

MINING THE DIVERSITY OF INBOUND STUDENTS BASED ON TIME SERIES ANALYSIS

DIAN-FU CHANG AND WEN-CHING CHOU*

Department of Education and Futures Design
Tamkang University

No. 151, Yingzhuang Road, Tamsui District, New Taipei City 251301, Taiwan
140626@mail.tku.edu.tw; *Corresponding author: sauichou1226@gmail.com

Received July 2021; accepted September 2021

ABSTRACT. *Global student mobility has become an emerging issue in higher education development. Various countries assumed international students could be an important source of income and have a disproportionate influence on the economic and innovation systems. In this sense, this study aims to explore the trend and diversity of inbound students. Taking Taiwan as an example, we collected series of data from the Ministry of Education from 1956 to 2020. Becker's D and Blau index were used to transform the data with index formats to interpret the diverse phenomena and their patterns. This study conducts ARIMA to build related predicting models. The findings suggest the diversity of gender has changed fundamentally in this system. The female inbound students will grow into a critical mass in the future. Regarding the program diversity, the slight dropping of the Blau index may reflect the selection of major programs that could be more balanced in the future. The predicted values suggest the increase of inbound students will demonstrate a positive long-term trend in the system. The findings may provide useful information for related policy makers. The design of detecting diversity and trends for inbound students can be extended to explore similar issues in higher education settings.*

Keywords: Higher education, Higher education management, Inbound students, Internationalization, Time series analysis

1. Introduction. Global student mobility in higher education for academic, economic, or political purposes has provided a new map of the world. OECD defines an international mobile student as an individual who has physically crossed an international border between two countries with the objective to participate in educational activities in a destination country where it is different from his or her country of origin [1]. Traditionally, student mobility referred to two different types: inbound and outbound mobilities. Inbound mobile students mean students who move into a host country for the purpose of study or study-related activities. Outbound mobile students indicate the students who leave their country to another country for the purpose of study or traineeship in the context of study [2]. Inbound and outbound mobilities might be intertwined in the number of students. For example, Organization for Economic Co-operation and Development countries had more inbound mobility compared to outbound mobility in tertiary education. Specifically, 89% of OECD citizens study abroad in another OECD country, and 70% of international students in OECD countries come from a non-OECD country [3]. The number of international and foreign tertiary students has grown on average by 4.8% per year between 1998 and 2018 in OECD [4]. For host countries, international students may be an important source of income and have a disproportionate effect on their economic and innovation systems. For many countries, attracting international students to higher education has become an essential economic target under the World Trade Organization framework. For example, Australia, New Zealand, the United Kingdom, and the United

States have shared a large number of international students in the world. In 2018, the ratio of international or foreign students exceeds 10 : 1 in Australia, New Zealand, the United Kingdom, and the United States [4]. The other ambitious project in China has shown that China's higher education will develop towards the internationalization of higher levels. The Chinese government has made relevant policies to increase the number of inbound students [5]. A previous study suggests there is a brain drain issue in the higher education system based on the net flow ratio in China [6]. In the longer run, highly educated mobile students are likely to integrate into domestic labor markets, contributing to innovation and economic performance [4]. In this sense, there are various countries that have moved their higher education into the high participate stage in the global competitive framework. The issues of attracting inbound students remain being addressed in higher education settings. Regarding the trend analysis with innovative approaches is still limited, this study tried to integrate the diversity index and time series analysis to explore the phenomena.

Moreover, globalization is both external to education and a threat to local places, thus requiring a defensive response to the phenomena of global mobility [7], especially the global influent event, like War and COVID-19 pandemic which might cause travel constraints and impact the international studying or study intention of students. In 2020, higher education institutions around the world closed down to control the spread of the COVID-19 pandemic potentially affecting more than 3.9 million international and foreign students studying in OECD countries [8]. In the short term, the COVID-19 lockdown has severely affected educational systems around the world, especially international student exchange [9]. However, how long the impact of the pandemic in higher education will be? It is still unclear. Taking Taiwan as an example, this study explores the trends of inbound mobile students and projects their future trends. Previous studies indicated the decline of enrollment in higher education has caused oversupply issues in Taiwan [10,11]. Attracting more international students has become a crucial strategy in a declining higher education setting. The Taiwanese Ministry of Education has set an ambitious goal of attracting approximately 130,000-140,000 foreign students by 2020, which equates to 10% of all students in the higher education system [12,13]. Considering the students' origins, Southeast Asian countries have been targeted for attracting inbound mobile students. This study will focus on the trend of inbound students, especially the pattern transformation during higher education moving from mass to universal stages [14]. In addition, the raw data themselves are not easy to realize the transformation of inbound students. This study intends to utilize the related indexes to interpret the pattern transformation among the inbound students. The result will provide clear patterns of gender, program diversity, and inbound mobility of international students in the next decade in Taiwan. For these purposes, this study explores the following research questions.

- a) What is the trend of inbound students in the targeted higher education system?
- b) What are the gender diversity patterns of inbound students in the higher education system?
- c) Did they show diversities of inbound students participating in major programs with the trend analysis?
- d) What is the trend of inbound mobility ratio in the higher education system?

The rest part of the paper will be presented as follows: Firstly, we will address the method section, which includes the definition of the terms, evaluating the diversities and time series models; Secondly, the results will be displayed; Finally, we will suggest a fitted strategy for future development in higher education and the conclusion will be drawn.

2. Methods. This study collected inbound student data from the Ministry of Education [15]. The data sets include the total number of inbound students, art and humanities, social science, STEM, and others. We selected the inbound students enrolled in art and

humanities, social science, STEM programs from 1956 to 2020. The series data cover 65 periods. We transformed the data for Becker's D [16] and Blau index [17] to determine their diversities and conducted a time series analysis for their future development.

2.1. Definition of the terms. In this study, the technical terms include inbound students, inbound mobility ratio, and major programs.

- Inbound students refer to the full-time international students in the target country. The data were based on the data bank in the Ministry of Education, Taiwan.
- Diversity of gender refers to the differences of male and female inbound students participating in higher education.
- Major programs refer to art and humanity, social science, and STEM (science, technology, engineering, and mathematics) programs which did not include law, education, life science, medicine, and agriculture programs in this study. We will calculate the inbound students enrolled in the major programs respectively for evaluating their diversity.
- Inbound mobility ratio refers to the number of international students with the total tertiary enrollment in that country. The calculation will present as [18]: Inbound mobility ratio = $100 \times [\text{Total number of students from abroad studying in a given country (inbound students)}] / [\text{Total tertiary enrollment in that country}]$.

2.2. Becker's coefficient of discrimination (D). This study applied Becker's formula of D to representing the gender diversity of the inbound students [16]. The series data cover 65 periods from 1956 to 2020. The D is defined as follows:

$$D_i = (EM_i/EF_i) - 1$$

EM_i refers to the male international students enroll in higher education; EF_i refers to female international students enroll in higher education.

The positive calculated D implies the females' enrollment is less than that of males, whereas, the negative D means that the females are preferred to select the related programs in the target higher education system [16,19]. If D becomes zero or nearly zero, it means the diversity of males and females are equal. The calculated D will be used to represent the index of gender diversity in the higher education system.

2.3. Blau index. Blau index has been used to assess heterogeneity or diversity in wide settings. Diversity has been an important concept applied in various ways across fields like ecology, demography, sociology, economics, and psychology [17,20-24]. Blau index computes proper theoretical upper bound for Blau index as well as a normalized measure that allows researchers to get a measure that ranges from 0 to 1. This index reaches its minimum value of 0 when there is no variety. A high index value indicates a high degree of heterogeneity. The various categories can be standardized to conduct the index. The standardized Blau index is defined as follows [17,25]:

$$1 - \sum_{i=1}^k p_i^2$$

where p_i corresponds to the proportion of group members in the i th category and k denotes the number of categories.

This index quantifies the probability that two members randomly selected from a population will be in different categories if the population size is infinite or if the sampling is carried out with a replacement. This study defines the Blau index for tackling the inbound student diversity in specific programs of the higher education system. Blau_Programs refers to the Blau index for humanity, social science, and STEM programs according to the change of participation ratios.

2.4. Time series analysis. Time series analysis has been used for wide settings, for example, environment, medicine, and social-related issues [26-28]. Various studies discuss the innovative approach to transform the time series data [29,30]. This study applied time series analysis to the flow of international students in higher education. We conducted time series analysis to explore the patterns of D, Blau index, and inbound mobility ratio and build fitted models for projecting their future trends. Before the model selection, this study will verify the series data sets which belong to seasonal or non-seasonal data. Then, the fitted model will follow the criteria of ARIMA(p, d, q) selection. The p as the order of the autoregressive part in terms of the number of autoregressive terms (AR) while d as the difference; the q as the order of the moving average, i.e., the number of lagged forecast errors in the prediction equation (MA) [31-34]. The white noise examination will use the Box-Pierce Chi-square statistics test to determine whether the model met the assumptions that the residuals were independent [35,36]. This study verifies the robustness of the series with the fitted ARIMA model for the next decade.

3. Results. In this higher education system, the inbound students have shown increasing from 39 in 1956 to 64,268 in 2019. The average increasing ratio is 14% from 1956 to 2020. In 2019, the COVID-19 has caused the number of inbound students to drop significantly. There are 18% of inbound students dropped in 2020. The trend of the total number of inbound students in the higher education system has shown in Figure 1.

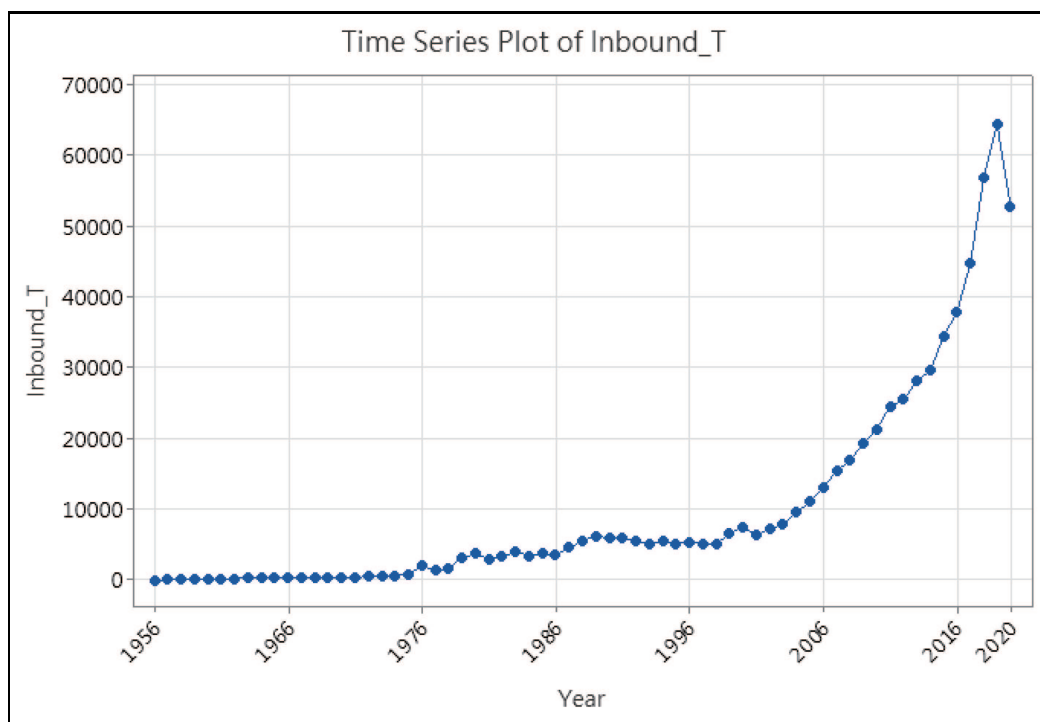


FIGURE 1. The trend of inbound students from 1956 to 2020

3.1. Trend of gender participation. Based on the number of inbound students, both males and females are similar with increasing patterns while the patterns of gender participation have shown changed. Before 1988, males are the critical mass in the inbound students. During 1999-2013, the male and female inbound students are equivalent, sometimes male-dominated, the other time female-dominated in the system. After 2014, the system becomes female inbound students dominated. Compared to the gender differences of inbound students, both are impacted by COVID-19 after 2019 with a significant dropping as seen in Figure 2.

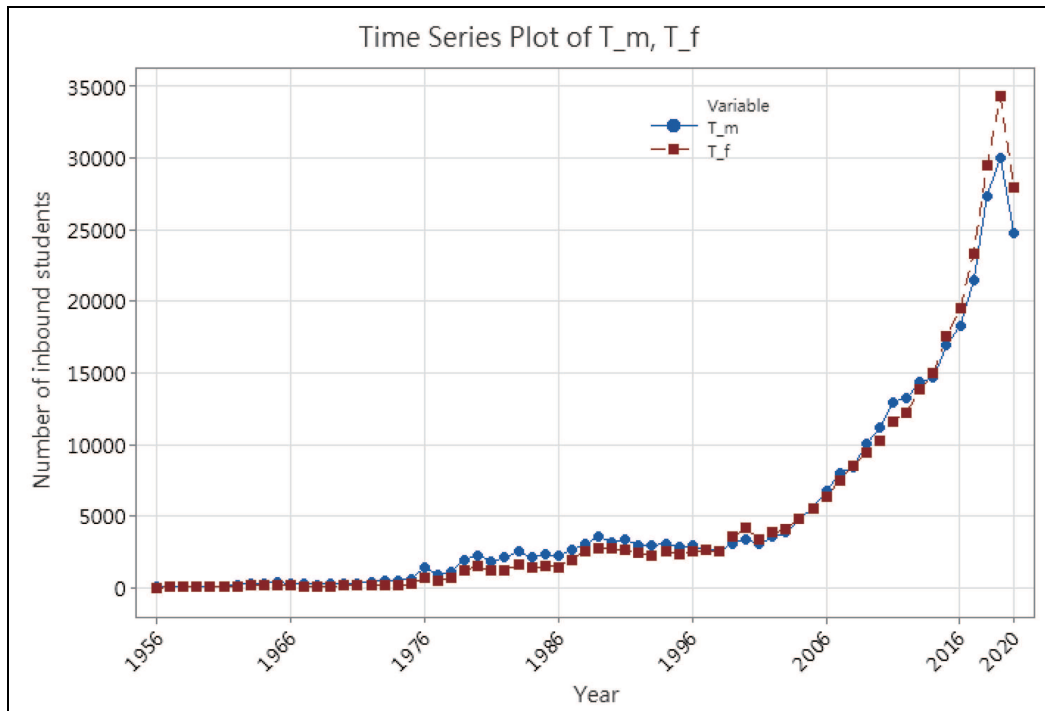


FIGURE 2. Changing of inbound students with gender participation

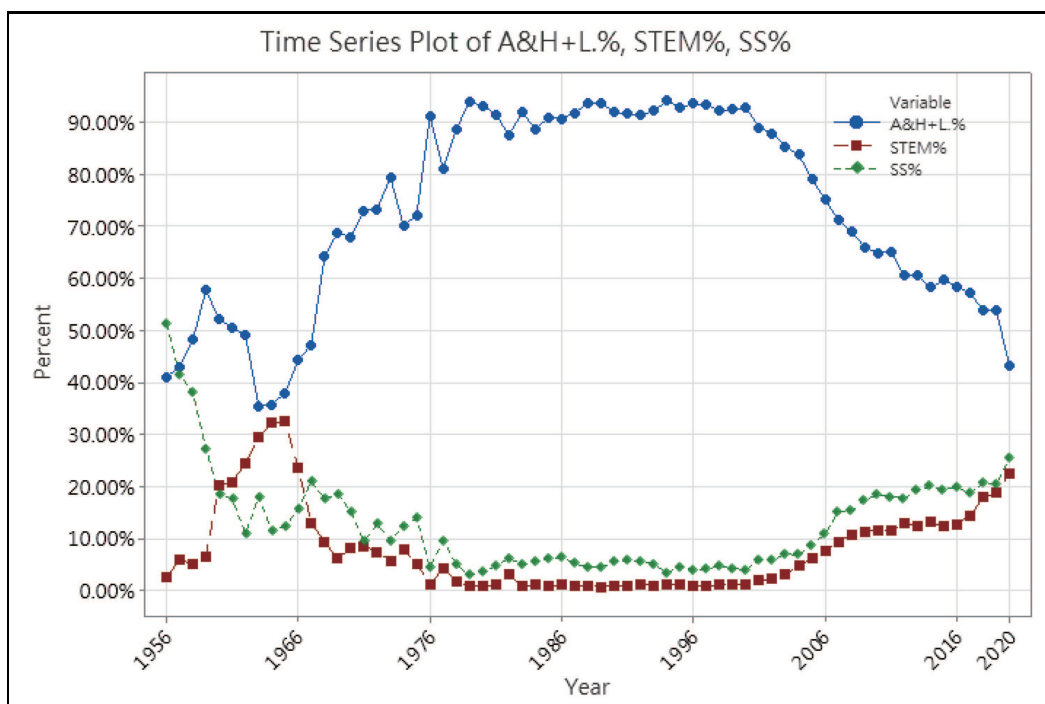


FIGURE 3. The percentage of inbound students in art & humanity, social science, and STEM

3.2. The trend of inbound students participating in major programs. The result reveals the art & humanity with a large share of inbound students. While the percentage of inbound students in art & humanity drops significantly in last two decades. Social science and STEM programs have shown increasing their inbound students in the last two decades. This phenomenon may reflect the government’s science-oriented policy can attract more international students to select their study journey in this higher education system. Basically, the transformed patterns may also reflect the effect of higher education

expansion during the last decades. Since 1999, the higher education system has moved into the universal stage in terms that the gross entrance ratio has reached 50% [14,19]. The expansion provides much more capacity for international students.

3.3. Diversity of gender. This study transformed the gender differences of inbound students with the D index. The D has been projected to next decade with the ARIMA model. The result reveals the ARIMA(1, 1, 1) is the fitted model to project D. The final estimates of parameters show $AR(1) = -0.997, p = .000$; $MA(1) = -0.827, p = .000$. The modified Box-Pierce Chi-square statistic for testing while noise shows the p -value in lag 12 = 0.413, in lag 24 = 0.660, in lag 36 = 0.970, and in lag 48 = 0.998. The error terms are not significant. It implies the errors are randomized. Based on the suggested model, the forecast D will decrease from -0.202 to -0.875 in the next decade, as seen in Table 1 and Figure 4. The result suggests the number of female inbound students in higher education might increase significantly in the future.

TABLE 1. Forecasts of D from 2021 to 2030

Period	Year	Forecast	95% limits	
			Lower	Upper
66	2021	-0.202	-0.88241	0.47794
67	2022	-0.266	-1.15026	0.61826
68	2023	-0.354	-1.46965	0.76106
69	2024	-0.418	-1.66864	0.83234
70	2025	-0.506	-1.92949	0.91677
71	2026	-0.570	-2.10179	0.96119
72	2027	-0.658	-2.33368	1.01683
73	2028	-0.722	-2.49080	1.04591
74	2029	-0.811	-2.70458	1.08361
75	2030	-0.875	-2.85162	1.10244

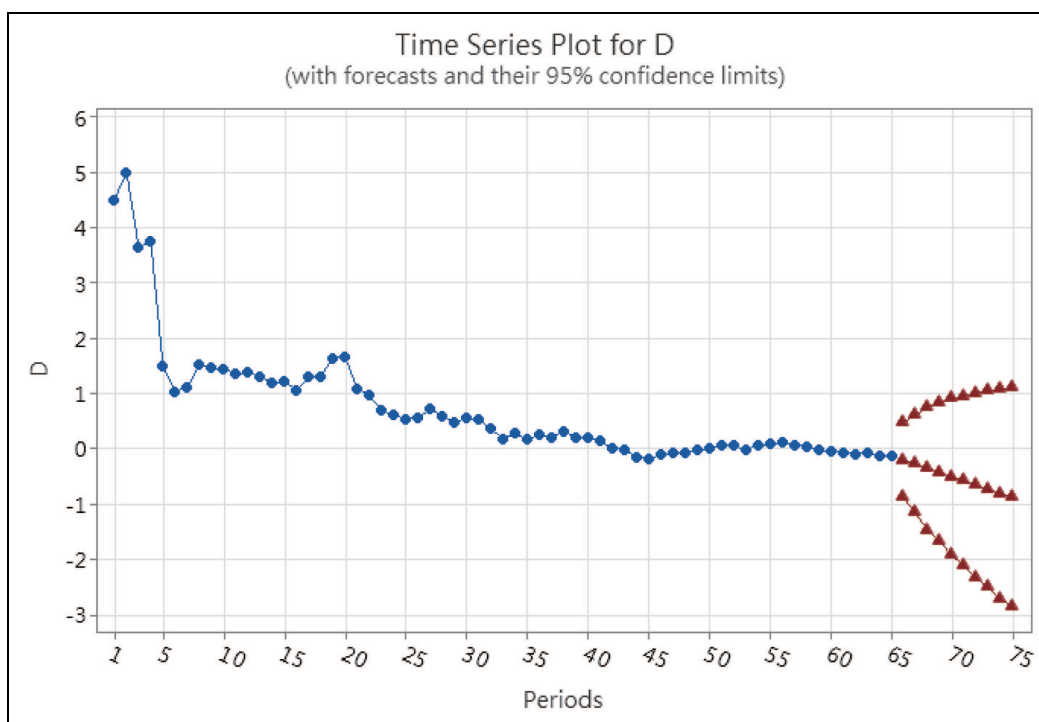


FIGURE 4. Forecasting D from 2021 (period 66) to 2030 (period 75)

3.4. Diversity of enrolling in major programs. Based on the calculation of the Blau index with the diversity of major programs, this study suggests ARIMA(0, 1, 1) could be the fitted model. After one difference, the series data of the Blau index show stationary. The final estimate of the parameter of MA(1) shows coefficient is -0.510 , $p = .001$. The modified Box-Pierce Chi-square statistic shows p -values are larger than .05. It suggests the error terms are randomized. Table 2 reveals the projected Blau index will drop from 0.333 in 2021 to 0.226 in 2030. The slight dropping of the Blau index may reflect the selection of major programs that could be more balanced in the future. Figure 5 displays the result of forecasting the trend of the Blau index with participating in major programs.

TABLE 2. Forecasts of Blau index for participation in major programs

Period	Year	Forecast	95% limits	
			Lower	Upper
66	2021	0.333	0.243174	0.422182
67	2022	0.296	0.183745	0.408727
68	2023	0.238	0.099490	0.377445
69	2024	0.237	0.031739	0.441589
70	2025	0.235	-0.019449	0.489172
71	2026	0.233	-0.062497	0.528614
72	2027	0.231	-0.100456	0.562967
73	2028	0.229	-0.134844	0.593749
74	2029	0.228	-0.166547	0.621847
75	2030	0.226	-0.196137	0.647832

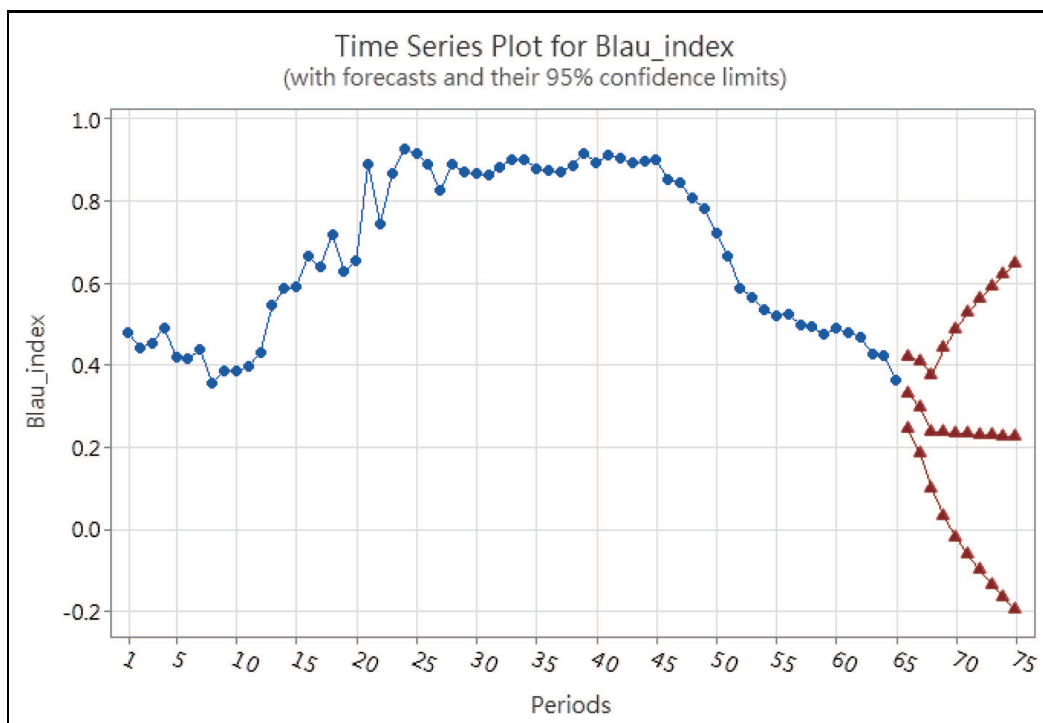


FIGURE 5. Blau index of participating in major programs

3.5. Future trend of inbound mobility ratio. The inbound mobility ratio refers to the proportion of inbound students in the total tertiary students. The result reveals ARIMA(0, 1, 3) could be the fittest model to project the inbound mobility ratio. After one difference, the suggested model shows the coefficient of MA(3) is 0.472 , $p = .000$. The

modified Box-Pierce Chi-square statistic shows p -values are larger than .05. It suggests the error terms are randomized. Based on the fittest ARIMA(0, 1, 3) model, the findings suggest the forecasts of inbound mobility ratio will increase from 0.0345 in 2021 to 0.0431 in 2030, as seen in Table 3. It implies the forecasts of inbound mobility ratio will increase to 4.31% in the future. The result reveals the impact of COVID-19 in 2019-2020 which results in the decline of the inbound students. The inbound mobility ratio has also shown significantly dropped in 2020. The trend of the inbound mobility ratio from 1956 to 2030 is displayed in Figure 6.

TABLE 3. Forecasts of inbound mobility ratio from 2021 to 2030

Period	Year	Forecast	95% limits	
			Lower	Upper
66	2021	0.0345	0.0300460	0.0388660
67	2022	0.0347	0.0257688	0.0437150
68	2023	0.0392	0.0274878	0.0508272
69	2024	0.0397	0.0268741	0.0525746
70	2025	0.0403	0.0263600	0.0542223
71	2026	0.0409	0.0259240	0.0557919
72	2027	0.0414	0.0255512	0.0572983
73	2028	0.0420	0.0252310	0.0587521
74	2029	0.0426	0.0249555	0.0601613
75	2030	0.0431	0.0247184	0.0615320

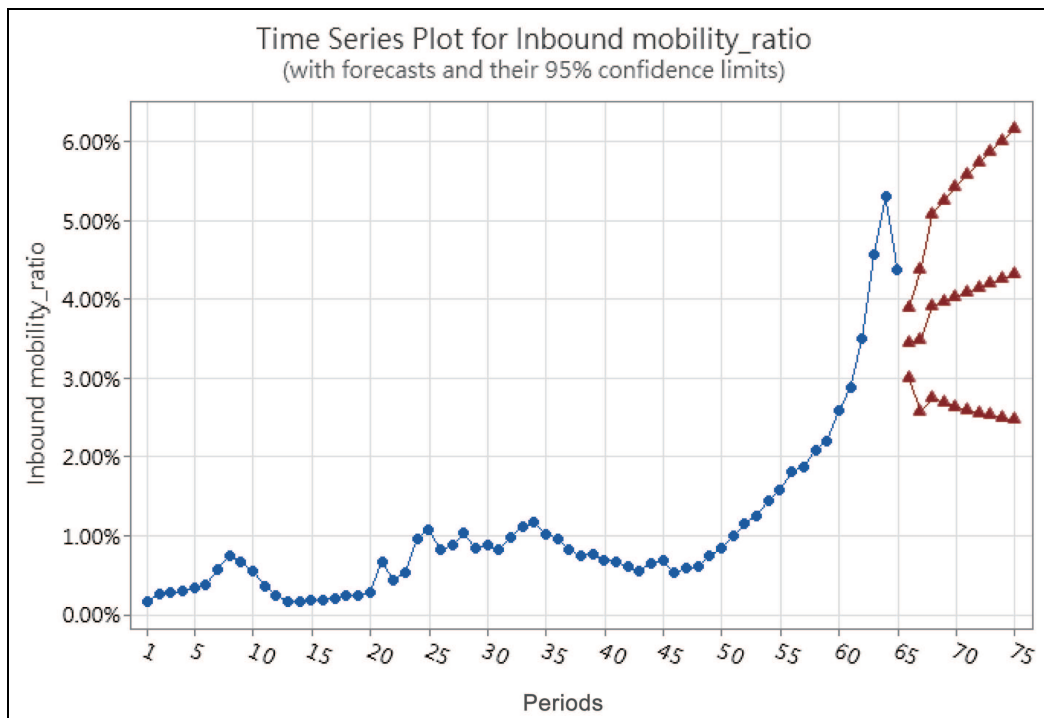


FIGURE 6. The trend of inbound mobility ratio from 1956 (period 1) to 2030 (period 75)

4. **Conclusion.** The study considered the diversity of gender and programs among the inbound students in Taiwan. The time series data has been transformed into an index format to conduct trend analyses. This study found the inbound students increase very quickly during the last decades. The impact of COVID-19 in 2019 resulted in the decline

of the inbound students significantly. However, we assume it is a short-term phenomenon only. The predicted values suggest the increase of inbound students will demonstrate a positive long-term trend in the system. Even though the inbound mobility ratio has shown to increase steadily and reaches 5%, the increasing inbound mobility ratio still did not fulfill the government's policy agenda reaching 10% in 2020. Moreover, the higher education system has attracted more female international students. The diversity of gender has displayed a fundamentally changed pattern in this system. The female inbound students will become the critical mass in the future. Regarding the program diversity, this study found the slight dropping of the Blau index may reflect the selection of major programs that could be more balanced in the future. This study provides a design with index transformation and ARIMA model building strategies to mine the issues among inbound students. We found the Becker's D and Blau index are useful to detect the diversity in a specific topic. The time series can widen the explanation to a long-term trend. The design of detecting diversity and trends for inbound students can be extended to explore similar issues in higher education settings. For further studies, we encourage the data collected can be extended to outbound studies, which can interpret the net flow of students in the target higher education. In addition, the design of the study can be used for international comparison, for example, Korea, Japan, or other countries, if the data are available.

Acknowledgment. This work is supported by Ministry of Science and Technology (MOST 108-2410-H-032-020-SS2), Taiwan. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

REFERENCES

- [1] OECD, Definitions and classifications of the OECD international education statistics, in *OECD Handbook for Internationally Comparative Education Statistics: Concepts, Standards, Definitions and Classifications*, Paris, OECD Publishing, 2017.
- [2] U. Teichler, Student mobility in Europe: The informational value of official statistics and graduate surveys, in *European Higher Education at the Crossroads: Between the Bologna Process and National Reforms*, A. Curaj, P. Scott, L. Vlasceanu and L. Wilson (eds.), Dordrecht, Netherland, Springer, 2012.
- [3] OECD, *Education at a Glance: OECD Indicators*, Paris, France, 2015.
- [4] OECD, *Education at a Glance: OECD Indicators*, OECD Publishing, Paris, DOI: 10.1787/69096873-en, 2020.
- [5] W. Wen, D. Hu and J. Hao, International students' experiences in China: Does the planned reverse mobility work?, *International Journal of Educational Development*, vol.61, pp.204-212, 2018.
- [6] F. Xia and D.-F. Chang, Forecasting student mobility flows in higher education: A case study in China, *ICIC Express Letters, Part B: Applications*, vol.12, no.6, pp.525-532, 2021.
- [7] A. Chang, D.-F. Chang and W. S. Zhu, Reviewing implementation of global mobility policy for youth from the perspectives of stakeholders, *Educational Policy Forum*, vol.23, no.3, 2020.
- [8] UNESCO, *COVID-19 Educational Disruption and Response*, <https://en.unesco.org/covid19/educationresponse>, 2020.
- [9] S.-J. Wu, D.-F. Chang and F.-R. Sun, Exploring college student's perspectives on global mobility during the COVID-19 pandemic recovery, *Education Sciences*, vol.10, no.9, DOI: 10.3390/educsci10090218, 2020.
- [10] D.-F. Chang and Y.-L. Huang, Detecting the effect of policy intervention for oversupply higher education system, *ICIC Express Letters, Part B: Applications*, vol.8, no.11, pp.1489-1495, 2017.
- [11] S.-J. Wu, D.-F. Chang and H. Hu, Detecting the issue of higher education over-expanded under declining enrollment times, *Higher Education Policy*, DOI: 10.1057/s41307-019-00163-z, 2019.
- [12] MOE, *Higher Education in Taiwan 2012-2013*, <http://www.edu.tw>, 2012.
- [13] MOE, *Toward World-class Universities Plan*, http://140.113.40.88/edutop/modules/catalog1/education_top_about/upload/20111130192552.pdf, 2012.
- [14] M. Trow, Problems in the transition from elite to mass higher education, *ERIC*, ED 091 983, <http://files.eric.ed.gov/fulltext/ED091983.pdf>, 1973.

- [15] MOE, *Educational Statistics 2021 (International Students Section)*, <https://depart.moe.edu.tw/ED4500/cp.aspx?n=1B58E0B736635285&s=D04C74553DB60CAD>, 2021.
- [16] G. S. Becker, *The Economics of Discrimination*, 2nd Edition, The University of Chicago Press, Chicago, IL, 1971.
- [17] P. M. Blau, *Inequality and Heterogeneity: A Primitive Theory of Social Structure*, Free Press, New York, NY, 1977.
- [18] O. Labe, Key indicators on tertiary education: Calculation and interpretation, *UIS Workshop on Education Statistics Windhoek*, UNESCO, 2010.
- [19] D.-F. Chang, Effects of higher education expansion on gender parity: A 65-year trajectory in Taiwan, *Higher Education*, vol.76, no.3, pp.449-466, 2018.
- [20] A. R. Solow and S. Polasky, Measuring biological diversity, *Environmental and Ecological Statistics*, vol.1, pp.95-103, 1994.
- [21] L. H. Pelled, Demographic diversity, conflict and work group outcomes: An intervening process theory, *Organization Science*, vol.7, pp.615-631, 1996.
- [22] S. Lieberman, Measuring population diversity, *American Sociological Review*, vol.34, pp.850-862, 1969.
- [23] J. C. Dissart, Regional economic diversity and regional economic stability: Research results and agenda, *International Regional Science Review*, vol.26, pp.423-446, 2003.
- [24] N. E. Betz and L. F. Fitzgerald, Individuality and diversity: Theory and research in counseling psychology, *Annual Review of Psychology*, vol.44, pp.343-381, 1993.
- [25] A. Solanas, R. M. Selvam, J. Navarro and D. Leiva, *Some Common Indexes of Group Diversity: Upper Boundaries*, <http://diposit.ub.edu/dspace/bitstream/2445/45784/1/623311.pdf>, 2021.
- [26] H. Qi, et al., COVID-19 transmission in Mainland China is associated with temperature and humidity: A time-series analysis, *Science of the Total Environment*, vol.728, 2020.
- [27] F. J. Elgar, A. Stefaniak and M. J. A. Wohl, The trouble with trust: Time-series analysis of social capital, income inequality, and COVID-19 deaths in 84 countries, *Social Science & Medicine*, vol.263, DOI: 10.1016/j.socscimed.2020.113365, 2020.
- [28] R. Salgotra, M. Gandomi and A. H. Gandomi, Time series analysis and forecast of the COVID-19 pandemic in India using genetic programming, *Chaos, Solitons & Fractals*, vol.138, DOI: 10.1016/j.chaos.2020.109945, 2020.
- [29] D. Shan, W. Lu and J. Yang, The data-driven fuzzy cognitive map model and its application to prediction of time series, *International Journal of Innovative Computing, Information and Control*, vol.14, no.5, pp.1583-1602, 2018.
- [30] I. Fernandez et al., NATSA: A near-data processing accelerator for time series analysis, *IEEE 38th International Conference on Computer Design (ICCD)*, pp.120-129, DOI: 10.1109/ICCD50377.2020.00035, 2020.
- [31] G. E. P. Box and G. M. Jenkins, *Times Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, CA, 1976.
- [32] S. Bisgard and M. Kulachi, *Time Series Analysis and Forecasting by Example*, Wiley & Sons, New York, NY, 2011.
- [33] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*, Springer, New York, NY, 2016.
- [34] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, OTexts, Melbourne, Australia, 2013.
- [35] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, 5th Edition, Wiley, New York, NY, 2015.
- [36] G. M. Ljung and G. E. P. Box, On a measure of lack of fit in time series models, *Biometrika*, vol.65, no.2, pp.297-303, 1978.