Center for Economic Institutions
Working Paper Series

No. 2020-3

"Structural Changes in World Population and Economy, 1950 to 2100: Visualization and Analyses"

Yasushi Nakamura

August, 2020
(Revised November 16, 2020)

Center for Economic Institutions

Working Paper Series

Institute of Economic Research
Hitotsubashi University
2-1 Naka, Kunitachi, Tokyo, 186-8603 JAPAN
http://cei.ier.hit-u.ac.jp/English/index.html
Tel:+81-42-580-8405/Fax:+81-42-580-8333

Visualisation and Analysis of Structural Changes in the World's Population and Economy, 1950 to 2100

October 2020

Yasushi NAKAMURA*
Yokohama National University
Graduate School of International Social Sciences

Abstract
This study promotes the use of the economic centroid method, which calculates the mean of the coordinates of the regions in an area weighted by a socioeconomic variable. It applies this method to world population and economic data from 1950 to 2100. Supported by the deviations from proportional growth approach and a newly developed Quantum GIS Python tool, it successfully visualises structural changes in the world population and economy. The economic centroid method needs further development to add the ability to minimise the influence of the geographical location of a region on the centroid's movement.

Running title: Changes in the World's Population and Economy
Keywords: World population, Productivity, Economic centroid, QGIS-Python
JEL: I20, O11, O44, Q56

---

* nakamura-yasushi-vr@ynu.ac.jp
ORCiD: 0000-0002-0026-8715

1. Introduction

Data visualisation is an important scientific tool. This study seeks to develop a tool that visualises the spatial and temporal changes in a socioeconomic variable of multiple regions, that is, in spatial panel data. It is often difficult to identify the spatial and temporal changes in time series data from multiple regions, even if the data do not qualify as 'big data'. This seems to be because the structure of spatial panel data inevitably includes more than four dimensions: the x and y spatial axes, the temporal axis, and at least one axis for the socioeconomic variable. Visualisation itself may do little to deepen the analysis; however, it undoubtedly contributes to unearthing latent structures and changes from spatial panel data.

One method for visualising the spatial and temporal changes in a socioeconomic variable in multiple regions is to calculate the mean of the coordinates of the geographical points that represent the regions, weighted by the value of the relevant variable. For simplicity, this paper refers to this method as the *economic centroid* method. National statistical or census offices, including those in the United States (Plane and Rogerson, 2015; USCB, 2020), Japan (MIC, 2017), and Australia (ABS, 2020), calculate one type of economic centroid known as the centre of the population. However, they rarely calculate other types of economic centroids such as productivity centroids, centroids for gross regional product and regional sectoral production, and education centroids, even though this is now easy using geographical information system (GIS) software. This is probably because the economic centroid method is not constructive for analysing the underlying socioeconomic variable in detail, as discussed in Section 2. However, this method is a powerful tool for visualising the spatial and temporal changes in spatial panel data. This study reassesses the economic centroid method, develops a supporting tool that employs Python and GIS software, namely, Quantum GIS (QGIS), and presents the results of applying this method to the world population and economic data from 1950 to 2100.

The remainder of this paper is structured as follows. Section 2 reviews the economic centroid methodology and proposes a Deviations from Proportional Growth (DPG) approach to identify structural changes in the variable under investigation. This section also briefly explains the newly developed QGIS Python tool for supporting economic centroid analysis. Section 3 presents the data on population, production, productivity, and education by country for 1950 to 2100, which are then applied to illustrate the application of the economic centroid method. Section 4 describes the past and future structural changes to the world in which the economic centroid method visualises. Finally, Section 5 points out the shortcomings of the economic centroid method and advocates reworking it to visualise changes in spatial panel data after controlling for the effects of the geographical locations of regions.

2. Economic Centroid Methodology

2.1 Basic concept of an economic centroid

We define an economic centroid as the mean of the coordinates of the geographical points ('point features' in GIS terminology) in an area, weighted by the values of a socioeconomic variable at those points:

$$\begin{pmatrix} c_{m,x} \\ c_{m,y} \end{pmatrix} = \begin{pmatrix} x_1, x_2, \ldots, x_n \\ y_1, y_2, \ldots, y_n \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}. \tag{1}$$

$$\mathbf{C_m} = \mathbf{C} \cdot \mathbf{W}$$

Here, we assume that there are n geographical points. $\mathbf{C_m}$ is the coordinate vector of the economic centroid for the n geographical points and $\mathbf{C}$ is the $2 \times n$ matrix of the coordinates of the n geographical points. $\mathbf{W}$ is the vector of the weights; that is, the ratio of the value of the socioeconomic variable at the given geographical point to the total value for that variable across the area. The coordinates can be three-dimensional and include the longitude, latitude, and elevation of a geographical point. However, this study uses two-dimensional coordinates because the world economic centroid examples do not include information on elevation.

A geographical point may represent a sub-region of the area. For the world economic centroid examples in this study, the area is the entire world and the sub-regions are countries. This study defines the geographical point representing a country as its area centroid, which is calculated using the world vector map provided by GADM (2020) and QGIS's *centroid* function. There are other ways to represent a country with a geographical point such as using the coordinates of its capital city. The area centroid is used because it is important to represent countries using geographical points whose coordinates do not change during the observation period to identify structural changes in the world's population and economy.

A variety of socioeconomic variables can be used as weights, w_i. Taking a population centroid as an example, the weight, w_i, is defined as

$$w_i = \frac{p_i}{\sum_i^n p_i}. \qquad (2)$$
$$\sum_i^n w_i = 1$$

Here, p_i is the population of sub-region i and n is the total number of sub-regions. The world economic centroid example in this study also includes normalised variables such as real gross domestic product (RGDP) per capita (RGDP/P) and the mean years of education. For normalised variables, the weights are interpreted differently. Taking RGDP/P as an example, w_i, the weight of sub-region i, is

$$w_i = \frac{g_i}{\sum_1^n g_i} = \frac{\frac{1}{n} g_i}{\frac{1}{n}\sum_1^n g_i} = \frac{1}{n} \cdot \frac{g_i}{g^*}. \qquad (3)$$

Here, g_i is the RGDP/P in sub-region i and n is the total number of sub-regions. In this case, the weight, w_i, is the normalised ratio of the RGDP/P of that region to the mean RGDP/P of the world, g_*. Normalisation with $1/n$ is necessary to ensure the sum of all the w_is equals one, and this does not change the position of sub-region i in the RGDP/P ranking.

2.2 Visualisation and analysis

The economic centroid method is a powerful visualisation tool; it reveals the spatial and temporal changes in the relevant socioeconomic variables in multiple regions on a map. However, this method also has drawbacks. First, economic centroids are not rich in information on the underlying socioeconomic variables. As Equation (1) indicates, the calculation of the economic centroid employs the values of the underlying socioeconomic variable as the normalised weights. That is, the economic centroid method does not fully use the information on the socioeconomic variable. Consider the population centroid of an area with, from west to east, three sub-regions labelled A, B, and C. Assume that the centroid is at B and moves eastwards. For simplicity, we also assume that sub-region B has no influence on the movement of the centroid. An eastward movement of the population centroid means that the weight of sub-region C, w_C, increases in relation to the weight of sub-region A, w_A. This relative change can occur in three classes of cases: when w_A increases and w_C decreases, when the increase in w_C is larger than the increase in w_A, and when w_C is smaller than the decrease in w_A. Moreover, these three classes can occur during any of the three possible cases of change to the total population: unchanged, increased, or decreased. It is impossible to comprehend, from the movement of an economic centroid on a map, how the actual populations in the sub-regions change.

Second, as shown in Equation (1), it is not the weights but the moments, the weight times its distance from the moment's centre, which define the location of the centroid. Moreover, the moment centre can be arbitrarily set anywhere. Equation (4) provides the decomposition of a centroid movement using the symbols from Equation (1) and showing only the x-axis:

$$\begin{aligned} c_{m,x}^{t+1} - c_{m,x}^{t} &= \sum_i^n \left[w_i^{t+1} \left(M_x^t + x_i - M_x^t \right) \right] - \sum_i^n \left[w_i^t \left(M_x^t + x_i - M_x^t \right) \right] \\ &= \sum_i^n \left[\left(w_i^{t+1} - w_i^t \right) \left(x_i - M_x^t \right) \right] \end{aligned} . \tag{4}$$

Here, M_x^t is the x-coordinate of the moment centre at time t. We use the fact that sub-region i's distance from the origin, x_i, does not change and that the sum of the weights is always one to rewrite the equation. Equation (4) indicates that the sign and magnitude of the moment of an individual sub-region change depending on the locations of the sub-regions relative to the moment centre, M_x^t. Therefore, depending on their distances from the moment centre, identical changes to the weights in the two sub-regions can influence the location of the economic centroid differently. Equation (4) also indicates that M_x^t can be any real constant number without changing the sum of the moments at time t and, thus, the location of the centroid at t. However, the choice of the placement of the moment centre changes the moment of an individual sub-region and thus the contribution of an individual sub-region to the movement of the economic centroid.

These issues seem to explain the unpopularity of the economic centroid method. The following subsection discusses an approach that makes this method more useful and turns it into a tool for exploring the causes of economic centroid movements.

2.3. Economic centroids and DPG approach

Although the economic centroid method contributes little to analysing its underlying socioeconomic variable, it is useful for identifying structural changes because it summarises and visualises changes in the variable. We can use the economic centroid method in a two-step procedure. First, the method visually identifies the time and place where structural changes occur. Second, it focuses on the changes to the socioeconomic variable in more detail, concentrating on the time and place of the structural change. To perform this procedure, it seems useful to adopt the DPG approach.

Before discussing the practical aspects of the DPG approach, we examine a case in which the economic centroid moves linearly. It is complicated to define the linear movement of an economic centroid on a map because the surface it moves on is roughly spherical. The trajectory of economic centroids varies depending on the map projection method one chooses. In this study, we consider an economic centroid to have moved linearly if it moves linearly on a map in which the latitude (x) and longitude (y) axes are orthogonal and have the same scale; all the maps in this study therefore employ EPSGE:3857-WGS84/Pseudo-Mercator as the coordinate reference system. Nevertheless, as seen later in this subsection, it is not essential if the centroid appears to move linearly on a map.

Many economic centroids tend to move monotonously and linearly. For example, the population centroid of the United States moved west almost monotonously and linearly from Maryland in 1790 to Missouri in 2010 (USCB, 2020). The same centroid in the United Kingdom moved roughly monotonously and linearly south from Rodsley in 1901 to Appleby Parva in 2000 (Appleby Magna, 2002). The one in Japan moved almost monotonously and linearly east–southeast in the Gifu prefecture from 1920 to 2017, although it is difficult to say if a structural change existed before and after World War II because of a lack of data during that war (Nakamura, Zhang, and Arikawa, 2019).

From Equation (1) and the general definition of a straight line using vector notation, Equation (5) must hold at any time t for a linearly moving economic centroid:

$$\mathbf{C}_m^1 - \mathbf{C}_m^0 = \theta^t \left(\mathbf{C}_m^{t+1} - \mathbf{C}_m^t \right) \tag{5}$$

Here, θ^t is any real constant. Rewriting Equation (5) using the symbols from Equation (1), we obtain

$$\begin{cases} \sum_i^n \left[\left(\Delta w_i^1 - \theta^t \Delta w_i^{t+1} \right) x_i \right] = 0 \\ \sum_i^n \left[\left(\Delta w_i^1 - \theta^t \Delta w_i^{t+1} \right) y_i \right] = 0 \end{cases}. \tag{6}$$

Here, we define Δw_i^{t+1} as $\Delta w_i^{t+1} = w_i^{t+1} - w_i^t$. Equation (6) indicates that $\Delta w_i^1 = \theta^t \Delta w_i^{t+1}$ must hold for any i and t because x_i and y_i, the coordinates of sub-region i, are constant. It is mathematically possible for Equation (6) to hold without satisfying $\Delta w_i^1 = \theta^t \Delta w_i^{t+1}$ only if any changes in the sub-regions cancel each other out. However, in socioeconomic terms, it is nonsense to assume that this coincidence would repeat. Moreover, the same geographical points can be expressed with the different values of the coordinates depending on the choice of the measuring units, such as the sexagesimal degree, decimal degree, yard, or metre. No case other than $\Delta w_i^1 = \theta^t \Delta w_i^{t+1}$ can therefore satisfy Equation (6) for all i and t. Equation (6) also points out that the essential part of the linear movement of a centroid is not the linear appearance of the centroid's trajectory on a map, but the condition $\Delta w_i^1 = \theta^t \Delta w_i^{t+1}$.

The condition $\Delta w_i^1 = \theta^t \Delta w_i^{t+1}$ indicates that the pattern of changing weights is persistent. θ^t defines the length of the span that the centroid moves in one period. If θ^t is negative, the centroid moves in the opposite direction on the same line. One special instance is when $\Delta w_i^1 = \Delta w_i^{t+1} = 0$ for all i and t, in which case the centroid stays at its initial location. Another special case is when $\theta^t = 1$, where the centroid moves linearly while maintaining a constant span. Because the sum of every Δw_i^t is zero at any time t, if an economic centroid moves linearly, there are always sub-regions that decrease their shares and sub-regions that increase their shares. Moreover, a sub-region whose share decreased in the first period will continue to decrease its share, while a sub-region with an increased share in the first period will continue to increase its share. This is an expected result, as it simply states that some socioeconomic phenomena have stable tendencies. However, it is surprising that the pattern of population change has changed little in the United States, the United Kingdom, and Japan over the last century or two, as they have experienced massive socioeconomic changes and periods of rapid increases and decreases in their populations. In terms of economics, we need to examine whether we can interpret this persistent pattern of change as an ongoing equilibrium process.

Next, we move onto the DPG approach. The essence of this approach is to set an expected linear movement for an economic centroid as the baseline and then measure the deviations of the centroid's actual locations from that baseline. Figure 1 summarises this idea. The black points from *a0* to *am* and the white points from *e0* to *em* indicate the actual and expected locations of the centroid at times from *0* to *m*, respectively. The fat black vectors from each black point to the next indicate the actual one-period movements of the centroid. The thin black dotted-line vector from *a0* to *ak* indicates the cumulative movement over the period from *0* to *k*; for simplicity, the figure does not illustrate the other movements. Figure 1 sets the baseline as the line from *a0* to *am*, the line connecting the actual locations of the centroid at times *0* and *m*. Under this assumption, the fat dotted-line vectors from each white point to the next indicate the expected one-period movement of the centroid. The centroid's expected one-period movement corresponds to the mean change in the weights from time *0* to *m*; consequently, θ^t is 1 for all *t*. It is not difficult to set θ^t to a variable and estimate it under certain conditions such as when minimising the sum of one-period deviations; however, it has little meaning. As discussed in Section 2.2, it is more productive to model the changes in the weights, that is, the changes in the underlying socioeconomic variable, not θ^t. Modelling the moving distance of an economic centroid may be meaningful only in such cases where we know the driver of the centroid movement as that mining booms shifted the income centroid in Canada (Breau et al., 2018).

In Figure 1, the vector from *aj* to *j** provides an example of a one-period deviation from the expected change, while the vector from *ek* to *ak* is an example of the cumulative deviation from the expected change over the period from *0* to *k*. Note that *j** is only plotted for the convenience of displaying the deviation vector. The position of *ek* is defined as three times the vector from *e0* (*a0*) to *e*1 plus the vector from the origin to *e0* (*a0*). The baseline can change freely to suit the purpose of the analysis. The thin dotted line with crossing marks is another baseline example: its expected one-period change is set equal to the actual movement from *ak* to *al*.

The DPG approach is simple, but the analysis procedure may require many resources. A map must display the trajectories of both the actual and the expected centroids. The one-time and cumulative deviations, in terms of both weights and moments, may need to be displayed in graphs under various baseline settings. It may be desirable to calculate at least two types of moments; those moment centres are the origin of the coordinate system and the centroid from one period earlier. It is also useful to have graphs in which the sub-regions are sorted by various ways: by their locations on the *x*-axis and the *y*-axis, by their weight sizes, and by their moment sizes. To automatise this economic centroid analysis, the Economic Centroid Analysis Tool (ECAT) has been developed in

QGIS Python. All the results presented in Section 4 used the ECAT. Supplementary material A provides the script file and additional information on the ECAT.

Figure 1. Deviations from proportional growth

3. Data

The world economic centroid analysis example uses data on population (POP), RGDP, RGDP/P, RGDP per employee (RGDP/E), and the mean years of education for the female population aged 30 to 34 (FMYE) for 1950 to 2100. We use projection data for 2020 to 2100, which IIASA (2018, 2020) adopts to study the impacts of global climate change because they provide a good example of visualisations and analyses that employ the economic centroid method. The series encompasses over 98% of the world's population, RGDP, and employment; the numbers of countries and periods covered, however, vary by series.

3.1 Population and employment

The population data covering 201 countries from 1950 to 2020 are estimates from the United Nations Department of Economic and Social Affairs (DESA, 2019) and those from 2020 to 2100 are the medium fertility variants of the DESA's population prospects. The Penn World Table version 9.1 (Feenstra, Inklaar and Timmer, 2015; PWT, 2020) is the principal basis for the employment data from 1960 to 2017; the PWT employment data cover 155 countries after 1990, excluding Anguilla after 2002, Bermuda after 2007, and Dominica after 2001. We calculate the missing figures for those three countries using the most recently available DESA (2019) population data and the employment to total population ratio. We also apply this simple estimation method to 36 countries whose employment data are unavailable for some years between 1960 and 1989. These 36 countries are mostly small; the sum of estimated employment in all of them was 21 million in 1960, accounting for only 1.7% of the estimated employees in all 155 countries. One exception is Nepal, which had an estimated employment of 4.6 million in 1960. Supplementary material B supplies the employment data for the former Soviet Union and Eastern European countries before 1990.

3.2 RGDP

PWT (2020) is the primary basis for the RGDP data of the 155 countries from 1960 to 2017. Supplementary material B provides the RGDP data for the former Soviet Union and Eastern European countries before 1990. The RGDP projections for 184 countries from 2010 to 2100 are the simulation results from the ENV-Growth model of the Organisation for Economic Co-operation and Development (OECD) (IIASA, 2020; OECD, 2020). IIASA (2020) also includes other RGDP projections, but we employ the OECD ENV-Growth model projections because they include actual RGDP figures from 2010 to 2015, which enables us to compare them with the PWT-based RGDP series from 1960 to 2017. As Figure 6 demonstrates, this makes little difference to that of the two series, which is used to calculate the locations of the RGDP centroids from 2010 to 2015.

We use three projected RGDP series under the three shared socioeconomic pathway (SSP) scenarios: SSP1, SSP2, and SSP3. SSPs are narrative frameworks that set possible future situations of worldwide coordination and governance; they are closely related to designing and implementing policies for global climate change (Riahi et al., 2017). There are five SSPs. SSP1 represents the sustainable pathway, which envisions successful worldwide governance over global climate change problems and structural changes towards decarbonisation. SSP3 is the inequality pathway, which roughly represents the pathway the world has been following so far. In SSP3, rich countries may achieve some structural changes that lead to decarbonisation, while developing countries will continue to rely on traditional energy sources. SSP2 is the middle pathway, representing a future that sits between the SSP1 and SSP3 worlds. SSP4 is the regional rivalry pathway, where worldwide governance on global climate change problems is poor and each country pursues its own profits. SSP5 oriented towards rapid economic growth, relies on traditional energy sources, and proposes solving global climate change problems only through the development of new technologies. This paper's Appendix summarises the features of the five SSPs that relate to economic growth. Note that the RGDP projections do not consider the consequences of global climate change (Dellink et al., 2017). For example, the largest projected RGDP for 2100 is in SSP5; however, this is an unacceptable situation since, by that time, the costs of mitigation and adapting to global climate change problems would have become too large for most people. We do not use the RGDP series projected for SSP4 and SSP5, which are rather extreme scenarios.

3.3 Education

Education centroids are shown because IIASA (2018), Lutz (2019), Lutz, Cuaresma, and Sanderson (2008), and Lutz and Samir (2011) assert that the spread of female education is the key factor to sustainable growth. The data on FMYE for the 185 countries from 1950 to 2100 come from the Wittgenstein Centre for Demography and Global Human Capital education database (WCD, 2020). This variable may represent the spread of female education and its possible effect on birth and productivity rates. However, it is a reasonable choice for three reasons: the mean year of education seems to be close to saturation point in the age group, the time lag between an improvement in female education and a change in the mean year of education must be relatively short for the age group, and the female population in the age group seems neither too young nor too old to affect the country's birth and productivity rates.

3.4 Data adjustments for calculating the baselines

In some cases, to prevent the expected shares from taking negative values, we adjust the expected shares (weights) in an ad-hoc manner. If we assume linear movements for a centroid and thus constant changes in shares, the expected shares of some small countries would be negative in future periods. In those cases, we change the negative expected shares to zero and simply rescale other countries' shares to keep the sum of the shares equal to one. The ECAT performs this adjustment by default. These adjustments have only a negligible influence because they only exclude countries with small shares; nevertheless, the adjustments may make the expected centroid trajectories non-linear, especially near their endpoints.

4. Results

4.1. Population, productivity, and education from 1950 to 2017

Figure 2 illustrates the movements of the centroids for POP, RGDP, RGDP/E, and FMYE before 2017. POP also displays the population centroids for 2020 to 2100. The eight marker sizes reflect the value indices of the underlying variables. The base of each index is set to the corresponding variable value in 1960 as 100. The smallest marker size includes all index values up to 100 and the largest includes all index values over 800. The employment centroid is presented for reference, with the fixed black dot markers.

Figure 2. Population, productivity, and education

4.1.1 Population

Figure 2 reveals that the world population centroid moved relatively little compared with the other economic centroids even though the world's population increased threefold, from 2.5 billion in 1950 to 7.5 billion in 2017. Figure 2 also indicates that the trajectories of the population and employment centroids changed direction notably, while the other centroids did not change direction as much. A close look at POP reveals that the population centroid shifted almost monotonously from north to south from 1950 to 2017 on the south–north (y) axis, and it continued to move south on the south–north (y) axis until 2100. By contrast, it changed its movement from east to west on the west–east (x) axis around 1990. After this change, the population centroid moved west until 2015, and it was projected to continue to move west until 2100.

The black plus symbols in Figure 2 indicate the expected population centroids from 1950 to 2100 based on the mean expected changes in population shares from 1970 to 1990 (for simplicity, we hereafter use *the 1970–1990 population baseline* to refer to the mean changes in expected shares in the given period). The expected locations of the population centroid between the DESA projection and 1970–1990 population baseline projection did not differ greatly along the south–north axis, but were different along the west–east axis.

What caused this change in the population centroid's movement around 1990? Figure 3 compares the 1970–1990 (black bars) and 1995–02015 (white bars) population baselines. The black bars in Figure 3 indicate that India's increase and the United States' decrease were the most significant factors in driving the population centroid eastwards before 1990. The increase in Indonesia's and Pakistan's population shares contributed to shifting the population centroid east, while the decreases in Russia's, Japan's, and China's population shares contributed to pushing it west. The declining population shares of Germany, the United Kingdom, Italy, and France to the west of the population centroid pushed it east. In terms of the moments, the United States contributed by far the most to shifting the population centroid east (see Supplementary material C for the moment and other graphs). China's population share in 1990 was smaller than that in 1970, although it recorded its largest population share around 1975. China was an important factor in determining the location of the population centroid; however, it was a relatively weak factor in moving the population centroid west and not east during 1970–1990 because China, located in the east of the population centroid, decreased its moment during the period.

The white bars in Figure 3 reveal that from 1995 to 2015, the decrease in China's population share became notably larger, while that in the United States' population share fell substantially. China and the United States were the two primary factors in shifting the movement of the population centroid west. The influences of China and the United States were more considerable in terms of the moments because both were far from the population centroids during that period. The 1995–2015 population baseline indicated that the smaller decreases in Germany, the United Kingdom, France, and Italy as well as the smaller increases in India and Indonesia also contributed to shifting the population centroid west; their influences were, however, far smaller than those of China and the United States. After 1995, the increases in the population shares of African countries such as Nigeria, the Democratic Republic of Congo, and Ethiopia became visible and these pulled the population centroid west. However, their influence was not yet remarkable. In short, the smaller increases, or even decreases, in the population shares of China and other East Asian countries shifted the population centroid trajectory from east to west, assisted by the smaller decreases in the population shares of the United States and major western European countries.

Looking to the future, the pattern of population change represented by the 1995–2015 population baseline should hold until 2100; the 1995–2015 population baseline extended out to 2100 does not differ greatly from the population trajectory from 2020 to 2100. The westward movement of the population centroid will accelerate in the mid-21st century: the primary cause will be accelerated decreases in China's population share. Further, after around 2060, the population shares of most East Asian countries, including India, will decrease and the decrease in the United States' population share will slow, and even turn into an increase. In addition, the increasing population share of African countries will become more meaningful. This will accelerate the eastward movement in the population centroid in the middle of the 21st century. It is interesting that the population centroid appears to be travelling towards the mean coordinates of the 201 sample countries, which is in eastern Niger (the red cross in Figure 2).

[Figure 3. Changes in the population trend (shares, in point)]()

In contrast to its movements along the west–east axis, the population centroid shifted almost monotonously southwards along the south–north axis during 1950–2017 and it will continue to move south until 2100. Figure 3 indicates that the population change patterns do not differ meaningfully if we sort the countries from south to north; instead, most countries to the north of the population centroid decreased their relative shares, while most to the south increased them. Comparing the 1970–1990 and 1995–2015 population baselines, we expect the United States, Germany, the United Kingdom, France, and Italy to slow their decreases and contribute to drawing the population centroid north again after 1995. Brazil and Indonesia's shifts to decreasing population shares are also important factors in the northern shift of the population centroid, as they are south of this centroid. However, China's accelerated decrease in its population share, in addition to African countries' increases, compensates for the forces drawing the population centroid north. India's smaller increase in its population share is the most important driver after China and the United States. However, describing India's influence on the population centroid's movements over the entire period is somewhat complicated because the population centroid is expected to move south across the latitude of India's area centroid around 2020. India's moments from the population centroids are small, but its sign will change from positive (moving the centroid north) to negative (moving the centroid south). China's decreasing population share dwarfed the influences of other countries on the population centroid's movement after 1995.

4.1.2 RGDP and productivity

Unlike the population centroids, the RGDP and RGDP/E centroids did not shift drastically from 1960 to 2017; in general, the RGDP centroid moved east and the RGDP/E centroid moved east-northeast. The trajectory of the RGDP centroid underwent some disturbances from 1990 to 2000; it moved south from 1990 to 1995 and then east from

1995 to 2000. Figure 4 compares the 1970–1990 and 1995–2015 RGDP baselines. The 1970–1990 RGDP baseline reveals that China recorded the largest increase, followed by Japan, Indonesia, Brazil, Korea, and India, whereas the United States recorded the largest decrease, followed by Germany. These countries excluding Brazil contributed to shifting the RGDP centroid east. The countries' moments from the RGDP centroids were also large, excluding Germany. Figure 5 indicates the differences between the expected and actual changes in the RGDP shares: the expected shares are defined by the 1970–90 RGDP baseline. Figure 5 reveals that Russia caused the disturbances primarily from 1990 to 1995. The collapse of the Russian economy after the fall of the Soviet Union pushed the RGDP centroid west and south. The United States' unexpected increase in the GDP share in the period of 1995–2000 probably was a cause of the westward movement of the centroid in the period. When the Russian economy started to recover in the late 1990s and the United States returned to its decreasing trend, the RGDP centroid shifted east again. Japan also recorded the unexpectedly large decreases; it was, however, not the cause of the disturbances from 1990 to 2000 because Japan constantly recorded its unexpected decreases over the period of 1990 to 2017.

After 1995, the eastward movement of the RGDP centroid appears to be accelerated: the 1995–2015 RGDP baseline in Figure 4 indicates that China's increase in its RGDP share, and the United States' decrease were the two primary factors in shifting the RGDP centroid east after 1995. Both countries also had large RGDP moments. Following China, India became the notable factor to move the RGDP centroid east after 1995.

Figure 4. Changes in the RGDP trend (shares, in point)

Figure 5. The actual and expected changes in the RGDP shares

Figure 2 also presents the results of an approximate RGDP decomposition. The weight of the E60*(G/E) centroid in each year is the country's hypothetical RGDP calculated from its employment and RGDP/E in 1960, while the weight of the E*(G/E60) centroid in each year is the country's hypothetical RGDP calculated from the country's RGDP/E and employment in 1960. The trajectories of the E60*(G/E) and E*(G/E60) centroids intersect almost orthogonally. This suggests that the increases in RGDP/E—that is, productivity—in eastern countries was almost exclusively responsible for the eastward shift in the RGDP centroids, while the increases in employment were almost exclusively responsible for the southward shift in the RGDP centroids. The movement of the RGDP/E centroid matched the movements of the E*(G/E60) and E60*(G/E) centroids. The RGDP/E centroid shifted almost monotonously northwest and intersected almost orthogonally with that of the population centroids before 1990. This implies that an increasing number of people worked with relatively low productivity and living standards. After 1990, the RGDP/E centroid moved in almost the opposite direction from the population and employment centroids; the gap between the centroids of population and RGDP/E narrowed on the west–east axis, while the gap between them widened on the south–north axis after 1990.

In summary, the RGDP centroid moved east not only due to growth in East Asian countries, but also because the United States' share of the world's RGDP decreased; the economic growth in East Asia may have accompanied increases in productivity and living standards. This transpired because the RGDP shares of East Asian economies have generally increased and are still increasing, while their populations are growing more slowly. A similar situation does not seem to have occurred in the southern part of the south–north axis.

4.1.3 Education

The education centroid (FMYE) moved east/southeast, roughly parallel to the RGDP centroid, until the mid-1970s when it changed direction slightly to the south. The eastward movement of the FMYE centroid may have reflected an improvement in female education in East Asia, accompanied by increasing RGDP shares. The FMYE centroid appeared to shift in parallel with the productivity centroid (RGDP/E) along the west–east axis. On the south–north axis, it is encouraging that the FMYE centroid veered to the south and moved towards the simple mean coordinates for all 184 sample countries (blue cross in Figure 2). The south–north gap between the FMYE and RGDP/E centroids, however, widened over 1960–2015. The southward drift of the FMYE centroid after the mid-1970s did not appear to improve productivity in the countries south of the FMYE and RGDP/E centroids.

4.2. Structural changes for sustainable future growth

Figure 6 displays the population, RGDP, RGDP/P, and FMYE centroids projected under the three SSP scenarios. The circles (hexagons for FMYE) with dots in their centres represent the SSP1 trajectories, circles (hexagons for FMYE) with crosses in their centres represent the SSP2 trajectories, and circles (hexagons for FMYE) with triangles in their centres represent the SSP3 trajectories. The circles representing the population centroids have solid lines, those of the RGDP centroids dashed lines, and those of the RGDP/P centroids dotted lines. The eight marker sizes reflect the value indices of the underlying variables. The base of the indices is set to the value of the corresponding variable in 2010 as 100. The smallest marker size included all index values up to 100 and the largest included all index values over 800 for total world RGDP and mean world RGDP/P as well as all the index values over 170 for the world population and the simple mean of the countries' FMYE. Figure 6 also illustrates the population centroids for 1950–2020 as black crosses, the RGDP/E centroids for 1960–2017 as small white triangles, and the RGDP/P centroids for 1960–2017 as black triangles as well as the FMYE centroids for 1950–2015 as black points for reference. These are the same centroids used in Figure 2. The RGDP/P centroids from 1960 to 2017 represent 154 countries, while the RGDP/P centroids for 2010–2100 represent 184 countries (see Section 3). The black stars represent the 2010–2020 RGDP/P baseline for the SSP1 scenario; there are few differences between the 2010 and 2020 RGDP/P baselines of the three SSP scenarios.

Figure 6. Population, productivity, and education towards 2100

The SSP projections assume that advancements in female education curb population growth, increase productivity, and improve the awareness of sustainable growth; consequently, they increase RGDP/P. The FMYE value was already close to its peak in advanced countries, including in East Asia, by 2010. Therefore, the FMYE centroid will primarily move south. The symbol sizes and locations of the FMYE centroids clearly represent the differences between the three SSP scenarios. The simple mean of FMYE for the 184 countries in SSP1 and SSP2 will be larger, by 67% and 55%, respectively in 2100 than the actual value of 9.2 years in 2010. However, that in SSP3 will be only 11% larger in 2030 than it was in 2010 and there will be no meaningful improvement by 2100 (see the SSP3 FMYE centroid in Figure 5). In the SSP1 scenario, advancements in female education will effectively curb population growth, as Figure 5 illustrates. Under the SSP1 scenario, the world's population will peak in the middle of the 21st century and then decrease to less than 7.5 billion in 2100. The population under the SSP2 scenario will peak at 9.6 billion around 2070 and then decrease to 9.2 billion in 2100. The population under the SSP3 scenario will continue to grow, reaching 13.5 billion in 2100. The DESA population centroid trajectory (the black crosses) is similar to the SSP3 population centroid trajectory, although DESA predicts a population of 10.9 billion in 2100.

In the SSP1 scenario with worldwide cooperation and coordination, advancements in female education and the enhanced awareness of sustainability would result in an increase in productivity, but a decrease in population, the development of green technologies, and rising RGDP/P globally. In the SSP3 scenario, where female education would continue to lag in developing economies and thus the population would continue to increase rapidly, we would not expect a rapid increase in productivity. Technology may continue to progress in advanced economies, but its spread to the rest of the world would be slow under the SSP3 scenario. In most developing economies, RGDP/P would only grow slowly (see the sizes of the RGDP and RGDP/P centroid markers in Figure 5). It is important to reiterate that these economic growth projections do not consider any mitigation or adaptation costs due to global climate change, which would undoubtedly be greater in the SSP3 scenario than in the SSP1 scenario. Further, all the projected productivity centroid trajectories appear to be different from previous productivity centroid trajectories. The RGDP/E and RGDP/P centroids for 1960–2017 in Figure 5 moved east/northeast and there is no sign that their directions will change. Even the 2010–2020 RGDP/P baseline extended to 2100 (the black stars in Figure 5) is different from the SSP1 and SSP2 RGDP/P centroid trajectories, and it also differs from the SSP3 RGDP/P centroid trajectory.

Figure 7 illustrates the differences in the 2010–2025 and 2040–2080 RGDP/P baselines of the SSP1 and SSP3 scenarios. We chose these baseline periods because Figure 5 demonstrates that the differences between these baselines are small for all three SSP scenarios from 2010 to 2025 and all three scenarios anticipate relatively stable RGDP/P trajectories after 2040. Figure 7 sorts the countries on the *x*-axis of the top and middle panels according to their values in the 2010–2025 RGDP/P baseline for SSP1, while it sorts countries in the bottom panel according to their values in the 2040–2080 RGDP/P baseline for SSP1. The top panel of Figure 7 indicates that countries' RGDP/P will converge under the SSP1 scenario in the latter half of the 21st century. Developed countries during the 2010 to 2025 period, which are mostly to the left on the *x*-axis, will slow their decreases, while rapidly growing countries, which are mostly to the right on the *x*-axis, will slow their increases. Underdeveloped countries, which are in the middle of the *x*-axis, will increase their shares. Again, note that a share (weight) of a normalised variable such as RGDP/P indicates a position relative to the world mean (see Section 2). The middle panel indicates that the convergence would occur slowly in the SSP3 scenario; the countries in the middle of the *x*-axis would catch up relatively slowly (see the black bars in the middle panel). The bottom panel of Figure 7 clearly illustrates the differences between the SSP1 and SSP3 scenarios in the latter half of the 21st century. In the SSP1 scenario (white bars), developed countries, most of which are on the right side of the *x*-axis, would considerably decrease their relative positions compared with the mean world RGDP/P, while developing countries, most of which are on the left side of the *x*-axis, would increase their relative positions. The five countries that would increase their positions would be Liberia, Saint Vincent and the Grenadines, the Republic of the Sudan, Guyana, and Lesotho, with a number of African countries following them. Figure 7 also reveals that some oil-producing countries would still enjoy rapid growth in their RGDP/P until the end of the 21st century under the SSP3 scenario.

It is beyond the scope of this study to examine the validity of the projections. What this study established is that the economic centroid method can visualise the magnitude of the challenges the world is facing. If we want to live in a sustainable world, the trajectory of the productivity centroid over the past 50 years must bend by almost 180°.

Figure 7. Expected changes to RGDP/P in SSP1 and SSP3

5. Discussion

This study revisited the economic centroid method and discovered that it is not useful for analysing its underlying socioeconomic variable. First, the calculation of the economic centroid uses only a proportion of the available information, namely, the ratio of the shares of the sub-regions to the total. Second, the contribution of an individual sub-region to the centroid movement varies depending on the location of the moment centre, which can be set arbitrarily. The economic centroid method is, however, undoubtedly a powerful visualisation tool for discovering the spatial and temporal structural changes in spatial panel data. Applying the DPG approach and using the newly developed QGIS Python tool, ECAT, makes it easy to visualise, identify, and quantify the spatial and temporal structural changes in a socioeconomic variable in multiple regions. Applying the economic centroid method to the world population and economy revealed that the method can visualise both continuities and changes in the past as well as the magnitudes of the challenges we face in the future.

One problem with the economic centroid method is that the moments of the regions define the economic centroid's movements on a map. The moments of the two regions vary depending on their distances to the moment centre, even if the values of the variable that calculate the economic centroid are identical in the regions. In terms of geography, this is not a problem. The economic centroid method displays the geographical location of the centroid as well as the influences of the regions, tied to their geographical locations, on the movement of the centroid. In terms of socioeconomic studies, however, it is sometimes desirable to visualise the influence of a change in a socioeconomic variable in a region on the economic centroid, which can control the effect of the region's geographical location. It seems impossible to eliminate all the locational attributes of regions if we want to visualise the temporal changes in a socioeconomic variable in the regions on a two-dimensional plane; however, it may be possible to develop a method that minimises or neutralises the influence of the geographical location of the regions on the movement of the economic centroid. We are working on elaborating a methodology and developing another QGIS Python tool that visualises the pure effect of a change in a socioeconomic variable in a region on the movement of the economic centroid; however, this is a task for the future.

The economic centroid method has its shortcomings in its partial use of the available information and complications due to moments. However, it is still a powerful visualisation tool that assists in finding and analysing the temporal and spatial changes in spatial panel data. Moreover, the method is easy to use. We thus expect the economic centroid method to be applied to a wide range of socioeconomic spatial studies in the future.

Funding

This work was supported by the JSPS under Grant-in-Aid for Challenging Research (Pioneering) 19K2168930.

Supplementary materials

Supplementary material A. *Economic Centroid Analysis Tool (ECAT).*

https://osf.io/du46q/?view_only=e32aaa6b9516499c91f8eacae27a0df7

The material explains how to install and use the QGIS processing script, ECAT. It also explains the naming convention for the html, jpg, and json files of the various graphs and the csv data files that ECAT outputs.

Supplementary material B. *Soviet Macroeconomic Data*.

https://osf.io/792tr/?view_only=684489155b5b4204941898450d6d2ace

World and Soviet Bloc GDP and Employment. The material includes the excel data file of the real GDP and the employment in the former Soviet Union and East European countries before 1990. It also explains the sources and the compilation method.

Supplementary material C. *Structural Changes in the World's Population and Economy. An Economic Centroid Analysis*.

https://osf.io/knyb7/?view_only=65272fda223e415aa90fbb767e0e5886

The material includes the html files of Figures 1, 3, and 4, and the QGIS map layer (gpkg) files and data (csv) files to reproduce the maps of Figure 2 and 5 in QGIS.

Reference

Appleby Magna (2002) Appleby Parva: Center of the country! *Daily Mail*, 2 July 2002. http://www.applebymagna.org.uk/population_centre.htm.

Australian Bureau of Statistics (ABS, 2020) *Regional population growth, Australia*, 2018–19. https://www.abs.gov.au/AUSSTATS/abs@.nsf/Latestproducts/3218.0Main%20Features52018-19?opendocument&tabname=Summary&prodno=3218.0&issue=2018-19&num=&view=.

Breau, S., Toy, B., Brown, M., Macdonald, R., and Coomes, O. (2018) In the Footsteps of Mackintosh and Innis: Tracking Canada's Economic Centre of Gravity since the Great Depression, *Canadian Public Policy*, 44(4), 356–367. DOI: 10.3138/cpp.2018-015.

Dellink, R., Chateau, J., Lanzi, E., Magne, B. (2017) Long-term economic growth projections in the Shared Socioeconomic Pathways. *Global Environmental Change*, 42. DOI: 10.1016/j.gloenvcha.2015.06.004.

Feenstra, R. C., Inklaar, R., Timmer, M. P. (2015) The Next Generation of the Penn World Table. *American Economic Review*, 105(10), 3150-3182.

GADM (2020) *GADM maps and data*. https://gadm.org/download_world.html.

International Institute for Applied Systems Analysis (IIASA, 2018) *Transformations to Achieve the Sustainable Development Goals*. Report prepared by the World in 2050 initiative. DOI: 10.22022/TNT/07-2018.15347.

International Institute for Applied Systems Analysis (IIASA, 2020) *The SSP-database*. https://tntcat.iiasa.ac.at/SspDb.

Lutz, W. (2019) World population trends and the rise of homo sapiens literata. *IIASA working paper*, WP-19-012.

Lutz, W., Cuaresma, J. C., Sanderson, W. (2008) The Demography of Educational Attainment and Economic Growth. *Science*, 319: 1047. DOI: 10.1126/ science.1151753.

Lutz, W., Samir, K. C. (2011) Global Human Capital: Integrating Education and Population. *Science*, 333: 587-592.

Ministry of Internal Affairs and Communication (MIC, 2017) *Waga Kuni no Jinkou Jyushinn*. https://www.stat.go.jp/data/kokusei/topics/pdf/topics102.pdf#page=2.

Nakamura, Y., Zhang, Y., Arikawa, K. (2019) *Mystery of Economic Centroid Movements*, paper presented at the annual convention of Japanese Association for Comparative Economic Studies, at Hitotsubashi University, Tokyo, on June 18, 2019.

Organisation for Economic Co-operation and Development (OECD, 2020) *Environmental-economic modelling*. http://www.oecd.org/environment/indicators-modelling-outlooks/modelling.htm.

Penn World Table (PWT, 2020) *Version 9.1*, Groningen Growth and Development Centre. https://www.rug.nl/ggdc/productivity/pwt/.

Plane, D. A., Rogerson, P. A. (2015) On Tracking and Disaggregating Center Points of Population. *Annals of the Association of American Geographers*, 105(5): 968–986. DOI: 10.1080/00045608.2015.1066742.

Riahi, K. et al. (2017) The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change*, 42: 153-168. DOI: 10.1016/j.gloenvcha.2016.05.009.

United Nations, Department of Economic and Social Affairs, Population Division (DESA, 2019). *World Population Prospects 2019*, Online Edition. Rev. 1.

United States Census Bureau (USCB, 2020) *Mean Center of Population for the United States: 1790 to 2010*. https://www2.census.gov/geo/pdfs/reference/cenpop2010/centerpop_mean2010.pdf.

Wittgenstein Centre for Demography and Global Human Capital (WCD, 2020). *Wittgenstein Centre Data Explorer Version 2.0*. www.wittgensteincentre.org/dataexplorer.

Figure 1. Deviations from proportional growth

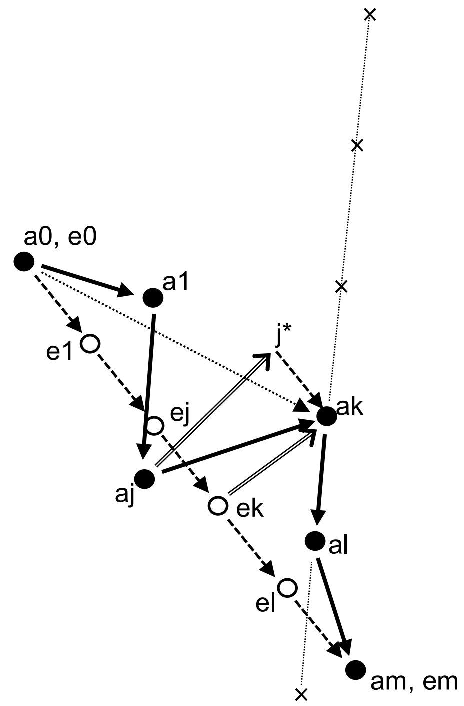
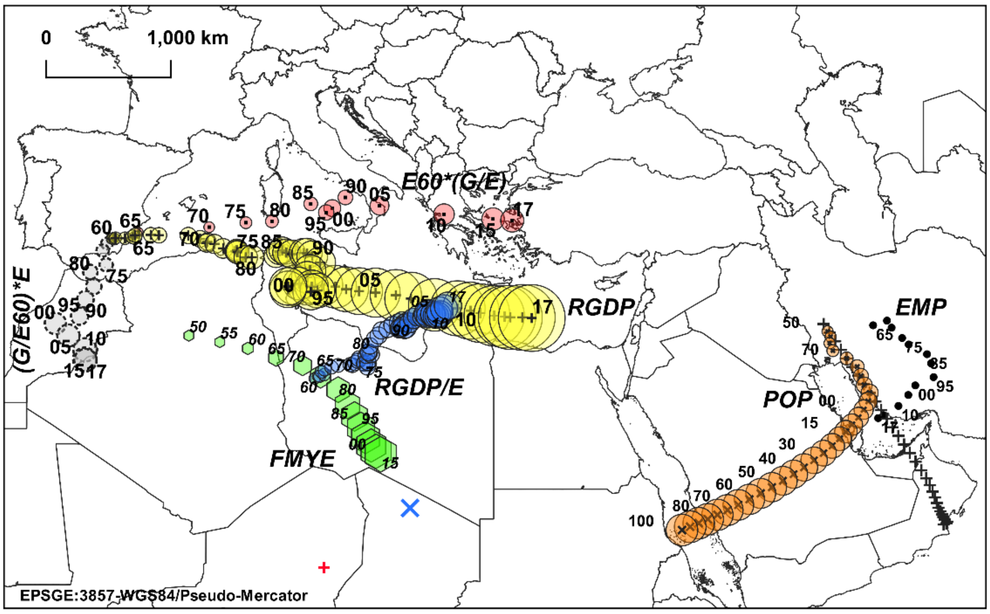
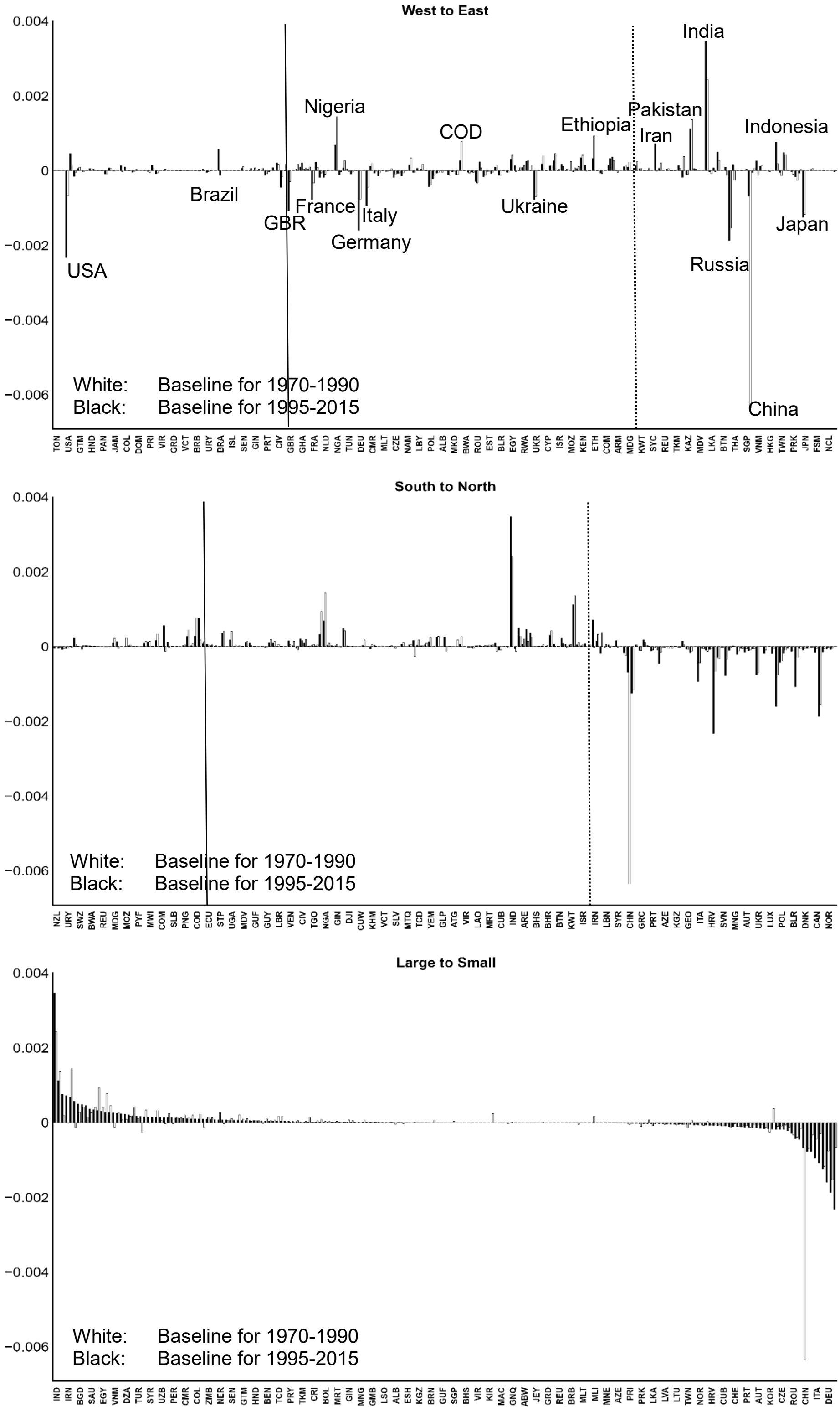


Figure 2. Population, RGDP, productivity, and education



Note: See the text for an explanation of the symbols. Supplementary material C includes all the QGIS map and data files necessary to reproduce the map.
Source: Author's calculation.

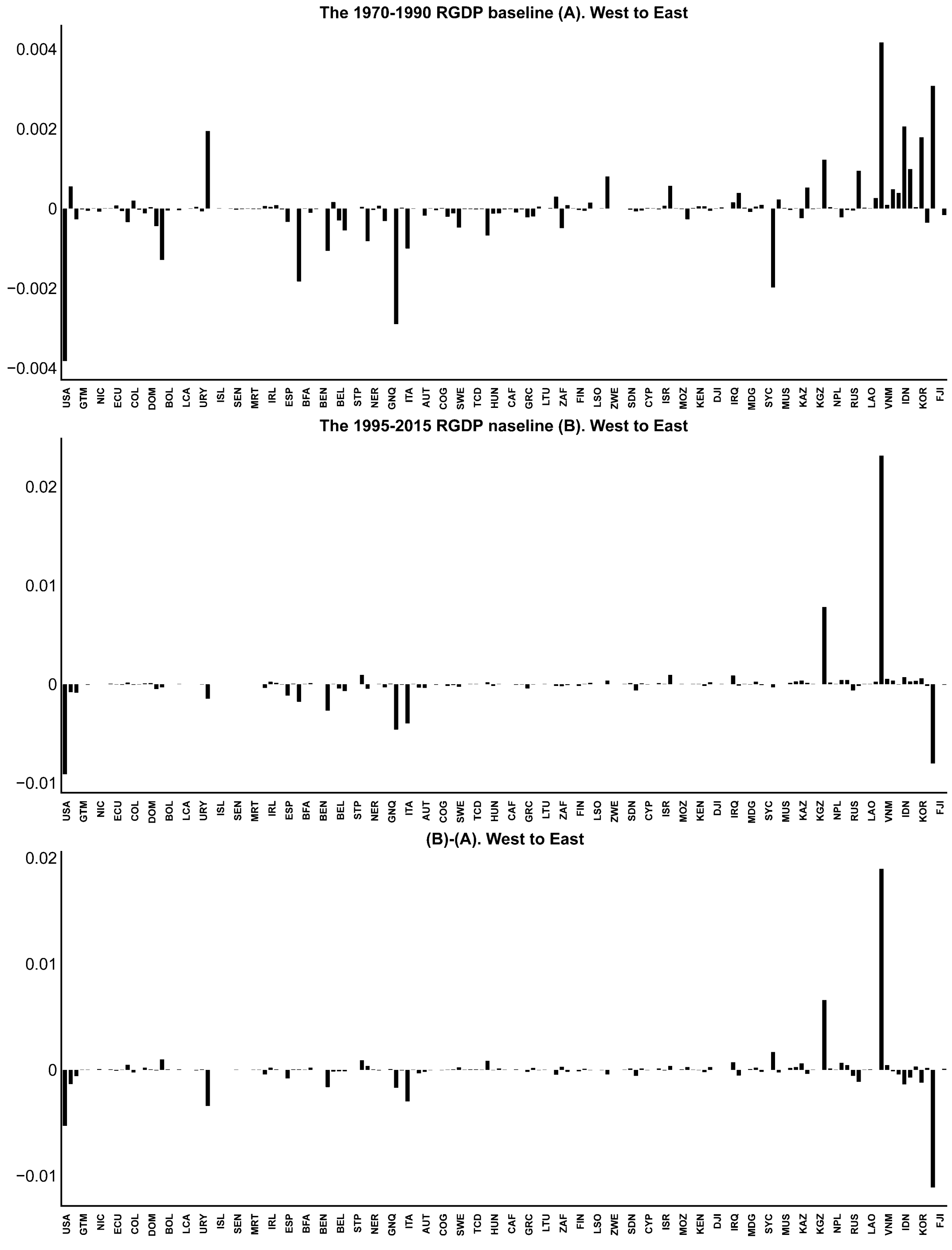
Figure 3. Changes in population trends (shares, in points)



Note: This figure compares the 1970–1990 and 1995–2015 population baselines (expected changes in shares). The baselines are the mean changes in shares during each period. The title of the subplot in each panel indicates how each panel sorts the countries along the *x*-axis. Owing to space limitations, country

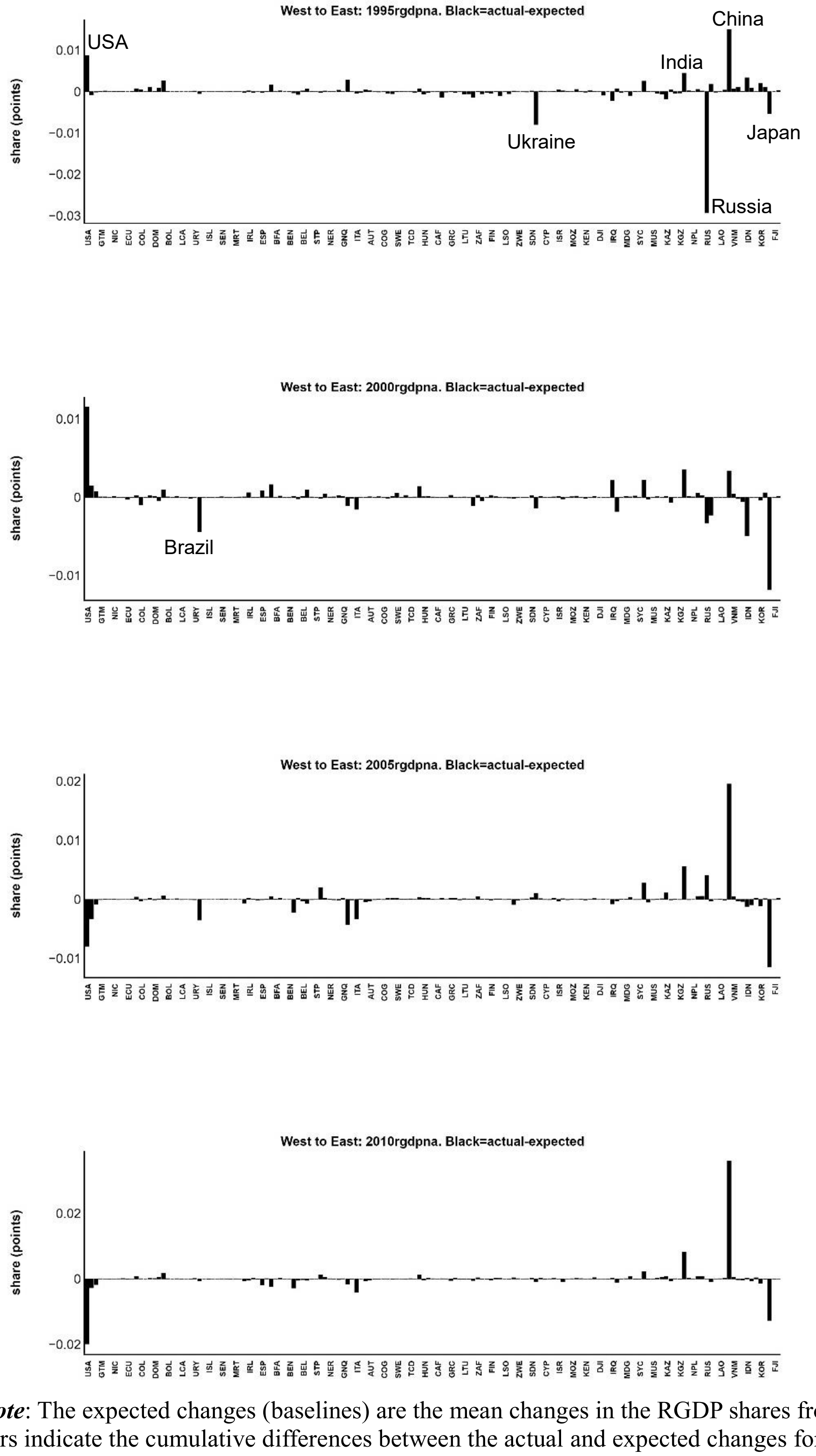
names in ISO 3166-1 A3 have been partially hidden. See Supplementary material C for the full interactive graph, which also provides the html, json, and jpg files for this and other graphs as well as the data (csv) files.
Source: Author’s calculation.

Figure 4. Changes in the RGDP trend (shares, in points)



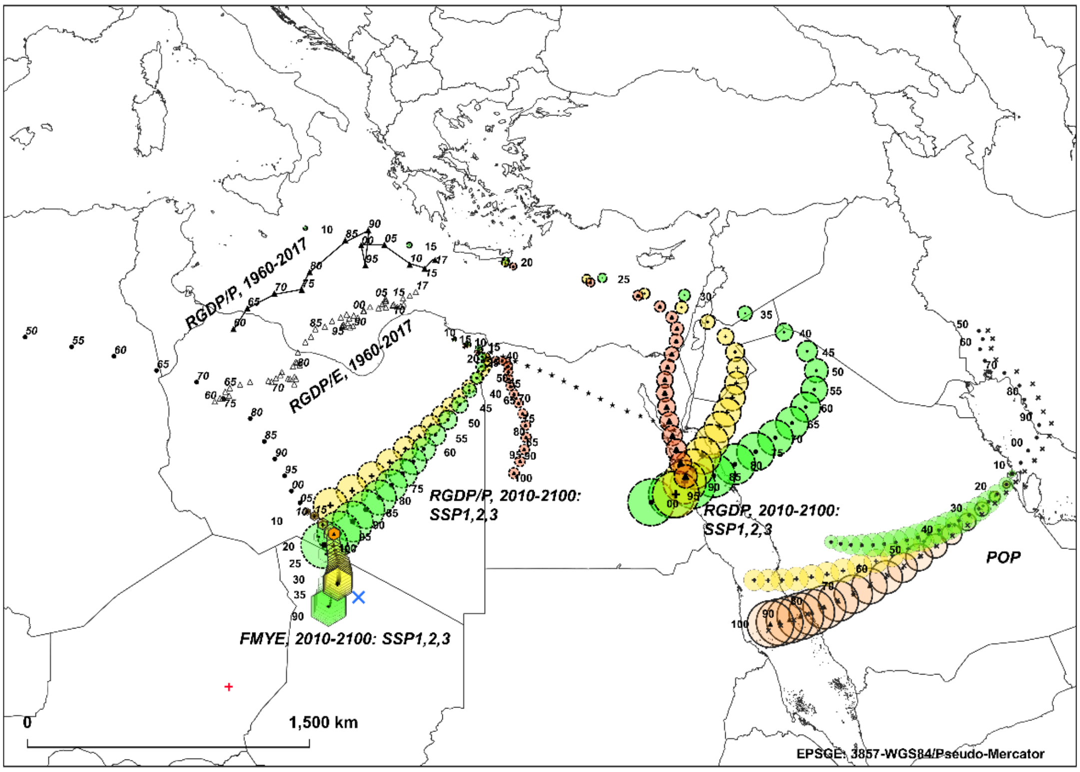
Note: For visibility, the bars display the cumulative changes over five years, calculated simply as five times the annual change. Owing to space limitations, the country names in ISO 3166-1 A3 are partially hidden. Supplementary material C provides the full interactive graph.
Source: Author’s calculation.

Figure 5. The actual and expected changes in the RGDP shares



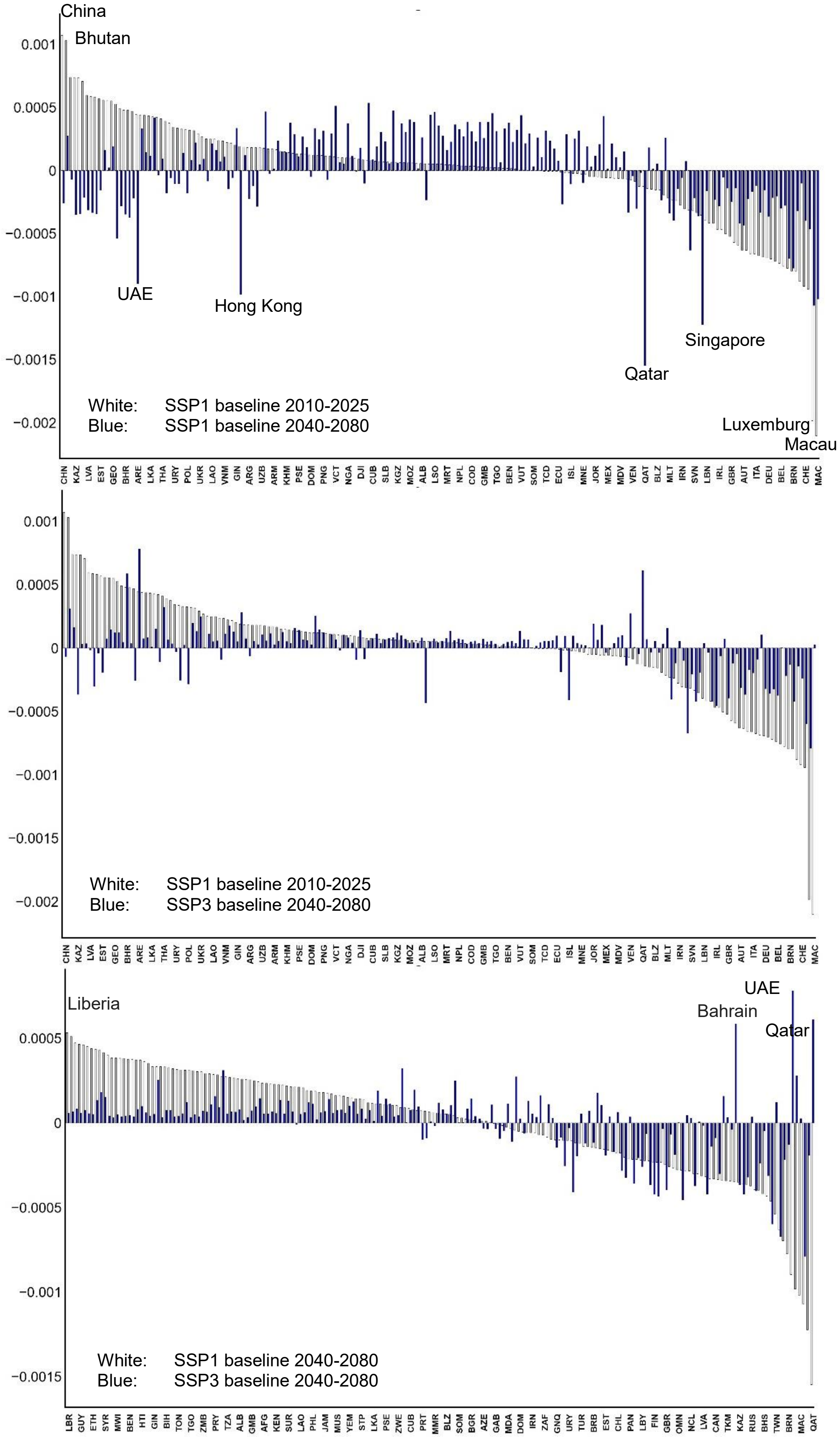
Note: The expected changes (baselines) are the mean changes in the RGDP shares from 1970 to 1990. The bars indicate the cumulative differences between the actual and expected changes for the five years before the year in the subtitle. The countries on the *x*-axis are sorted as per the subplot titles. Owing to space limitations, not all country names in ISO 3166-1 A3 are shown on the *x*-axis. See Supplementary material C for the full interactive graph, which also includes the html, json, and jpg files for this and other graphs as well as the data (csv) files.
Source: Author's calculation.

Figure 6. Population, productivity, and education from 2010 to 2100



Note: See the text for an explanation of the symbols. Supplementary material C provides the QGIS and data files necessary to reproduce the map.
Source: Author's calculation.

Figure 7. Expected changes to RGDP/P in SSP1 and SSP3



Note: The baselines are the mean changes in the share of RGDP/P from 2010 to 2025 and 2040 to 2080. The RGDP/P share indicates the country's position relative to the mean of world RGDP/P (see the Methods section). For visibility, the bar displays the cumulative changes over five years, calculated simply as five times the annual change. The countries on the *x*-axis have been sorted according to their 2010–2025 RGDP/P values in the SSP1 baseline for the top and middle panels and according to their 2040–2080 RGDP/P values in the SSP1 baseline for the bottom panel. Owing to space limitations, the country names in ISO 3166-1 A3 have been partially hidden. See Supplementary material C for the full interactive graph, which also includes the html, json, and jpg files for this and other graphs as well as the data (csv) files.
Source: Author's calculation.