



TAX DECLARATION RATES VIA AUDITS: A PREDICTION USING MARKOV MODEL

DOI: 10.17261/Pressacademia.2017.499

JBEF- V.6-ISS.2-2017(5)-p.97-111

Ersin Kiral¹, Can Mavruk²

¹Çukurova University, Department of Econometrics, Adana, 01330, Turkey. ekiral@cu.edu.tr

²Ömer Halis Demir University, School of Social Sciences, Niğde, 51100, Turkey. can.mavruk@ohu.edu.tr

To cite this document

Kiral, E., C. Mavruk (2017). Tax declaration rates via audits: a prediction using markov model. Journal of Business, Economics and Finance (JBEF), V.6, Iss.2, p. 97-111.

Permant link to this document: <http://doi.org/10.17261/Pressacademia.2017.499>

Copyright: Published by PressAcademia and limited licenced re-use rights only.

ABSTRACT

Purpose- This article predicts tax declaration rates by Markov chain model.

Methodology- Four Markov models are constructed for the declaration rates of three tax revenues. Declaration probabilities for the year 2017 are estimated by constructing probability matrices of transitions between classes described for every model. Declaration rates are predicted by the product of the initial probability matrix and transition probability matrix. Limiting matrices of predictions are found. The best Markov model was found by estimating the sum of mean square errors for every model.

Findings- Main results of this study are i) transitions of tax declarations are declining in higher states and improving in lower states, ii) in its best model corporate tax declaration rate is predicted to be between 47% and 64.5% in 2017 with a probability of 78% and would be stable in the same interval at a probability of 60% in 2038, iii) in the long run income tax would decrease 20,3 billion with a probability of 42% and Value Added Tax would decrease 26,2 billion with a probability of 40% and iv) expected declaration rates of income tax, corporate tax and value added tax in 2017 are 52.3%, 61.6% and 62.8%, respectively.

Conclusion- Income tax and Value Added Tax payoffs may substantially decrease from 2015 to 2038 and 2023 respectively. This may cause a revenue deficiency to Turkish Revenue Administration. Therefore Income tax and Value Added Tax audits should be increased. Even though tax revenues increase over time, the declaration rates show a decreasing to stationary or increasing to stationary behavior.

Keywords : Tax declaration rates, tax auditing, tax administration, transition probabilities, Markov analysis

JEL Codes: C51, H21, M42

1. INTRODUCTION

There is no study in the literature about tax declaration rates in connection with Markov modelling. Markov modelling of time series data does not provide sharp predictions. However it certainly provides accurate predictions and stationary probabilities for finite states. In a similar study thirteen tax revenues are predicted using Markov model in the paper by Kiral and Mavruk (2016).

The tax declaration is a decision under uncertainty Allingham and Sandmo (1972). The uncertainty of taxpayers decision on tax declaration has a great importance for tax administration revenues. Therefore tax declaration rates are important indicators of tax revenues. Every year tax administration audits a number of taxpayers and as a result of the audits a base difference is found between actual income and declared income when tax is underdeclared or evaded. Tax administration can determine evaded tax only on a small sample of taxpayers being audited. Therefore an increase in audit rate would always increase declared income. In this study tax declaration rates are calculated and predicted for three tax items such as income tax, corporate tax and value added tax since approximately 64% of tax revenues in 2015 are from these three taxes.

The objective of this paper is to find the best model predicting tax declaration rates using four Markov Models for each tax. The best of the four has the least sum of the mean square errors. Predictions of tax declaration rates are expected to be stationary and to have a limiting matrix. The rest of this study is organized as follows. In section 2 related literature is given. In section 3 methods for constructing transition probabilities, Markov models and predictive probability matrices, and for finding better model of Markov models and statistical significance of the models are given. In section 4 data of this study

along with description of the symbols is provided. In section 5 calculations for the three taxes are given. In section 6 a comparison of 2017 predictions of the three taxes, in section 7 a better model for each tax item, in section 8 a comparison of initial probability matrices and limiting matrices, and in section 9 statistical significance of the models are given. In section 10 results of this study are given and discussed, and the study ends with conclusion remarks in section 11.

2. LITERATURE REVIEW

Allingham and Sandmo (1972) analyzes income tax evasion by theoretical analysis approach. They analyze taxpayers decision under uncertainty on evading tax in static and dynamic cases. In static case they find the conditions for maximizing taxpayer utility function. They reach the first result that there exists an interior maximum solution. They concluded that an increase in audit rate and fine rate would always increase declared income, and an increase in tax rate would increase tax evasion. Baasch et. al (2010) used Markov models to quantify transitions between successional stages. They presented a solution for converting multivariate ecological time series into transition matrices and demonstrate the applicability of this approach for a data set that resulted from monitoring the succession of sandy dry grassland in a post-mining landscape. They analyzed five transition matrices, four one-step matrices referring to specific periods of transition (1995–1998, 1998–2001, 2001–2004, 2004–2007), and one matrix for the whole study period (stationary model, 1995–2007).

Büyüktatlı et. al (2013) used initial allocations of investment program with actual spending percentages from the years of 1998-2009 of Turkish Atomic Energy Institute (TAEK) to predict annual allowances from Ministry of Energy and Natural Resources. An estimated percentage of realization of investment program for 2011 and results are interpreted with Markov analysis. Cavers and Vasudevan (2015) directed graph representation of a Markov chain model to study global earthquake sequencing leads to a time series of state-to-state transition probabilities that includes the spatio-temporally linked recurrent events in the recordbreaking sense. A state refers to a configuration comprised of zones with either the occurrence or non-occurrence of an earthquake in each zone in a pre-determined time interval.

Grimshaw and Alexander (2011) used a Markov chain model to forecast outstanding balance of loans in each delinquency state. For that they used a markov chain X_n as the delinquency state of a loan in month n and a Markov Chain model for loan accounts that are 'current' this month having a probability of moving next month into 'current', 'delinquent' or 'paid-off' states. They forecasted 'one month ahead' portfolio delinquency balance for a portfolio of loans where each loan is n_i months from origination this month $i=1, \dots, N$. Lazri et.al (2015) adopted a Markovian approach to discern the probabilistic behaviour of the time series of the drought. A transition probability matrix was constructed from drought distribution maps. Markov transition probability formula for four states and a simulation model with an initial probability vector was used to calculate the drought distribution area in the future.

Lukić et. al. (2013) used the stochastic method based on a Markov chain model to predict the annual precipitation in the territory of South Serbia for the period 2009-2013. For this purpose, the precipitation data rainfall recorded on the four synoptic stations were used for the period 1980-2010. Kiral and Mavruk (2016) used one step Markov chain model to predict tax revenues in Turkey for the period 2000-2014. Four Markov models were used for each tax and best predictive model was found by the sum of mean square errors. Usher (1979) discussed that complex non-random or Markovian processes are likely to characterize ecological successions, the transition probability matrix elements not being constant but being functions either of the abundance, or of the rate of change of abundance, of a recipient class.

3. DATA AND METHODOLOGY

3.1. Methodology

A stochastic process $X = \{X_n; n \geq 0\}$ on a countable set S is a Markov Chain if, for any $i, j \in S$ and $n \geq 0$,

$$P\{X_{n+1} = j | X_0, \dots, X_n\} = P\{X_{n+1} = j | X_n\} \quad (1)$$

$$P\{X_{n+1} = j | X_n = i\} = p_{ij}. \quad (2)$$

The p_{ij} is the probability that the Markov chain jumps from state i to state j . These transition probabilities satisfy

$$\sum_{j \in S} p_{ij} = 1 \quad \forall i \in S$$

and the matrix $P = (p_{ij})$ is the transition matrix of the chain. Condition (1), called the Markov property, says that, at any time n , the next state X_{n+1} is conditionally independent of the past X_0, \dots, X_{n-1} given the present state X_n . In other words, the next state is dependent on the past and present only through the present state. Condition (2) simply says the transition probabilities do not depend on the time parameter n ; the Markov chain is therefore "time-homogeneous". (Serfozo, 2009, p.2). Four Markov models are constructed for each tax and best one for each tax is found. For the four models raw data of tax declaration rates are categorized into five, four, three and two categories by a grouped frequency distribution for each

model. Transition probability matrix is constructed and multiplied by the initial probability matrix to predict next year's probability matrix. Continuing on this procedure stationary probability matrix is found.

3.1.1. Construction of Transition Probabilities

Transition probability matrices are estimated for 2000-2015 for income tax, corporate tax and value added tax declaration rates. The estimator of the transition probabilities is the relative frequency of the actual transitions from phase i to phase j , i.e. the observed transitions have to be divided by the sum of the transitions to all other phases (Lipták, 2011, p.141).

In this paper,

$$p_{ij} = n_{ij} / \sum_j n_{ij} \text{ for } i, j = A, B, C, D, E$$

where n_{ij} is the number of observed transitions from i to j , $\sum_j n_{ij}$ is the sum of observed transitions from i to j .

Frequency distribution of the declaration rate intervals must be mutually exclusive (nonoverlapping) and class width must be equal for each interval (Bluman, 2014, p.45-46). Transition probabilities from X_i to X_j , $i, j = 0, 1, 2, \dots, m$, can be constructed at time n as the following matrix $0, 1, 2, \dots, m$, can be constructed at time n as the following matrix

$$P_n = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1m} \\ P_{21} & P_{22} & \dots & P_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ P_{m1} & P_{m2} & \dots & P_{mm} \end{bmatrix} \quad (3)$$

3.1.2. Prediction

Declaration rates can be estimated for every year in the period of the study and can be predicted for the years after the last year in the period of the study. Given that data at time n is in state X_0 and that the data will be in one of states $X_n \in \{1, 2, \dots, m\}$ at time n , then the data at time $n+1$ can be predicted. Given that transition probability matrix up to the year n is P_n and initial probability matrix for the year n is Q_n for $n = 0, 1, 2, \dots$, next year's probability distribution matrix can be estimated by

$$Q_{n+1} = Q_n P_n \quad (4)$$

Initial probability matrices for four Markov models are $1 \times j$ row matrices. For after the last year in the period declaration rates can be predicted by

$$\hat{Q}_{n+1} = \hat{Q}_n P_n \quad (5)$$

Prediction matrices have a limiting (stationary) matrix \hat{Q} , which can be written as

$$\lim_{n \rightarrow \infty} \hat{Q}_{n+1} = \hat{Q}. \quad (6)$$

3.1.3. Best of Four Markov Models

For every year of the sample and for every Markov model, mean square error (mse) is calculated by

$$\frac{1}{m} \sum_{i=1}^m (r_i - \hat{r}_i)^2 \quad (7)$$

where m : the number of states and $m = 2, 3, 4, 5$, r_i : realized declaration rate in state i in the year $n+1$ and

\hat{r}_i : estimated declaration rate in state i in the year $n+1$.

Estimated matrix at time $n+1$ can be found by

$$Q_{n+1} = Q_n P_n = [\hat{r}_1 \quad \hat{r}_2 \quad \hat{r}_3 \quad \dots \quad \hat{r}_m]$$

where $Q_n = [r_1 \quad r_2 \quad \dots \quad r_m]$: the initial probability matrix at time n for m states.

The sum of all mse (smse) for each model is calculated by

$$\sum_{n+2}^k \frac{1}{m} \sum_{i=1}^m (r_i - \hat{r}_i)^2 \quad (8)$$

where n is the first year and k is the last year in the data.

The least sum MSE gives the best Markov model.

3.1.4. Statistical Significance of the Models

Variations between observed and expected frequencies can be tested by constructing a contingency table of frequency distribution of transitions between the states at 0.05 significance level with a degree of freedom.

To validate Markov model, for every year, the value of the χ^2 statistic is computed based on the null hypothesis, H_0 : model is valid. At 0.05 level of significance and with the degrees of freedom, the χ^2 critical value and χ^2 test value are estimated. The null hypothesis is not rejected whenever χ^2 test value is less than the critical value. Test values are calculated by

$$\chi^2 = \sum_i (r_i - \hat{r}_i)^2 / \hat{r}_i \quad (9)$$

where i : the number of categories and $i=1, \dots, m$, r_i : actual probability, \hat{r}_i : expected probability

3.2. Data

According to Turkish Revenue Administration (TRA) Activity Report 2015, total tax revenue is 464,886,790,000 ₺, income tax collected is 105,393,946,000 ₺, corporate tax collected is 37,009,483,000 ₺, collected Value Added Tax (VAT) Included is 79,188,287,000 ₺ and collected VAT on imports is 74,387,089,000 ₺. (http://www.gib.gov.tr/fileadmin/user_upload/VI/GBG/Tablo_22.xls.htm accessed 08.08.2016). This data shows that the total of the three taxes is approximately 64% of the total tax revenue. Raw data of this study is extracted from inspection results in TRA Activity Reports for the years 2000-2015 (<http://www.gib.gov.tr/kurumsal/stratejik-yonetim/faaliyet-raporlari>). Data is available as inspected tax base and base difference. Inspected tax base is the declared tax base or declared income. The quotient of declared income and actual income gives tax declaration rates in Table 2. Missing data for the years 2002 and 2010 are completed by regression in SPSS missing value analysis. Inspection results for the years 2000-2015 are used to construct Tables 3, 4 and 5 for tax declaration rates of income tax, corporate tax and VAT respectively. 2015 inspection results in Table 2 are used to find 2015 tax declaration rates (X/B) in Tables 3, 4 and 5.

3.2.1. Income Tax Declaration Rates

Income tax base inspected (declared amount) and base difference (difference between actual income and declared tax base) and declaration rates for the years 2000-2015 are given in Table 3. In the last sixteen years the highest rate in income tax declaration was 98.9% in 2011 and the lowest realized was 4.2% in 2006. Significant ups and downs are observed in the rate data over the given period.

3.2.2. Corporate Tax Declaration Rates

Corporate tax income declarations, base difference and declaration rates for years 2000-2015 are given in Table 4. In the last sixteen years the highest declared tax rate was 99.6% in 2011 and the lowest was 11.8% in 2007. Corporate tax declaration rates has significant ups and downs over the given period.

3.2.3. Value Added Tax Declaration Rates

VAT declarations, base difference and declaration rates for years 2000-2015 are given in Table 5. In the last sixteen years the highest declared tax rate was 96.5% in 2011 and the lowest was 23.7% in 2009. VAT declaration rates has significant jumps and falls over the given period.

Table 1 shows the descriptions of some symbols used in this article.

Table 1: Description of Symbols

Symbols	Description
M1	Markov model 1
M2	Markov model 2
M3	Markov model 3
M4	Markov model 4
Code	A state of declaration rate
A	Highest state of a declaration rate
P_n	Transition probability matrix for the

	year n
SY	Stationary year of tax declaration
df	Degree of freedom
\hat{Q}	Stationary probability matrix
Q_n	Initial probability matrix
$\chi^2_{0,05}$	Ki-square test value at 0.05 level of significance

Table 2: 2015 Tax Audit Results According To Turkish Tax Administration Audits

Tax Type	# of Tax Payers Audited	Declared Income X (000₺)	Base Difference B-X (000₺)	Tax Difference (000₺)	Calculated Tax (000₺)	Rate Declared X/B
Income	1,779	52,051	48,157	11,232	13,153	0.519
Corporate	491	38,046	27,709	5,190	5,529	0.579
VAT	2,104	1,229,173	464,869	41,539	57,313	0.726
Special	38	3,353	754	290	359	0.816
Bank Ins.	13	851	529	11	34	0.617
Stamp	44	59,354	110,599	1,871	1,888	0.349
Advance	1,066	32,711	81,930	8,874	9,069	0.285
Others	4,367	120,152	157,605	4,861	4,966	0.433
Total	9,902	1,535,689	892,152	73,868	92,310	0.6325

Source: TRA, Activity Reports, <http://www.gib.gov.tr/kurumsal/stratejik-yonetim/faaliyet-raporlari> (5.1.2016) (TRA 2000-2015).

Table 3: Income Tax Declaration Rates

Year	Declared Income X (₺) (1)	Base Difference B - X (₺) (2)	Declaration Rate (%) (1) / (1)+(2)
2000	21,448,826	131,855,163	14
2001	16,029,993	112,399,579	12.5
2002	16,029,993*	131,855,163*	10.8
2003	53,664,347	98,700,897	35.2
2004	95,385,634	117,625,705	44.8
2005	82,221,251	187,773,866	30.5
2006	36,155,770	834,488,383	4.2
2007	73,393,391	553,781,491	11.7
2008	165,504,453	432,502,457	27.7
2009	52,452,224	217,014,074	19.5
2010	16,029,993*	18,476,344*	46.5
2011	1,586,197,618	18,430,305	98.9
2012	96,143,976	25,582,141	79
2013	19,523,920	18,476,344	51.4
2014	34,742,733	26,826,589	56.4
2015	52,050,730	48,157,004	51.9

Source: Own Elaboration, * Estimated by SPSS Missing Value Analysis- Regression

Table 4: Corporate Tax Declaration Rates

Year	Declared Income X (₺)	Base Difference B – X (₺)	Declaration Rate (%)
2000	29,443,990	170,079,633	14.8
2001	71,369,019	144,037,269	33.1
2002	71,369,019*	312,024,421*	18.6
2003	91,985,678	312,024,421	29.5
2004	827,247,790	1,300,020,818	38.9
2005	516,441,597	1,292,700,622	28.5
2006	381,244,134	877,826,392	30.2
2007	222,611,118	1,658,821,068	11.8
2008	474,464,700	2,141,819,149	18.1
2009	2,246,755,090	5,217,618,792	30.1
2010	9,888,342*	10,625,315*	51.2
2011	4,251,799,781	18,600,315	99.6
2012	15,694,708	10,625,304	59.6
2013	9,888,342	9,435,351	51.2
2014	39,222,657	12,231,106	76.2
2015	38,045,708	27,709,318	57.9

Source: Own Elaboration, * Estimated by SPSS Missing Value Analysis- Regression

Table 5: VAT Declaration Rates

Year	Declared Income X (₺)	Base Difference B-X (₺)	Declaration Rate (%)
2000	243,428,422	105,167,009	69.8
2001	127,302,400	172,138,684	42.5
2002	243,428,422*	334,385,008*	42.1
2003	662,220,809	246,897,979	72.8
2004	1,385,551,886	1,530,445,020	47.5
2005	1,049,336,007	1,482,289,093	41.4
2006	1,746,829,450	507,541,737	77.5
2007	1,786,847,392	668,431,089	72.8
2008	399,139,277	481,729,811	45.3
2009	2,218,678,268	7,132,560,056	23.7
2010	1,385,551,886*	246,897,979*	84.9
2011	9,136,386,241	334,385,008	96.5
2012	3,746,260,371	274,356,674	93.2
2013	800,703,515	221,229,335	78.4
2014	2,246,873,937	255,122,508	89.8
2015	1,229,173,026	464,869,057	72.6

Source: Own Elaboration, * Estimated by SPSS Missing Value Analysis – Regression

4. FINDINGS AND DISCUSSIONS

Transition probability matrices with respect to Markov models, predictions and stationarity are calculated for the three taxes.

4.1. Income Tax

Income tax declaration rates from largest to smallest are classified as A, B, C, D, E in model 1, as A, B, C, D in model 2, as A, B, C in model 3 and as A and B in model 4. Income tax declaration rates, codes and transitions for all Markov models are shown in Table 7 for the years 2000-2015.

4.1.1. Markov Models and Transition Probability Matrices

Classification of income tax declaration rates and transition probability matrices are given in Table 6. This table shows that declaration rates are over 52% in three categories of model 1, in two categories of model 2, in two categories of model 3 and in one category of model 4.

4.1.2. Prediction

Given that 2015 income tax declaration rate which is 51.9% is in state C and that income tax will be in one of states A, B, C, D or E in 2016, income tax declaration rates for 2017 and later years are predicted by (5) and the predictions are given in Table 8.

4.1.3. Stationarity

Limiting matrices are estimated in Excel by formula (6) for all models. According to four models, all probabilities become stationary in 2038, 2017, 2039 and 2036 respectively.

4.1.4. Statistical Significance of The Model

In model 1 of income tax, variations between observed and expected frequencies can be tested by constructing a contingency table of frequency distribution of transitions between the states at 0.05 significance level with 16 df. Since chi square test value 23.87 is less than critical value 26.296, H_0 is not rejected. This shows that there is no significant variations. The values in paranthesis in the Table 9 are expected frequencies which are found from (row sum x column sum)/total. Table 9 shows that transitions in higher realization states are declining and those in lower states rates are improving in model 1.

4.1.5. Technology

All calculations in this study are done using Excel and TI-36XPro. Prediction matrices and limiting matrices are estimated using Excel. The sum of mean square errors, classes of Markov models and all other simple calculations are done using TI-36XPro.

Table 6: Income Tax Markov Models and Transition Probability Matrices

Model 1		Transition Matrix					Model 2		Transition Matrix			
Classes (%)		A	B	C	D	E	Classes (%)	A	B	C	D	
$80.2 \leq r$	A	0	1	0	0	0	$75.3 \leq r$	A	1/2	0	1/2	0
$61.2 \leq r \leq 80.1$	B	0	0	1	0	0	$51.6 \leq r \leq 75.2$	B	0	1	0	0
$42.2 \leq r \leq 61.1$	C	1/4	0	1/2	1/4	0	$27.9 \leq r \leq 51.5$	C	1/5	1/5	2/5	1/5
$23.2 \leq r \leq 42.1$	D	0	0	1/3	0	2/3	$r \leq 27.8$	D	0	0	2/7	5/7
$r \leq 23.1$	E	0	0	1/6	1/3	1/2						
Model 3		Transition Matrix			Model 4		Transition Matrix					
Classes (%)		A	B	C	Classes (%)	A	B					
$67.4 \leq r$	A	1/2	1/2	0	$51.6 \leq r$	A	3/4	1/4				
$35.8 \leq r \leq 67.3$	B	1/4	1/2	1/4	$r \leq 51.5$	B	1/11	10/11				
$r \leq 35.7$	C	0	2/9	7/9								

Source: Own Elaboration

Table 7: Classification of Income Tax Declaration Rates Based On Inspections By Audit Officers and Transitions Between States

Year	Declaration Rate X/B (%)	M1 Code	M1	M2 Code	M2	M3 Code	M3	M4 Code	M4
2000	14	E		D		C		B	
2001	12.5	E	EE	D	DD	C	CC	B	BB
2002	10.8	E	EE	D	DD	C	CC	B	BB
2003	35.2	D	ED	C	DC	C	CC	B	BB
2004	44.8	C	DC	C	CC	B	CB	B	BB
2005	30.5	D	CD	C	CC	C	BC	B	BB
2006	4.2	E	DE	D	CD	C	CC	B	BB
2007	11.7	E	EE	D	DD	C	CC	B	BB
2008	27.7	D	ED	D	DD	C	CC	B	BB
2009	19.5	E	DE	D	DD	C	CC	B	BB
2010	46.5	C	EC	C	DC	B	CB	B	BB
2011	98.9	A	CA	A	CA	A	BA	A	BA
2012	79	B	AB	A	AA	A	AA	A	AA
2013	51.4	C	BC	C	AC	B	AB	B	AB
2014	56.4	C	CC	B	CB	B	BB	A	AA
2015	51.9	C	CC	B	BB	B	BB	A	AA

Source: Own Elaboration

Table 8: Income Tax Declaration Rates Predictions for 2017

Declaration Interval (%)	M1	Declaration Interval (%)	M2	Declaration Interval (%)	M3	Declaration Interval (%)	M4
$r \leq 23.1$	0.17	$r < 27.8$	0	$r \leq 35.7$	0.32	$r \leq 51.5$	0.52
$23.2 \leq r \leq 42.1$	0.125	$27.9 \leq r \leq 51.5$	0	$35.8 \leq r \leq 67.3$	0.43	$51.6 \leq r$	0.48
$42.2 \leq r \leq 61.1$	0.33	$51.6 \leq r \leq 75.2$	1	$67.4 \leq r$	0.25		
$61.2 \leq r \leq 80.1$	0.25	$75.3 \leq r$	0				
$80.2 \leq r$	0.125						

Source: Own Elaboration

Table 9: Contingency Table of Observed and Expected Income Tax Declaration Rates of Model 1

	A	B	C	D	E	Total
A	0 (0.07)	1 (0.07)	0 (0.33)	0 (0.33)	0 (0.33)	1
B	0 (0.07)	0 (0.07)	1 (0.33)	0 (0.33)	0 (0.33)	1
C	1 (0.27)	0 (0.27)	2 (1.33)	1 (0.80)	0 (1.33)	4
D	0 (0.20)	0 (0.20)	1 (1)	0 (0.60)	2 (1)	3
E	0 (0.40)	0 (0.40)	1 (2)	2 (1.20)	3 (2)	6
Total	1	1	5	3	5	15

Source: Own Elaboration

4.2. Corporate Tax

Corporate tax declaration rates, codes and transitions for four Markov models are shown for the years 2000-2015 in Table 11.

4.2.1. Markov Models and Transition Probability Matrices

Table 10 shows that in three categories of model 1, in two categories of model 2 and model 3, in one category of model 4, declaration rates are over 58% between 2000 and 2015. Classification of corporate tax declaration rates and transition probability matrices are given in Table 10.

4.2.2. Prediction

Given that 2015 corporate tax declaration rate 57.9% is in state C and that corporate tax will be in one of states A, B, C, D or E in 2016, corporate tax declaration rates for 2017 and later years are predicted by (5) and the predictions for 2017 are given in Table 12.

4.2.3. Stationarity

Limiting matrices are estimated in Excel by formula (6) for all models of corporate tax. According to four models, all probabilities become stationary in 2038, 2041, 2037 and 2030 respectively.

4.2.4. Statistical Significance of The Model

In model 1 of corporate tax declarations, variations between observed and expected frequencies can be tested by constructing a contingency table of frequency distribution of transitions between the states at 0,05 significance level with 16 df. Since chi square test value 22,85 is less than critical value 26,296, H_0 is not rejected. This shows that there is no significant variations. The values in paranthesis in the Table 13 are expected frequencies which are found from (row sum X column sum)/total. Table 13 shows that transitions in higher realization states are declining and in lower states except D are improving in model 1.

Table 10: Corporate Tax Declaration Rates Markov Models and Transition Probability Matrices

Model 1		Transition Matrix					Model 2		Transition Matrix			
Classes (%)		A	B	C	D	E	Classes (%)	A	B	C	D	
$82.2 \leq r$	A	0	0	1	0	0	$77.8 \leq r$	A	0	1	0	0
$64.6 \leq r \leq 82.1$	B	0	0	1	0	0	$55.8 \leq r \leq 77.7$	B	0	1/2	1/2	0
$47 \leq r \leq 64.5$	C	1/3	1/3	1/3	0	0	$33.8 \leq r \leq 55.7$	C	1/3	1/3	0	1/3
$29.4 \leq r \leq 46.9$	D	0	0	1/5	1/5	3/5	$r \leq 33.7$	D	0	0	2/9	7/9
$r \leq 29.3$	E	0	0	0	4/5	1/5						

Model 3		Transition Matrix			Model 4		Transition Matrix	
Classes (%)		A	B	C	Classes (%)	A	B	
$70.4 \leq r$	A	0	1	0	$55.8 \leq r$	A	2/3	1/3
$41.1 \leq r \leq 70.3$	B	2/3	1/3	0	$r \leq 55.7$	B	1/6	5/6
$r \leq 41$	C	0	1/10	9/10				

Source: Own Elaboration

Table 11: Classification of Corporate Tax Declaration Rates Based On Inspections By Audit Officers and Transitions Between States

Year	Declaration Rates (%)	M1 Code	M1	M2 Code	M2	M3 Code	M3	M4 Code	M4
2000	14.8	E		D		C		B	
2001	33.1	D	ED	D	DD	C	CC	B	BB
2002	18.6	E	DE	D	DD	C	CC	B	BB
2003	29.5	D	ED	D	DD	C	CC	B	BB
2004	38.9	D	DD	C	DC	C	CC	B	BB
2005	28.5	E	DE	D	CD	C	CC	B	BB
2006	30.2	D	ED	D	DD	C	CC	B	BB
2007	11.8	E	DE	D	DD	C	CC	B	BB
2008	18.1	E	EE	D	DD	C	CC	B	BB
2009	30.1	D	ED	D	DD	C	CC	B	BB
2010	51.2	C	DC	C	DC	B	CB	B	BB
2011	99.6	A	CA	A	CA	A	BA	A	BA
2012	59.6	C	AC	B	AB	B	AB	A	AA
2013	51.2	C	CC	C	BC	B	BB	B	AB
2014	76.2	B	CB	B	CB	A	BA	A	BA
2015	57.9	C	BC	B	BB	B	AB	A	AA

Source: Own Elaboration

Table 12: Corporate Tax Declaration Rate Predictions for 2017

Declaration Interval (%)	M1	Declaration Interval (%)	M2	Declaration Interval (%)	M3	Declaration Interval (%)	M4
$r \leq 29.3$	0	$r < 33.7$	0.21	$r \leq 41$	0	$r \leq 55.7$	0.58
$29.4 \leq r \leq 46.9$	0	$33.8 \leq r \leq 55.7$	0.25	$41.1 \leq r \leq 70.3$	0.48	$55.8 \leq r$	0.42
$47 \leq r \leq 64.5$	0.78	$55.8 \leq r \leq 77.7$	0.46	$70.4 \leq r$	0.52		
$64.6 \leq r \leq 82.1$	0.11	$77.8 \leq r$	0.08				
$82.2 \leq r$	0.11						

Source: Own Elaboration

Table 13: Contingency Table of Observed and Expected Corporate Tax Declaration Rates of Model 1

	A	B	C	D	E	Total
A	0 (0.07)	0 (0.07)	1 (0.27)	0 (0.33)	0 (0.27)	1
B	0 (0.07)	0 (0.07)	1 (0.27)	0 (0.33)	0 (0.27)	1
C	1 (0.20)	1 (0.20)	1 (0.80)	0 (1)	0 (0.80)	3
D	0 (0.33)	0 (0.33)	1 (1.33)	1 (1.67)	3 (1.33)	5
E	0 (0.33)	0 (0.33)	0 (1.33)	4 (1.67)	1 (1.33)	5
Total	1	1	4	5	4	15

Source: Own Elaboration

4.3. Value Added Tax

For the years 2000-2015, VAT declaration rates, codes and transitions for four Markov models are shown in Table 15.

4.3.1. Markov Models and Transition Probability Matrices

Classification of VAT declaration rates and transition probability matrices are given in Table 14. This table shows that in two categories of model 1 and model 2, in one category of model 3 and model 4, declaration rates are over 72.6% between 2000 and 2015.

4.3.2. Prediction

Given that 2015 VAT declaration rate 72,6% is in state B and that income tax will be in one of states A, B, C, D or E in 2016, income tax declaration rates for 2017 and later years are predicted by (5). The predictions for 2017 are given in Table 16.

4.3.3. Stationarity

Limiting matrices are estimated in Excel by formula (6) for all models of VAT tax. According to four models, all probabilities become stationary in 2027, 2023, 2019 and 2021 respectively.

4.3.4. Statistical Significance of the Model

In model 2 of VAT, variations between observed and expected frequencies can be tested by constructing a contingency table of frequency distribution of transitions between the states at 0.05 significance level with 9 df. Since chi square test value 12.478 is less than critical value 16.919, H_0 is not rejected. This shows that there is no significant variations. The values in the paranthesis in Table 17 are expected frequencies which are found from (row sum X column sum)/total. Table 17 shows that transitions in higher declaration states are declining and in lower states are improving in model 2.

Table 14: VAT - Markov Models and Transition Probability Matrices

Model 1		Transition Matrix					Model 2		Transition Matrix			
Classes (%)		A	B	C	D	E	Classes (%)	A	B	C	D	
$82.1 \leq r$	A	1/2	1/2	0	0	0	$78.6 \leq r$	A	1/2	1/2	0	0
$67.5 \leq r \leq 82$	B	1/5	1/5	0	3/5	0	$60.3 \leq r \leq 78.5$	B	1/5	1/5	3/5	0
$52.9 \leq r \leq 67.4$	C	0	0	0	0	0	$42 \leq r \leq 60.2$	C	0	1/4	1/4	1/2
$38.3 \leq r \leq 52.8$	D	1/5	1/5	0	2/5	1/5	$r \leq 41.9$	D	1/2	1/2	0	0
$r \leq 38.2$	E	1	0	0	0	0						
Model 3		Transition Matrix			Model 4		Transition Matrix					
Classes (%)		A	B	C	Classes (%)	A	B					
$72.3 \leq r$	A	3/4	0	1/4	$60.2 \leq r$	A	2/3	1/3				
$48 \leq r \leq 72.2$	B	0	0	1	$r \leq 60.1$	B	1/2	1/2				
$r \leq 47.9$	C	1/2	0	1/2								

Source: Own Elaboration

Table 15: Classification of VAT Declaration Rates Based On Inspections By Audit Officers and Transitions Between States

Year	Declaration Rate (%)	M1 Code	M1	M2 Code	M2	M3 Code	M3	M4 Code	M4
2000	69.8	B		B		B		A	
2001	42.5	D	BD	C	BC	C	BC	B	AB
2002	42.1	D	DD	C	CC	C	CC	B	BB
2003	72.8	B	DA	B	CB	A	CA	A	BA
2004	47.5	D	BD	C	BC	C	AC	B	AB
2005	41.4	D	DD	D	CD	C	CC	B	BB
2006	77.5	B	DB	B	DB	A	CA	A	BA
2007	72.8	B	BB	B	BB	A	AA	A	AA
2008	45.3	D	BD	C	BC	C	AC	B	AB
2009	23.7	E	DE	D	CD	C	CC	B	BB
2010	85.9	A	EA	A	DA	A	CA	A	BA
2011	96.5	A	AA	A	AA	A	AA	A	AA
2012	93.2	A	AA	A	AA	A	AA	A	AA
2013	78.4	B	AB	B	AB	A	AA	A	AA
2014	89.8	A	BA	A	BA	A	AA	A	AA
2015	72.6	B	AB	B	AB	A	AA	A	AA

Source: Own Elaboration

Table 16: VAT Declaration Rate Predictions for 2017

Declaration Interval (%)	M1	Declaration Interval (%)	M2	Declaration Interval (%)	M3	Declaration Interval (%)	M4
$r \leq 38.2$	0.12	$r \leq 41.9$	0.30	$r \leq 47.9$	0.33	$r \leq 60.1$	0.39
$38.3 \leq r \leq 52.8$	0.36	$42 \leq r \leq 60.2$	0.27	$48 \leq r \leq 72.2$	0	$60.2 \leq r$	0.61
$52.9 \leq r \leq 67.4$	0	$60.3 \leq r \leq 78.3$	0.29	$72.3 \leq r$	0.67		
$67.5 \leq r \leq 82$	0.26	$78.6 \leq r$	0.14				
$82.1 \leq r$	0.26						

Source: Own Elaboration

Table 17: Contingency Table of Observed and Expected VAT Declaration Rates of Model 1

	A	B	C	D	Total
A	2 (1.07)	2 (1.33)	0 (1.07)	0 (0.53)	4
B	1 (1.33)	1 (1.67)	3 (1.33)	0 (0.67)	5
C	0 (1.07)	1 (1.33)	1 (1.07)	2 (0.53)	4
D	1 (0.53)	1 (0.67)	0 (0.53)	0 (0.27)	2
Total	4	5	4	2	15

Source: Own Elaboration

- **Comparison of Predictions of Tax Declarations for 2017**

Given that 2015 tax declaration rate is in state A, B, C, D or E and will be in one of these states in 2016, declaration rate matrices are predicted for 2017 and later years by formula (5). Predictions for the three tax declaration rates are combined in Table 18.

- **A Better Model For Tax Revenues**

Sum of mean square errors for a better model of each tax is given in Table 19. Error values in bold indicates the better model.

- **Comparison of Initial Probability Matrices and Stationarity of Tax Declarations**

Stationary matrices for tax declarations are found for every model and are given in Table 20. Q_0 is initial probability matrix for the 2015 declaration rates of tax revenues and SY is the stationarity year when probability matrix becomes stable.

- **Statistical Significance of Markov Model**

In the present study, the validity of model is checked for the years 2011 and 2015. The degrees of freedom (df), χ^2 critical values, test values and hypothesis are given in Table 20. The null hypothesis is not rejected since χ^2 test value is less than the critical value. The values of the χ^2 test are less than χ^2 critical values for the years 2011 and 2015, which implies that the estimated declaration rates of tax revenues and the actual declaration rates of tax revenues are not significantly different. Table 21 results show that Markov model is valid.

Table 18: 2017 Prediction of Tax Declaration Rates Based on Inspections by Audit Officers

		A	B	C	D	E						
Income Tax	M1	0.125	0.25	0.33	0.125	0.17	Corporate Tax	0.11	0.11	0.78	0	0
	M2	0	1	0	0			0.08	0.46	0.25	0.21	
	M3	0.25	0.43	0.32				0.52	0.48	0		
	M4	0.48	0.52					0.42	0.58			
VAT	M1	0.23	0.29	0	0.36	0.12						
	M2	0.14	0.29	0.27	0.30							
	M3	0.67	0	0.33								
	M4	0.61	0.39									

Source: Own Elaboration

Table 19: Tax Declarations SMSE

Tax Revenues	Sum of Mean Square Errors (SMSE)			
	Model 1	Model 2	Model 3	Model 4
Income tax	3.44	3.68	4.15	7.59
Corporate tax	3.45	4.71	5.02	6.56
Vat	3.86	3.32	3.66	4.47

Source: Own Elaboration

Table 20: Stationary Matrices of Tax Declarations

Tax Revenues	Markov Model 1				Markov Model 2								
	Q_n	$\lim_{n \rightarrow \infty} \hat{Q}_{n+1} = \hat{Q}$			SY	Q_n	$\lim_{n \rightarrow \infty} \hat{Q}_{n+1} = \hat{Q}$			SY			
Income tax	0 0 1 0 0	.10	.10	.39	.18	.24	2038	0 1 0 0	0	1	0	0	2017
Corporate tax	0 0 1 0 0	.20	.20	.60	0	0	2038	0 1 0 0	.08	.32	.24	.36	2041
Vat	0 1 0 0 0	.35	.29	0	.29	.06	2031	0 1 0 0	.27	.33	.27	.13	2023
Tax Revenues	Markov Model 3				Markov Model 4								
	Q_n	$\lim_{n \rightarrow \infty} \hat{Q}_{n+1} = \hat{Q}$			SY	Q_n	$\lim_{n \rightarrow \infty} \hat{Q}_{n+1} = \hat{Q}$			SY			
Income tax	0 1 0	.19	.38	.43	2039	1 0	0,27	0,73	2036				
Corporate tax	0 1 0	.40	.60	0	2037	1 0	0,33	0,67	2030				
Vat	1 0 0	.67	0	.33	2020	1 0	0,60	0,40	2021				

Source: Own Elaboration

Table 21: Validity of Tax Declarations

Tax Revenues	Df	2011			H ₀ : Valid	df	2015			H ₀ : Valid
		X ² _{0,05} Critical Value	X ² _{0,05} Test Value				X ² _{0,05} Critical Value	X ² _{0,05} Test Value		
Income tax	1	3.841	1	Accept	2	5.991	2	Accept		
Corporate tax	1	3.841	2	Accept	1	3.841	1	Accept		
Vat	2	3.841	3	Accept	2	5.991	1	Accept		

Source: Own Elaboration

Table 22: Tax Declaration Predictions For 2017 According to State B of Better Models

Tax Declarations	Better Markov Model	Declaration Rate r (%)	Probability (%)	1 – Probability (%)
Income tax	1	B or higher $61.2 \leq r$	37.5	62.5
Corporate tax	1	B or higher $64.6 \leq r$	22	78
Vat	2	B or higher $60.3 \leq r$	43	57

Source: Own Elaboration

Table 23: Comparison of 2017 Predictions To Stationary Matrices According to Better Models

Tax Revenues	Better Markov Model	2017 Prediction	Stationary Matrix	$E(\hat{Q}_{17})$	SY
Income tax	1	.125 .25 .33 .125 .17	.10 .10 .39 .18 .24	52,3	2038
Corporate tax	1	.11 .11 .78 0 0	.20 .20 .60 0 0	61,6	2038
Vat	2	.14 .29 .27 .30	.27 .33 .27 .13	62,8	2023

Source: Own Elaboration

Original data shows that the highest base differences for income tax and VAT are 834,488,383 ₺ and 7,132,560,056 ₺ which occur at the lowest declaration rates 4.2% in 2006 and 23.7% in 2009 respectively. The highest base difference for corporate tax is 5,217,618,792 ₺ which occurs at the declaration rate 30.1% in 2009. According to contingency tables 9, 13 and 17 transitions of income tax, corporate tax and VAT declaration rates are declining in higher states and improving in lower states. 2017 predictions of declaration rates with respect to state B of the better models are given in Table 22.

According to model 1 of income tax declaration rates, the probabilities of five states will be stable in 2038. Income tax declaration rates more likely will be less than 61.2% in 2017. However, in the long run income tax declaration rates would be in states D or E (less than 2015 realization rate) with a probability of approximately 42%. A decrease of at least 19.2% which is approximately 20,267,260,000 ₺ from 2015 to 2038. This may cause a revenue deficiency to TRA. Probability of income tax declaration rate greater than 80.2% is declining from 12.5% in 2017 to a stable 10%, that of between 61.2% and 80.1% is declining from 25% in 2017 to a stable 10% and that of $42.2 \leq r \leq 61.1$ is improving from 33% to a stable 39%. Similarly, other tax revenues predictions are compared in Table 23.

A similar discussion for VAT can be given. VAT declaration rates more likely will be less than 60.3% in 2017. However, in the long run VAT declaration rates would be in states C or D (less than 2015 realization rate) with a probability of approximately 40%. A decrease of at least 17.1% which is approximately 26,248,500,000 ₺ from 2015 to 2023. This may also cause a revenue deficiency to TRA. Therefore income tax and VAT audits should be increased by TRA. This was proposed as the only measure to be taken by TRA in "Payment Strategies in Corporate Tax: A Mathematical Analysis" study by Mavruk and Kiral (2017). However a stability or improvement is expected in corporate tax declarations from 2015 to 2038.

Income tax declaration rate was realized in state C in 2015. We predict that it improves to state B in 2017 with a probability of 25% which declines to a stable 10% probability in 2038. Corporate tax declaration rate was also realized in state C in 2015. It improves to state B in 2017 with a probability of 11% which improves to a stable 20% probability in 2038. VAT declaration rate was realized in state B in 2015. It would stay in the same state in 2017 with a probability of 29% which improves to a stable 33% probability in 2023.

The probability of income tax declaration rate greater than or equal to 0.612 would be stable at 10% by the year 2038, the probability of corporate tax declaration rate greater than or equal to 0.646 would be stable at 20% by the year 2038 and the probability of VAT declaration rate greater than or equal to 0.603 would be stable at 27% by the year 2023. Expected declaration rates of income tax, corporate tax and VAT in 2017 are estimated to be 52.3%, 61.6% and 62.8%, respectively.

5. CONCLUSION

Tax declaration rates are predicted with respect to the results of audits provided by TRA based on 2000-2015 data. In the original data base differences has the highest values at the lowest declaration rates for two taxes. Therefore, the more tax declared by the taxpayer, the more tax revenues will be collected by the tax administration. Transition matrices show that transitions of tax declarations are declining in higher states and improving in lower states. According to prediction results all tax declaration rates more likely will be in state C or lower in 2017 according to better Markov models, and more likely will be stationary in state C or lower except VAT. Income tax and VAT payoffs may substantially decrease from 2015 to 2038 and 2023 respectively. This may also cause a revenue deficiency to TRA. Therefore income tax and VAT audits should be increased by TRA. Even though tax revenues increase over time, the declaration rates show a decreasing to stationary or increasing to stationary behavior. Observed and expected declaration rates were not significantly different from each other according to Markov model 1 of income and corporate tax and Markov model 2 of VAT. Validity of the models were checked for the years 2011 and 2015 and models are found to be valid. Declaration rate predictions can be used by the TRA to take some measures against tax evasion and for tax policy revision. Central government can take the advantages of this study in the planning and improvement of tax collection process. For future studies tax payoff rates can be analyzed and predicted. This study can also be used to predict tax declaration rates of other tax revenues.

REFERENCES

- Allingham, M.G. & Sandmo, A. 1972, "Income Tax Evasion: A Theoretical Analysis", *Journal of Public Economics*, vol.1, pp. 323-338.
- Usher, M.B. 1979, "Markovian Approaches to Ecological Succession", *Journal of Animal Ecology*, vol.48, no.2, pp. 413-426
- Serfozo, R. 2009, "Basics of Applied Stochastic Processes, Probability and its Applications", Springer-Verlag, Berlin Heidelberg, 2
- Yeh, H.W., Chan, W., Symanski, E. & Davis, B.R. 2010, "Estimating Transition Probabilities for Ignorable Intermittent Missing Data in a Discrete-Time Markov Chain", *Communications in Statistics-Simulation and Computation*, vol.39, no.2, pp. 433-448.
- Baasch, A., Tischew, S. & Bruelheide, H. 2010, "Twelve years of succession on sandy substrates in a post-mining landscape: A Markov chain analysis", *Ecological Applications*, vol. 20, no. 4, pp. 1136-1147.
- Grimshaw, S.D. & Alexander, W.P. 2011, "Markov Chain Models for Delinquency: Transition Matrix Estimation and Forecasting", *Applied Stochastic Models in Business and Industry*, vol. 27, pp. 267-279.
- Lipták, K. 2011, "The Application Of Markov Chain Model To The Description Of Hungarian Labor Market Processes", *Zarządzanie Publiczne*, vol. 4, no. 16, pp. 133-149.
- Büyüktatlı, F., İşbilir, S. & Çetin, E.İ. 2013, "Markov Analizi ile Yıllık Ödeneklere Bağlı Bir Tahmin Uygulaması", *Uluslararası Alanya İşletme Fakültesi Dergisi*, vol. 5, pp. 1-8
- Lukić, P., Gocić, M. & Trajković, S. 2013, "Prediction of Annual Precipitation on the Territory Of South Serbia Using Markov Chains", *Bulletin of the Faculty of Forestry*, vol.108, pp. 81-92.
- Bluman, A.G. 2014, "Elementary Statistics", McGraw Hill Education, New York.
- Cavers, M.S. & Vasudevan, K. 2015, "Brief Communication: Earthquake Sequencing: Analysis of Time Series Constructed from the Markov Chain Model", *Nonlinear Process Geophysics*, vol. 22, pp. 589-599
- Lazri, M., Ameer, S., Brucker, J.M., Lahdir, M. & Sehad, M. 2015, "Analysis of Drought Areas in Northern Algeria Using Markov Chains", *J. Earth Syst. Sci.*, vol. 124, no. 1, pp. 61-70.
- Kiral, E. & Mavruk, C. 2016, "Prediction of Central Government Budget Revenues Using Markov Model", *Journal of Cukurova University Social Sci. Inst.*, vol. 25, no. 2, pp. 41-56
- Mavruk, C. & Kiral, E. 2017, "Payment Strategies in Corporate Tax: A Mathematical Analysis", Manuscript to be submitted for publication.
- TRA Activity Reports. 2000-2015, <http://www.gib.gov.tr/kurumsal/stratejik-yonetim/faaliyet-raporlari>, 5.1.2016