Institut de hautes études en administration publique
Swiss Graduate School of Public Administration
Institut universitaire autonome

Florian Chatagny &
Nils C. Soguel

Tax Revenue Forecasting in the Swiss Cantons : a Time Series Analysis

Working paper de l'IDHEAP
Chaire de Finances publiques

Tax Revenue Forecasting in the Swiss Cantons : A Time Series Analysis

Florian Chatagny*
Nils C. Soguel†

IDHEAP‡ Working Paper 2009, Chair of Public Finance

2 avril 2009

Abstract : Forecasting tax revenue accurately is a critical point of an efficient fiscal policy. Overestimating tax revenue ex ante may contribute to create or increase fiscal deficit ex post, while underestimating tax revenue ex ante may contribute to create fiscal surplus ex post. Both situations may be inefficient. Empirical studies and indicators suggest that most governments face difficulties in forecasting and budgeting tax revenue accurately. Using new data about tax revenue forecasting errors in the Swiss cantons over the fiscal years between 1945 and 2006, we show that tax revenue are significantly and systematically underestimated in a majority of cantons. These results are robust for different timespans. This suggests possibilities of improvement and appeal for further analysis. Thus, by fitting ARIMA models on actual tax revenue and simulating new tax revenue forecasts, we show that for a majority of cantons, unbiased forecasts can be generated by using simple regression techniques. Such a result gives support to the idea that prediction of tax revenue can be improved and thus budgeting process and fiscal policy in Swiss cantons become more efficient.

Keywords : Fiscal policy, tax revenue, forecasting, time series analysis.

JEL Classification Numbers : C10, E62.

*Contact author (Please do not quote without authors' permission). Chair of Public Finance, Swiss Graduate School of Public Administration, University of Lausanne. Address : Route de la Maladière 21 - CH-1022 Chavannes-Lausanne,Switzerland. e-mail : *florian.chatagny@idheap.unil.ch*.

†Chair of Public Finance, Swiss Graduate School of Public Administration, University of Lausanne. e-mail : *nsoguel@idheap.unil.ch*. Authors are grateful to the Swiss National Fund for Scientific Research for its financial support. We also thank Sven Jari Stehn (IMF) for his useful comments.

‡Swiss Graduate School of Public Administration

1 Introduction

Forecasting and budgeting government revenue and expenditure accurately is a key component of an efficient fiscal policy. Misforecasting and misbudgeting either revenue or expenditure may create or amplify fiscal imbalances, i.e. deficits or surpluses. Depending on the current fiscal position of the government, both may not be desirable. In particular, forecasting and budgeting tax revenue is the first step in the budgetary decision-making process. It is supposed to set the budgetary limits within which public spending should remain to reach fiscal balance. Thus, tax revenue forecasting and budgeting errors may have strong repercussions on the whole budget process and fiscal outcome. While underestimated tax revenues may lead to a situation where taxpayers are overtaxed with respect to the quantity of public goods they receive, overestimated tax revenues may lead to a situation where governments overproduce public goods with respect to the available financial resources. Numerous contributions suggest that, in fact, most governments face difficulties in forecasting and budgeting tax revenue accurately. For instance, the International Monetary Fund (IMF) has paid attention to problems in the revenue-forecasting process faced by some low-income countries from sub-Saharan Africa, the Middle East and Developing Asia (Danninger and Kyobe, 2005). Danninger et al. analyse interferences that may occur during the budgeting process and thus alter revenue forecast. Furthermore several empirical studies show that not only low-income or less-developed countries are concerned with such problems. Indeed a lot of empirical studies have been performed on this subject in the North American (United States and Canada) context. Among others, we can mention Auerbach, 1999, Bretschneider and Gorr 1992, Cassidy et al. 1989, Feenberg et al. 1989, Jones et al. 1997, Mocan and Azad 1995, Rodgers and Joyce 1996 and Campbell and Ghysels, 1997. Some of these studies clearly show that tax revenue prediction can be highly inaccurate and thus can create budgeting problems. These studies suggest that no country is sheltered from problems in forecasting and budgeting tax revenue accurately. In Switzerland, indicators on tax revenue forecasting errors for recent years suggest that the national government (Confederation) as well as the sub-national governments (cantons) often make important errors in predicting tax revenue (Soguel, 2008). Further available data also indicate that these errors tend to be due to underestimation rather than overestimation. This would mean that forecast errors are systematically biased, and room for improvement in the forecast of tax revenue and hence in the budgeting process does exist. Since Ammann (2001) already analysed accuracy of tax revenue forecasts at the federal level, it appeared more relevant to focus on the cantonal level. Since Switzerland is composed of 26 cantons, this level of jurisdiction is also statistically more interesting to handle with.

In this study we focus on a precise carateristic that tax revenue forecasts observed in Swiss cantons seem to exhibit, i.e. systematic underestimation. To assess this caracteristic, we collected new data about budgeted and actual tax revenues at the cantonal level. These data especially pertained to a high number of periods and allow us to statistically corroborate what has already been observed for the most recent years. The aim of this paper is twofold. On the one hand we seek to show that observed tax revenue forecast errors effectively suffer from systematic underestimation ; on the other, we try to show how it is possible to improve forecasting, in the sense that biasedness may be reduced using univariate time-series regression methods. Thus by fitting ARIMA models on actual tax revenues, we expect to show that forecasts generated by a simple econometric model exhibit lower bias or no bias at all. Such a result would give support to the idea that prediction of tax revenue could be easily improved and thus budgeting process and fiscal policy in Swiss cantons become more efficient.

The current paper is organized as follows : in the second section we present the methodology we used to assess accuracy of observed tax revenue forecasts and to generate new forecasts. Section 3 presents the data we used to perform our analysis. Section 4 presents the results for two different timespans, 1945-2006 and 1979-2006, and Section 5 provides the conclusions.

2 Methods

2.1 Measuring observed forecast error

To find an appropriate measure of the observed forecast error in Swiss cantons, we need to keep our objective in mind. Since we seek to test whether tax revenue forecasts are (upward or downward) biased and since we want to compare cantons to eachother, the indicator to be chosen should exhibit two properties : first, it must give information about the sign of the error and, second, since cantons strongly differ in their financial size, it must wash out potential size effects to allow comparability. The first property automatically rules out all the indicators based on absolute values or quadratic terms since they would not take into account the sign of the errors. The second property lead us to consider the family of relative indicators which divide the error through a measure of the financial size of the cantons. Given these desirable properties an indicator is expected to have, we chose the following one :

$$PFE_o = \frac{R_f - R_a}{R_a} \times 100 \tag{1}$$

where PFE means *Percentage Forecast Error*, o states for *observed*, R means *Revenue*, a subscript states for *actual* and f subscript states for *forecast*. As the numerator is neither an absolute value nor a quadratic term, this indicator will give us information about the sign of the forecast error. As a denominator it seems natural to choose actual tax revenue. Indeed it makes sense that forecasting errors are compared to reality rather than a potentially erroneous forecast.[1] Using this indicator, three general situations may arise :

$$R_f = R_e \tag{2}$$

$$R_f > R_e \tag{3}$$

$$R_f < R_e \tag{4}$$

Situation (2) describes perfect forecast with zero error. Situation (3) describes an overestimation of tax revenue and situation (4) describes an underestimation. The natural question to ask is which one of these three situations has been observed on average for a given period of time and/or for a given canton or group of cantons. The method we used to answer this question and the results we get are presented in Subsection 2.3 and Section 4, respectively.

2.2 Simulating errors using univariate time-series

One important question we then seek to address in this paper is whether we are able, using simple univariate time-series methods, to improve (in terms of bias) tax revenue predictions made by Swiss cantons. We proceeded in two steps. First we fitted time-series models on actual tax revenue. Then we used these models to "reforecast" tax revenue and generate new forecasting errors. Basically, the type of model we fitted for each canton is an ARIMA(p,1,q), i.e. an order one integrated autoregressive of order p and moving average

[1] Using actual tax revenue as denominator is equivalent to putting more weight on overestimation. However, according to our hypothese and despite this intrinsic bias towards overestimation, we expect to find a tendency to underestimate tax revenue

of order q model. Three reasons motivated us for choosing this type of model. First, it appears reasonable to think that future tax revenue are highly correlated to their lagged values. Second, this kind of model turns out to perform quite well in predicting the future path of economic variables and often outperforms more complicated structural models (Verbeek, 2004, p.288). Third, these models forecast future values of the dependent variable by strictly using its past values, which have the huge advantage of being easily available and being almost costless information. To illustrate the procedure we followed, let us take the ARIMA(1,1,1) case. Such a model can be expressed as follows :

$$R_{a_t} = \alpha + \rho R_{a_{t-1}} + \epsilon_t + \theta \epsilon_{t-1} \tag{5}$$

where ρ and θ are the parameters of interest, α is a constant, ϵ is an i.i.d error term with zero mean and t is the time subscript[2].

An important condition for these kind of models to be valid is that the series is stationary. In this particular case, stationarity means that ρ must be less than one in modulus (Hamilton, 1994, p.53). Thus before using the models to generate new forecasts and new forecast errors we have tested whether the estimate of ρ, let us call it $\hat{\rho}$, is significantly less than one by performing Dickey-Fuller (1979) testing procedure[3]. Then, given stationarity, we used the fitted model to simulate new tax revenue forecasts :

$$\hat{R}_{f_{t+1}} = \hat{\alpha} + \hat{\rho} R_{a_t} + \hat{\theta} \hat{e}_t \tag{6}$$

where $\hat{\alpha}$, $\hat{\rho}$ and $\hat{\theta}$ are the estimated coefficients and $\hat{e}$ are the regression's residuals. Using the new forecast $\hat{R}_{f_t}$ and actual revenue R_{a_t} we generated a new forecasting error defined as

$$PFE_{s_t} \equiv \frac{\hat{R}_{f_t} - R_{a_t}}{R_{a_t}} \tag{7}$$

Given our observed and simulated forecasting errors, we now turn to the testing procedure for biasedness.

2.3 Testing for biasedness

Once observed and simulated forecast errors have been computed and generated respectively, we then have to test whether these are biased and, if yes, whether they are downward or upward biased. A simple procedure to assess whether tax revenue forecast suffer from any kind of bias is to test whether the average forecast error is statistically different from zero. The natural statistical test to use is the t-test. Formally, we run the test for $H_0 : \overline{PFE}_{o/s} = 0$, where $\overline{PFE}_{o/s} \equiv \frac{1}{T}\sum_{t=1}^{T} PFE_{o_t/s_t}$, against the alternative hypotheses that it is $>$ or $<$ than 0. We use the t-test since we do not know the variance of the population and we want to evaluate the statistical significance of the value of the mean against the hypothetical value given by the theory (zero = perfect forecast on average). Under H_0, the test statistic is given by :

$$tstat = \frac{\hat{\beta}}{se(\hat{\beta})} \tag{8}$$

[2]Although we present the model in levels for the sake of simplicity, note that, as we did, it is better to work with logarithms because it stabilizes variance. Levels can then be easily recovered.

[3]Since all series exhibit stationarity we will not go into the detailed results. Results for $\hat{\rho}$ and D-F tests are reported in Appendices (A-1) and (A-2).

where $\hat{\beta} \equiv \frac{1}{T}\sum_{t=1}^{T} PFE_t$, t is the time subscript and $se(\cdot)$ is the standard error [4] associated to $\hat{\beta}$ (Verbeek, 2004, pp.23-24). Critical values for confidence interval and the p-values can be directly computed from the normal distribution. This is possible due to the fact that this test is asymptotically normal and thus valid. However, to have in-sample validity we need the data to be normally distributed. Thus, it is necessary to perform a normality test on the data[5].

The normality test is performed using the methodology developed by Jarque and Bera (1987). The general idea is to test whether the skewness and the kurtosis of the empirical distribution are statistically different from the values of skewness and kurtosis of the normal distribution, i.e. zero for skewness and three for kurtosis. The test statistic is computed as follows

$$JB = [N/6]\left[S^2 + (K-3)^2/4\right] \tag{9}$$

,where N denotes the number of observations, S the skewness and K the kurtosis of the empirical distribution (Pindyck and Rubinfeld, 1998, pp. 47-48). This statistic follows a chi squared distribution with two degrees of freedom which allows us to compute critical values and p-values. Note that this test relies on asymptotic standard errors and does not correct the sample size. Thus to assess the robustness of the test, we have performed the test suggested by D'Agostino et al. which makes two adjustments for the sample size[6].

Under (asymptotic) normality, we expect our t-tests to show that, in a majority of cantons, tax revenue predictions are downward biased. Conversely, we expect forecast errors generated by univariate time-series model to be statistically equal to zero on average. Before discussing the results, we now describe the data on which these tests have been performed.

3 Data

3.1 Source of the Data

To perfectly assess cantonal tax revenue forecasting accuracy, the ideal situation would be to rely on features issued at every step of the budgeting process. This would not only have allowed us to assess directly the forecast made by the administration officers but also to evaluate in which sense this forecast might have been altered by the executive and/or the legislature during the budget discussions. However most of the Swiss cantons do not produce documents containing such information. Hence, the best data we could find for forecast amounts -our R_f in (1)- is the amount reported in public accounts. As this feature comes out at the end of the budgeting process, we might think that it is highly proned to be biased since many actors might have had an influence or an impact on it. Accordingly, the feature chosen for actual revenue (R_a) is the amounts figuring in the public accounting at the end of the pertained year. Some of these data were either made available to us by cantonal administrations in numeric format or directly collected in the cantonal administrations. The remainder was collected in a centralized manner in the National Library.

[4]Note that performing t-test directly on mean percentage forecast error is equivalent to regressing the percentage forecast error on a constant only and an error term : $PFE_t = \alpha_t + \epsilon_t$.

[5]In-sample validity is necessary in our case especially when we perform the test on the shorter timespan (1979-2006). In this case, the sample for each canton has less than 30 observations. Thus asymptotics cannot really be relied on and t-test may lead to invalid inferential statements (Jarque and Bera, 1987, p.164). Further, even if the timespan is longer, having in-sample validity makes results more robust.

[6]For details about these adjustments see D'Agostino et al. 1990 and Royston 1991.

3.2 Timespan and level of aggregation

In this study we consider a timespan going from the end of WW2 up to now. To understand this choice, let us come back to the context of the study. The test of the hypotheses mentioned earlier is meant to be performed using panel data estimation methods. As the last Swiss Canton has been created in 1979 and as the database of independent variables already available spreads over 1979-2006, this timespan should have been our natural choice. However, we decided to consider a longer time period in order to use some pure times-series regression techniques and confront the results to these given by panel data models. But, as we will show, the timespan has sometimes been chosen according to the availability of the data. To explain the choice of the level of aggregation we refer to the taxonomy of the standardized accounting system which has been used in most of the Swiss cantons from the early eighties up to now. Table (1) shows how this accounting system classifies public revenues.

TAB. 1 – Classification of Revenues

Account No	Types of Revenue
6	Investment Revenue
40	Tax Revenue
42	Assets and Financial Revenue
43	Royalties and Concessions
44	Share to Revenue without Allotment
45	Reimbursement of Jurisdictions
46	Received Grants
47	Grants to be Distributed
48	Withdrawals from Reserves
49	Cross Charges

First, investment revenue is not taken into account in the current analysis. At the cantonal level, investment revenues are mainly constituted by grants from the federal level. Thus, this category of revenue is not really problematic in terms of forecasting. Once conditions to get a grant are fulfilled, cantonal government knows the amount of money it will get. Since these revenues are affected to particular investment projects, they are anyway not available to the government for yearly expenditure decisions. Further, they influence current deficits (surpluses) only indirectly through the reduction of debt interest and amortisation installments. Among all the other types of revenue we focused exclusively on the Number 40 since it is the most important source of revenue for the cantons to finance public services[7].

Then for practical reasons we had to restrict our focus further. During the collection process, Tax Aggregate 40 has been revealed to be very heterogeneous among the cantons and through time since it is composed of numerous kinds of particular taxes. This poor comparability spoke against the choice of this level of aggregation. Furthermore, due to insufficient data or some intertemporal changes in the accounting system, the Number 40 could not be calculated for some time periods and/or some cantons. These elements led us to consider only the Numbers 400 and 401 of the accounting system, i.e. the personal tax and corporate tax revenue[8]. Again, it has not been possible to distinguish both 400 and 401 separately for every time period

[7]According to the Federal Statistical Office, in 2006 for all cantons Tax Revenue (No 40) accounts for 50% of total revenue, excluding investment revenues and cross charges. This ratio increases up to 70% if we only consider own revenue of the period into account(no 40+42+43). See http ://www.bfs.admin.ch/bfs/portal/fr/index/themen/18/02/blank/key/einnahmen_von_bund0/kantone.html

[8]Note that this is not much harmful to the relevance of the current analysis since 400 and 401 represent an important proportion of total tax revenue. According to the Federal Statistical Office, in 2006, direct taxes (400+401) represented 87,2% of the total of tax revenue for all cantons. It varies from a minimum of 74% for the canton of AI (Appenzell-Innerrhoden) to 95% for the canton of ZG (Zug). See http ://www.bfs.admin.ch/bfs/portal/fr/index/themen/18/02/blank/key/einnahmen_von_bund0/gesamt.html

TAB. 2 – Availability of data

Cantons	400+401	400 vs 401
AG		1944-2006
AI	1944-2006	
AR	1944-2006	
BE		1944-2006
BL	1944-1971	1972-2006
BS		1944-2006
FR	1944-1975	1976-2006
GE		1944-2006
GL	1944-1984	1985-2006
GR	1944-1966	1967-2006
JU		1979-2006
LU	1944-1987	1988-2006
NE		1978-2006
NW	1944-1972	1973-2006
OW	1944-2003	2003-2006
SG		1944-2006
SH	1944-2006	
SO	1944-1975	1976-2006
SZ	1944-1986	1987-2006
TG	1944-1981	1982-2006
TI	1944-1975	1976-2006
UR		1944-2006
VD		1948-2006
VS		1944-2006
ZG	1944-52 + 56-73	1953-55 + 74-2006
ZH	1944-2006	

in every canton. Table (2) shows for every canton[9] the timespans for which data could be collected separately (400 vs 401) and the timespans for which this could not be done (400+401). Clearly, for a majority of cantons, distinction between personal and corporate tax is not available for the whole timespan (1945-2006).

Given the constraints on the data, we finally restricted our attention to the addition of both corporate and personal tax (called direct taxes) for all cantons and all periods. The next subsection presents summary statistics on both observed and simulated forecast errors.

3.3 Summary statistics

In this subsection we seek to shed light on the most important empirical characteristics of our series. For each canton, we have computed a range of summary statistics (relative number of negative errors, mean, standard deviation, median, skewness and kurtosis). Over 1945-2006, the statistics are presented in table (3). We now comment on the facts that appear to bethe most interesting given our goal to test for biasedness. As a first observation note that, except for the canton of JU, observed errors (see Column 3 in Table (3)) were negative (= underestimation) for more than 50% of the considered years for all cantons, ranging from 60% for BE to 90% for GR, FR and AI. When considering simulated errors (Column 4) only six cantons have negative errors strictly more than 50% of the time. This already suggests that, for all cantons except JU, tax revenue forecast might be systematically underestimated and also that univariate time-series model might improve the forecast in terms of biasedness. This statistic alone is however not sufficient since positive and negative errors could exactly offset eachother in terms of size thus leading to a perfect forecast (in terms of bias) on average. Thus, it is also necessary to look at the mean percentage forecast error (designated by $\overline{PFE}$ in Table (3)). For all cantons without exception, $\overline{PFE}_o$ (Column 5) turns out to be negative. This provides even more support to the idea that tax revenue might be systematically underestimated. However, also note that $\overline{PFE}_o$ exhibits strong heterogeneity among cantons with $-0,0712\%$ for SH down to $-9,6438\%$ for UR as shown in Table (3). Whereas biasedness seems obvious in cases like UR, AI or OW, no clear conclusion can be drawn for cantons like SH, JU or ZH. Thus some formal statistical tests are clearly needed. Note further that $\overline{PFE}_s$ provide further support to the idea that biasedness (if confirmed) of tax revenue forecasts might

[9]For a complete list of the names of the Swiss cantons see Appendix Table A-3

TAB. 3 – Summary statistics for observed and predicted errors 1944-2006

Canton	N		no PFE<0		$\overline{PFE}$		sd		median		sk		kt	
	obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
SH	62	61	61%	49%	−0.0712	−0.1604	8.1994	8.5378	−1.4029	0.0671	1.8843	0.0808	8.7197	5.8571
JU	28	27	50%	37%	−0.6923	0.1394	2.8478	4.2672	−0.1628	0.6177	−0.4893	−0.5462	2.3255	2.9211
ZH	63	62	68%	45%	−1.6506	0.0269	5.1052	7.4818	−1.8796	0.8638	−0.3077	−0.5798	3.4123	3.8357
BE	63	62	60%	40%	−2.5201	0.2704	5.8871	7.1739	−1.6568	1.3960	−0.9528	−0.4135	4.1067	3.5104
GE	63	62	68%	47%	−3.1900	0.3842	7.0183	8.7597	−3.0960	0.2391	−0.1331	−1.1387	3.8025	7.0859
SG	62	61	77%	57%	−3.6931	−0.5911	4.6416	7.0430	−3.4169	−0.7389	−0.3700	0.2789	3.4945	3.4953
TG	63	62	73%	52%	−3.9734	0.2420	6.3937	6.0586	−3.5447	−0.2692	−0.3763	0.1196	2.9627	4.4818
AR	63	62	79%	55%	−4.1889	−0.3265	4.9009	8.4432	−4.0791	−1.8340	0.1034	1.2894	2.8061	5.9484
VD	59	58	73%	47%	−4.4271	0.2079	6.0627	6.6882	−4.8265	0.3286	−0.4008	−0.1756	4.1243	2.8716
AG	63	62	81%	52%	−4.7514	0.2199	5.6345	7.5807	−4.3578	−0.3950	−0.3046	−0.0637	2.8224	3.1814
LU	63	62	81%	48%	−4.7862	−0.1827	5.4081	6.6651	−4.6528	0.3545	−0.1842	−0.2520	2.9867	3.3417
BL	63	62	75%	42%	−4.7995	0.6072	7.6958	9.8268	−3.4404	2.1755	−1.0038	−0.8560	3.6426	5.7848
SO	63	62	76%	50%	−5.1566	0.0868	7.7296	7.8041	−4.6492	0.7374	−0.3591	−0.1813	3.8354	3.3835
GR	63	62	90%	45%	−5.2106	0.1649	4.8576	7.5013	−4.6304	1.4632	−1.3893	−0.0771	6.3789	2.9264
BS	63	62	83%	50%	−5.4941	0.1826	6.0448	7.3764	−5.4190	0.2139	−0.2193	−0.5968	2.7707	6.2861
NE	58	56	78%	46%	−6.6315	−0.2775	10.3671	8.7278	−4.1532	1.0009	−0.2863	0.0308	5.2793	6.6147
NW	50	47	82%	36%	−6.6695	−0.2892	6.2043	9.8286	−6.8260	3.2891	−0.2714	−1.1277	2.4566	3.9309
FR	63	62	90%	44%	−6.8847	0.0919	6.2552	7.2935	−5.7997	0.9596	−0.7039	−0.5536	3.2241	4.1823
SZ	59	55	86%	53%	−6.9757	−0.6191	6.8173	9.1646	−6.8454	−0.6394	0.6180	0.6098	3.3652	6.3384
ZG	61	59	84%	47%	−7.2603	−0.0756	8.2755	8.3473	−6.1989	0.7058	−0.0028	−0.1237	3.1438	2.7530
VS	63	62	87%	45%	−7.4786	0.2477	7.6082	8.2922	−6.4409	1.0119	−0.1116	−0.6276	2.2565	4.9472
TI	63	62	87%	52%	−8.3113	0.0907	8.0994	8.2684	−8.9522	−0.5610	0.1841	0.0573	2.6732	2.5737
GL	59	56	85%	43%	−8.4319	0.0943	7.0758	6.8715	−9.2428	1.3544	−0.0222	−1.1069	2.5017	4.5939
OW	63	62	84%	35%	−9.1609	0.6763	10.4007	13.6924	−8.1967	2.0963	−0.3713	−1.7854	3.5675	9.2875
AI	63	62	90%	42%	−9.4869	0.5939	8.7022	11.3062	−8.9381	1.0461	0.1652	−0.2916	4.8309	5.0818
UR	63	62	81%	39%	−9.6438	1.9360	12.5092	13.5666	−6.9741	2.2413	−0.7736	0.3361	3.0276	8.4687

be easily reduced using simple univariate models. Indeed, except for canton SH, $\overline{PFE}_s$ is closer to zero than $\overline{PFE}_o$ for all cantons.

A particularly interesting case to focus on is the case of JU. This canton was created later than the others (1979) by partition from canton Bern. Consequently, the timespan of this canton is restricted to $1979 - 2006$ and the number of observations to 28. Since JU has a $\overline{PFE}_o$ of $-0,6923\%$ which is quite close to zero, this good performance lead us to think that some cantons may have clearly improved (in terms of bias) their predictions over time. To analyse this question, we have recomputed summary statistics for a shorter timespan going from 1979 to 2006. They are presented in Table (4)[10]. First we can see that still in 24 cantons more than 50% of the forecast errors were negative. Furthermore, for all cantons except BE, $\overline{PFE}_o$ is still negative. However, it is important to note that 23 cantons have "improved" their predictions (in the sense that $\overline{PFE}_o$ is closer to zero). Only SH and ZG have a worse performance on the shorter timespan. This suggests that Swiss cantons have improved their direct tax revenue predictions over time. However, table (4) suggests these predictions could have been even better (in terms of bias) since $\overline{PFE}_s$ is closer to zero than $\overline{PFE}_o$ for all cantons without exception. Interestingly we also notice that PFE_o of canton BE has now become positive on average, which is the only such case among all cantons and over both periods. However, even in this case, $\overline{PFE}_s$ is closer to zero.

Of course all the comments and observations that have been made on these statistics are fragile since they do not rely on any statistical test. Especially, to know whether $\overline{PFE}_o$ is statistically different from zero or $\overline{PFE}_s$ is statistically equal to zero, we have to perform t-tests as explained in Section 2. We also mentioned that for t-tests to be in-sample reliable, we need the data to be normally distributed. For this reason we have reported in Table (3) and (4) median, skewness and kurtosis. By definition, the normal distribution has mean=median, skewness = 0 and kurtosis = 3. Just by looking at the median, skewness and kurtosis

[10]Note that in the case of GE, we could not find any significant ARIMA model. Consequently, we could not generate any simulated prediction for GE.

TAB. 4 – Summary statistics for observed and predicted errors 1979-2006

Canton	N		no PFE<0		$\overline{PFE}$		sd		median		sk		kt	
	obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.	obs.	sim.
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
BE	28	27	29%	41%	1.4267	0.2676	3.0542	6.0373	1.8845	1.3621	−0.2600	−1.5727	2.7492	6.6547
UR	28	27	57%	52%	−0.4705	0.1234	5.8681	5.2524	−0.8289	−0.2578	0.1960	0.1863	2.4818	3.0022
ZH	28	27	57%	41%	−0.5229	0.1754	4.4836	6.3102	−0.4531	0.4456	−0.8865	−0.3675	3.7932	3.2704
SH	28	27	61%	48%	−0.6644	−0.5181	2.8511	4.1914	−0.7452	0.1635	0.4991	0.0283	2.2555	3.7164
JU	28	27	50%	37%	−0.6923	0.1394	2.8478	4.2672	−0.1628	0.6177	−0.4893	−0.5462	2.3255	2.9211
SO	28	27	61%	56%	−0.7495	−0.5367	5.7709	6.9363	−1.1578	−1.8469	0.6715	0.3196	4.1734	2.7449
NE	28	27	61%	44%	−1.0332	0.0279	3.6837	4.0635	−1.8700	0.4369	0.6179	−0.8595	2.6650	4.6394
GE	28		54%		−1.0357		5.6905		−0.3487		0.3593		3.2822	
AR	28	27	71%	56%	−1.2049	0.1275	3.2119	4.7813	−1.9277	−0.1352	0.9303	0.1780	4.0641	2.5424
TG	28	27	68%	59%	−1.4436	−0.3342	3.9784	4.4527	−1.2969	−1.3001	0.2768	0.1008	2.5105	2.2793
BS	28	27	61%	59%	−1.6088	−0.1818	4.8158	3.9226	−1.4343	−0.7530	−0.1279	0.0477	2.0033	2.1477
VD	28	27	64%	44%	−1.7489	0.2288	4.7342	6.3808	−2.9800	1.1934	0.8480	−0.3399	2.8974	2.7151
LU	28	27	64%	52%	−1.8285	−0.2111	4.0150	6.1992	−1.8743	−0.6334	−0.1682	−0.9244	2.2569	6.7688
AG	28	27	75%	48%	−2.0440	0.2241	4.3138	4.1034	−2.4835	0.3240	0.3921	−0.1597	2.7601	2.7296
SG	28	27	75%	0%	−2.2300	0.1106	3.3596	4.7512	−2.7070	0.5315	0.6462	−0.4890	3.1358	4.1319
VS	28	27	86%	52%	−3.2110	−0.5101	4.1574	4.6962	−3.2235	−0.3872	0.8966	0.0681	4.6718	2.7373
BL	28	27	75%	48%	−3.5503	0.2285	4.7756	5.9488	−3.4664	0.2756	−1.0023	−0.4046	4.2031	2.7383
GR	28	27	89%	59%	−3.7291	0.1105	2.4726	4.6560	−3.8736	−0.9203	0.4881	0.7407	2.9056	2.8892
GL	27	25	67%	56%	−3.8356	−0.3729	5.5717	3.8660	−3.6438	−0.6623	0.0030	0.1143	1.8687	2.2467
NW	27	25	70%	28%	−4.0195	1.0699	5.2494	10.3641	−3.4124	4.6265	−0.7786	−1.5015	3.6129	4.7036
FR	28	27	89%	56%	−4.1864	0.0381	3.5852	4.1830	−3.9074	−0.5659	−0.5131	−0.0939	3.1664	2.8276
OW	28	27	79%	56%	−4.3629	−0.4039	6.8493	7.7431	−5.6270	−0.6470	0.9520	0.1188	3.8837	3.6877
SZ	28	27	79%	56%	−4.5390	0.2235	6.8596	8.9568	−6.3901	−1.0387	0.9779	1.6159	3.0852	7.2290
TI	28	27	79%	56%	−4.5740	−0.4244	7.5619	8.2964	−5.3709	−0.9364	0.3904	0.2374	2.3677	2.7568
ZG	28	27	89%	44%	−7.3403	−0.2950	7.1158	6.0842	−5.9713	0.7008	−0.3201	−0.4201	2.7313	3.6139
AI	28	27	89%	44%	−7.8229	0.7537	6.1600	11.7330	−9.1769	0.8701	0.1452	−0.4756	2.4285	5.2684

in the tables, it is sometimes possible to tell that PFE is normally distributed. See, for example, PFE_o of LU over $1945-2006$ (Columns 5, 9, 11 and 13 in Table (3)). Clearly, it is not possible to do so for every canton and/or over every timespan. Consequently, we had to perform a formal test of normality as described in Section 2. Results of the different tests are presented in the next Section.

4 Results

Results are summarized in Table (5) for the whole period and in Table (6) for the shorter timespan (1979-2006). Both tables contain symbols of the cantons in the first column ranked according to the observed mean (Column No 1), p-values of the t-tests (Columns 2-4 and 6-8) and p-values of both normality tests (Columns 9-12) for observed, as well as for simulated errors.

4.1 Biasedness tests on forecast errors over 1945-2006

Results for biasedness and normality tests are reported in Table (5). For observed and simulated errors, Table (5) reports the mean and then p-values of the t-test associated with each alternative hypothesis (H_0 being mean PFE=0). The last four columns report the p-values of the normality tests. Tests show that, except for SH and JU, tax revenues appear to be significantly underestimated on average in all the cantons, i.e. we neither can reject the hypothesis that $\overline{PFE}_o \neq 0$ nor that $\overline{PFE}_o < 0$, whereas $\overline{PFE}_o > 0$ is strongly rejected. This gives strong evidence in favor of the hypothesis we seek to test. Of course, these results need to be differentiated. First, because of non-normality (the nul hypothesis of normality is strongly rejected), results for BE, BL and GR cannot be relied on in-sample and are weaker than for the other cantons. On a lesser extend, this comment also applies to NE. However, in this case S-K test shows that the hypothesis of normality cannot be rejected at the 2% level of significance. The same kind of comment applies to AI,UR,

Tab. 5 – T-tests and normality tests on errors 1944-2006

Canton	t-test								Normality Tests			
	Obs.				Sim.				J-B		S-K	
	$\overline{PFE}$	$\overline{PFE}$<0	$\overline{PFE}\neq 0$	$\overline{PFE}$>0	$\overline{PFE}$	$\overline{PFE}$< 0	$\overline{PFE}\neq 0$	$\overline{PFE}$>0	Obs.	Sim.	Obs.	Sim.
	1	2	3	4	5	6	7	8	9	10	11	12
SH	−0.0712	0.4729	0.9457	0.5271	−0.1604	0.4419	0.8838	0.5581	0.0000	0.0000	0.0000	0.0190
JU	−0.6923	0.1046	0.2092	0.8954	0.1394	0.5667	0.8665	0.4333	0.4386	0.5093	0.3626	0.3452
ZH	−1.6506	0.0064	0.0127	0.9936	0.0269	0.5112	0.9775	0.4888	0.4867	0.0714	0.3139	0.0549
BE	−2.5201	0.0006	0.0012	0.9994	0.2704	0.6162	0.7676	0.3838	0.0017	0.2953	0.0060	0.1726
GE	−3.1900	0.0003	0.0006	0.9997	0.3842	0.6345	0.7310	0.3655	0.3913	0.0000	0.2705	0.0001
SG	−3.6931	0.0000	0.0000	1.0000	−0.5911	0.2573	0.5147	0.7427	0.3594	0.4930	0.2164	0.3126
TG	−3.9734	0.0000	0.0000	1.0000	0.2420	0.6229	0.7542	0.3771	0.4746	0.0545	0.3921	0.0933
AR	−4.1889	0.0000	0.0000	1.0000	−0.3265	0.3809	0.7618	0.6191	0.8999	0.0000	0.9351	0.0001
VD	−4.4271	0.0000	0.0000	1.0000	0.2079	0.5932	0.8137	0.4068	0.0960	0.8446	0.0815	0.8228
AG	−4.7514	0.0000	0.0000	1.0000	0.2199	0.5900	0.8201	0.4100	0.5894	0.1269	0.5565	0.7623
LU	−4.7862	0.0000	0.0000	1.0000	−0.1827	0.4149	0.8299	0.5851	0.8366	0.6195	0.7538	0.4254
BL	−4.7995	0.0000	0.0000	1.0000	0.6072	0.6858	0.6283	0.3142	0.0029	0.0000	0.0079	0.0015
SO	−5.1566	0.0000	0.0000	1.0000	0.0868	0.5347	0.9305	0.4653	0.2033	0.6978	0.1283	0.4865
GR	−5.2106	0.0000	0.0000	1.0000	0.1649	0.5684	0.8632	0.4316	0.0000	0.9630	0.0000	0.9338
BS	−5.4941	0.0000	0.0000	1.0000	0.1826	0.5770	0.8461	0.4230	0.7250	0.0000	0.7353	0.0030
NE	−6.6315	0.0000	0.0000	1.0000	−0.2775	0.4064	0.8128	0.5936	0.0013	0.0000	0.0281	0.0106
NW	−6.6695	0.0000	0.0000	1.0000	−0.2892	0.4205	0.8410	0.5795	0.5409	0.0029	0.5222	0.0075
FR	−6.8847	0.0000	0.0000	1.0000	0.0919	0.5393	0.9213	0.4607	0.0694	0.0337	0.0589	0.0396
SZ	−6.9757	0.0000	0.0000	1.0000	−0.6191	0.3092	0.6184	0.6908	0.1298	0.0000	0.0858	0.0038
ZG	−7.2603	0.0000	0.0000	1.0000	−0.0756	0.4724	0.9448	0.5276	0.9740	0.8605	0.8151	0.9100
VS	−7.4786	0.0000	0.0000	1.0000	0.2477	0.5926	0.8148	0.4074	0.4534	0.0010	0.2669	0.0112
TI	−8.3113	0.0000	0.0000	1.0000	0.0907	0.5343	0.9314	0.4657	0.7275	0.7775	0.7718	0.8500
GL	−8.4319	0.0000	0.0000	1.0000	0.0943	0.5407	0.9186	0.4593	0.7352	0.0002	0.7791	0.0023
OW	−9.1609	0.0000	0.0000	1.0000	0.6763	0.6506	0.6987	0.3494	0.3178	0.0000	0.1890	0.0000
AI	−9.4869	0.0000	0.0000	1.0000	0.5939	0.6597	0.6806	0.3403	0.0106	0.0024	0.0548	0.0548
UR	−9.6438	0.0000	0.0000	1.0000	1.9360	0.8672	0.2656	0.1328	0.0432	0.0000	0.0477	0.0011

VD, FR and SZ. For all other cantons except JU and SH, i.e. 15 cantons over 26, we have extremely strong evidence of systematically underestimated tax revenue. The key question is : were we able to reduce, using simple univariate time-series techniques, biasedness significantly ? From table (5), we can distinguish different groups of cantons. JU, BE, SG, VD, AG, LU, SO, GR, ZG and TI are cantons for which $\overline{PFE}_s$ is clearly closer to zero than $\overline{PFE}_o$, t-tests never reject $\overline{PFE}_o = 0$ and normality is clearly not rejected. For these nine cantons we can clearly improve tax revenue predictions by reducing biasedness[11]. Then we find ZH, TG, FR, NE, VS and AI. Although results for these cantons are also very strong, it is possible to find some levels of significance for which normality is rejected. For GE, AR, BL, BS, NW, SZ, VS, GL, OW and UR normality is clearly rejected by both test procedures. Thus, results for these cantons have to be taken carefully although, given the number of observations, we might be confident that, asymptotically, tests indicate that biasedness could be decreased. We now turn to the cases of JU and SH. For both cantons the hypothesis that $\overline{PFE}_o = 0$ cannot be rejected. This suggests that, on average, SH and JU have neither underestimated nor overestimated tax revenue. Results for canton JU are the strongest since normality cannot be rejected. This provides strong evidence that canton JU, in forecasting tax revenue, made no error on average. Note that our results clearly show that, although JU made no error on average, predictions could have been improved. Conversely, results for SH are clearly less reliable since normality is strongly rejected. We now perform tests over a shorter time span for all the cantons. Results are presented in the next subsection.

[11] As already mentioned JU was created by partition from canton BE. Thus, in the year 1979 we might expect to find a structural break in the data of canton Bern. Taking into account this break could even improve further the results for BE further.

4.2 Biasedness tests on forecast errors over 1979-2006

In this section, we test whether Swiss cantons improved forecasting accuracy over time by reducing biasedness. Results are shown in Table (6). It reports the same information as Table (5). First note that

TAB. 6 – T-tests and normality tests on errors 1979-2006

Canton	t-test								Normality Tests			
	Obs.				Sim.				J-B		S-K	
	$\overline{PFE}$	$\overline{PFE}<0$	$\overline{PFE}\neq 0$	$\overline{PFE}>0$	$\overline{PFE}$	$\overline{PFE}<0$	$\overline{PFE}\neq 0$	$\overline{PFE}>0$	Obs.	Sim.	Obs.	Sim.
	1	2	3	4	5	6	7	8	9	10	11	12
BE	1.4267	0.9900	0.0200	0.0100	0.2676	0.5902	0.8196	0.4098	0.7817	0.0000	0.8444	0.0008
UR	−0.4705	0.3373	0.6747	0.6627	0.1234	0.5481	0.9038	0.4519	0.1856	0.9249	0.1163	0.7684
ZH	−0.5229	0.2712	0.5423	0.7288	0.1754	0.5569	0.8862	0.4431	0.4386	0.7082	0.3626	0.4252
SH	−0.6644	0.1141	0.2282	0.8859	−0.5181	0.2632	0.5263	0.7368	0.3734	0.7479	0.1958	0.3910
JU	−0.6923	0.1046	0.2092	0.8954	0.1394	0.5667	0.8665	0.4333	0.1564	0.5093	0.0708	0.3452
SO	−0.7495	0.2489	0.4978	0.7511	−0.5367	0.3455	0.6909	0.6545	0.5706	0.7661	0.4148	0.7073
NE	−1.0332	0.0747	0.1494	0.9253	0.0279	0.5141	0.9718	0.4859	0.0412	0.0418	0.0266	0.0292
GE	−1.0357	0.1720	0.3440	0.8280					0.4868		0.1485	
AR	−1.2049	0.0287	0.0574	0.9713	0.1275	0.5546	0.8908	0.4454	0.6783	0.8277	0.6561	0.8925
TG	−1.4436	0.0327	0.0655	0.9673	−0.3342	0.3499	0.6997	0.6501	0.7272	0.7298	0.7512	0.7468
BS	−1.6088	0.0442	0.0884	0.9558	−0.1818	0.4058	0.8116	0.5942	0.0765	0.6612	0.0409	0.5654
VD	−1.7489	0.0305	0.0610	0.9695	0.2288	0.5732	0.8536	0.4268	0.0300	0.7367	0.0244	0.6819
LU	−1.8285	0.0115	0.0230	0.9885	−0.2111	0.4304	0.8609	0.5696	0.5392	0.0001	0.2878	0.0049
AG	−2.0440	0.0092	0.0185	0.9908	0.2241	0.6106	0.7789	0.3894	0.1107	0.9062	0.0540	0.9131
SG	−2.2300	0.0008	0.0016	0.9992	0.1106	0.5477	0.9046	0.4523	0.7868	0.2840	0.8597	0.1162
VS	−3.2110	0.0002	0.0004	0.9998	−0.5101	0.2887	0.5773	0.7113	0.3843	0.9519	0.2879	0.9731
BL	−3.5503	0.0003	0.0005	0.9997	0.2285	0.5783	0.8434	0.4217	0.1069	0.6657	0.0699	0.5789
GR	−3.7291	0.0000	0.0000	1.0000	0.1105	0.5486	0.9028	0.4514	0.5323	0.2890	0.3079	0.1676
GL	−3.8356	0.0007	0.0014	0.9993	−0.3729	0.3170	0.6339	0.6830	0.7549	0.7241	0.7041	0.7310
NW	−4.0195	0.0002	0.0005	0.9998	1.0699	0.6948	0.6105	0.3052	0.4047	0.0020	0.3093	0.0047
FR	−4.1864	0.0000	0.0000	1.0000	0.0381	0.5187	0.9627	0.4813	0.2070	0.9641	0.0896	0.9309
OW	−4.3629	0.0011	0.0023	0.9989	−0.4039	0.3943	0.7885	0.6057	0.0686	0.7425	0.0374	0.3874
SZ	−4.5390	0.0008	0.0016	0.9992	0.2235	0.5511	0.8978	0.4489	0.6755	0.0000	0.5877	0.0005
TI	−4.5740	0.0017	0.0035	0.9983	−0.4244	0.3963	0.7925	0.6037	0.5549	0.8521	0.5132	0.8203
ZG	−7.3403	0.0000	0.0000	1.0000	−0.2950	0.4015	0.8031	0.5985	0.7063	0.5438	0.4254	0.2454
AI	−7.8229	0.0000	0.0000	1.0000	0.7537	0.6294	0.7412	0.3706	0.8233	0.0333	0.7889	0.0434

for the canton of GE, we could not significantly fit any kind of ARIMA model. Consequently we could not simulate new forecasts and compute new errors. For this reason Table (6) exhibits blanks instead of simulated values for GE. Then, although results of normality tests on observed errors are to some extent weaker for cantons of NE, BS, BL, VD and OW, normality cannot be rejected for any canton (see Columns 9 and 11 in Table (6)). Thus t-test on $\overline{PFE}_o$ are highly reliable. Looking at Table (6), Columns 1-4, allows one to distinguish four "groups" of cantons. The first "group" of cantons is only the canton BE. For this canton, t-test shows that tax revenue are significantly overestimated. In the second group, we find ZH, UR, SO, SH, GE and JU. For these cantons, the hypothesis that mean forecast error is equal to zero cannot be rejected. Hence it appears that these canton did not make systematic errors in predicting tax revenue over 1979-2006. Then we find NE, AR, TG, BS, VD and LU. For these cantons, although we find some significance level for which $\overline{PFE}_o < 0$ can be rejected, t-tests nevertheless show that they underestimated tax revenue over 1979-2006. For the 13 other cantons, tests strongly show that direct tax revenue has been systematically underestimated. Although Table (6) provides evidence that prediction of tax revenue has been improved over time, still tax revenue was significantly underestimated in 19 cantons over 26. Thus, one might ask whether predictions over 1979-2006 could have been improved in terms of bias by using simple univariate time-series methods. PFE_s is reported in Table (6). Note first that normality can be clearly rejected for BE, LU, NW and SZ. Given the small number of observations in the sample, we cannot rely on the t-tests for these cantons. For all other cantons, we can assume normality and t-tests strongly show that $\overline{PFE}_o = 0$ cannot be rejected. This means that, even over 1979-2006, tax revenue predictions could have been improved (in term of biasedeness) in 21 cantons over 26 by using simple univariate time-series model.

4.3 Synthesis of the results

Table (7) summarizes the results that we have found. It classifies the cantons depending on whether under-, over- or perfect estimation has been observed respectively simulated on average and depending on the statistical significance of the results[12]. Over the period 1945-2006, Table (7) shows that tax revenue

Tab. 7 – Synthesis of the results

	45-06						79-06				
	Obs.			Sim.			Obs.		Sim.		
	Strong	Weak	NN	Strong	Weak	NN	Strong	Weak	Strong	Weak	NN
Overest.							BE				
Zero Error	JU		SH	JU BE SG VD AG LU SO GR ZG TI	ZH TG FR NE VS	GE AR BL BS NW SZ VS GL OW UR SH	ZH UR SO SH GE	JU	UR ZH SH JU SO NE AR TG BS VD AG SG VS BL GR GL FR OW TI ZG AI		LU NW SZ BE
Underest.	ZH GE SG TG AR AG LU SO BS NW ZG VS TI GL OW	NE AI UR VD FR SZ	BE BL GR				AG SG VS GR GL NW FR SZ TI ZG AI	NE AR TG BS VD LU BL OW			
Total	16	6	4	10	5	11	17	9	21		4

predictions suffered from underestimation in 24 cantons over 26. Although prediction errors turned out to be non-normal for three cantons, results can be considered as asymptotically valid. We also showed that for a majority of cantons, predictions were improved in terms of bias by using simple ARIMA models. For all cantons, simulated prediction errors turned out to be zero on average. Then Table (7) shows that, over 1979-2006, a smaller number of cantons did underestimate tax revenues systematically, although it was still the dominant case (19 cantons). Nevertheless, we showed that room for improvement in tax predictions also exists when considering a shorter timespan. Table (7) shows that simulated tax revenue forecasting errors statistically do not suffer from any bias in a majority of cantons. These results suggest that Swiss cantons could easily reduce problems of bias in tax revenue by implementing some simple regression techniques that do not require costly information. However, a direct use of the models that have been derived in this paper seems difficult because of some important limitations in the analysis. These limitations are observed in the final section.

5 Conclusions

To conclude, let us first recall the objectives and results of this paper. Then we will shed light on some limitations of our analysis. These caveats should be preliminary steps to some further research.

In this paper we sought to test whether predictions of tax revenue in Swiss cantons exhibit systematic downward bias. Then, using simple univariate time-series models, we tried to generate new tax revenue predictions that outperform (in terms of bias) predictions made by Swiss cantons. Using t-tests and Jarque-Bera normality tests, we have shown that tax revenue predictions are downward biased in a majority of Swiss cantons. Although a majority of cantons have clearly reduced this bias over time, we still have strong evidence of biasedness for the shorter timespan. Furthermore, we showed that, for a majority of cantons over both timespans, using simple univariate methods allowed us to improve average forecast error by reducing biasedness. In spite of these statistically strong and interesting results, the current analysis still suffers from important limitations.

[12]Cantons have been classified in columns weak or strong depending on the strength of the statistical tests. NN in Table (7) states for non-normal and classifies all the cantons for which errors were clearly non-normal

A first limitation of the model we have used to simulate forecast errors is that, in most of the cantons, R_{a_t} is not (entirely) known at the time $R_{a_{t+1}}$ is predicted. Thus cantons would be forced to predict with $R_{a_{t-1}}$ or with an estimate of R_{a_t} instead of R_{a_t}, making predictions less accurate than our simulations show. Note that this limitation does not apply to cantons for which AR(2) models are significant[13]. This point also emphasizes the importance for cantonal administrations to be as quick as possible in taxing taxpayers in order to get as much information about R_{a_t} as possible[14]

The most important limitation of the analysis comes from the fact that we strictly focused on the problem of the bias. Actually we did not perform any analysis on the variance. In any forecasting methods, variance turns out to be a key determinant of the accuracy of predictions. Even if our predictions errors are centered on zero, it is desirable that they have a variance as small as possible. One would easily understand that, even if tax revenue forecast errors are zero on average, huge variations year after year could create important problems in the elaboration of public budgets. Unfortunately, ARIMA models did not allow us to reduce variance of simulated prediction errors compared to observed errors (see standard deviation in Tables (3) and (4)). This is clearly not satisfying. There may be several ways to improve this point of our analysis. On the one hand, one could try to better exploit information contained in the series. For instance, tests for structural breaks in the variance could be performed. Taking such breaks into account if they exist can clearly contribute to improve forecasts. Note also that our ARIMA models have been fitted using maximum likelihood procedure. However, we know that in forecasting, under ergodicity of second moments of the time series process, OLS provide us with a consistent estimate of the best linear projection of the variable we seek to forecast and the linear projection turns out to produce the smallest mean squared error among the class of linear forecasting rules (Hamilton, pp.74-76). Thus, under some regularity conditions, we could expect to reduce variance by using OLS estimator instead of ML. In this respect, an interesting way of improvement could be to consider a model with Newey-West standard errors, which are consistent with non-normal errors. Such a model would allow us to both take into account the information given when errors are non-normal and simultaneously fit the model through OLS.

On the other hand, we could improve prediction by adding regressors such as GDP, unemployment rate, inflation, etc. in the model. We can reasonably expect such regressors to significantly reduce variance of prediction errors. A practical drawback of including additional regressors is that this information is costly to collect and, in some cases, has to be forecast as well and may hence generate more bias.

To conclude, one last point that is worth emphasizing is that our analysis shows the existence of a bias in cantonal tax revenue forecasts but does not tell anything about the origin of such a bias - that was not the aim of the current analysis. Biasedness may be due to numerous different factors : intrinsic characteristics of the cantons (institutions, budget process...), bias in other forecasts (GDP, Inflation...), behavioural factors (asymmetric loss function, strategic behaviour...). Understanding these factors may clearly be helpful in trying to improve tax revenue forecasting and budgeting process in the Swiss cantons. To address this issue, panel data models appear to be the most appropriate method to use. Such a panael data analysis of tax revenue forecasting errors is a research avenue that we are currently exploring.

[13] This is the case for SZ, AR and NE over 1979-2006 (see appendix Table A-2).

[14] This observation gives clearly support to the requirements of IPSAS 23 (IPSAS = International Public Sector Accounting Standard). According to this standard, governments should ideally book tax receipts for year t on the basis of the tax base of year t. Thus governments have no choice but shorten the gap between the time of tax collection and the tax base creation (IFA, 2008, pp. 693-710).

Références

Y. Amman, Est-il possible d'améliorer les prévisions budgétaires de la Confederation : Seco, (2001), Bern.

A.J.Auerbach, On the performance and use of Government Revenue Forecasts. National Tax Journal, Vol 52, No 4, (1999), pp 767-782

S. Bretschneider & W. Gorr, Economic organizational and political influences on biases in forecasting state sales tax receipts. International Journal of Forecasting, Vol 7, No 4, (1992) pp. 457-466.

B. Campbell & E.Ghysels, An empirical analysis of the canadian budget process. Canadian Journal of Economics, Vol 30, No 3, (1997), pp. 553

G. Cassidy , M.S. Kamlet & D.S. Nagin, An empirical examination of bias in revenue forecasts by state governments. International Journal of Forecasting, Vol 5, No 3, (1989), pp. 321-331

R. B. D'Agostino, A. Balanger & R. B. D'Agostino Jr., A suggestion for using powerful and informative tests of normality, American Statistician, Vol 44, (1990), 316-321.

S. Danninger, Revenue Forecasts as Performance Targets. IMF Working Paper Series(05/14), (2005).

S.Danninger, and A.J. Kyobe, The Political Economy of Revenue-Forecasting Experience from Low-Income Countries. IMF Working Paper Series(05/02), (2005).

D.A.Dickey & W.A.Fuller, Distribution of the estimators for autoregressive time-series with a unit root. Journal of the American Statistical Association, Vol 74, (1979), pp. 427-431.

D.R. Feenberg, D. Gilroy & H.S. Rosen, Testing Rationality of State Revenue Forecasts. The Review of economics and Statistics, Vol 71, No 2, (1989), pp. 300-308.

J.D.Hamilton, Times Series Analysis. Princeton University Press, Princeton NJ, 1994.

International Federation of Accountants, Handbook of International Public Sector Accounting Pronouncements, New York, vol. II, 2008.

C.M. Jarque & A.K. Bera, A Test for Normality of Observations and Regression Residuals, International Statistical Review, Vol. 55, No 2, (Aug., 1987), pp. 163-172.

H.N.Mocan & S. Azad, Accuracy and rationality of state General Fund Revenue forecasts : Evidence from panel data. International Journal of Forecasting, Vol 11, No 3, (1995), pp. 417-427.

R.S. Pindyck & D.L. Rubinfeld, Econometric Models and Economic Forecasts. McGraw-Hill, New York NY, 1998.

R. Rodgers & P.Joyce, The Effect of Underforecasting on the Accuracy of Revenue Forecasts by State Governments. Public Administration Review. Vol 56, 1, (1991), pp. 48-56.

P. Royston, Comment on sg 3.4 and an improved D'Agostino test. Stata Technical Bulletin 3, (1991), pp 23-24

N. Soguel, Comparatif 2007 des finances cantonales et communales, Cahier de l'IDHEAP, no 242, (2008).

M.J.C.M. Verbeek, A Guide to Modern Econometrics, 2nd edition. Chichester : John Wiley and Sons, (2004).

APPENDIX

Tab. A-1 – ARIMA model results 1944-2006

canton	method used	cons	coeff. AR(1)	coeff(MA1)	DF-test(p-value)
ZH	AR(1)MA(1) nocons		0.9894596***	-0.859967***	0.0000
			(0.0179641)	(0.1020279)	
BE	AR(1)MA(1)	0.061843***	-0.9577577***	0.8299076***	0.0000
		(-0.0099776)	(0.0582297)	(0.1365259)	
LU	AR(1)MA(1) nocons		0.9872651***	-0.8514741***	0.0000
			(0.0227859)	(0.1142254)	
UR	AR(1)MA(1) nocons		0.9931184***	-0.9096825***	0.0000
			(0.0176075)	(0.1157158)	
SZ	AR(1)MA(1) nocons		0.955726***	-0.7013401***	0.0000
			(0.0549996)	(0.1420258)	
OW	AR(1)MA(1) nocons		0.9932304***	-0.9275322***	0.0000
			(0.0237506)	(0.1177072)	
NW	AR(1)MA(1) nocons		0.9941228***	-0.9085342***	0.0000
			(0.0188012)	(0.0939459)	
GL	AR(1)MA(1) nocons		0.9798416***	-0.7902278***	0.0000
			(0.0276081)	(0.1305887)	
ZG	AR(1)MA(1) nocons		0.9899208***	-0.8313499***	0.0000
			(0.012801)	(0.0803177)	
FR	AR(1)MA(1)nocons		0.9914453***	-0.8647267***	0.0000
			(0.0177276)	(0.0894988)	
SO	AR(1)MA(1) nocons		0.9887953***	-0.8624168***	0.0000
			(0.0163265)	(0.0824084)	
BS	AR(1)MA(1) nocons		0.9852106***	-0.7778171***	0.0000
			(0.0158463)	(0.1147834)	
BL	AR(1)MA(1) nocons		0.9951843***	-0.9145683***	0.0000
			(0.0110469)	(0.0863745)	
SH	AR(1)MA(1) nocons		0.9861978***	-0.879345***	0.0000
			(0.0218255)	(0.0838988)	
AR	AR(1)MA(1) nocons		0.9714549***	-0.7671259***	0.0000
			(0.0330086)	(0.1215421)	
AI	AR(1)MA(1)	0.088236***	0.1318599	-0.2644418	0.0000
		(-0.0147544)	(1.409142)	(1.337208)	
SG	AR(1)MA(1) nocons		0.9653123***	-0.6889746***	0.0000
			(0.0327365)	(0.1211078)	
GR	AR(1)MA(1)	0.0644336***	0.8911417***	-0.7652912***	0.0000
		(-0.0196649)	(0.1735057)	(0.2484318)	
AG	AR(1)MA(1) nocons		0.9894985***	-0.8436429***	0.0000
			(0.018097)	(0.0967774)	
TG	AR(1)MA(1)	0.0715281***	0.5794583***	-0.1083805***	0.0002
		(-0.0165608)	(0.2602717)	(0.2979022)	
TI	AR(1)MA(1) nocons		0.985456***	-0.814234***	0.0000
			(0.0194067)	(0.0922969)	
VD	AR(1)MA(1)	0.07801***	-0.9776691***	0.8748302***	0.0000
		(-0.0099415)	(0.033677)	(0.0797914)	
VS	AR(1)MA(1) nocons		9917771***	-0.8741514***	0.0000
			(0.0135404)	(0.1013881)	
NE	AR(1)MA(1) nocons		0.9991894***	-0.9695241***	0.0000
			(0.0056523)	(0.1003446)	
GE	AR(1)MA(1) nocons		0.9941285***	-0.8965158***	0.0000
			(0.0116182)	(0.0946365)	
JU	AR(1)MA(0)		-0.5459671***		0.0000
			(0.2396568)		

Tab. A-2 – ARIMA model results 1979-2006

	canton	method used	cons	coeff. AR(1)	coeff(MA1)	coeff. AR(2)	coeff(MA2)	DF test(p-value)
1	ZH	AR(1)MA(0)	0.0356238***	-0.2069145***				0.0000
			(0.0120848)	(0.2827397)	0.5501833			
2	BE	AR(1)MA(1)	0.0423546***	-0.8140161**	(0.6031119)			0.0000
			(0.0155909)	(0.3569043)	-0.7480326***			
3	LU	AR(1)MA(1) nocons		0.9668378***	(0.3054635)			0.0008
				(0.076877)				
4	UR	AR(1)MA(0)	0.0330502***	0.2187822**				0.0023
			(0.0127478)	(0.182202)	0.9999***			
5	SZ	AR(2) MA(1)	0.0418061***	-0.836896***		0.256745***		0.0000
			(0.0286644)	(0.216501)		(0.231522)		
6	OW	AR(1) MA(1)	0.0569233***	0.668328***				0.0002
			(0.0038817)	(0.2470474)				
7	NW	AR(1) MA(1)		0.9969204***				0.0026
				(0.0225981)	-0.7889315***			
8	GL	AR(1)MA(1)nocons		0.9655815***	(0.1760703)			0.0016
				(0.0534989)	-1.000001			
9	ZG	AR(1) MA(1)	0.0568639***	0.473269*				0.0000
			(0.0034666)	(0.2646777)	-0.8141014***			
10	FR	AR(1) MA(1)nocons		0.9916915***	(0.1509705)			0.0001
				(0.016424)	-0.826704***			
11	SO	AR(1) MA(1) nocons		0.9677798***	(0.1683543)			0.0001
				(0.0537482)	-0.9999944			
12	BS	AR(1) MA(1)	0.0319789***	0.7556258*	(3087.556)			0.0000
			(0.0033346)	(0.3618537)				
13	BL	AR(1) MA(0)	0.0488855***	-0.496255***				0.0000
			(0.009027)	(0.2571266)	-0.709285***			
14	SH	AR(1) MA(1) nocons		0.9528165***	(0.2063211)			0.0024
				(0.0579583)	0.183185***			
15	AR	AR(2)MA(2)	0.441982***	-0.142419***	(0.53085)	0.427437***	0.046982	0.0000
			(0.169185)	(0.465471)		(0.459318)	(0.538428)	
16	AI	AR(1)MA(0)	0.0655644***	-0.3013547				0.0000
			(0.0210241)	(0.3193722)				
17	SG	AR(1)MA(0)	0.0561276***	0.0081785				0.0000
			(0.0099364)	(0.2484244)				
18	GR	AR(1) MA(0)	0.0434481***	0.3847971**				0.0204
			(0.0165102)	(0.1908247)				
19	AG	AR(1)MA(1) nocons		0.9938046***	-0.860008***			0.0000
				(0.0143722)	(0.1787145)			
20	TG	AR(1)MA(1) nocons		0.960564***	-0.721493***			0.0044
				(0.0558615)	(0.1954423)			
21	TI	AR(1)MA(1) nocons		0.999854***	-0.9931472			0.0000
				(0.0836552)	(1.964649)			
22	VD	AR(1)MA(0)	0.048456***	-0.3014725				0.0000
			(0.0100213)	(0.2010633)				
23	VS	AR(1) MA(1)	0.0394637***	0.7932158***	-1			0.0000
			(0.0040149)	(0.2199898)				
24	NE	AR(2) MA(2)	0.049767***	-0.6955735	0.9074634	-0.8575301***	0.4597144	0.0000
			(0.0286644)	(0.695545)	(0.8452741)	(0.4051926)	(0.856234)	
26	JU	AR(1)MA(0)		-0.5459671***				0.0000
				(0.2396568)				

Tab. A-3 – List of Swiss cantons

Name	Acronym
Aargau	AG
Appenzell Ausserrhoden	AI
Appenzell Innerrhoden	AR
Basel-Landschaft	BE
Basel-Stadt	BL
Bern	BS
Fribourg	FR
Geneva	GE
Glarus	GL
Graubunden	GR
Jura	JU
Lucerne	LU
Neuchâtel	NE
Nidwalden	NW
Obwalden	OW
SanktGallen	SG
Schaffhausen	SH
Schwyz	SO
Solothurn	SZ
Thurgau	TG
Ticino	TI
Uri	UR
Valais	VD
Vaud	VS
Zug	ZG
Zurich	ZH