ICIC Express Letters
Part B: Applications
Volume **9**, Number **12**, December **2018**
 ISSN 2185-2766
pp. **1241–1248**

ESTIMATING THE GAPS BETWEEN POTENTIAL ENROLLMENT NUMBERS AND KINDERGARTEN CAPACITIES BY USING TIME SERIES MODELS

Dian-Fu Chang[1], Huei-Ting Huang[2,*] and Hsiao-Chi ChangTzeng[3]

[1]Graduate Institute of Educational Policy and Leadership
[2]Doctoral Program of Educational Leadership and Technology Management
Tamkang University
No. 151, Yingzhuan Road, Tamsui District, New Taipei City 25137, Taiwan
140626@mail.tku.edu.tw; *Corresponding author: a.tingting.523@gmail.com

[3]Extension of Education
University of California
1156, High Street, Santa Cruz 95064, USA
hcachang@ucdavis.edu

Received May 2018; accepted August 2018

Abstract. *This study selected Taiwan as a target to tackle the issue of gap between potential enrollment numbers and kindergarten capacities (KC). The purposes of this study are to realize the effect of 0-year babies (ZYB) declining existing in Taiwan, to project the future development of ZYB and to estimate the gap between ZYB and KC in the system. This study selects the ARIMA model to verify the robustness of forecasting ZYB. Employing the projecting data of ZYB, this study calculates the gaps between ZYB and estimated KC in next decade. Based on the fittest ARIMA model and the gap estimation, this study found the one-year model of KC is oversupply and the gaps will increase steadily. The gaps between demands and supplies in two-year model of KC are diminishing in next decade. This study provides a practical method to detect whether it is an issue of oversupply in the kindergarten system. Both predict and estimate models provide useful knowledge to reinvent management strategies to ameliorate the issue of the system.*
Keywords: ARIMA, Kindergarten capacity, Management strategy, Newborn babies, Time series analysis

1. **Introduction.** The total fertility in Taiwan has begun to decline since 1950. As the trend is moving ahead, in 1984, it fell below 2.1 children and dropped to 1.05 in 2008. In 2010, it even dropped to 0.89, which made Taiwan's birth rate the lowest in the world [1]. At this stage, early childhood education is sure to be affected by the declining. Because of the low birth rate, the newborn babies have dropped from 410,000 in 1981, to 270,000 in 1998 to 167,000 in 2010, which has achieved the lowest level in the last 50 years [2]. According to the data from Ministry of Interior, the newborn babies are decreasing from 1974 (328,461) to 2016 (196,973) which means 40% drop [3]. Previous studies or government document only provide the facts but neglect the impact and possible solutions. This study employs the concept of data mining to tackle the declining birthrate issues and find new solutions.

Recently, providing adequacy kindergarten capacity has become an ambiguous management strategy in Taiwan. Under the free-market policy, it has shown that the government's engagement in this field is not enough for a long time. Obviously, when the demand is getting strong in this market, the private new kindergartens spring up like mushrooms in past decades. The public kindergartens are still limited. In 2017, there are 6,323 kindergartens

DOI: 10.24507/icicelb.09.12.1241

in Taiwan, while only 2041 public kindergartens. It implies 67.73% of kindergartens are private [4]. Over 69% of kids enroll in private sector and pay more tuitions [4,5]. Based on the notion of public good, the government initiated subside policy in preschool level in 2000 and expanded the free kindergarten policy for economic disadvantaged kids in 2011 [6,7]. This policy will promote 19,000 economic disadvantaged kids to enroll the private kindergartens. Totally, the 5-year kids enroll in kindergartens has increased to 94.5%, and the enrollment of economic disadvantaged kids has reached 95.35% [7].

Unfortunately, the oversupply issues are emerging under the declining of birthrate. For example, the numbers of closed kindergartens in private sector are increasing and up 70 in 2017 [4]. However, the government announced a new public kindergarten policy which implies the government will expand the capacity of kindergartens by opening 1,000 new classes to accept another 30,000 kids from 2017 to 2020 [8]. Moreover, the government has announced NT$30 billion per year investment plan to expand the capacity of pubic or non-profit kindergartens in next few years [9]. Under current political environment, the government has implemented the expanded policy in public or non-profit kindergartens. The teachers and managers of private sector are worry about the negative impact of declining of birthrate. Therefore, how to balance the demand and supply has become a crucial issue in this turning point. Reviewing related studies in kindergarten area, this study found the previous studies or related reports did not cover the real gap issues between potential enrollment numbers and kindergarten capacities. Furthermore, the previous related studies are lack of predicted function to foresee the future development in this issue. Therefore, this study selected Taiwan as the target to tackle the issue of gap between potential enrollment and kindergarten capacity (KC) with autoregressive integrated moving average (ARIMA) model.

The purposes of this study are as follows: a) to realize the effect of 0-year babies (ZYB) declining existing in Taiwan; b) to project the future development of ZYB in the system; c) to estimate the gap between ZYB and the KC. Given these purposes, the structure of this paper is as follows: Section 1 addresses the current issues of newborn babies (NBB) declining. Section 2 provides a brief description of the method related to data transformation and forecasting. Section 3 verifies the robustness of predictive method and its results for estimating the gaps. Finally, the conclusions with suggestions are presented in Section 4.

2. **Design of Research.** This study employs the concept of data mining. First, define the data set in the data warehouse (government's data set) to fit the time series analysis. Second, select the tool to transform the data. Third, select the fittest model to interpret the data. Fourth, verify the robustness of forecasting model to estimate. Finally, explain the results and interpret the meanings. The process of mining the data set is displayed in Figure 1.

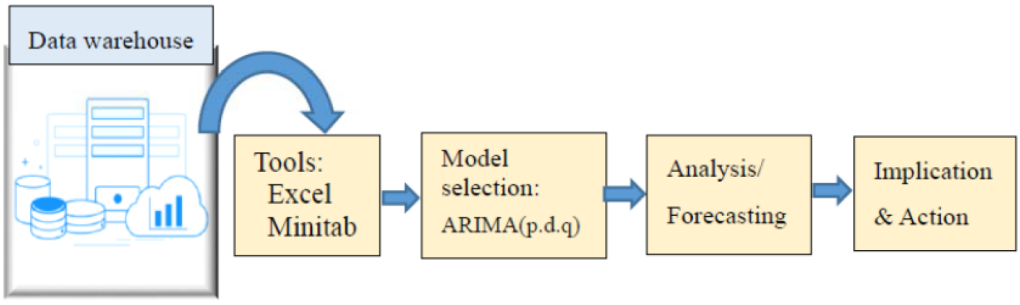

FIGURE 1. Research framework

2.1. **Defining the data set.** First, this study selects the newborn baby series data from Ministry of Interior, Taiwan [3]. This study found the series of the NBB and ZYB are different. There is a gap between both data sets, see Figure 2. It is about 19,731 babies per year having the gap between the two series. For better interpretation, this study decides to apply the ZYB data to predict their future trend.

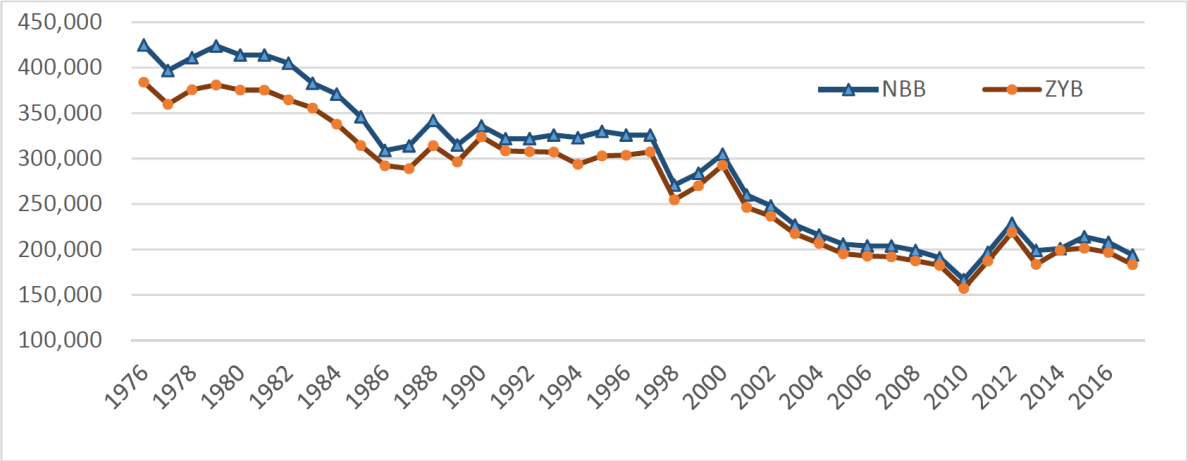

FIGURE 2. The gaps between newborn babies (NBB) and 0-year babies (ZYB)

Kindergarten capacities (KC) in terms of the number of kindergarten kids, the data come from Ministry of Education, Taiwan [4,5]. In 2012, the policy leads to integrate the data of kindergartens and day care centers. This study calculates KC by using total preschool kids minus the kids in day care centers after 2012. This estimation might have little difference from the really enrolled kids in the kindergartens. After the scope of data warehouse is defined, this study selects ARIMA models to estimate the trend for further interpretation.

2.2. **Building an ARIMA model.** The Box-Jenkins method was proposed by George Box and Jenkins in their seminal 1970 textbook [10]. Box and Jenkins have introduced a stochastic difference equation which is able to model stochastic disturbances. This equation has been used to forecast one-step-ahead disturbances, according to the data characteristics of stationary or non-stationary. The stochastic equation is expressed in the form of ARIMA. The ARIMA(p,d,q) model indicates p as the order of the autoregressive part, d as the amount of difference and q as the order of the moving average part [11-14]. Box-Jenkins' model refers to the process as a stochastic model building and it is an iterative approach that consists of the following 3 steps [15]:

a) Identification: Use the data and all related information to help select a sub-class of model that may best summarize the data;

b) Estimation: Use the data to train the parameters of the model (i.e., the coefficients);

c) Diagnostic checking: Evaluate the fitted model in the context of the available data and check where the model may be improved.

This study follows the process of ARIMA model. First, detecting the series is seasonal or non-seasonal. Second, the fittest ARIMA(p,d,q) model was selected by using difference, ACF and PACF [16,17]. Third, using the fitted ARIMA model predicts the trend of ZYB for next ten years.

2.3. **Building the GAP model.** This study conducts the gap estimation with ZYB and KC data. In this first stage, the gaps are calculated by using the formula: $\text{ZYB}_i - \text{KC}_i$. The data are modified by a reasonable way to fit the formula. Then, $\text{GAP}_i = \text{ZYB}_i - \text{KC}_i$, $i = 1$ to 42 periods in terms of the data from 1976 to 2017. In the second stage, this

study employs the projecting data of ZYB and estimates the data of KC to calculate the GAPs in next decade (from 2018-2027).

3. Results.

3.1. Projecting the 0-year babies.

This study found the ARIMA(2,2,2) model is better than ARIMA(1,2,1), ARIMA(1,2,2), and ARIMA(2,2,1) comparing their coefficients and residual terms. The decision making follows the criteria of AR, MA and Ljungd-Box, see Table 1.

Table 1. Comparison of ARIMA models

ARIMA models	AR	MA	Ljung-Box	Decision making
ARIMA(1,2,1)	X	•	★	
ARIMA(1,2,2)	X	★	X	
ARIMA(2,2,2)	★	★	•/★	◎
ARIMA(2,2,1)	★	•/★	X	

Note. • = excellent, ★ = fair, X = unacceptable.

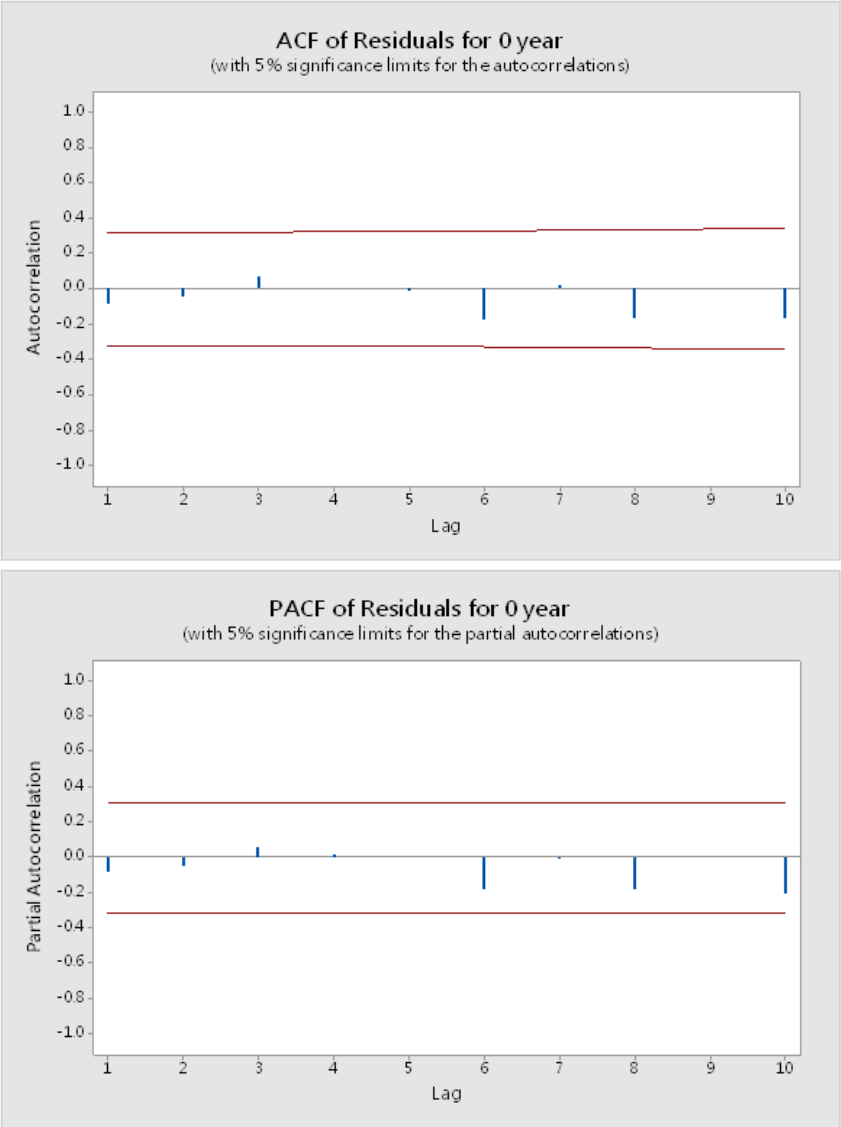

Figure 3. Plots of ACF and PACF for series of ZYB

3.2. **Verification of ARIMA(2,2,2) model.** Tests for ACF (autocorrelation) and PAC F (partial autocorrelation) indicated that the model ARIMA(2,2,2) can be used to predict the series of ZYB, see Figure 3. Evidences also exist to support that the residuals follow a white noise process and the ARIMA(2,2,2) is a robust representation of the observed time series of ZYB, see Figure 4.

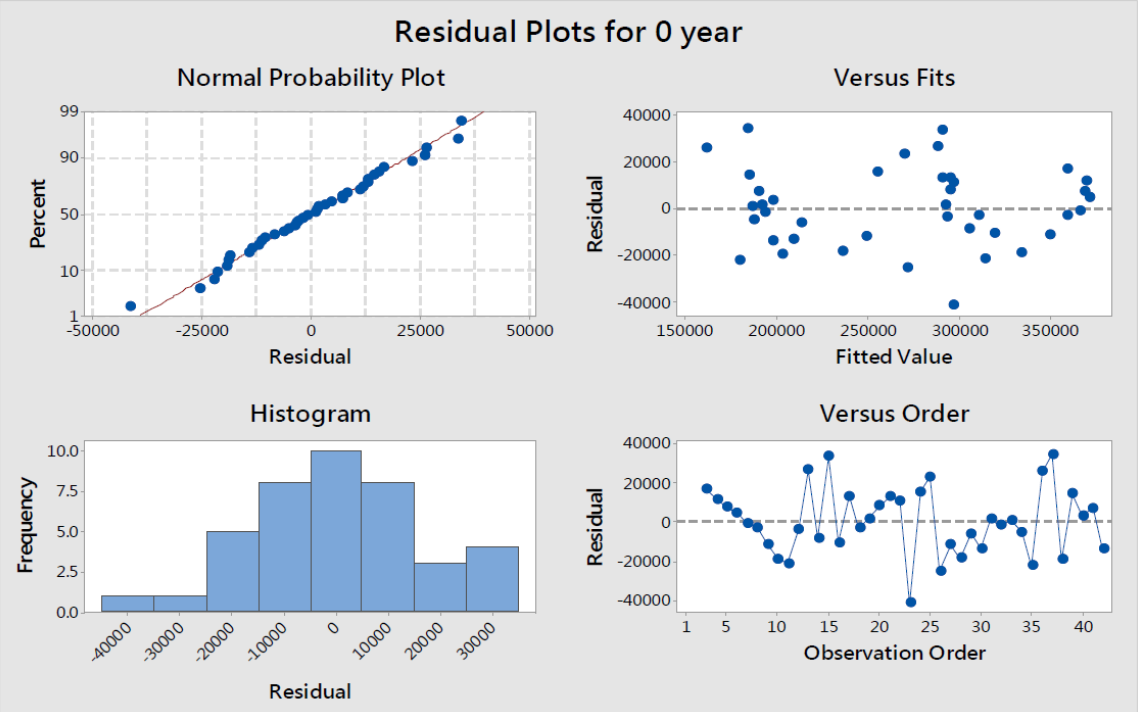

FIGURE 4. The versus fits for ZYB

Based on ARIMA models with 2 times differences, this study found the ARIMA(2,2,2) is a better fit model. This study also considered the Box-Pierce Chi square statistics to check the residual, it appears to be low and diminish as the number of lags increases [11-17]. In the ARIMA(2,2,2) model, the final estimates of parameters are presented in Table 2. The projection of ZYB in next ten years ahead has shown in Figure 5. The numbers of ZYB in next ten years in terms of 2018-2027 will decrease as the forecasting, see Table 3.

TABLE 2. Estimates of parameters for ZYB

Type	Coef	SE Coef	T-Value	P-Value
AR 1	−0.936	0.272	−3.44	0.002
AR 2	−0.257	0.169	−1.52	0.138
MA 1	0.252	0.289	0.87	0.390
MA 2	0.888	0.323	2.75	0.009
Constant	129.7	20.0	6.50	0.000

Modified Box-Pierce Chi-Square Statistic

Lag	12	24	36	48
Chi-Square	14.76	27.03	34.10	*
DF	7	19	31	*
P-Value	0.039	0.104	0.321	*

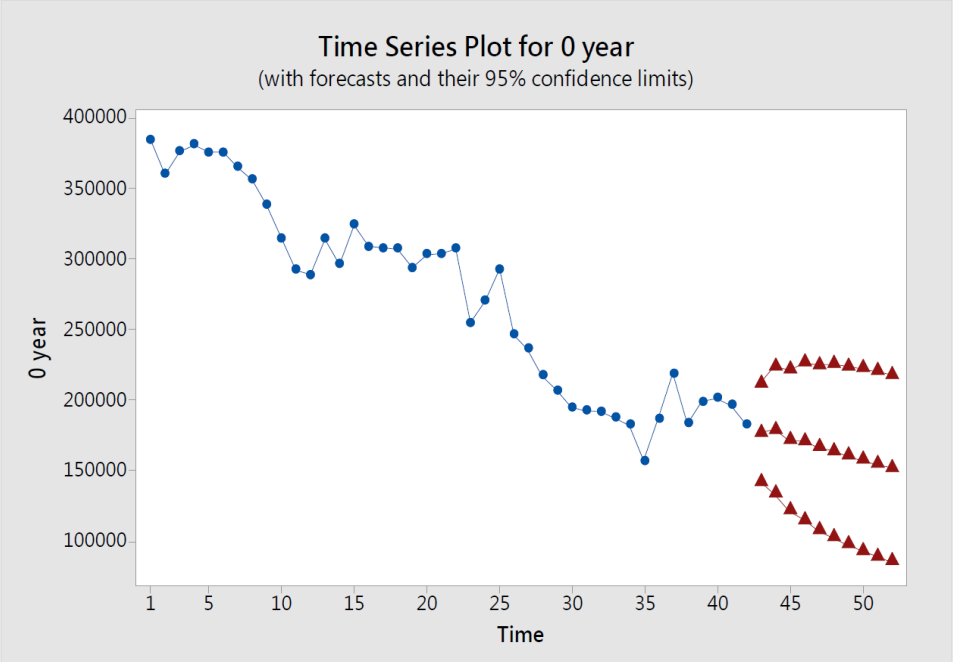

FIGURE 5. Projecting ZYB ten years ahead based on ARIMA(2,2,2)

TABLE 3. Forecasts from period 43 (2018)

Period	Year	Forecast	95% Limits	
			Lower	Upper
43	2018	177,268	142,263	212,273
44	2019	179,132	134,024	224,241
45	2020	171,743	122,102	221,384
46	2021	171,080	115,608	226,551
47	2022	166,627	108,282	224,973
48	2023	164,124	102,942	225,306
49	2024	160,899	97,919	223,878
50	2025	157,979	93,625	222,332
51	2026	155,088	89,828	220,348
52	2027	152,221	86,392	218,050

3.3. **Estimating the gap of ZYB and KC.** This study calculates the one-year stay model and two-year model following the KC in 2017 and the forecasting data of ZYB in next decade. Under declining the newborn babies, will the kindergartens become an oversupply system? To answer the question, this study carries the one-year enrollment and two-year enrollment models in current kindergarten system. The result reveals the previous one-year enrollment model is oversupply, while the gaps between demand and supply in two-year enrollment model are diminishing in next decade. The results have been presented in Table 4 and Figure 6.

4. **Conclusions.** This study provides a practical method to detect the issues of oversupply in kindergartens in Taiwan. Following the process of ARIMA model, this study detects whether the series is seasonal or non-seasonal, selecting ARIMA(p,d,q) model by using difference and visualization of ACF and PACF, and verifying the robustness of ZYB with ARIMA model for next ten years. The ARIMA model is a useful way to tackle the series issues.

TABLE 4. Comparing the gaps between ZYB and KC with different models

Year	ZYB	KC	Gap of one-year model	Gap of two-year model
2018	177,268	228,385	51,117	−126, 151
2019	179,132	228,385	49,253	−129, 879
2020	171,743	228,385	56,642	−115, 101
2021	171,080	228,385	57,305	−113, 775
2022	166,627	228,385	61,758	−104, 869
2023	164,124	228,385	64,261	−99, 863
2024	160,899	228,385	67,486	−93, 413
2025	157,979	228,385	70,406	−87, 573
2026	155,088	228,385	73,297	−81, 791
2027	152,221	228,385	76,164	−76, 057

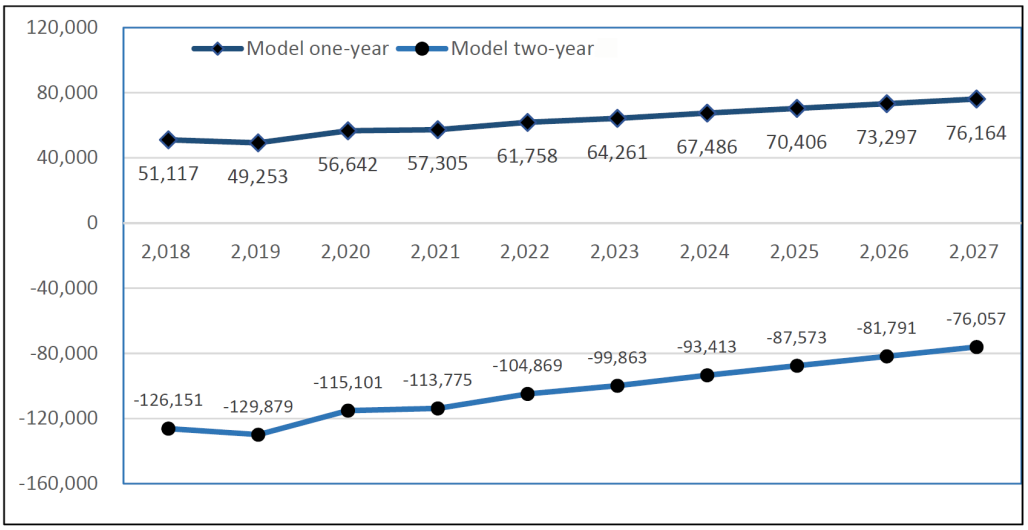

FIGURE 6. Projecting the gaps of model one and model two

For the purpose of realizing the effect of ZYB, this study found the linkages of ZYB and kindergarten capacities are tight. The declining ZYB will impact the enrollment of kindergartens directly in the system.

For the second purpose, what is the trend of ZYB in the system? Following the process of ARIMA model, this study found the result of projecting the declining numbers of ZYB has provided a clear picture to realize the demand and supply issue in the kindergarten system.

The third purpose is to estimate the gap between ZYB and the KC. This study calculates the one-year model and two-year model following the KC in 2017 and the forecasting data of ZYB in next decade. This study found the one-year model of KC is oversupply and gaps will increase steadily. The gaps between demand and supply in two-year model of KC are diminishing in next decade.

Based on the result of study, the suggestions for policy makers are as follows.

a) It needs to reconsider the long-term investment in the extended kindergarten capacities. The declining of ZYB displays that the demand of kindergartens will drop in next few years. The long-term investment should be balanced to the real demands.
b) The substantial funds for economic disadvantaged kids are welcomed, but the related subsidy policy as the government report should consider the feasibility of government's budget.

c) How to transfer the private kindergartens into non-profit kindergartens and how many percent of private kindergartens should be transformed? It should be listed in the future policy agenda.

Finally, the study found both predict and estimate models will provide useful knowledge to build new management strategies to ameliorate the issue of the system. For further studies, this study suggests extending more useful variables in the models to avoid the system overloading or oversupply. Furthermore, this study may provide a workable way to tackle similar issues in different settings.

REFERENCES

[1] Taiwan Ministry of the Interior, *Fertility Rates of Childbearing Age Women*, Department of Statistics Year Report, 2017.
[2] D.-S. Chen, *Higher Education in Taiwan: The Crisis of Rapid Expansion*, 2010.
[3] Department of Statistics, Ministry of Interior, *The Main Directory of Dynamic Query Statistics*, 2018.
[4] Department of Statistics, Ministry of Education, *Summary of Preschool – By Public or Private*, 2018.
[5] Ministry of Education, *The Republic of China Educational Statistics 2017*, Taipei, 2018.
[6] National Academy for Educational Research, *The Republic of China Education Yearbook 2016*, New Taipei City, 2017.
[7] MOE, *Mid-Term Policy Implementation (2013-2016).*
[8] K-12 Education Administration, Ministry of Education, *Expanded the Capacity of Pre-School Services*, 2018.
[9] UDN, *Is It Useful to Spend NT$30 Billion on Expanding Public Ownership of Child-Care and Early-Education Services Per Year?*, 2018.
[10] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control*, Holden-Day, San Francisco, 1970.
[11] S. Bisgard and M. Kulachi, *Time Series Analysis and Forecasting by Example*, Wiley & Sons, New York, 2011.
[12] R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, OTexts, Melbourne, Australia, 2013.
[13] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*, Springer-Verlag, 1991.
[14] P. J. Brockwell and R. A. Davis, *Introduction to Time Series and Forecasting*, Springer, Netherlands, 2016.
[15] G. E. P. Box, G. M. Jenkins, G. C. Reinsel and G. M. Ljung, *Time Series Analysis: Forecasting and Control*, 5th Edition, Wiley, New York, 2016.
[16] R. Davis, T. Coole and D. Osipyw, The application of time series modelling and Monte Carlo simulation: Forecasting volatile inventory requirements, *Applied Mathematics*, no.5, pp.1152-1168, 2014.
[17] P. R. Junior, F. L. R. Salomon and E. de O. Pamplona, ARIMA: An applied time series forecasting model for the Bovespa stock index, *Applied Mathematics*, no.5, pp.3383-3391, 2014.