International Journal of Business and Economics, 2018, Vol. 17, No. 2, 95-112

Using Hierarchical Cluster Analysis as a Tool to Fit Aggregate Production Functions

Rafael Solis*
Department of ISDS Craig School of Business, California State University, Fresno

K.C. Tseng
Department of Finance, Craig School of Business, California State University, Fresno

Abstract

In this study, we discuss the use of clustering algorithms to group countries based on a set of macroeconomic variables and their effect on the estimation of the parameters of a Cobb-Douglas type production function. The findings suggest that the utilization of clustering methodologies could play an important role when grouping of countries may be needed without adversely affecting the meaning of the parameters of the fitted production functions.

Key words: hierarchical clustering; Cobb-Douglas production function; macroeconomics

JEL classification: C31; C38; D24; C51; C82

1. Introduction

Since Robert Solow (1957) published his pioneering work on technical change and aggregate production function, the empirical estimation of aggregate production functions has been controversial. A typical production function shows the maximum outputs to be produced when all inputs are efficiently utilized. Estimating the parameters of a production function from time series data for a single product firm, or Using cross-sectional data for firms within the same industry manufacturing the same product, is rather straight forward. On the other hand, the fitting of a production function for various products requires combining different inputs across an economy. This is the case when various inputs (and outputs) such as labor and capital need to be aggregated. As Fisher (1993) points out, "the aggregate production function will exist if and only if every firm's production function is additively separable in capital and labor". In his book (p. 10-35), he discusses in detail the difficulties of aggregations of capital, labor, and outputs. If the ratios of capital

*Correspondence to: Craig School of Business, California State University, Fresno, 5245 N Backer Ave, Fresno, CA 93740, USA. E-mail: rafaels@csufresno.edu.

goods are constant, capital can be aggregated. Similar rules apply to labor or output aggregation. Since the stringent conditions for aggregating inputs and outputs is rarely met, the aggregate production function can only be roughly estimated. Even with the problems of aggregation, aggregate production functions can only approximately depict the relationship between aggregate inputs and aggregate outputs. In this sense, aggregate production functions can still be used to identify productivity, to do forecasting, and to predict capital and labor's shares of national product. In a cross-sectional study, an aggregate production function can be used to detect the differences in technology and quantify the shares of various inputs among different countries. Differences in technology and factor contributions to national products among countries may also help us understand and predict trade flows between countries. Unfortunately, in recent years there has been no cross-sectional study to examine the important issues just mentioned. The major thrust of this research is to fill the gap so that the empirical findings from this study may shed some light on the essential economic issues.

Most significant empirical studies on aggregate production functions were conducted in the early days. For example, a path-breaking study by Solow (1957) using U.S. data from 1909 to 1949 found that technical change (which he defines as "any kind of shift in production over time") focused on how neutral and aggregate product increased about one percent per year for the first half of the period and accelerated to two percent per year during the second half. About 7/8 of the increase in output per man-hour was the result of technical progress, and 1/8 was attributable to increased use of capital. In another study (Zellner et al., 1966) it was established that when random errors are normally distributed and independent across equations, the least squares estimators are consistent and unbiased for the Cobb-Douglas type of production function. Beckman and Sato's study (1969) was the most extensive based on German, Japanese, and U.S. data and ranging from 1850-1959 for Germany, 1909-1960 for the U.S. and 1930-1960 for Japan. They discussed various types of production functions and their relationship to technical progress. They found that the following types of neutral technical progress performed well: "In the United States, labor-combining and capital combining technology; in Japan, labor-decreasing and labor additive technology; in Germany, both labor and capital decreasing, and capital-additive technology" (p. 95). They also found that the estimated production functions are most close to the Cobb-Douglas or constant elasticity of substitution production functions regardless the types of technical change being specified. The only cross-sectional study known to us was done by Arrow et al. (1961). In this study, they included 19 countries with annual data ranging from 1949 for Brazil to 1955/56 for Australia and New Zealand. The number of industries varies from two for Iraq and 6 for S. Rhodesia to 24 for several countries such as the U.S. They found that "the elasticity of substitution between capital and labor in manufacturing may typically be less than unity. There are weaker indications that in primary production this conclusion is reversed" (p. 401). Their assumption that: "international differences in efficiency are approximately neutral in their incidence of capital and labor" (p. 415) was supported by their study.

All these studies except the last one cited are based on time series analysis, and their focus is on the test of neutrality of technical progress.

2. Purpose and Objective

The proposed empirical study is based on cross-sectional data (macroeconomic data for different countries at approximately the same time period, 2014). There are three principal objectives of this study. First, we want to see if the proposed Cobb-Douglas estimation results will differ when multivariate clustering analysis is applied to group countries versus the traditional univariate grouping used by many economists such as using gross domestic product per capita to segregate economies into traditional groups such as developed, economies in transition, and developing countries. As far as we are aware, cluster analysis has not been applied to segment data in any production function study to date. Second, it is motivating to see if the technical differences among the group of developed countries are significantly different from those among the less developed countries. If there are differences as we expect, the trade and investment flows are expected to continue in addition to continuous technical transfers. Third, it is meaningful to also examine if the contributions to GDP by labor and capital among developed countries are significantly different from those among the less developed ones. Under competitive markets, labor and capital shares will reflect the marginal productivities of labor and capital, respectively. The differences of labor and capital productivities among countries have important implication for factor movements, particularly among the developed and developing countries.

3. Methodology and Data Sources

Since we have not found any cross-sectional empirical study on aggregate production functions at the national level in recent years, we have decided to apply the well-known Cobb-Douglas (1928) production function (1) to the present study.

$$Q = \alpha K^{\beta_1} L^{\beta_2} . \tag{1}$$

In this format, the inputs are Capital (K) and Labor (L), and Q is the quantity of production. The sum of the exponents β_1 and β_2 determine the returns to scale on the factor inputs, and α is a scaling factor. When $\beta_1+\beta_2=1$ the Cobb-Douglas function is said to be homogeneous, meaning that a doubling of the two inputs (simultaneously) results in a doubling of the output. β_1 and β_2 are also the respective partial "output elasticities" of the production function, meaning that they measure the percent increase in production resulting from a percent increase in the particular input factor.

In addition, this special function can be transformed into log-linear form with linearity in parameters so that standard statistical hypothesis testing can be conducted by using the ordinary least squares estimation method. As pointed out by

Zellner et al. (1966), the least squares estimators are unbiased and consistent if random errors are normally distributed and independent across equations. The estimated parameters are the elasticities of aggregate output on the inputs included in the equation. With properly defined input and output variables and under perfect competition and constant returns to scale, the exponent associated with each input indicates the factor share of national income of that particular factor of production. This is derived from the fact that under competitive markets, there exist no excess profit and all factors of production will exhaust national product. Furthermore, the sum of the exponents associated with all inputs included in the equation shows the returns to scale. For example, if the sum of exponents is greater (less) than one, there exist economies (diseconomies) of scale. The estimated parameter (exponent) is the output elasticity on a particular input (factor). Finally, based on several time series empirical studies cited above and others, the Cobb-Douglas production function has performed well.

In previous empirical studies, proxy variables to Q, K, and L have been used. For example, Beckman and Sato (1969), in their comparative study of the United States, Germany, and Japan, GNP and capital stock were in constant dollar, mark, and yen respectively, and labor was in billion hours worked. In Solow's (1957) study of the U.S. economy, he used GNP, capital stock, and labor (defined as % of labor force employed multiplied by capital stock). More recently, McMillin and Smyth (1994) used output/capital ratio, energy price, the government capital/the private capital ratio, and inflation rate. Finally, Basu and Fernald (1997) used Dale Jorgenson's data for 34 industries in the U.S. Since different studies used different proxy variables in their production functions, it is difficult and not very meaningful to directly compare the empirical findings from various studies.

Due to the limitation of available and comparable data, the Cobb-Douglas production function with two inputs was utilized. Gross domestic product per capita is the output variable. We used total employment as a proxy variable for labor as one of the two input variables. This proxy (EMP) was calculated as the total labor force (1- unemployment rate). The second input variable is electricity production (ELECP), which is a proxy for capital. In a recent study by Strauss-Kahn (2004) the electricity consumption was used to proxy capital. We consider the application of either electricity production or consumption as a proxy for capital will not make any significant difference, and only electricity production is available when we use the USA Central Intelligence Agency (CIA) data in this study. Fisher (1993) has discussed in detail the existence of an aggregate capital stock and that of labor and output aggregates. We realize that using electricity production as a proxy for capital may not be the most appropriate, but this was the closest proxy variable we could find in the CIA's World Factbook data retrieved from https://www.cia.gov/library/publications/the-world-factbook/print/textversion.html /. In addition, there is no aggregation problem with this proxy variable. From the above and in terms of the proxy variables, the production function in this study is specified as:

$$GDP = \alpha ELECP^{\beta_1} EMP^{\beta_2} \varepsilon \,. \tag{2}$$

Where GDP is gross domestic product (note: most of the country data is from 2014), ELECP is the yearly production in millions of KwH, and EMP is the actual number of the labor force actively employed. The random error term is indicated by ε. For the purpose of estimation, the above equation was transformed into the log-linear form as:

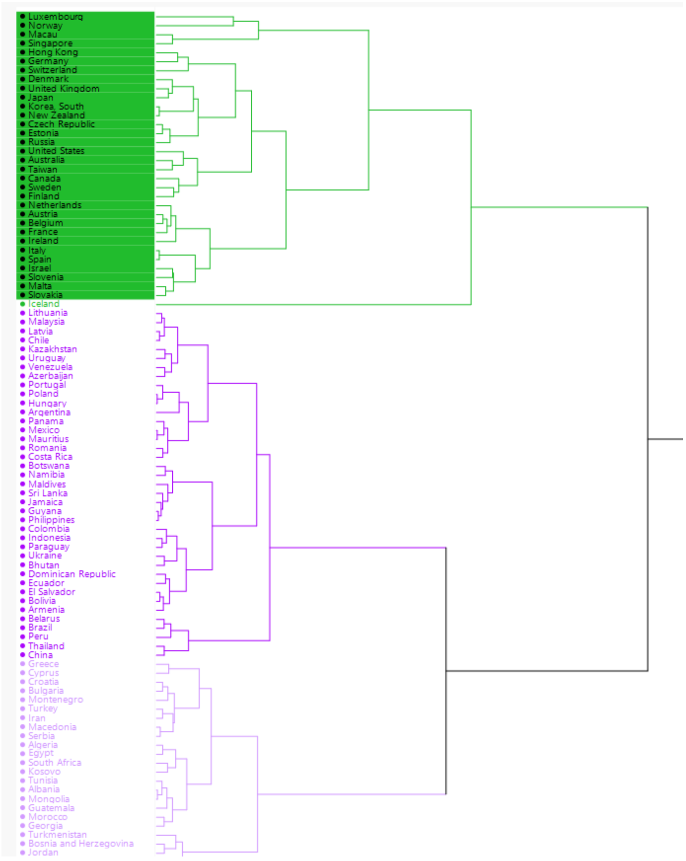
$$Ln(GDP) = Ln(\alpha) + \beta_1 Ln(ELECP) + \beta_2 Ln(EMP) + Ln(\varepsilon) \,. \tag{3}$$

Where the random error term $Ln(\varepsilon)$ is assumed to be normally distributed with zero mean and constant variance σ^2. Under this assumption, the random error in (1) is log-normally distributed with mean $e^{\sigma^2/2}$, and variance $e^{\sigma^2(\sigma^2-1)}$. Since Equation (2) is linear in its parameters, the ordinary least squares method can be applied to estimate these parameters, and conducting standard statistical tests to them.

For this study, 95 countries were selected in such a way that there were no missing values among the variables used and that the chosen countries were not too small in the economic sense (we eliminated countries below the 25th percentile for GDP per capita or those less than \$6,040). To test the validity of model 3, we estimated the parameters based on our sample of 95 countries (we call this fit WORLD) then we partitioned these 95 countries based on GDP per capita and used the two subsamples to fit the same equation. The WORLD sample was also divided into two groups based on a cluster analysis algorithm to be discussed in more detail later. The second set of estimates was based on dividing the WORLD sample into two groups of countries based on per capita GDP. The top group of countries (GDP per capita) was at the 50th percentile (\$19,039) and above and the ones below the 50% percentile as the second group, which mostly included the developing countries. Note that the above partitioning is based on a single variable (GDP per capita). Grouping the countries based on a multivariate approach can be done (among other methods) using Clustering algorithms. Cluster Analysis is a well-known statistical technique for discovering structure within complex bodies of data. A typical problem involves finding groups or clusters among m objects, in this case, the countries, which had been subject to k measurements (variables). The data can be represented by a (mxn) matrix of data X in which the entries of each row vector $\vec{X}_i$, i=1,…,m in X, contain the values of l (macroeconomic) variables for each of the m objects (countries). From this data matrix, the objective is then to group the objects into clusters such that their elements have a high degree of "similarity" within each cluster and at the same time, the clusters are relatively "distinct" from one another. The need for clustering this kind of multivariate data arises in a natural way in many fields of study. A large body of literature exists in life sciences, earth sciences, medicine, engineering, economics, marketing, operations research, etc. Good introductions include Anderberg (1973), Hartigan (1975), Gordon (1981), Kauffman and Rousseuw (1990). Applications of cluster analysis in Economics include: Fisher (1969), Hirschberg et al. (1991) and Szilagyi (1991).

Using a hierarchical clustering algorithm using Ward's minimal increase of sum-of-squares (MISSQ) method (Ward, 1963) we obtained three distinct clusters shown in Figures 2 and 3 below. The clusters were obtained by utilizing three (standardized) variables: Electricity production (ELECP) and Percent of labor force actively employed (EMP), and GDP per capita (GDP_Cap).

Figure 2. Hierarchical Clustering Dendrogram of 95 Countries



In Figure 3, the lower right-hand cluster (labeled H1 in the discussion that follows) closely corresponds to countries with a GDP per capita above $27,529 (32 countries) plus Russia with a GDP per capita of $25,052 (in total 33 countries in this cluster). The rest of the countries are broken down into two distinct clusters. The top left cluster (H2) consisting of 39 countries and the top right cluster with 23 countries (H3). It is of significance importance to note that the two top clusters consist of a breakdown of countries with a GDP per capita under $27,529 but not necessarily in any order. Classifying countries based on per capita GDP above or below $19,039 (50 percentile) and the hierarchical clustering method, produces a different classification of countries. For example, 21 countries with GDP per capita below $19,039 are included in H2 and H3, while 7 countries with GDP per capita above that threshold are classified as H1. Summary statistics resulting from the univariate and hierarchical clustering methods are shown in Table 1.

Figure 3. Hierarchical Clustering (Constellation Graph)

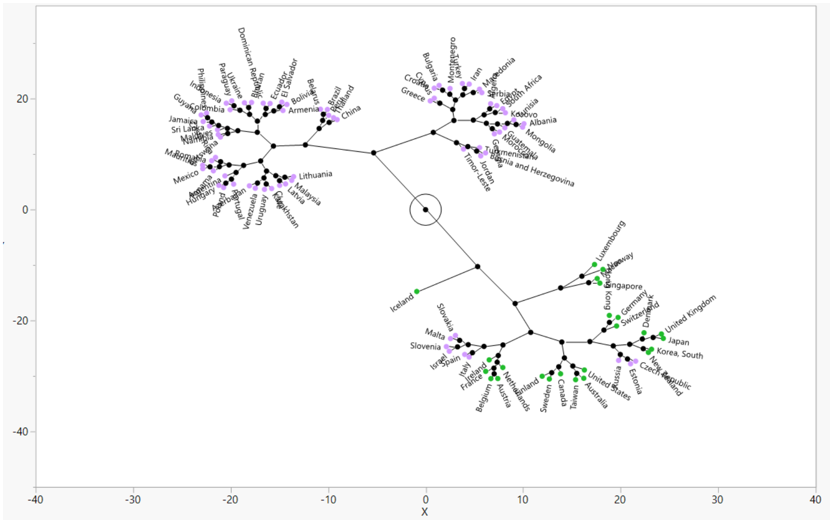


Table 1. Clusters Sample Statistics

Countries	N	GDP/Capita (US$)	Electr' Prod./ Capita (Kwh/year)	EMP Employment (% of Labor Force)
World	95	$25,018($17,742)	4,957(5,968)	41.29(9.29)
GDP/Cap. ≥ $19,039	48	$37,910($16545)	7,920((7,972)	44.93(7.69)
GDP/Cap. < $19,039	47	$11,851($3,608)	2,461(2,047)	37.58(9.38)
H1	33	$44,649($15,783)	9,616(8,098)	47.42(6,43)
H2	39	$15427($6,272)	2,172(1,185))	43.77(4.91)
H3	23	$13,115($4,994)	2,581(1483)	28.32(4.89)

(Entries are Mean, Std. Dev. K1, K2, K3 are the Hierarchical Method Clusters. A List of the Clusters can be found in the Appendix)

4. Empirical Findings

The Cobb-Douglas fitting results for each type of cluster are shown in Table 2. All six fitted equations were found statistically significant with (adjusted) R-squares ranging from 0.9292 to 0.9716, which are considered very high in cross-sectional studies. All estimated parameters were significant at 5% level for all equations classified by a clustering method and for the whole sample (except for electricity production of the countries in the top 50% percentile for GDP per capita and the H1 hierarchical cluster which contain almost the same countries). The intercept term representing the technical differences among countries within a cluster or among all countries were highly significant for the six models ($p<.001$). The technical differences among all countries (n95) amounted to only 0.005 in 2014. The technical transfer between the developed and the developing countries, the convergence of technology among countries in the past three decades, the widespread of Internet, world wide web, and mobile technology (Spence, 2011) may account for the smaller difference in technology among countries. In addition, the technical differences among the developed countries are greater than those among the less developed countries as evidenced by the larger intercept term estimated for the developed countries than that for the developing countries, about 0.00005 for H1 versus 0.0011 for H3 based on the cluster analysis groups, and 0.0004 versus 0.0002 using GDP per capita to classify the countries. The higher value of the technical coefficient among the countries classified by GDP per capita may be due to the inclusion of highly technologically-advanced countries such as the United States, Japan, Germany, United Kingdom, and Switzerland on the one hand, and the less technically inclined countries such as Uruguay, Panama, Kazakhstan, Chile, and Hungary. At the same time, the technical differences among either the developed or the developing countries are somehow twice as different for both groups of countries. Among the hierarchical clustering groups of countries, the smallest technical coefficient corresponds to the truly highly technical countries in the world found in H1 with the highest value for this coefficient found for the less advanced (and also less developed countries). Except for the group of the higher income countries (applying GDP per capita to classify countries and the group containing H1 obtained through hierarchical clustering), the sum of the regression coefficients of labor and electricity production (the proxy for Capital) is approximately one. This is consistent with constant returns to scale for the dependent variable (GDP per capita). Under the assumptions of perfect competition and constant returns to scale and based on the H2 and H3 groups obtained by cluster analysis, labor contributes between 72 and almost 82 percent to GDP, while electric production contributes about 28 and 18 percent for the same group of countries. For the most developed countries in H1, the coefficient for electricity was insignificant and meaningless negative while the proxy for labor was significant and almost 1 (100%). This may be consistent with the thought that for developed countries the injection of capital may not be as important as the creation of jobs to increase the GDP. Using GDP per capita as a grouping criterion, for less developed countries, labor contributes 68 percent, and electric production (capital) contributes almost 30 percent. For the developed countries labor contributes only about 86 percent, whereas electricity contributes almost negative 11 percent, which is statistically insignificant and also economically meaningless. On the basis of the regression results, it is clear that the multivariate clustering method (the countries in each cluster) are more homogenous and therefore provides more refined groups (in a more holistic economic sense) than the traditional classification using GDP per capita. In this study, applying cluster analysis to countries grouped together makes a difference in both the grouping of countries and the estimated production functions. In a more complex study with a larger number of variables and cases the cluster analysis is expected to play a more important role. Our empirical findings appear to be consistent with economic implications of the Cobb-Douglas type aggregate production function.

Table 2. Least Squares Estimates for the Cobb-Douglas Function[a]

Clustering Method	Cluster	α	Ln(α)	β_1	β_2	
		Technical Coefficient	Technical Coefficient	Electric Production	Labor	Adj. R^2
Whole Sample	World (n=95)	0.0049	-5.309 (***)	.3636 (***)	.6241 (***)	.9292
Using GDP per Capita	GDP > $19, 039 (n=47)	0.0004	-7.775 (***)	-.107 (NS)	.862 (***)	.9582
	GDP < $19,039 (n=46)	0.0002	-8.363 (***)	.2995(**)	.6795 (***)	.9632
Hierarchical Clusters	H1(n=33)	5.57E-05	-9.794 (***)	-0.0871 (NS)	1.0561 (***)	.9716
	H2(n=39)	0.0003	-8.061 (***)	.1707 (**)	.8177 (***)	.9489
	H3(n=23)	0.0011	-6.727 (***)	.2670 (***)	.7216 (***)	.9489

a. Entries are: Parameter Estimates (*** =P<.001, * *=P<.05, NS=Not Significant) and R^2 (adjusted R^2).

In brief, the Cobb-Douglas production function is a good fit with the cross-sectional data compiled by the CIA and the study conducted by Arrow et al.

(1961). In fact, the Cobb-Douglas production function also fits well with time series data in a few of the studies cited previously. Upon the application of clustering method, the estimated production functions for both the developed and the less developed countries show approximately constant returns to scale. Labor contributes much more than electricity to GDP in both the developed and the developing countries. Further, the technical differences among countries appear to be rather moderate, which is in agreement to smooth and continuous technical transfusion, trade, and global economic convergence as envisioned by Spence (2011). Finally, the cluster analysis introduced in this study does show its applicability and usefulness to this type of analysis.

5. Summary and Conclusions

The study applies the well-known and tested Cobb-Douglas production function to cross-sectional data from a reliable CIA data source. We have utilized multivariate clustering techniques to classify countries into similar clusters so that countries within the same cluster will be less heterogeneous and more suitable to fit into a production function. We have discussed the advantages and disadvantages of clustering analysis since this technique has never been applied to production function studies. Based on the multivariate hierarchical clustering techniques all 95 countries are put into three groups. In addition, the traditional univariate approach used by economists based on a single variable such as GDP or income per capita is also used to separate the developed from the developing countries. The empirical findings from both approaches are somewhat different.

The empirical findings from the Cobb-Douglas production function are consistent with theoretical expectations. First, Labor contributes much more than electricity production to GDP for both the developed and the developing countries and also for the countries in the three hierarchical clusters. Second, technical differences among the developed countries are greater than those among the developing countries although the technical differences among the developed or the developing countries are quite moderate. Third, the multivariate clustering techniques are a useful method for grouping countries into similar clusters/groups, which makes it more holistic than utilizing a single variable. More applications can be tested particularly in more complex studies with large number of variables and observations. Finally, since there is a lack of cross-sectional studies on production functions, we hope this study may induce similar works in the future. It is worth pointing out that the Cobb-Douglas production function, introduced and tested a long time ago, is still a valuable and applicable model. Indeed, the implications from this type of studies are significant for international trade, factor movements and technical transmission, and global economic growth convergence. It is worthwhile continuing further researches along our approach in the future.

Appendix. GDP/Capita ≥ $19,039 (n=48)

COUNTRY	GDP_CAP	ELECTP(Bkwh)	LABOR_FO	EMPLOYED	UNEMPLOY
Luxembourg	88820.38	2.211	253600	235594.4	0.071
Macau	87189.64	0.4137	367800	360811.8	0.019
Singapore	78456.64	47.95	3557000	3489417	0.019
Norway	65192.07	144.7	2724000	2631384	0.034
Hong Kong	56097.75	39.97	3871000	3750999	0.031
Switzerland	54753.67	65.54	5046000	4884528	0.032
United States	54330.09	4048	156000000	146328000	0.062
Australia	48349.49	239.7	12370000	11627800	0.06
Netherlands	47091.37	98.57	7893000	7324704	0.072
Ireland	45929.27	26.09	2174000	1928338	0.113
Canada	44985.97	644.1	19210000	17884510	0.069
Germany	44784.2	575.9	44760000	42522000	0.05
Austria	44648.06	69	3778000	3607990	0.045
Denmark	44557.89	29.85	2771000	2626908	0.052
Sweden	44298.82	162.9	5124000	4719204	0.079
Taiwan	43647	235	11500000	11063000	0.038
Iceland	41606.66	17.19	185900	177534.5	0.045
Belgium	41248.77	74.13	5225000	4780875	0.085
Finland	40442.42	67.69	2665000	2435810	0.086
France	38870.83	532	29870000	26972610	0.097
United Kingdom	37994.5	365.7	32590000	30732370	0.057
Japan	37874.35	963	65930000	63556520	0.036
Korea, South	36363.49	494.7	26430000	25557810	0.033
New Zealand	35756.18	42.91	2452000	2307332	0.059
Italy	33400.63	286.2	25510000	22321250	0.125
Israel	33332.03	64.44	3784000	3534256	0.066
Malta	32321.57	2.221	190500	179260.5	0.059
Spain	31861.33	276.5	22930000	17358010	0.243
Slovenia	30523.16	14.76	913500	789264	0.136
Czech Republic	28154.48	81.71	5416000	4988136	0.079
Estonia	27974.9	11.66	669800	612197.2	0.086
Slovakia	27529.71	26.42	2730000	2383290	0.127

Appendix. GDP/Capita ≥ $19,039 (n=48) (Continued)

COUNTRY	GDP_CAP	ELECTP(Bkwh)	LABOR_FO	EMPLOYED	UNEMPLOY
Lithuania	27371.06	12.27	1454000	1292606	0.111
Greece	26383.58	54.98	3910000	2862120	0.268
Portugal	25495.81	44.27	5271000	4522518	0.142
Russia	25052	1054	75250000	71562750	0.049
Malaysia	24474.13	118	14010000	13603710	0.029
Latvia	24457.58	5.891	1014000	917670	0.095
Poland	24412.51	150.9	18260000	15940980	0.127
Hungary	24238.34	34.28	4388000	4076452	0.071
Chile	23434.65	61.85	8514000	7960590	0.065
Kazakhstan	23164.46	90.53	9103000	8638747	0.051
Argentina	21352.98	123.2	17310000	15977130	0.077
Panama	21041.7	7.642	1563000	1492665	0.045
Cyprus	20972.13	4.443	356700	299984.7	0.159
Uruguay	20880.38	10.16	1712000	1597296	0.067
Croatia	19552.75	14.24	1714000	1354060	0.21
Turkey	19039.4	228.1	27560000	24969360	0.094

GDP/Capita < $19,039 (n=47)

COUNTRY	GDP_CAP	ELECTP(Bkwh)	LABOR_FO	EMPLOYED	UNEMPLOY
Venezuela	18640.19	127.6	14340000	13221480	0.078
Belarus	17852.51	30.33	5000000	4950000	0.01
Romania	17838.72	56.71	9945000	9248850	0.07
Mexico	17603.55	277.6	52900000	50413700	0.047
Mauritius	17479.87	2.628	600200	552184	0.08
Azerbaijan	17217.44	22.99	4821000	4560666	0.054
Bulgaria	17156.23	46.65	2513000	2236570	0.11
Turkmenistan	15703.19	19.97	2300000	920000	0.6
Iran	15692.17	225.8	28400000	25474800	0.103
Botswana	15402.81	0.35	1017000	835974	0.178
Brazil	15044.56	530.4	110900000	104800500	0.055
Costa Rica	14791.83	9.889	2257000	2065155	0.085
Montenegro	14679.95	2.557	251300	203301.7	0.191

GDP/Capita < $19,039 (n=47) (Continued)

COUNTRY	GDP_CAP	ELECTP(Bkwh)	LABOR_FO	EMPLOYED	UNEMPLOY
Thailand	14565.35	173.3	39510000	39114900	0.01
Algeria	13974.96	48.05	12190000	11007570	0.097
Colombia	13751.5	59.22	23670000	21492360	0.092
Macedonia	13077.2	5.676	959300	690696	0.28
Dominican Republic	12950.01	13.09	4996000	4266584	0.146
China	12892.28	5169	801600000	768734400	0.041
South Africa	12726.46	257.9	20230000	15172500	0.25
Serbia	12585.01	37.65	3140000	2320460	0.261
Peru	12373.13	38.4	16550000	15292200	0.076
Ecuador	11469.34	22.85	7214000	6853300	0.05
Tunisia	11334.37	15.23	3950000	3349600	0.152
Maldives	10817.46	0.2652	159700	142133	0.11
Egypt	10684.01	152	28260000	24473160	0.134
Namibia	10663.08	1.331	1168000	847968	0.274
Albania	10121.22	6.987	1098000	951966	0.133
Indonesia	9976.81	172.7	124300000	117214900	0.057
Mongolia	9926.8	4.472	1037000	945744	0.088
Bosnia and Herzegovina	9847.29	12.93	1468000	817676	0.443
Sri Lanka	9844.25	11.8	8916000	8541528	0.042
Jordan	9826.84	17.26	1959000	1718043	0.123
Kosovo	9027.35	5.847	800000	552800	0.309
Paraguay	8531.28	57.05	3260000	3022020	0.073
Ukraine	8397.58	198.1	22110000	20164320	0.088
El Salvador	8288.08	5.992	2752000	2581376	0.062
Jamaica	8229.92	4.745	1311000	1132704	0.136
Guatemala	7956.3	8.361	4576000	4388384	0.041
Armenia	7937.49	7.075	1489000	1252249	0.159
Bhutan	7907.87	7.55	345800	335771.8	0.029
Morocco	7634.44	23.65	12000000	10848000	0.096
Guyana	7478.01	0.725	313100	278659	0.11
Georgia	6949.59	9.694	1959000	1667109	0.149
Philippines	6877.34	66.01	41680000	38679040	0.072
Timor-Leste	6793.84	0.1317	247500	201960	0.184
Bolivia	6516.13	7.375	4881000	4524687	0.073

Hierarchical Cluster H1 (n=33)

COUNTRY	GDP_CAP	ELECTP(Bkwh)	LABOR_FO	EMPLOYED	UNEMPLOY
Luxembourg	88820.38	2.211	253600	235594.4	0.071
Macau	87189.64	0.4137	367800	360811.8	0.019
Singapore	78456.64	47.95	3557000	3489417	0.019
Norway	65192.07	144.7	2724000	2631384	0.034
Hong Kong	56097.75	39.97	3871000	3750999	0.031
Switzerland	54753.67	65.54	5046000	4884528	0.032
United States	54330.09	4048	1.56E+08	1.46E+08	0.062
Australia	48349.49	239.7	12370000	11627800	0.06
Netherlands	47091.37	98.57	7893000	7324704	0.072
Ireland	45929.27	26.09	2174000	1928338	0.113
Canada	44985.97	644.1	19210000	17884510	0.069
Germany	44784.2	575.9	44760000	42522000	0.05
Austria	44648.06	69	3778000	3607990	0.045
Denmark	44557.89	29.85	2771000	2626908	0.052
Sweden	44298.82	162.9	5124000	4719204	0.079
Taiwan	43647	235	11500000	11063000	0.038
Iceland	41606.66	17.19	185900	177534.5	0.045
Belgium	41248.77	74.13	5225000	4780875	0.085
Finland	40442.42	67.69	2665000	2435810	0.086
France	38870.83	532	29870000	26972610	0.097
United Kingdom	37994.5	365.7	32590000	30732370	0.057
Japan	37874.35	963	65930000	63556520	0.036
Korea, South	36363.49	494.7	26430000	25557810	0.033
New Zealand	35756.18	42.91	2452000	2307332	0.059
Italy	33400.63	286.2	25510000	22321250	0.125
Israel	33332.03	64.44	3784000	3534256	0.066
Malta	32321.57	2.221	190500	179260.5	0.059
Spain	31861.33	276.5	22930000	17358010	0.243
Slovenia	30523.16	14.76	913500	789264	0.136
Czech Republic	28154.48	81.71	5416000	4988136	0.079
Estonia	27974.9	11.66	669800	612197.2	0.086
Slovakia	27529.71	26.42	2730000	2383290	0.127
Russia	25052	1054	75250000	71562750	0.049

Hierarchical Cluster H2 (n=39)

COUNTRY	GDP_CAP	ELECTP(Bkwh)	LABOR_FO	EMPLOYED	UNEMPLOY
Lithuania	27371.06	12.27	1454000	1292606	0.111
Portugal	25495.81	44.27	5271000	4522518	0.142
Malaysia	24474.13	118	14010000	13603710	0.029
Latvia	24457.58	5.891	1014000	917670	0.095
Poland	24412.51	150.9	18260000	15940980	0.127
Hungary	24238.34	34.28	4388000	4076452	0.071
Chile	23434.65	61.85	8514000	7960590	0.065
Kazakhstan	23164.46	90.53	9103000	8638747	0.051
Argentina	21352.98	123.2	17310000	15977130	0.077
Panama	21041.7	7.642	1563000	1492665	0.045
Uruguay	20880.38	10.16	1712000	1597296	0.067
Venezuela	18640.19	127.6	14340000	13221480	0.078
Belarus	17852.51	30.33	5000000	4950000	0.01
Romania	17838.72	56.71	9945000	9248850	0.07
Mexico	17603.55	277.6	52900000	50413700	0.047
Mauritius	17479.87	2.628	600200	552184	0.08
Azerbaijan	17217.44	22.99	4821000	4560666	0.054
Botswana	15402.81	0.35	1017000	835974	0.178
Brazil	15044.56	530.4	1.11E+08	1.05E+08	0.055
Costa Rica	14791.83	9.889	2257000	2065155	0.085
Thailand	14565.35	173.3	39510000	39114900	0.01
Colombia	13751.5	59.22	23670000	21492360	0.092
Dominican Republic	12950.01	13.09	4996000	4266584	0.146
China	12892.28	5169	8.02E+08	7.69E+08	0.041
Peru	12373.13	38.4	16550000	15292200	0.076
Ecuador	11469.34	22.85	7214000	6853300	0.05
Maldives	10817.46	0.2652	159700	142133	0.11

Hierarchical Cluster H2 (n=39) (Continued)

COUNTRY	GDP_CAP	ELECTP(Bkwh)	LABOR_FO	EMPLOYED	UNEMPLOY
Namibia	10663.08	1.331	1168000	847968	0.274
Indonesia	9976.81	172.7	1.24E+08	1.17E+08	0.057
Sri Lanka	9844.25	11.8	8916000	8541528	0.042
Paraguay	8531.28	57.05	3260000	3022020	0.073
Ukraine	8397.58	198.1	22110000	20164320	0.088
El Salvador	8288.08	5.992	2752000	2581376	0.062
Jamaica	8229.92	4.745	1311000	1132704	0.136
Armenia	7937.49	7.075	1489000	1252249	0.159
Bhutan	7907.87	7.55	345800	335771.8	0.029
Guyana	7478.01	0.725	313100	278659	0.11
Philippines	6877.34	66.01	41680000	38679040	0.072
Bolivia	6516.13	7.375	4881000	4524687	0.073

Hierarchical Cluster H3 (n=23)

Costa Rica	14791.83	9.889	2257000	2065155	0.085
Greece	26383.58	54.98	3910000	2862120	0.268
Cyprus	20972.13	4.443	356700	299984.7	0.159
Croatia	19552.75	14.24	1714000	1354060	0.21
Turkey	19039.4	228.1	27560000	24969360	0.094
Bulgaria	17156.23	46.65	2513000	2236570	0.11
Turkmenistan	15703.19	19.97	2300000	920000	0.6
Iran	15692.17	225.8	28400000	25474800	0.103
Montenegro	14679.95	2.557	251300	203301.7	0.191
Algeria	13974.96	48.05	12190000	11007570	0.097
Macedonia	13077.2	5.676	959300	690696	0.28
South Africa	12726.46	257.9	20230000	15172500	0.25
Serbia	12585.01	37.65	3140000	2320460	0.261
Tunisia	11334.37	15.23	3950000	3349600	0.152

Hierarchical Cluster H3 (n=23) (Continued)

Egypt	10684.01	152	28260000	24473160	0.134
Albania	10121.22	6.987	1098000	951966	0.133
Mongolia	9926.8	4.472	1037000	945744	0.088
Bosnia and Herzegovina	9847.29	12.93	1468000	817676	0.443
Jordan	9826.84	17.26	1959000	1718043	0.123
Kosovo	9027.35	5.847	800000	552800	0.309
Guatemala	7956.3	8.361	4576000	4388384	0.041
Morocco	7634.44	23.65	12000000	10848000	0.096
Georgia	6949.59	9.694	1959000	1667109	0.149
Timor-Leste	6793.84	0.1317	247500	201960	0.184

References

Anderberg, M. R., (1973), *Cluster Analysis for Applications*, Academic Press.

Arrow, K. J., H. B. Chenery, B. S. Minhas, and R. M. Solow, (1968), “Capital-Labor Substitution and Economic Efficiency,” in *Review of Economics and Statistics* (1961), 225-250, A. Zellner (ed.), Readings in Economic Statistics and Econometrics, Little, Brown and Co. 364-408.

Basu, S. and J. G. Fernald, (1997), “Returns to Scale in U.S. Production: Estimates and Implications,” *Journal of Political Economy*, 105(2), 249-283.

Beckman, M. J. and R. Sato, (1969), “Aggregate Production Functions and Types of Technical Progress: A Statistical Analysis,” *American Economic Review*, 59(1), 88-101.

CIA WorldFact Book: https://www.cia.gov/library/publications/the-world-factbook/print/textversion.html .

Cobb, C. W. and P. H. Douglas, (1928), “A Theory of Production,” *American Economic Review* 18(1), 139-165.

Fisher, F. M., (1993), “Aggregation: Aggregate Production Functions and Related Topics,” Collected Papers of Franklin M. Fisher, J. Monz (ed.), Cambridge: MIT Press.

Fisher, W. D., (1969), *Clustering and Aggregation in Economics*, Johns Hopkins Press, Baltimore.

Gordon, A. D., (1981), *Classification: Methods for the Exploratory Analysis of Multivariate Data*, Chapman and Hall.

Hartigan, J.A., (1975), *Clustering Algorithms*, John Wiley and Sons, New York.

Hirschberg, J. G., E. Maasoumi, and D. J. Slottje, (1991), “Cluster Analysis for

Measuring Welfare and Quality of Life Across Countries," *Journal of Econometrics*, 50(1-2), 131-150.

Kauffman, L. and P. J. Rousseuw, (1990), *Finding Groups in Data*, Wiley.

McMillin, W. D. and D. J. Smyth, (1994), "A Multivariate Time Series Analysis of the United States Aggregate Production Function," Empirical Economics, 19(4), 659-673.

Solow, R. M., (1968), "Technical Change and the Aggregate Production Function," Review of Economics and Statistics, (1957), 312-320, A. Zellner (ed.), Readings in Economic Statistics and Econometrics, Little, Brown and Co., 349-363.

Spence, M., (2011), *The Next Convergence: The Future of Economic Growth in a Multispeed World*, Picador, New York.

Strauss-Kahn, V., (2004), "The Role of Globalization in the Within-Industry Shift Away from Unskilled Workers in France" R. E. Baldwin and L. A. Winters (eds.), *Challenges to Globalization: Analyzing the Economics*, University of Chicago Press, 209-234.

Szilagyi, G. J., (1991), "Cluster Analysis in International Comparisons," *Journal of Official Statistics*, 7, 425- 435.

Ward, J. H., (1963), "Hierarchical Grouping to Optimize an Objective Function," *Journal of the American Statistical Association*, 58(301), 236-244.

Zellner, A., J. Kmenta, and J. Dreze, (1968), "Specification and Estimation of Cobb-Douglas Production Models," Econometrica, (1966), 784-795, A. Zellner (ed.), Readings in Economic Statistics and Econometrics, Little, Brown and Co., 323-335.