Table B1 p. 96. For 1820-1950 we used Banks for urban-rural population shares. Banks begins with year 1862. Issawi (1969) states that the urban share in the middle 18^{th} century was no higher than at the start of the century. So we assumed that the 1862 value holds for 1820, 1830, 1840 and 1850. We assumed a rural 15-64 labor force participation rate of 56%, and an urban 15-64 labor force participation rate of 45%. We assume this holds for all years. We compute the ratio of this labor force with the overlapping years from Maa and *WDR*. The root mean ratio for 1956-2010 is 0.995. The 1956 and 1966 values are .931 and 1.022, respectively. Labor force figures for Iran 2000 and 2010 come from *WDR*.

Real GNP comes from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1955-1990 and intraperiod average gross capital formation and income for 1991-1998 from Maa Table J1, p. 1029 and *WDR* (various years). For 2010 we used intraperiod average investment rates from S & H. We used Esfahani and Pesaran (2008) for sectoral output shares for 1900, 1940 and 1950. We interpolated for 1910, 1920 and 1930. We used Mitchell for sectoral output shares 1956-2000. The *WDI* provides sectoral output shares for 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1820-1890. We used the US 1850-1960 capital - sectoral output ratios, applying the 1850 values for 1820, 1830 and 1840, to construct our capital stock estimates for 1820-1956. We used perpetual inventory for 1960-2010.

We used Easterlin for enrollment rates for 1900 , 1910, and Mitchell for 1920, 1930, 1940. For 1950 we used the 1954 *UN Statistical Yearbook*. Enrollments in primary and secondary schools from 1960-1993 come from Maa Table I1 pp. 983 and 987. For 1900 and 1910, we used Easterlin (1981) Appendix Table 1. The 1920, 1930, 1940 and 1950 observations for primary school enrollment rates are: 1.7%, 6.3%, 14.6% and 33.4%, respectively. For primary school enrollment rates covering 1820-1890 we assumed 1.7%. This produces a time series of primary exposed workers that matches Morrisson & Murtin in 1940 and 1950, and is below their 1870-1930 values, but is consistent with the Easterlin, Lindert and Mitchell time series. The Mitchell observations for secondary school enrollment rates for this period are: .2% (1920), .6% (1930), 1.7% (1940) and 2.5% (1950). Finally, for these years, the Lindert higher education enrollment rates are: .1% (1920), .1% (1930), .1% (1940), .1% (1950). For years prior to 1920, we assumed secondary enrollment rates of .2 (1900-1910) and .1 percent (1820-1890) and higher education enrollment rates of .01 percent (1820-1860) and .05 percent (1870-1910). For secondary exposure and higher education, our time series fit Morrisson & Murtin from 1870-present. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1948-1991 come from Maa Table I2 p. 1003. For 2010 we used *HDR*. We interpolated for all 2000 enrollment rates. Our 1870-1920 time series of years of schooling in the labor force is: .02 (1870-1913), .03 (1920). For Morrisson & Murtin, their time series of years of schooling is: .29 (1870), .30 (1880), .31 (1890), .32 (1900), .34 (1910), .35 (1920).

10.3 Iraq (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1913, 1920, 1930, 1940, 1950, 1957, 1965 and 1977 comes from Maa Table A1 p. 8. Population for 1990 comes from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distribution for 1957, 1965, 1977 and 1990 come from Maa Table A2 p.21. The age distribution for year prior to 1957 is extrapolated from the age distribution in 1957. Age distribution for Iraq 2000 and

2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2001 age distribution, and for 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1957, 1977 and 1987 come from Maa Table B1 p. 96. Labor force for Iraq for 1965 is interpolated from the 1957 and 1977 figures. Labor force figures for Iraq 2000 and 2010 come from *WDR*. We used Clio Infra for urbanization rates for 1820-1950, and *WDR* and *WDI* for 1960-2010. We interpolated between 1800 and 1850 for 1820-1840, and we interpolated between 1850 and 1900 for 1860-1890, and we interpolated between 1900 and 1950 for 1910-1940. We assumed a rural 15-64 labor force participation rate of 60% for 1820-2000 and 50% for 2010. We assumed an urban 15-64 labor force participation rate of 46% for 1820-1957, 50% in 1965, 57% in 1977, 37% in 1990, 40% for 2000 and 35% for 2010. We constructed the ratio of this labor force to that from *WDI* and *WDR*. The root mean ratio for the overlapping years 1957-2010 is 1.002. The 1957 and 1965 values are 1.004 and .986. The range of the ratio is .986 to 1.010.

Real GNPs come from Maddison. Physical capital investment rates come from the intraperiod average gross capital formation and income for 1950-1998 from Maa Table J1, p. 1030 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. Mitchell provides sectoral output shares for 1950-2000. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1820-1940. We used the US 1850-1970 capital - sectoral output ratios, applying 1850 to 1820, 1830 and 1840, to construct our 1820-1965 capital estimates. We used perpetual inventory for 1977-2010.

Enrollments in primary and secondary schools from 1930-1993 come from Maa Table I1 pp. 983 and 987. For 1900-1920 primary school enrollment rates, we used Benavot Riddle (1988). The higher education enrollment rates for these years are: 1% and 1%. For years prior to 1900 we assumed that the primary enrollment rate was 1.4 percent, the 1900 value. For secondary school enrollment rates for 1900-1920, we used the average ratio of secondary school enrollment rates to primary school enrollment rates for 1930, 1940, 1950 and 1957, .16. For years prior to 1900, we assumed the constant 1900 value. These produce an 1870 exposure rate of 1.5 percent, 1.1 percent primary exposure, .3 percent secondary exposure. Morrisson & Murtin report 2 percent exposure, all of which are primary exposure. We assumed for higher education enrollment rates were 75 percent of the higher education enrollment rates of the succeeding decade. The higher education enrollment rate time series is: .02% (1820), .03% (1830), .04% (1840), .06% (1850), .08% (1860), .10% (1870), .13% (1880), .18% (1890), .24% (1900), .32% (1910), .43% (1920), .57% (1930), .76% (1940) and 1.0% (1950 datum). Our higher education exposure series matches Morrisson & Murtin from 1870-1950. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1940-1988 come from Maa Table I2 pp. 1001 and 1003. For 2010 we used *HDR*. For all 2000 enrollment rates we interpolated. Our time series of years of schooling in the labor force for 1870-1920 is: .10 (1870-1910), .20 (1920). Morrisson & Murtin's time series of years of schooling is: .10 (1870), .12 (1880), .14 (1890), .15 (1900), .17 (1910), .19 (1920).

10.4 Jordan (1820-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. We assumed the 1820-1940 age

distributions were identical to the 1950 age distribution. Age distribution for Jordan 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1960, 1970, 1980, 1990, 2000 and 2010 come from *WDR* (various years). For 1820-1950 we used the following procedure. We used Clio Infra for urbanization rates for 1820-1950, and *WDR* and *WDI* for 1960-2010. We interpolated between 1800 and 1850 for 1820-1840, and we interpolated between 1850 and 1900 for 1860-1890, and we interpolated between 1900 and 1950 for 1910-1940. We assumed a rural 15-64 labor force participation rate of 40% for 1820-1990, 41% for 2000 and 44% for 2010. We assumed an urban 15-64 labor force participation rate of 30% for 1820-1970, 35% for 1980, 25% for 1990, 41% for 2000 and 44% for 2010. We constructed the ratio of this labor force to that from *WDI* and *WDR*. The root mean ratio for the overlapping years 1960-2010 is 1.002. The 1960 and 1970 values are 1.004 and 1.003, respectively. The range of the ratio is 1.000 to 1.004.

Real output come from Maddison. The 1950-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1970-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1820-1960. We used US 1850-1950 capital - sectoral output ratios, applying the 1850 values for 1820, 1830 and 1840, to construct our capital estimates for 1820-1950. We used perpetual inventory for 1960-2010.

Enrollments in primary and secondary schools for 1950-1990 come from Maa Table I1 p. 987. For primary school enrollment rates in 1920, 1930 and 1940 we used Benavot and Riddle. The 1920 value is 5.5%. We assumed a 2% enrollment rate for years 1820-1910. For secondary schools, the 1950 data show that secondary school enrollment rates were about one eighth the contemporaneous primary school enrollment rates. We used this ratio to produce secondary school enrollment rates from 1820-1940. We assumed the primary school and secondary school ages are 6-11 and 12-17. The tertiary school enrollments are from Maa Table I2 p. 1004. The 1960 observation for higher education enrollment rate is .07%. We assumed the 1820-1950 higher education enrollment rate series is constant at .05%. For 2010 we used *HDR*. We interpolated for all 2000 enrollment rates.

10.5 Kuwait (1950-2010)

Population for 1950, 1960, 1970 come from Maddison. Populations for 1980 and 1990 come from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. Age distribution for Kuwait 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 1998 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1960, 1970, 1980, 1990, 2000 and 2010 come from *WDR* (various years). For 1950 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for 1960-2010. We assumed rural 15-64 labor force participation rates of 60% for 1950-1970, 50% for 1980-2010. We assumed urban 15-64 labor force participation rates of 95% for 1950-1960, 78.35% for 1970, 60.375% for 1980, 64.22% for 1990, 68.9% for 2000, 50.12% for 2010. We constructed the ratio of this labor force to that from *WDR*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real output come from Maddison. The extremely high values prior to 1970 are consistent with the fact of 1938 as the peak year of oil discovery as well as applying 2000 terms of trade on the early years. The 1980-2000 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). For 2010 we used the average investment rate from 2000-2009 from S & H. The *WDI* provides sectoral output shares for 1980-2010. We used Sabillon (2005) information on manufacturing and aggregate growth rates for Bahrain to construct sectoral output shares for 1950-1970. We assumed the 1980 value for farming, .004, was constant for 1950-1970. We used the US 1950-1970 capital - sectoral output ratios to construct our capital estimates for 1950-1970. We used Nehru and Dhareshwar (1993) for 1980 and 1990. We used perpetual inventory for 2000 and 2010.

Enrollment rates in primary and secondary schools for 1960-1990 come from *WDR* (various years). For 1950 enrollment rates we used the 1954 *UN Statistical Yearbook*. We assumed the primary school and secondary school ages are 6-9 and 10-17. The tertiary school enrollments are from *WDR* (various years). We used *HDR* for all 2010 enrollment rates. We interpolated for our 2000 enrollment rates.

10.6 Lebanon (1820-2010)

Population for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1913, 1920, 1930, 1940, 1950, 1960, 1970 come from Maddison. Populations for 1980 and 1990 come from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. We assumed the age distributions prior to 1950 were identical to the 1960 age distribution. Age distribution for Lebanon 2010 comes from the *Demographic Yearbook*. For 2000 we interpolated, and for 2010 we adjusted the 2007 age distribution. We adjust by assuming the same share by age category as in 2007.

Labor force figures for 1960, 1970, 1980, 1990, 2000 and 2010 come from *WDR* (various years). We used Clio Infra for urbanization rates for 1820-1950, and *WDR* and *WDI* for 1960-2010. We interpolated between 1800 and 1900 for 1820-1890, and we interpolated between 1900 and 1950 for 1910-1940. We assumed a rural 15-64 labor force participation rate of 40% for 1820-1990, 41% for 2000 and 44% for 2010. We assumed an urban 15-64 labor force participation rate of 57% for 1820-1980, 47% for 1990 and 60% for 2000 and 2010. We constructed the ratio of this labor force to that from *WDI* and *WDR*. The root mean ratio for the overlapping years 1960-2010 is 1.010. The 1960 and 1970 values are 1.029 and 1.035, respectively. The range of the ratio is .980 to 1.035.

Real output come from Maddison, except for 2010. For 2010 we used the 2008 Maddison value and applied the real PPP per capita income growth rates from *WDR 2011* and *WDR 2012*. The 1970-1990 investment rates are the intraperiod average investment rate taken from Mitchell. For 2000 and 2010 we used the average investment rate from 1990-2009 from S & H. The *WDI* provides sectoral output shares for 2000 and 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1820-1990. We used the US 1850-1960 capital - sectoral output ratios, applying 1850 values to 1820, 1830 and 1840, to construct our capital estimates for 1820-1960. We used perpetual inventory for 1970-2010.

Enrollment rates in primary, secondary and higher education for 1960-1990 come from *WDI* (various years). For 1920-1940 primary enrollment rates we used Benavot Riddle (1988). For 2010 we used *WDR*. For 2000 we interpolated our enrollment rates. The 1950 and 1960 observations for secondary school enrollment rates are: 6.4% and 19%, respectively. The 1950 and 1960 higher education enrollment rates

are: .48% and 5%., respectively. For years 1900-1913 we assumed that primary, secondary and 1900-1930 for higher education enrollment rates were 95 percent of the primary, secondary and higher education enrollment rates of the succeeding decade. For 1820-1890, we assumed that primary, secondary and higher education enrollment rates were 75 percent of the primary, secondary and higher education enrollment rates of the succeeding decade. The primary school enrollment rate time series for 1820-1950 is: 2.9% (1820), 3.8% (1830), 5.1% (1840), 6.8% (1850), 9.1% (1860), 12.1% (1870), 16.2% (1880), 21.6% (1890), 28.8% (1900), 30.3% (1913), 31.9% (1920 datum). The secondary school enrollment rate time series for 1820-1950 is: .1% (1820), .1% (1830), 0.2% (1840), 0.2% (1850), 0.3% (1860), 0.4% (1870), 0.5% (1880), 0.7% (1890), 0.9% (1900), 1.0% (1913), 1.0% (1920 datum). The higher education enrollment rate time series for 1820-1950 is: .04% (1820), .05% (1830), .07% (1840), .09% (1850), .12% (1860), .16% (1870), .21% (1880), .28% (1890), .37% (1900), .39% (1913), .41% (1920), .44% (1930), .48% (1940 datum). We assumed the primary school and secondary school ages are 6-10 and 11-17.

10.7 Oman (1950-2010)

Populations for 1950, 1960, 1980 and 1990 come from S & H online. For 1970, population comes from *Population Pyramid*, https://populationpyramid.net/oman/1970/, accessed July 24, 2016. We used the 1989 value for the 1990 observation. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 come from KF. Age distributions for Oman 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2001 age distribution, and for 2010 we interpolated using the 2009 and 2012 age distributions.

Labor force figures for 1980, 1990, 2000 and 2010 come from *WDR* (various years). For 1950, 1960 and 1970 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for 1960-2010. We assumed rural 15-64 labor force participation rates of 52.75% for 1950-1980, 50% for 1990, 49.72% for 2000, and 50% for 2010. We assumed urban 15-64 labor force participation rates of 95% for 1950-1980, 80.95% for 1990, 49.72% for 2000, 59.05% for 2010. We constructed the ratio of this labor force to that from *WDR*. The root mean ratio for the overlapping years 1980-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real output come from Maddison. The values prior to 1970 are consistent with the fact of 1962 as the peak year of oil discovery as well as applying 2000 terms of trade on the early years. The 1980-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1961-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1960 and 1950. For 1950 we renormalized sectoral output shares as the calculations produced negative share for services of -.001. We imposed that .001 was the service share in 1950 and renormalized farming and manufacturing. We used the US 1950 and 1960 capital - sectoral output ratios to produce our 1950 and 1960 capital estimates. We used perpetual inventory for 1970-2010.

Enrollment rates for primary and secondary schools for 1970-2010 come from *WDR*, (various years). For 1950 and 1960 we assumed the same enrollment rates as for 1970. Thus the 1950 and 1960 primary and secondary school enrollment rates are: 3% and 0%, respectively. We assumed the primary school and secondary school ages are 6-11 and 12-17. The tertiary school enrollments are from *WDR*, (various years).

10.8 Qatar (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from S & H online. We used the 1989 value for the 1990 observation. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Qatar 2010 comes from the *Demographic Yearbook*. For 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in 2012.

Labor force figures for 1960, 1970 and 1980 come from *WDI* (various years). For 1990, 2000 and 2010 we used *WDR*. For 1950 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for 1960-2010. We assumed rural 15-64 labor force participation rates of 50% for 1950-1980, 103.6% for 1990, 50% for 2000, 2010. We assumed urban 15-64 labor force participation rates of 78.12% for 1950 (which is the average urban 15-64 labor force participation rate for 1960-2010), 56.8% for 1960, 68.65% for 1970, 68.2% for 1980, 103.6% for 1990, 82.85% for 2000, 88.64% for 2010. We constructed the ratio of this labor force to that from *WDR*. The root mean ratio for the overlapping years 1960-2010 is 1.000. The 1960 and 1970 ratios are 1.000 and 1.000, respectively. The range of the ratios is .999 to 1.000.

Real output come from Maddison. The extremely high values prior to 1970 are consistent with the fact of 1940 as the peak year of oil discovery as well as applying 2000 terms of trade on the early years. The 1990, 2000, 2010 investment rates are the intraperiod average investment rate taken from S & H online. The *WDI* provides sectoral output shares for 2000 and 2010. We used Sabilllon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1950-1990. For farming, we assumed that the 1980 sectoral output share was .007, compared with .004 in 1990 and .004 in 2000. We then applied the relative farming compared to aggregate growth rate from 1970s, 1960s and 1950s. We used the 1950-1980 US capital - sectoral output ratios to produce our 1950-1980 capital stock estimates. We used perpetual inventory for 1990-2010.

Enrollment rates for primary, secondary and higher education schools for 1960-1990 come from *WDI* (various years). For 1950 enrollment rates we used the 1954 *UN Statistical Yearbook*. For 2000 we used UNESCO *UIS Global Database*. For 2010 we used *HDR*. We assumed the primary school and secondary school ages are 6-11 and 12-17.

10.9 Saudi Arabia (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from S & H online. We used the 1989 value for the 1990 observation. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distributions for Saudi Arabia 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2012 age distribution. We adjust by assuming the same share by age category as in 2012.

Labor force figures for 1960, 1970, 1980, 1990, 2000 and 2010 come from *WDR* (various years). For 1950 we used Banks for urban-rural population shares for 1950, 1960, and *WDI* for 1970, 1980, 1990, 2000 and 2010. We assumed a rural 15-64 labor force participation rate of 50% for 1950-1980, and 65% for 1990, 51% for 2000 and 52% for 2010. We assumed an urban 15-64 labor force participation rate of 50% for 1950-1980, and 65% for 1990, 51% and 52% for 2000 and 2010, respectively. We constructed the ratio of

this labor force to that from WDR. The root mean ratio is .998 for the overlapping years 1960-2010. The 1960 and 1970 ratios are 1.003, and 1.020, respectively. The range of values is .987 to 1.025.

Real output come from Maddison. The extremely high values prior to 1970 are consistent with the fact of 1948 as the peak year of oil discovery as well as applying 2000 terms of trade on the early years. The 1980-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The 1973 *UN Statistical Yearbook* provides sectoral output shares for 1960. The *WDI* provides sectoral output shares for 1970-2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1950. We used US 1950-1970 capital - sectoral output ratios to construct our capital estimates for 1950-1970. We used perpetual inventory for 1980-2010.

Enrollments in primary and secondary schools for 1950-1990 come from Maa Table I1 p. 990. We assumed the primary school and secondary school ages are 6-11 and 12-17. The tertiary school enrollments are from Maa Table I2 p. 1005. For 2010 we used *HDR* for all enrollment rates. We interpolated for all enrollment rates in 2000.

10.10 Syria (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1913, 1920, 1930, 1940, 1950, 1960, 1970 and 1981 from Maa Table A1 p. 9. Population in 1950 is interpolated from 1946 and 1960 values. Population for 1990 comes from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1960, 1970, 1981 and 1991 come from Maa Table A2 p. 26. The age distribution prior to 1960 is extrapolated from the age distribution in 1960. The age distributions for 1981 and 1991 are interpolated from the age distributions in 1970 and 1992. Age distributions for Syria 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2001 age distribution, and for 2010 we adjusted the 2011 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for 1960, 1970, 1981 and 1991 come from Maa Table B1 p. 99. Labor force figures for Syria 2000 and 2010 come from *WDR*. For 1820-1950 we used the following procedure. We used Clio Infra for urbanization rates for 1820-1950, and *WDR* and *WDI* for 1960-2010. We interpolated between 1800 and 1900 for 1820-1890, and we interpolated between 1900 and 1950 for 1913-1940. We assumed a rural 15-64 labor force participation rate of 60% for 1820-1960, 64% for 1970, 60% for 1981, 64% for 1991, 60% for 2000 and 45% for 2010. We assumed an urban 15-64 labor force participation rate of 30% for 1820-1975, 48% in 1980, 56% in 1990, 61% in 2000 and 2010. We constructed the ratio of this labor force to that from *WDI* and *WDR*. The root mean ratio for the overlapping years 1960-2010 is 1.022. The 1960 and 1970 values are 1.068 and 1.007. The range of the ratio is .999 to 1.068.

Real output come from Maddison. Physical capital investment rates come from the intraperiod average investment rate for 1970-1998 from S & H online and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. We used the 1961 *UN Statistical Yearbook* for 1950 and 1960 sectoral output shares. Mitchell provides sectoral output shares for 1970-1990. The *WDI* provides sectoral output shares for 2000 and 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1820-1940. We used US 1850-1960 capital - sectoral output ratios, applying 1850 values for 1820, 1830 and 1840, to construct our capital estimates for 1820-1960. We used perpetual inventory for 1970-2010.

Enrollments in primary and secondary schools from 1927-1993 come from Maa Table I1 pp. 985 and 990. For 1900 and 1910 we use Benavot and Riddle (1988). We interpolated for 1920. For secondary school enrollment rates from 1820-1920, we used the average ratio of secondary school enrollment rate to primary school enrollment rate for 1920 and 1930. For years prior to 1900 we assumed that the primary enrollment rate was 95% of the succeeding decade value. Our time series for primary school enrollment rates is: 4.0% (1820), 4.3% (1830), 4.5% (1840), 4.7% (1850), 5.0% (1860), 5.2% (1870), 5.5% (1880), 5.8% (1890), 6.1% (1900 datum). Our time series for secondary school enrollment rates is: 0.04% (1820), 0.07% (1830), 0.15% (1840), 0.29% (1850), 0.31% (1860), 0.32% (1870), 0.34% (1880), 0.36% (1890), 0.38% (1900), 0.58% (1913), 0.84% (1920), 1.2% (1930 datum). Our series produces 1870 education shares of 5 percent exposure, 4.8 percent primary exposure and .3 percent secondary exposure. This compares with Morrisson & Murtin series of 4 percent exposure, 3 percent primary and 1 percent secondary. Our time series for higher education enrollment rates is: 0.00004% (1820), 0.0001% (1830), 0.0002% (1840), 0.0003% (1850), 0.001% (1860), 0.001% (1870), 0.0003% (1880), 0.01% (1890), 0.01% (1900), 0.02% (1913), 0.04% (1920), 0.08% (1930), and 0.14% (1940 datum). Our higher education series matches Morrisson & Murtin for 1870-1950. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollment rates for 1937-1992 come from Maa Table I2 p. 1002 and 1005. For 2010 we used *HDR*. We assumed a 20% higher education enrollment rate for 2010. We interpolated for our 2000 enrollment rates. Our 1870-1930 time series of years of schooling in the labor force is: .31 (1870), .33 (1880), .35 (1890), .37 (1900), .39 (1913), .52 (1920) and .80 (1930). The Morrisson & Murtin time series of years of schooling is: .29 (1870), .38 (1880), .45 (1890), .53 (1900), .61 (1910), .69 (1920), and .77 (1930).

10.11 United Arab Emirates (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for UAE 2010 comes from the *Demographic Yearbook*. For 2010 we adjusted the 2005 age distribution. We adjust by assuming the same share by age category as in 2005.

Labor force figures for 1980, 1990, 2000 and 2010 come from *WDR* (various years). For 1950, 1960, 1970 we used the following procedure. We used Clio infra for 1950 urbanization rates, and *WDI* for 1960-2010. We assumed urban and rural 15-64 labor force participation rates of 95.9% for 1950-1980, 86.76% for 1990, 103.55% for 2000, 74.82% for 2010. We constructed the ratio of this labor force to that from *WDR*. The root mean ratio for the overlapping years 1980-2010 is 1.000. The 1980 and 1990 ratios are 1.000 and 1.000, respectively. The range of the ratios is 1.000 to 1.000.

Real output come from Maddison. The extremely high values prior to 1970 are consistent with the fact of 1964 as the peak year of oil discovery as well as applying 2000 terms of trade on the early years. However it does appear that 1950 could be high. The 1980-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). We used *WDI* for sectoral output shares for 2000 and 2010; we used the 2001 value for 2000 and the 2007 value for 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1950-1990. We used US 1950-1970 capital - sectoral output ratios to construct our 1950-1970 capital estimates. We used perpetual inventory for 1980-2010.

Enrollment rates for primary and secondary schools for 1970-1990 come from *WDR*, (various years). The 1970 observations for primary, secondary and higher education enrollment rates are: 93%, 22% and .5%, respectively. We assumed the primary school and secondary school ages are 6-11 and 12-17. The tertiary school enrollments are from *WDR* (various years). For 1950 and 1960 we assumed 25 percent and 50 percent enrollment rates in primary school, 10 percent each for secondary school enrollment rates, and .1 percent for each year for higher education enrollment rates. For 2010 enrollment rates we used *HDR*. We interpolated for our 2000 enrollment rates.

10.12 Yemen (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Yemen 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in 2013.

Labor force figures for 1960, 1970, 1980 come from *WDI* (various years). Labor force data for 1990, 2000 and 2010 come from *WDR*. For 1950 we used Banks for urban-rural population shares for 1950 and 1960. We used *WDR* and *WDI* for 1970, 1980, 1990, 2000 and 2010. We assumed an urban 15-64 labor force participation rate of 60% for 1950, 1960, 1970 and 1980, 41.25% for 1990 and 2000, and 47.5% for 2010. We assumed a rural 15-64 labor force participation rate of 64% for 1950, 1960, 1970 and 1980, and 50% for 1990, 2000 and 2010. We constructed the ratio of this labor force to that from *WDI* and *WDR*. The root mean ratio of the overlapping years 1960-2010 is 1.016, and the 1960 and 1970 values are .988 and 1.008, respectively.

Real output come from Maddison. The 1980-2010 investment rates are the intraperiod average investment rate taken from S & H online and *WDR* (various years). The *WDI* provides sectoral output shares for 1990 and 2000. The 1981 *UN Statistical Yearbook* provides sectoral output shares for 1970. We interpolated for 1980 sectoral output shares. We used Sabillon (2005) information on manufacturing and aggregate growth rates to construct our 1950 and 1960 manufacturing shares. For 1950 and 1960, we assumed that services were 30% of the non-manufacturing sector, and farming was 70% of the non-manufacturing sector. This is consistent with the information in 1970. We used the US 1950-1970 capital - sectoral output ratios to produce our 1950-1970 capital estimates. We used perpetual inventory for 1980-2010.

Our 1950 enrollment rates in all categories come from the 1954 *UN Statistical Yearbook*. Our 1960 enrollment rates in all categories come from the 1963 *UN Statistical Yearbook*. Enrollment rates for primary and secondary schools for 1970-1990 come from *WDR*, (various years). We assumed the primary school and secondary school rates are 6-13 and 14-19. The tertiary school enrollments, for 1970-1990, are from *WDR*, (various years). We used the *HDR* for all enrollment rates in 2010. We interpolated all enrollment rates in 2000.

11 North Africa

11.1 Algeria (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1913, 1920, 1930, 1936, 1948, 1954, 1966, 1980, 1987 come from Maa Table A2 p. 14. Population for 1948 and 1980 are interpolated from 1936 and 1954, 1966 and 1987 values. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1948, 1954, 1966, 1980 and 1987 come from Maa Table A2 p. 14. The age distribution for Algeria for 1980 is interpolated from 1966 and 1987 values. The age distribution prior to 1948 is assumed identical to the 1948 age distribution. Age distributions for Algeria 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2008 age distribution. We adjust by assuming the same share by age category as in 2008

Labor force figures for 1948, 1954, 1980 and 1987 come from Maa Table B1 p. 90. For 1960 and 1973 we used *WDR*. For 1966 we interpolated. Labor force figures for Algeria 2000 and 2010 come from *WDR*. For 1820-1956 we used the following procedure. We used Clio Infra for urbanization rates for 1820-1950, and *WDR* and *WDI* for 1960-2010. We interpolated between 1800 and 1850 for 1820-1840, and we interpolated between 1850 and 1900 for 1860-1890, and we interpolated between 1900 and 1950 for 1913-1940. We assumed a rural 15-64 labor force participation rate of 75% for 1820-1954, 55% for 1960, 1966 and 1973, and 50% for 1980-2010. We assumed an urban 15-64 labor force participation rate of 50% for 1820-2010. We constructed the ratio of this labor force to that from *WDI* and *WDR*. The root mean ratio for the overlapping years 1948-2010 is 1.019. The 1948 and 1954 values are .948 and .995. The range of the ratio is .948 to 1.075.

Real output come from Maddison. Physical capital investment rates come from the intraperiod average gross real capital formation and real income for 1950-1998 from Maa Table J1, p. 1010 and *WDR* (various years). For 2010 we used the average investment rate for 2000-2009 from S & H. Mitchell provides sectoral output shares for 1954, 1966, 1980-2000. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1820-1948, as well as for 1960 and 1973. We used the US 1840-1940 capital - sectoral output ratios, applying 1840 to 1820 and 1830, to produce our capital estimates for 1820-1936. We used *WDI* for our 2010 sectoral output shares. We used Nehru and Dhareshwar (1993) for capital stocks in 1948-1990. We used the 1950 capital output ratio to compute our 1948 value. We used perpetual inventory for 2000 and 2010.

Enrollments in primary and secondary schools from 1860 (primary) 1873 (secondary)-1993 come from Maa Table I1 pp. 966, 968 and 973. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1910-1993 come from Maa Table I2 pp. 996 and 997. Our primary school enrollment rate time series is: 2.70% (1820-1850), 2.70% (1860 datum). We assumed that each secondary school enrollment rate from 1820-1860 was equal to the 1870 secondary school enrollment rate relative to the the 1870 primary enrollment rate multiplied by the 1860 primary enrollment rate. Our secondary school enrollment rate time series is: .42% (1820-1860), .77% (1870 datum). These produce education exposure in 1870 of 3 percent, primary exposure of 3 percent and secondary school exposure of 0 percent. This is similar to Morrisson & Murtin values of 9 percent exposure, 9 percent primary exposure, 0 percent secondary exposure. Our higher education enrollment rate time series is: 0% (1820-1850), .01% (1860), .01% (1870), .04% (1880), .05% (1890), .06% (1900), .08%

(1910 datum). For 2010 we used *HDR* for all enrollment rates. We interpolated for our 2000 enrollment rates. Our time series of years of schooling in the labor force for 1870-1920 is: .23 (1870), .36 (1880), .45 (1890), .66 (1900), .86 (1913), .99 (1920). The Morrisson & Murtin time series of years of schooling is: .40 (1870), .44 (1880), .48 (1890), .50 (1900), .62 (1910), .72 (1920).

11.2 Egypt (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1910 come from Mitchell. Populations for 1917, 1927, 1937, 1947, 1960, 1966 and 1975 come from Maa Table A2 p. 15. Population for 1966 is geometrically interpolated from 1960 and 1975 values. Populations for 1980 and 1990 come from S & H online. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1917, 1927, 1937, 1947, 1960 and 1975 come from Maa Table A2 p. 15. The age distributions for 1820-1910 are assumed to be identical to the age distribution in 1917. The age distributions for Egypt for 1966, 1980, 1990 are interpolated from 1960, 1975, 1992 values. Age distributions for Egypt 2000 and 2010 come from the *Demographic Yearbook*. For 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in 2013.

Labor force figures for 1917, 1927, 1937, 1947, 1960, 1966, 1976 and 1986 come from Maa Table B1 p. 91. We used Saleh (2015) for labor force data for years 1850, 1870, 1900, 1910. We interpolated for 1860, 1880, 1890. Saleh provides labor force participation rates by gender for 1848, 1868, 1897, 1907, 1917. His data can be checked against Maa for years 1917, 1927 and 1937, and are quite similar. Labor force figures for Egypt 2000 and 2010 comes from *WDR*. For 1820, 1830 and 1840 we used the following procedure.[25] For 1820-1950 we used Clio Infra for urbanization rates. We used Clio Infra for urbanization rates for 1820-1950, and *WDR* and *WDI* for 1960-2010. We interpolated between 1800 and 1850 for 1820-1840, and we interpolated between 1850 and 1900 for 1860-1890, and we interpolated between 1900 and 1950 for 1913-1940. We assumed a rural 15-64 labor force participation rate of 56.75% for 1820-1840 (which is the average rural labor force participation rate of 15-64 we used for 1850-1900), 42.5% for 1850, 41.25% for 1860, 40% for 1870, 53.75% for 1880, 70% for 1890, 93% for 1900, 54.25% for 1910, 94.875% for 1917, 79.5% for 1927, 72% for 1937, 60.5% for 1947, 56.75% for 1960, 51.5% for 1966, 44.5% for 1973, 1980, 54.5% for 1990, 48.9% for 2000 and 53.75% for 2010. We assumed an urban 15-64 labor force participation rate of 40% for 1820-1870, 50% for 1880, 1890, 55% for 1900, 50% for 1910-1966, 44.5% for 1973, 43.5% for 1980, 50% for 1990, 48.9% for 2000 and 51% for 2010 . We constructed the ratio of this labor force to that from *WDI* and *WDR*. The root mean ratio for the overlapping years 1850-2010 is 1.001. The 1850 and 1860 values are 1.003 and 1.002. The range of the ratio is .998 to 1.003.

Real output come from Maddison. From 1960 onward we used the intraperiod average investment rate from S & H online and *WDR* (various years). Mitchell provides sectoral output shares for 1960-2000. The *WDI* provides sectoral output shares for 2010. The 1956 *UN Statistical Yearbook* provides sectoral output shares for 1947. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1820-1937. We used the US 1850-1950 capital - sectoral output ratios, applying the 1850 values for 1820, 1830 and 1840, to construct our capital estimates for 1820-1947. We used perpetual inventory for 1960-2010.

[25] Our 1900 value of 4.482 million workers is consistent with about 8% of the workforce in industry. An estimate, contained in Owen, of industrial workers in 1897 produces a value of 352 thousand industrial workers, Our estimated manufacturing share of GDP is 9.2% in 1900.

Enrollment rates for 1850-1890 come from Morris and Adelman. Our 1850 observations for primary school, secondary school and higher education enrollment rates are: 2%, .2%, and .01%, respectively. For 1820, 1830 and 1840, we assumed 2 percent primary enrollment rates, .2 percent secondary enrollment rates and 0 percent higher education enrollment rates. Enrollment rates for 1900, 1910 and 1917 come from Lindert. This produces 1870 education shares of 2 percent, 2 percent primary exposure and 0 percent secondary exposure. This compares with Morrisson & Murtin values of 2 percent, 1 percent primary exposure and 1 percent secondary exposure. Enrollments in primary and secondary schools from 1920-1993 come from Maa Table I1 pp. 968 and 974. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Prior to 1910 we assumed 0 percent enrollment rate for 1820-1840, 0.01 percent for 1850-1890, and .02 percent for 1900. Higher education enrollments for 1910-1993 come from Maa Table I2 pp. 996 and 997. Our higher education share match Morrisson & Murtin 1820-1950. For 2010 we used *HDR* for all enrollment rates. For 2000 we interpolated all enrollment rates. Our time series of years of schooling in the labor force for 1870-1930 is: .11 (1870), .12 (1880), .12 (1890), .13 (1900), .10 (1910), .54 (1917) and 1.05 (1927). The Morrisson & Murtin time series of years of schooling is: .15 (1870), .17 (1880), .23 (1890), .30 (1900), .42 (1910), .45 (1920) and .74 (1930).

11.3 Libya (1950-2010)

Populations for 1950, 1960, 1970, 1980 and 1990 come from Maa Table A5 p 53. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1950, 1960, 1970, 1980, 1990 and 2000 come from KF. Age distribution for Libya 2010 comes from the *Demographic Yearbook*. For 2010 we adjusted the 2006 age distribution. We adjust by assuming the same share by age category as in 2006.

Labor force figures for 1960, 1970, 1980, 1990, 2000 and 2010 come from *WDR* (various years). For 1950 we used Banks for urban-rural population shares for 1950 and 1960. We adjusted his 1952 value to reflect urban growth. We used *WDR* and *WDI* for urban-rural population shares for 1970, 1980, 1990, 2000 and 2010. We assumed a rural 15-64 labor force participation rate of 55% for 1950, 1960 and 1970, 80% for 1980, and 70% for 1990, 2000 and 2010. We assumed an urban 15-64 labor force participation rate of 50% for 1950, 1960, 1970, 1990 and 2010, 65% for 1980 and 2000. We constructed the ratio of this labor force with that from *WDR*. The root mean ratio is 1.007 for the overlapping years 1960-2010. The 1960 and 1970 values are 1.014 and 1.015, respectively. No values are below .975 and only one value is at 1.048.

Real output for 1950, 1960, 1970, 1980, 1990 and 2000 come from Maddison. Real output for 2010 comes from the 2000 Maddison value with the real PPP per capita income growth rates from 2000 to 2010 from *WDI* applied. The 1960-2000 investment rates are the intraperiod average investment rate taken from Maa Table J1 p. 1016 and *WDR* (various years). The *WDI* provides sectoral output shares for 2000 and 2010. We assumed that the 2000 farming share, .052, holds for 1990 and 1980. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1950-1990. We used the US 1950-1970 capital - sectoral output ratios to construct our capital estimates for 1950-1970. We used Nehru and Dhareshwar (1993) for 1980 and 1990. We used perpetual inventory for 2000 and 2010.

Enrollment rates in primary and secondary schools for 1960-1990 come from *WDR*, (various years). For 1950 we used Banks (1971). We assumed the primary school and secondary school ages are 6-11 and 12-17. The tertiary school enrollments are from *WDR* (various years). For 2010 we used *HDR* for all enrollment

rates. For 2000 we interpolated all enrollment rates.

11.4 Morocco (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1913, 1920, 1930, 1940, 1951, 1960, 1971 and 1980 come from Maa Table A2 p. 16. Population for 1990 comes from Maddison. Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1951, 1960, 1971, 1980 and 1990 come from Maa Table A2 p. 16. We assumed the age distribution prior to 1951 was identical to the age distribution in 1951. The age distribution for Morocco for 1980 and 1990 are interpolations using 1971, 1982 and 1993. Age distributions for Morocco 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 2002 age distribution, and for 2010 we adjusted the 2013 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for Morocco for 1982, 1990, 2000 and 2010 come from *WDR* (various years). For years prior to 1982 we used the following procedure. We used Clio Infra for urbanization rates for 1820-1900, and UN 1969 for urbanization rates for 1920-1960, and *WDR* and *WDI* for 1970-2010. We interpolated between years 1800 and 1850 for the 1820-1840 values, and interpolated between 1850 and 1900 for the 1860-1890 values. Finally we interpolated between 1900 and 1920 for the 1913 value. We assumed a rural 15-64 labor force participation rate of 72.5% for 1982, and 53% for 1990, 2000 and 2010. We assumed an urbal 15-64 labor force participation rate of 50% for 1982-2010. We constructed the ratio of this labor force to that from *WDR*. The root mean ratio is 1.001. The 1982 and 1990 values are .997, .998. The range of the ratios is .997 to 1.011.

Real output come from Maddison. Physical capital investment rates for 1960-2009 come from the intraperiod average investment rates from S & H online and *WDR* (various years). Mitchell provides sectoral output shares for 1951-2000. The *WDI* provides sectoral output shares for 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1820-1940. We used US 1840-1940 capital - sectoral output ratios, applying 1850 values to 1820, 1830 and 1840, to construct our 1820-1940 capital estimates. We used Nehru and Dhareshwar (1993) for capital stocks in 1951-1990. We used perpetual inventory for 2000 and 2010.

Enrollments in primary and secondary schools from 1914-1993 come from Maa Table I1 pp. 970 and 976. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1914-1993 come from Maa Table I2 pp. 996 and 999. For years prior to 1913 we assume primary, secondary and higher education enrollment rates that were 90 percent of the primary, secondary and higher education enrollment rates of the succeeding decade. Our primary school enrollment rate time series is: 1.4% (1820), 1.6% (1830), 1.7% (1840), 1.9% (1850), 2.1% (1860), 2.4% (1870), 2.6% (1880), 2.9% (1890), 3.3% (1900), 3.6% (1913 datum). Our secondary school enrollment rate time series is: .2% (1820), .2% (1830), .3% (1840), .3% (1850), .3% (1860), .4% (1870), .4% (1880), .4% (1890), .5% (1900), .55% (1913 datum). Our higher education enrollment rate time series is: .003% (1820), .003% (1830), .003% (1840), .004% (1850), .004% (1860), .005% (1870), .005% (1880), .006% (1890), .006% (1900), .007% (1913 datum). These produce nearly an exact match of the 1870-1950 education shares from Morrisson & Murtin. For all enrollment rates in 2010 we used *HDR*. We interpolated for all enrollment rates in 2000. Our time series of years of schooling in the labor force for 1870-1920 is: .12 (1870), .13 (1880), .15 (1890), .16 (1900), .19 (1913), ..20 (1920). The Morrisson & Murtin time series of years of schooling is:

.05 (1870), .06 (1880), ..08 (1890), .10 (1900), .15 (1910), .21 (1920).

11.5 Tunisia (1820-2010)

Populations for 1820, 1830, 1840, 1850, 1860, 1870, 1880, 1890, 1900, 1913, 1920, 1930, 1940, 1950, 1956, 1966 comes from Maa Table A2 p. 18. Population for 1973 is geometrically interpolated from 1966 and 1984 from Maa Table A2 p. 18. Populations for 1981 and 1990 come from S & H (1991). Population for 2000 comes from *Time Almanac 2001*. Population for 2010 comes from *Wikipedia*.

The age distributions for 1956 and 1966 come from Maa Table A2 p. 18. We assumed the same age distribution in year prior to 1956 as for 1956. The age distribution for Tunisia 1973 is geometrically interpolated from 1966 and 1984 from Maa Table A1 p. 18. The age distributions for Tunisia for 1990 comes from DK. Age distributions for Tunisia 2000 and 2010 come from the *Demographic Yearbook*. For 2000 we adjusted the 1998 age distribution, and for 2010 we adjusted the 2008 age distribution. We adjust by assuming the same share by age category as in the reference year.

Labor force figures for Tunisia 1966, 1975, 1984 and 1990 come from *WDI*. Labor force figures for Tunisia 2000 and 2010 come from *WDR*. For 1820-1956 we used the following procedure. We used Clio Infra for urbanization rates for 1820-1950, and *WDR* and *WDI* for 1960-2010. We interpolated between 1800 and 1850 for 1820-1840, and we interpolated between 1850 and 1900 for 1860-1890, and we interpolated between 1900 and 1950 for 1913-1940. We assumed a rural 15-64 labor force participation rate of 58% for 1820-2010, and an urban 15-64 labor force participation rate of 50% for 1820-2010. We constructed the ratio of this labor force to that from *WDI* and *WDR*. The root mean ratio for the overlapping years 1966-2010 is .987. The 1966 and 1973 values are .908 and 1.000. The range of the ratio is .908 to 1.051.

Real GNP come from Maddison. Physical capital investment rates for 1966-2010 come from the intraperiod average investment rates from S & H online and *WDR* (various years). Mitchell provides sectoral output shares for 1956-2000, and the *WDI* provides sectoral output shares for 2010. We used Sabillon (2005) information on farming, manufacturing and aggregate growth rates to construct sectoral output shares for 1820-1950. We used US 1850-1950 capital - sectoral output ratios, applying 1850 values for 1820, 1830 and 1840, to construct our 1820-1950 capital estimates. We used Nehru and Dhareshwar (1993) for capital stocks in 1956-1990; we used the 1960 capital output ratio for 1956. We used perpetual inventory for 2000 and 2010.

Enrollments in primary and secondary schools from 1895-1993 come from Maa Table I1 p. 967, 972 and 979. To calculate enrollment rates, we assumed 6-11 are primary school age and 12-17 are secondary school age. Higher education enrollments for 1943-1993 come from Maa Table I2 p. 996 and 1000. For years prior to 1900 we assumed that primary enrollment rates, secondary enrollment rates were 85 percent of the succeeding decade's primary, secondary, and for higher education enrollment rates we assumed 90 percent of the succeeding decade value. Our primary school enrollment rate time series is: 2.5% (1820), 2.8% (1830), 3.1% (1840), 3.5% (1850), 3.9% (1860), 4.3% (1870), 4.8% (1880), 5.3% (1890) (1900 datum). Our secondary school enrollment rate time series is: 0.3% (1820), 0.3% (1830), 0.4% (1840), 0.4% (1850), 0.5% (1860), 0.5% (1870), 0.6% (1880), .66% (1890) (1900 datum). Our higher education enrollment rate time series is: .1% (1820-1830), .2% (1840-1850), .3% (1860), .4% (1870), .6% (1880), .7% (1890), .1% (1900), .1% (1910), .2% (1920), .2% (1930), .3% (1940 datum). For 2010 we used *HDR* for all enrollment rates. For all 2000 enrollment rates we interpolated. These produce education shares that match Morrisson & Murtin for 1870-1900. Our time series for years of schooling in the labor force is: .25 (1870), .28 (1880),

.31 (1890), .35 (1900), .39 (1913) and .62 (1920). The Morrisson & Murtin time series of years of schooling is: .30 (1870), .33 (1880), .34 (1890), .37 (1900), .39 (1910) and .42 (1920).

References

Agnew, H. *The Czechs and the Lands of the Bohemian Crown*, Stanford, CA: Hoover Institution Press, 2004.

Alexander, H. B. "Brazilian and United States Slavery Compared,"Journal of Negro History 7, 1922: 349-364.

Allen, R.C. "Economic Structure and Agricultural Productivity in Europe, 1300-1800,"*European Review of Economic History* 3, 2000: 1-25.

Altuğ, S., Filiztekin, A., Pamuk, Ş. "Sources of Long-Term Economic Growth for Turkey, 1880-2005."*European Review of Economic History* 12, 2008: 393-430.

Astorga, P., Berges, A., and Fitzgerald, V. "The Standard of Living in Latin America During the Twentieth Century, "University of Oxford working paper, 2004.

Bairoch, P., Goertz, G. "Factors of Urbanisation in the Nineteenth Century Developed Countries: A Descriptive and Econometric Analysis,"*Urban Studies* 23, 1986: 285-305.

Banerjee, A., and Duflo, E. "Growth Theory through the Lens of Development Economics," in Philippe Aghion & Steven N. Durlaug (eds), *Handbook of Economic Growth*: Amsterdam: North-Holland, 2005.

Bank of Japan. *Primary Economic Statistics of Japan since the Meiji Era (Meiji-iko Honpo Shuyo Keizai Tokei)*, edited by Bank of Japan, Statistics Department. Tokyo: Nihon Keizai Shumbun, Inc. 2015, htt[://www.nikkei.co.jp/needs/senzen/contents/ accessed July 22, 2015.

Banks, A. S. *Cross-Polity Time-Series Data*. Cambridge, MA: MIT Press, 1971.

Benavot, A., Riddle, P. "The Expansion of Primary Education, 1870-1940: Trends and Issues "*Sociology of Education* 61, 1988: 191-210.

Bethel, K. "Educational Reform in the Bahamas."https://www.educoas.org/Portal/bdigital/contenido/interamer/BkIACD/Interamer/Interamerhtml/Millerhtml/mil_beth.htm accessed May 26, 2016.

Bohlin, J. "Swedish historical national accounts: The fifth generation", *European Review of Economic History* 7, 2003: 73-97.

Bolt, J., van Zanden, J. L. "The First Update of the Maddison Project: Reestimating Growth Before 1820 http://www.ggdc.net/maddison/maddison-project/home.htm, 2013 version accessed May 11, 2014.

Broadberry, S., Campbell, B. M. S., van Leeuwen, B. "The Sectoral Distribution of the Labour Force and Labour Productivity in Britain, 1381-1851,"London School of Economics working paper, 2011.

Broadberry, S., Guan, H., Li, D.D. "China, Europe and the Great Divergence: A Study in Historical National Accounting, 980-1850", London School of Economics working paper, 2014.

Broadberry, S., Custodis, J., Gupta, B. "India and the Great Divergence: An Anglo-Indian Comparison of GDP Per Capita, 1600-1871."*Explorations in Economic History* 55, 2015: 58-75.

Butlin, M., Dixon, R., Lloyd, P.J. "Statistical Appendix: Selected Data Series, 1800-2010"in *The Cambridge Economic History of Australia* eds. Simon Ville and Glenn Withers. Port Melbourne: Australia: Cambridge University Press, 2015.

Butlin, N.G. *Forming A Colonial Economy: Australia 1810-1850* New York: Cambridge University Press, 1994.

Cambridge Economic History of India: Volume 2: c. 1757-c.1970, edited by Dharma Kumar, New York: Cambridge University Press, 1983.

Cameron, R. *A Concise Economic History of the World: From Paleolithic Times to the Present* New York: Oxford University Press, 1993.

Central Intelligence Agency. *The World Factbook 2002*. Washington D.C.: Central Intelligence Agency, 2002.

Central Intelligence Agency. *The World Factbook 2011-2012*. Washington D.C.: Central Intelligence Agency, 2012.

Chan, V. personal internet correspondence for the Commissioner for Census and Statistics, approximate date 2002.

Christensen, R. P. "Population Growth and Agricultural Development," *Agricultural Economics Research* 18, 1966: 119-128.

Clarke, R. *Soviet Economic Facts, 1917-1970* New York: Macmillan, 1972.

Clio Infra. "The World Countries Urbanization Ratio Dataset Comprises Information on Urbanization Ratio's Per Each Country for the Period 1500-2000, "at Clio Infra : Reconstructing Global Inequality, https://www.clio-infra.eu/datasets/searchresults accessed October 1, 2016.

Cohen, J., Federico, G. *The Growth of the Italian Economy, 1820-1960* Cambridge: Cambridge University Press, 2001.

Crafts, N. F. R. "The Human Development Index and Changes in Standards of Living: Some Historical Comparisons "*European Review of Economic History* 1, 1997: 299-322.

Damianova, Z. "Bulgaria,"in *The Formation of Labour Movements 1870-1914: An International Perspective* edited by M. Van Der Linden and J. Rojhan. New York: E.J. Brill, 1990.

De Zwart, P. Van Zanden, P. L. "Labor, Wages, and Living Standards in Java, 1680-1914,"*European Review of Economic History* 19, 2015: 215-234.

Didenko, D., Földvari, P., Van Leeuwen, B. "The Spread of Human Capital in the Former Soviet Union Area in Comparative Perspective: Exploring a New Dataset." *Journal of Eurasian Studies* 4, 2013: 123-135.

Dorling Kindersley World Reference Atlas. New York: Dorling Kindersley Publishing, 1994.

Dower, P. "Re: Labor Force Estimates for Russia in the 19^{th} Century."Message to Robert Tamura. 31 Oct 2016. E-mail.

Easterlin, R. A. "Why Isn't the Whole World Developed? "*Journal of Economic History* 41, 1981: 1-19.

Easterlin, R.A. *Growth Triumphant: The Twenty-first Century in Historical Perspective*, Ann Arbor: University of Michigan Press, 1998.

Easterly, W., Fischer, S. "The Soviet Economic Decline" *The World Bank Economic Review* 9, 1995: 341-371.

Education Department. personal internet correspondence with Education Department Room 1420 Wu Chung House, 213 Queen's Road East, Wan Chai, Hong Kong.

"Education and Training,"Statistics Portal: Grand Duchy of Luxembourg: http://www.statistiques.public.lu/stat/ReportFolders/ReportFolder.aspx?IF_Language=eng&MainTheme=3&FldrName=6 accessed May 23, 2016.

Engerman, S., Sokolof, K. *Economic Development in the Americas since 1500: Endowments and Institutions* New York: Cambridge University Press, 2012.

Europa Publications. *Europa World Yearbook 2003*, New York: Europa Publications, 2003.

Esfahani, H.S., Pesaran, M.H. "Iranian Economy in the Twentieth Century: A Global Perspective."University of Illinois working paper, 2008.

Fajgelbaum, P., Redding, S.J. "External Integration, Structural Transformation and Economic Development: Evidence from Argentina 1870-1914"NBER Working paper 20217, 2014.

Federico, G. "Agriculture and Economic Development in Italy, 1870-1939"in *Agriculture and Economic Development in Europe Since 1870* edited by Pedro Lains and Vincente Pinilla, New York: Routledge, 2009.

Feinstein, C.H. *An Economic History of South Africa: Conquest, Discrimination and Development* New York: Cambridge University Press, 2005.

Francis, J.A. *The Terms of Trade and the Rise of Argentina in the Long Nineteenth Century* Unpublished Ph.D. Dissertation London School of Economics and Politics, 2013.

Federico, G. *Feeding the World: An Economic History of Agriculture, 1800-2000*, Princeton, Princeton University Press, 2010.

Geary, F., Stark, T. "Examining Ireland's Post-Famine Economic Growth Performance. *The Economic Journal* 112, 2002: 919-935.

Geohive. *World 2000 by sector*. http://www.geohive.com/charts/ec_sector.aspx, accessed May 11, 2016.

Gill, H. B. "Literacy in Virgina," *Colonial Williamsburg* Winter 2011.

Gold, C. *Private Girls Schools in Copenhagen 1790-1820* The Royal Library, Copenhagen & Museum Tusculanum Press 1996.

Goldsmith, R.W. *Comparative National Balance Sheets: A Study of Twenty Countries, 1688-1978* Chicago: University of Chicago Press, 1985.

Good, D. F. "Uneven Development in the Nineteenth Century: A Comparison of the Habsburg Empire and the United States, " *Journal of Economic History* 46, 1986: 137-151.

Grigg, D. *Population Growth and Agrarian Change: An historical perspective* New York: Cambridge University Press, 1980.

Grigg, D. "The Industrial Revolution and Land Transformation," *Land Transformation in Agriculture* Edited by M.G. Wolman and F.G.A. Fournier. New York: John Wiley & Sons, 1987.

Grubb, F. W. "Growth of Literacy in Colonial America: Longitudinal Patterns, Economic Models, and the Direction of Future Research " *Social Science History* 14, 1990: 451-482.

Harman, D. "Illiteracy: An Overview" *Harvard Educational Review* 40, 1970: 226-243.

Harris, C. A. "Where did all the women go? Labor Market Changes in a Settler Society: Argentina, 1860-1914," worldhistoryconnected.press.illinois.edu/4.3/harris.html accessed Sep 26, 2016.

Henriksen, I. "The Contribution of Agriculture to Economic Growth in Denmark, 1870-1939" in *Agriculture and Economic Development in Europe Since 1870* edited by Pedro Lains and Vincente Pinilla, New York: Routledge, 2009.

Hooley, R. "American Economic Policy in the Philippines, 1902-1940: Exploring a Dark Age in Colonial Statistics." *Journal of Asian Economics* 16, 2005: 464-488.

indexmundi. "Seychelles - Labor force participation rate, total (% of total population ages 15+) (national estimate)" http://www.indexmundi.com/facts/seychelles/indicator/SL.TLF.CACT.NE.ZS, accessed October 8, 2016.

Issawi, C. "Economic Change and Urbanization in the MIddle East," in *Middle Eastern Cities: Ancient, Islamic, and Contemporary Middle Eastern Urganism: A Symposium* Edited by Ira M. Lapidus, Berkeley, CA: University of California Press, 1969.

Issawi, C. *The Economic History of Turkey: 1800-1914* Chicago: University of Chicago Press, 1980.

Kessler, G., Valetov, T. "Occupation Change and Industrialization: from Russia to the Soviet Union (1897-1959)" working paper presented at European Social Science History Conference, Vienna, Austria 2014.

Keyfitz, N., and Flieger, W. *World Population Growth and Aging: Demographic Trends in the Late Twentieth Century*. Chicago: University of Chicago Press, 1990.

Koralka, J. "The Czech Workers' Movement in the Hapsburg Empire " in *The Formation of Labor Movements 1870-1914* edited by Marcel Van Der Linden and Jurgen Rojahn. Netherlands: E. J. Brill 1990.

Lains, P. "Portugal's Growth Paradox, 1870-1950" Universidade do Porto working paper, 2003.

Lains, P. "Agriculture and Economic Development in Portugal, 1870-1973" in *Agriculture and Economic Development in Europe Since 1870* edited by Pedro Lains and Vincente Pinilla, New York: Routledge, 2009.

Lamounier, M.L. *Between Slavery and Free Labour: Experiments with Free Labour and Patterns of Slave Emancipation in Brazil and Cuba c. 1830-1888* Unpublished London School of Economics and Political Science dissertation, 1993.

Lampe, J.R., Jackson, M.R. *Balkan Economic History, 1550-1950: From Imperial Borderlands to Developing Nations*. Bloomington: Indiana University Press, 1982.

Layne, A. "Secondary Education for the Masses: Critical Issues Then and Now. "St. Leonard's Boys Secondary School's 50^{th} Anniversary Lecture, October 10, 2002.

Lee, J-W., Lee, H. "Human Capital in the Long Run," *Journal of Development Economics* 122, 2016: 147-169.

Lewis, K. "Colonial Education: A History of Education of Belize"University of Illinois Urbana-Champaign working paper, 2000.

Liao, C-h. "Urbanization in Taiwan: 1900-1985,"*Journal of Population Studies* 11, 1988: 129-162.

Lindert, P. H. *Growing Public: Volume 2* New York: Cambridge University Press, 2004.

Llorca-Jaña, M., Navarrete-Montalvo, J. "The Chilean Economy during the 1810-1830s and its Entry into the World Economy,"*Bulletin of Latin American Research* 2016.

Lynch, J. "Every Man Able to Read: Literacy in Early America,"*Colonial Williamsburg* Winter 2011.

McEvedy, C., Jones, R. *Atlas of World Population History* London: Penguin Reference Books, 1978.

Maddison, A. "A Long-Run Perspective on Saving "*Scandinavian Journal of Economics* 94, 1992: 181-196.

Maddison, A. *Monitoring the World Economy: 1820-1992*. Paris: OECD, 1995.

Maddison, A. *Chinese Economic Performance in the Long Run*. Paris: OECD, 1998.

Maddison, A. *The World Economy: A Millennial Perspective*. Paris: OECD, 2001.

Maddison, A. *The World Economy: Historical Statistics*. Paris: OECD, 2003.

Malmberg, B. "Number of Students and Enrollment Rates in Extended Primary, Integrated-Upper Secondary, and Tertiary Education, Sweden 1768-2002."Insitute for Future Studies working paper, 2007.

Melliantsev, V.A. "Three Century's of Russia's Endeavors to Surpass the East and to Catch Up with the West: Trends, Factors, and Consequences "Unpublished Moscow State University working paper, 2002.

Meliantsev, V.A. "Russia's Comparative Economic Development in the Long Run "*Social Evolution & History* 3, 2004: 106-136.

Miller, T., Holmes, K.R., Feulner, E.J. *2012 Index of Economic Freedom*, Washington, D.C.: The Heritage Foundation, 2012.

Mishkova, D. "Modernization and Political Elites in the Balkans, 1870-1914 "University of Sofia 1994 working paper.

Mitch, D. "Underinvestment in Literacy? The Potential Contribution of Government Involvement in Elementary Education to Economic Growth in Nineteenth-Century England,"*The Journal of Economic History* 44, 1984: 557-566.

Mitch, D. "Education and Skill of the British Labour Force,"*The Cambridge Economic History of Modern Britain Volume I*. Cambridge: Cambridge University Press, 2004.

Mitchell, B.R. *International Historical Statistics: Africa, Asia Oceania: 1750-2000*. Fourth Edition, New York: Palgrave Macmillan, 2003.

Mitchell, B.R. *International Historical Statistics: The Americas: 1750-2000*. Fourth Edition. New York: Palgrave Macmillan, 2003.

Mitchell, B.R. *International Historical Statistics: Europe: 1750-2000*. Fourth Edition. New York: Palgrave Macmillan, 2003.

Moll-Murata, C. "Methodological Paper China 1800," working paper, 2011.

Morris, C., and Adelman, I. *Comparative Patterns of Economic Development*. Baltimore: Johns Hopkins University Press, 1988.

Morrisson, C., Murtin, F. "The Century of Education." *Journal of Human Capital* 3, 2009: 1-42.

Nakamura, J.I. *Agriculture Production and the Economic Development of Japan 1873-1922* Princeton, N.J.: Princeton University Press, 2015.

NationMaster. "Taiwan Labor Stats," http://www.nationmaster.com/country-info/profiles/Taiwan/Labor accessed October 4, 2016.

Nehru, V., Dhareshwar, A. "A New Database on Physical Capital Stock: Sources, Methodology and Results. *Revista de Análisís Económico* 8, 1993: 37-59.

Owen, R. *The Middle East in the World Economy: 1800-1914* New York: I.B. Tauris, 2002.

Pamuk, Ş. "Agriculture and Economic Development in Turkey, 1870-2000" in *Agriculture and Economic Development in Europe Since 1870* edited by Pedro Lains and Vincente Pinilla, New York: Routledge, 2009.

Petmezas, S.D. "Agricultural Change and Export Trade in Greece, ca. 1830-1914" University of Crete working paper, 2003.

Petmezas, S.D. "Agriculture and Economic Development in Greece, 1870-1973" in *Agriculture and Economic Development in Europe Since 1870* edited by Pedro Lains and Vincente Pinilla, New York: Routledge, 2009.

Pierenkemper, T., Tilly, R. *The Germany Economy During the Nineteenth Century* New York: Berghahn Books, 2004.

Piketty, T., and Zucman, G. "Capital is Back: Wealth-Income Ratios in Rich Countries 1700-2010" *Quarterly Journal of Economic* 129, 2014: 1255-1310.

Prados de la Escosura, L., Rosés, J.R. "The Sources of Long-Run Growth in Spain, 1850-2000" *The Journal of Economic History* 69, 2009: 1063-1091.

Rashin, A.G. *Naselenie Rossii za 100 let, 1811-1913: Statisticheskjie Ocherki*, (*The Population of Russia Over 100 Years* Moscow: Gos. Statisticheskoe izd-vo, 1956.

Resnick, D.P., Resnick, L.B. "The Nature of Literacy: An Historical Exploration" *Harvard Educational Review* 47, 1977: 370-385.

Richardson, P. *Economic Change in China, c. 1800-1950* Cambridge, UK: Cambridge University Press, 1999.

Roser, M. "Agricultural Employment," *Our World in Data* at https://ourworldindata.org/agricultural-employment accessed September 6, 2016.

Roser, M. "Literacy," *Our World in Data* at http://www.ourworldindata.org/data/education-knowledge/literacy/ accessed October 27, 2014.

Rozman, G. "East Asian Urbanization in the Nineteenth Century: Comparisons with Europe,"in *Urbanization in History: A Process of Dynamic Interactions* edited by A. van der Woude, J. de Vries, A. Hayami, New York: Clarendon Press, 1990.

Sabillon, C. *World Economic Historical Statistics* New York: Algora Publishing 2005.

Saito, O., Takashima, M. Estimating the Shares of Secondary- and Tertiary-Sector Output in the Age of Proto-Industrializsation: The Case of Japan, 1600-1874"Hitotsubashi University working paper 2014.

Saleh, M. "Occupational Structure in Eqypt in 1849-1996"Working paper Toulouse School of Economics, 2015.

Salvatore, R.D., Newland, C. "Between Independence and the Golden Age: The Early Argentine Economy,"in *A New Economic History of Argentina* edited by Gerardo della Paolera and Alan M. Taylor, New York: Cambridge University Press, 2003.

Salvucci, R. "Export-Led Industrialization,"in *The Cambridge Economic History of Latin America: Volume II, The Long Twentieth Century* editors Victor Bulmer-Thomas, John H. Coatsworth, and Roberto Cortés Conde, New York: Cambridge University Press 2006.

Schoellman, T. "The Causes and Consequences of Cross-Country Differences in Schooling Attainment," *Review of Economic Studies* 79, 2012: 388-417.

Schön, L. "Sweden - Economic Growth and Structural Change, 1800-2000"EH.Net Encyclopedia, edited by Robert Whaples. February 10, 2008. http://eh.net/encyclopedia/sweden-economic-growth-and-structural-change-1800-2000/.

Schulze, M-S. "Origins of Catch-Up Failure: Comparative Productivity Growth in the Habsburg Empire, 1870-1910."*European Review of Economic History* 11, 2007: 189-218.

Smits, J-P. Economic Growth and Structural Change in Sub Saharan Africa during the Twentieth Century: New Empirical Evidence"Unpublished Groningen Growth and Development Centre working paper, 2006.

Smits, J-P., Horlings, E., van Zanden, J.L. "Dutch GNP and Its Components, 1800-1913", *Groningen Growth and Development Centre* Monograph Series No. 5, 2000.

Spulber, N. *Russia's Economic Transitions: From Late Tsarism to the New Millennium* New York: Cambridge University Press, 2003.

Statistics Canada. "Series D1-7."at http://www.statcan.gc.ca/pub/11-516-x/sectiond/D1_7-eng.csv accessed July 11, 2016.

Steckel, R., Floud, R. *Health and Welfare During Industrialization*. Chicago: NBER University of Chicago Press, 1997.

Summers, R., Heston, A., Aten, B *Penn World Tables Version 6.3* Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Tafunell, X., Ducoing, C. "Non-Residential Capital Stock in Latin America, 1875-2008: New Estimates and International Comparisons."*Australian Economic History Review*, 56, 2016: 46-69.

Tilly, R.H. "Capital Formation in Germany in the Nineteenth Century,"in *The Cambridge Economic History of Europe Vol VII: The Industrial Economies: Capital, Labour, and Enterprise, Part I Britain, France, Germany and Scandinavia*, editors Peter Mathias and M.M. Postan, New York: Cambridge University Press 1978.

Time Almanac 2001 New York: Time, Incorporated Home Entertainment, 2000.

Turnbull, C.M. *A History of Singapore: 1819 - 1988* second edition New York, Oxford University Press, 1989.

United Nations. *Demographic Yearbook: 2003* New York, United Nations, 2004.

United Nations. *Demographic Yearbook: 2014* New York, United Nations, 2015.

United Nations. *Population Pyramids of the World from 1950 to 2100* https://populationpyramid.net, accessed July 24, 2016.

United Nations. *Statistical Yearbook: 1954* New York, United Nations, 1955.

United Nations. *Statistical Yearbook: 1961* New York, United Nations, 1962.

United Nations. *Statistical Yearbook: 1963* New York, United Nations, 1964.

United Nations. *Statistical Yearbook: 1966* New York, United Nations, 1967.

United Nations. *Growth of the World's Urban and Rural Population, 1920-2000* New York, United Nations, 1969.

United Nations. *Statistical Yearbook: 1973* New York, United Nations, 1974.

United Nations. *Statistical Yearbook: 1981* New York, United Nations, 1982.

United Nations. *Statistical Yearbook: 1996* New York, United Nations, 1997.

United Nations Development Programme. *Human Development Report 2010* New York, Palgrave Macmillan, 2010.

UNESCO. *Progress of Literacy in Various Countries: A Preliminary Statistical Study of Available Census Data Since 1900*. Paris, 1953.

UNESCO. *UIS Global Database* http://data.uis.unesco.org accessed May 3, 2016.

UNESCO. *World Illiteracy at Mid-Century: A Statistical Study* Westport, CT, Greenwood Press, 1957.

UNICEF. *Education Statistics: Eritrea* Division of Policy and Practice, Statistics and Monitoring Section, *www.childinfo.org* May 2008, accessed May 3, 2016.

U.S. Census Bureau. *The Statistical Abstract of the United States: 2000*. Washington, D.C., 2000.

U.S. Census Bureau. *The Statistical Abstract of the United States: 2012*. Washington, D.C., 2012.

U.S. Department of Commerce. *Historical Statistics of the United States: Millennial Edition*. Washington, D.C., PALGRAVE MACMILLAN, 2003.

van der Eng, P. "The Sources of Long-Term Economic Growth in Indonesia, 1880-2008." *Explorations in Economic History* 47, 2010: 294-309.

van Leeuwen, B., Földvári, P. "Capital Accumulation and Growth in Central Europe, 1920-2006", *Eastern European Economics* 51, 2013: 69-93.

van Waijenburg, M.F.M. *Living Standards in British Africa in a Comparative Perspective, 1880-1945: Is Poverty Destiny?* unpublished MA Thesis Utrecht University, 2010.

van Zanden, J. L. "Economic Growth in Java 1815-1939: The Reconstruction of the Historical National Accounts of a Colonial Economy"unpublished Maddison-Project Working Paper WP-3, 2012.

Warr, P. G. *Thailand's Macroeconomic Miracle: Stable Adjustment and Sustained Growth*, New York: World Bank, 1996.

Wikipedia. "Demographic History of Poland,"https://en.wikipedia.org/wiki/Demographic_history_of_Poland accessed July 24, 2016.

Wikipedia. "Demographics of Romania,"https://en.wikipedia.org/wiki/Demographics_of_Romania#Between_WWI_and_WWII accessed July 24, 2016.

Wikipedia. "List of countries by population in 2010,"http://en.wikipedia.org/wiki/List_of_countries_by_population_in_2010 accessed April 16, 2014.

Williamson, J.G. "Land, Labor and Globalization in the Pre-Industrial Third World,"NBER working paper 7784, 2000.

Wolf, N. "Local Comparative Advantage: Agriculture and Economic Development in Poland, 1870-1970,"Freie Universitat working paper, 2006.

World Bank. *2005 International Comparison Program: Tables of final results*, February 2008.

World Bank. *World Development Indicators* various issues. Washington D.C.: World Bank.

World Bank. *World Development Report 1982*. Washington D.C.: World Bank, 1981.

World Bank. *World Development Report, 2008*. Washington D.C.: World Bank, 2007.

World Bank. *World Development Report 2010*. Washington D.C.: World Bank, 2009.

World Bank. *World Development Report 2011*. Washington D.C.: World Bank, 2010.

World Bank. *World Development Report, 2012*. Washington D.C.: World Bank, 2011.

World Bank. *World Tables of Economic and Social Indicators, 1950-1981*. Ann Arbor: ICPSR, 1984.

WorldOMeters. "China Population,"http://www.worldometers.info/world-population/china-population/, accessed October 4, 2016.

WorldOMeters. "Taiwan Population,"http://www.worldometers.info/world-population/taiwan-population/, accessed October 4, 2016.

Wrigley, A. E. "Urban Growth and Agricultural Change: England and the Continent in the Early Modern Period."In Robert I. Rotberg and Theodore K. Rabb eds. (guest edited by Roger S. Schofield and E. Anthony Wrigley), *Population and Economy* Cambridge: Cambridge University Press.

Zhihong, S., Yuping, N., van Leeuwen, B. "Chinese National Income, ca. 1661-1933."Utrecht Universiity working paper, 2015.