

TOURNAMENTS IN THE PUBLIC SECTOR

A Thesis
Presented to
The Academic Faculty

by

Celso Vila Nova de Souza Junior

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in the
School of Economics

Georgia Institute of Technology
May 2008

TOURNAMENTS IN THE PUBLIC SECTOR

Approved by:

Dr. Emilson C. D. Silva, Advisor
School of Economics
Georgia Institute of Technology

Dr. Rehim Kilic
School of Economics
Georgia Institute of Technology

Dr. Haizheng Li
School of Economics
Georgia Institute of Technology

Date Approved: March 31, 2008

ACKNOWLEDGEMENTS

There are so many to thank for making this work possible. First and foremost I would like to thank my grandmother Luiza Bezerra de Oliveira, for her love, sacrifice, and support from the beginning since 1978 until nowadays. I am also indebted to my mother, Rosa Amelia Bezerra de Oliveira who has always been the foundation of my life. I want to express my greatest gratitude to my uncles, Madson Flavio, Pedro Afonso, Antonio Bezerra, and aunts, Ana Rosa, Lucia de Fatima and Marlene Louzada. Without their faith in me, to be here would not have been possible. I also would like to thank my wife Kelly Guedes for their love, and who have shared their wisdom and support through all these 10 years. In fact, all these sentiments extend to many other important people, in my family, who help me to be here and achieve my goals.

There is a very special person, Dr. Emilson C. D. Silva, who deserves my eternal thanks and respect for his support and guidance during my course of studies. Since my first course, I have learned a great deal from him and he inspired me to be a high level academic professional. He is my committee chair, academic advisor, and mentor for this work and tricky issues that I have faced. Not only was he always available to provide suggestions that largely improve my work, but also come up with creative ideas that offered me alternatives to understand the economic issues. I will never forget his support and endless patience. I also would like to express my gratitude to my committee members, Dr. Rehim Kilic and Dr. Haizheng Li. I wish to extend many thanks to excellent faculty at Tulane University, in special Dr. Claudiney Pereira and Catholic University of Brasilia, in special Dr. Paulo Roberto Amorin Loureiro.

I am very lucky and grateful to have wonderful and understanding friends in Brazil for their encouragement and their mentoring.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	III
LIST OF TABLES	V
SUMMARY.....	VI
CHAPTER 1 - INTRODUCTION AND LITERATURE REVIEW	1
1.1. Introduction	1
1.2. Literature Overview on the Tournaments Theory.....	5
CHAPTER 2 – DATA.....	9
2.1. Approach and Data	9
2.1.1. Variables in the core model.....	11
CHAPTER 3 – EMPIRICAL METHODOLOGY	19
3.1. GMM Estimators	19
3.2. Hypotheses.....	24
CHAPTER 4 – RESULTS	27
CHAPTER 5 - CONCLUSIONS	34
APPENDIX A	37
REFERENCES.....	39

LIST OF TABLES

Table 1: Group of Tax agencies	10
Table 2: Descriptive Statistics for Dependent and Explanatory Variables	14
Table 3: Detailed Descriptive Statistics for Tax Agencies Ranking	15
Table 4: Correlation Coefficients for Quantitative Variables.....	17
Table 5: Regression results	29
Table A.1: Alternative Estimates of the AR (1) specification	37
Table A.2: AR (1) Model Estimates	38

SUMMARY

Tournament theory shows that a firm may motivate employees by running competitors for rewards either for a group or individualistic schemes. The empirical literature on Tournaments has been grown. However, many studies use no appropriate data. This paper provides the first empirical evidence on two key assumptions in these models using a special case surrounding the incentives for workers in public sector. The dataset contains information from the Coordenação de Fiscalização (i.e., the Inspections Division) of the Secretaria da Receita Federal (SRF) on the bonus program created by the Brazilian government to compensate tax officials for their efforts in collecting taxes and uncovering tax violations. We constructed a larger unbalanced panel data Tax collection containing information upon 110 tax agencies distributed between 10 regions and 45 time period by month, which allowed us to support the predictions raised above. In order to examine the tournaments predictions we emphasize the dynamic of the process taking into account the unobserved heterogeneity and endogeneity problems using appropriate GMM techniques. This enable us to pondered the possible inertia for time adjustments within tax agency, possibly in determining strategies to improve the tax agency performance on the sources most valuable for collection, which supports the hypothesis of learning by doing. The results also demonstrated evidence to support the following tournaments hypothesis: (1) prizes motivate agents to exert effort; and (2) number of participants increased as the size of the prize increase.

CHAPTER 1 - INTRODUCTION AND LITERATURE REVIEW

1.1. Introduction

In the last few decades the Governments in developing countries often have poor performance in collecting tax system. A systematic review and evaluation of issues surrounding incentives for workers and managers in public sector agencies by using bonus and salary supplement systems to enhance effectiveness in tax collection agencies have been a topic of debate in recent policy discussion especially in developing countries where tax evasion is widespread.

In August 1989 the Brazilian Government created a bonus program called *Retribuição Adicional Variável (RAV)* to compensate tax officials based on their individual and group performance for their efforts in collecting taxes and discovering tax evaders. Therefore both types of reward increased with the amount of fines collected giving them incentives to improve their productivity.

An executive committee, called *Comissão de Administração da Retribuição Adicional Variável (CRAV)* created in March 1989 to manage the program, established the goals to be implemented by each tax agency and also supervised and evaluated the performance of each tax agency. There are 111 tax agencies in the country: one central agency, ten regional agencies and 100 local agencies. Kahn, Silva and Zilak (2001) investigated such reform and they found that previously, fine collections per inspection were relatively stable; using a panel data they found that after reform fine collections were about 75% above what would have been before the bonus program.

Over the last several years a literature has developed studies examining which theory of incentive schemes might dominate other forms of incentives. The theory of tournaments as incentive schemes reward players with prizes based upon relative performance. The best performer receives the largest prize; the worst performer

receives the smallest. Rank order tournament contracts make each agent's compensation a function solely of the ranking of his output relative to the outputs of other agents. In contrast, individual piece-rate contracts make each agent's compensation a function only of his own output. The early work of Lazear and Rosen (1981) compared these two payment schemes when agent's action are not observed by the principal, except for the case in which the principal is risk-neutral and the utility function of agents additively separable in income and action. Lazear and Rosen (1981) and Bhattacharya (1982) analyze the rank order contract that result in adjusting for the effects of correlated shocks on the observed performance monitors of all agents. In this context the second best optimality has been proved for a restricted scenario in which marginal productivity of effort is not affected by correlated shocks (Green and Stokey 1983), whereas Lazear and Rosen (1981) and Nalebuff and Stiglitz (1983) have compared tournaments with linear piece-rate contracts.

The limitation of such comparisons have been noted by Holmstrom (1982) who examined a situation where when the principal has many agents, it is not in general efficient to restrict the form of contract either to individual piece-rate contracts or to rank-order contracts. Rank order contracts remain incentive compatible even when information about agents' performance is known only to the principal because the total payment from the principal to all agents taken together is independent of the outcome that occurs. Bhattacharya (1983) has rationalized the use of tournaments even without correlated productivity shocks, but with principal's moral hazard, and compared their performance with that of termination-based incentives schemes. Carmichael (1983) and Malcomson (1984) found similar results that under the appropriate conditions, there is no loss of efficiency from the performance of agents not being public information. The essential feature of some tournament contracts is that the principal's wage bill is predetermined.

Nevertheless Bhattacharya (1984) and Malcomsom (1984) have found results that as long as agents can observe each other's compensation, rank-order contracts can be used to provide incentives for agents to perform well even in a situation of asymmetric information, i.e. where their outputs are not public information provided that the principal has at least some private information about relative performance. That private information need be no more formal than a supervisor's assessment of each agent's performance. Thus, they noted that if the principal has access only to some imperfect and unverifiable measure of performance with which rank agents under a rank order contract gets no financial gain from falsely reporting that ranking.

Rosen (1986) investigate the incentive properties of prizes in sequential elimination events, where rewards are increasing in survival. However he jointed in his work traditional statistic views relative comparisons as an experimental design for ranking and selecting contestants in order to determine the best contestants and promote survival of the fittest; and maintain the quality of play as the game proceeds through its stages. He shows that an elimination design requires an extra reward for the overall winner to maintain performance incentives throughout the game. The chief result is identifying a unique role for top-ranking prizes to maintain performance incentives in career and other games of survival and induce competitors to aspire to higher goals independent of the past achievements.

The empirical literature on Tournaments has been grown. However, many studies use no appropriate data from contexts other than employees in firms, such as baseball, golf, auto racing, agricultural production, and lab experiments.

Unexpectedly, little is known about the effects of the tournaments theory applied in some issues surrounding a specifics government's plans. So far for my best

knowledge the present study provides the first empirical evidence on two major predictions of theory using a rich government collecting tax system program data.

It is important to note that this paper does not attempt to answer the question under what circumstances tournament incentive systems might dominate other forms of incentive systems. The principal designs tournaments and prizes to motivate agents to exert effort. Since competitors exert effort to increase their chance of winning, the prize must compensate for their expending effort (Bognanno, 1994). Our focus is to demonstrate that prize is the main factor to motivate the competitors (Harbling & Irlenbusch, 2003). In this paper we also test the tournament prediction that as the number of participants increase, the size of the prize increase. According to the tournament theory the wage differentials directly compensate workers for higher effort at higher levels of the hierarchy, implying that differentials arise to maintain incentives to lower rungs of the tax agencies. The study also explicitly considers possible causes of the persistence of tax agencies performance over time (ranking inertia).

In order to examine the tournaments predictions of the bonuses paid over the salaries, we constructed an unbalanced dynamic panel data set from information obtained by the Coordenação de Fiscalização (i.e., the Inspections Division) of the Secretaria da Receita Federal (SRF), Brazilian tax authority, from August 1989 to April 1993. There are 111 tax agencies for each of ten tax regions. We have inedited rich data set containing information on fines collected, on the number of personnel involved in the collection, and Tax Agencies ranking.

We find evidences to support the hypothesis that prizes motivate agents, and larger prizes and larger first place prizes as the number of contents increased. We observed both the efficiency and tournament models were indicative that the AFTNs as a whole acted upon ranking pressuring them downwards as the number of AFTNs increased and as the number of tax agencies employing them increased. We also

pondered possible outstanding performance (inertia) for time adjustments within tax agency.

The paper is organized as follows. In the next section, we present a literature overview on the tournaments theory. Section 2, we discuss the data in the overview of the tax collection reform in Brazil. Section 3, we presents the econometric methodology and hypotheses. Section 4, we analyze the results supported by the empirical tournament literature, in the context of the Brazilian incentive reform. Section 5, the conclusions summarize the main findings.

1.2. Literature Overview on the Tournaments Theory

In this section, we discuss empirical studies on the tournament theory, focusing on issues related data, econometric method utilized and the results found.

The evidence of rank-order tournament has been provided some explicit empirical predictions about the hierarchical structure of organization pay. But these findings are not unanimously accepted in the literature and quite number of puzzles still rose. Furthermore, organizations theorists have long argued that economic models are too constrained and those noneconomic used factors affect the results found.

Bull, Schotter and Weight (1987) provides the first attempt to examination of rank-order tournaments and piece rates and the results – with data coming from 225 paid undergraduate student volunteer in a experimental tournament study repeated 12 times, although this theory deals with one-shot rather than repeated tournaments. Experimental parameters were choice from the experimental realized and they found evidences that there is a systematic behavior by agents when faced with tournament to support the theory of tournaments.

Main, O'Reilly and Wade (1993) reports the results of an empirical investigation upon 210 firms with up to five annual observations (1980-84) on the top executive

team giving 13,347 individual observations and 769 firm year observation for investigate two apparently contradictory issues emerged by developments in tournament theory regarding the structure of pay among top executives: (i) suggest the need for substantial variations in compensation among executives at the top levels of an organization, (ii) alternatively, arguments can also be made that comparatively compressed wages are more efficient since they reduce sabotage and promote cooperation and teamwork. They found evidences to support the operation for tournaments while not support for arguments in favor of the efficiency of pay wage equity at the top of corporations.

Knoeber and Thurman (1994) used an OLS model to test three predictions of the theory of tournaments that previously have not been subjected to an empirical test. Analyzing the broiler production for a sample of 1,174 flocks produced by 75 growers from November 30, 1981, to December 17, 1985 they provided evidence that supports the predictions raised by the tournament theory.

Lambert, Larcker and Weigelt (1999) provided insights into internal compensation schemes selected by organizations by examining the applicability of several alternative theories. From 303 large publicly traded U.S. firms in manufacturing and service markets included 42 different two-digit standard industrial classification codes. They used a cross sectional multiple regression, in which the level of compensation for an individual executives is expressed as a function of job size, *proxies* for managerial power, executive performance, and indicator variables for managerial position in the organizational hierarchy. Their results support the importance of tournament models for explaining levels of executive compensation and show that cash compensation differentials are an increasing function of organizational hierarchy, and the spread between adjacent levels increase as one moves up the organizational hierarchy.

Eriksson (1999) used a fixed effect panel models for a bulk of data comes from an unbalanced panel containing information about approximately 2,600 managers in about 210 Danish firms from 1992 to 1995 and tested multiple predictions from the tournament theory upon the same data set. The result reached support most of the prediction raised and adds to a very small literature on the effect of promotions and pay structures on performance of firms and conclude that almost all findings are consistent with tournaments model.

Bognanno (2001) investigated the determinants of pay in corporate hierarchies as well as the relationship between pay and promotion and presents a narrowly focused effort to determine whether the skewed pay structures at the top of large U.S. corporations result from an attempt to manage tournament incentives according to a specific tournament model. From a rich data base that tracks individual executives at more than 600 firms for up to 8 years between the years 1981 to 1988 contains information on the annual base pay and bonus, reporting level, and various personal and job characteristics for about 25,000 managers and executives per year. They found that the tournament theory predicts that the CEO promotion prize rises with the number of contestants for the position not occur for the companies in their study. Additionally, attempts to provide further evidence that firms intentionally manage promotion incentives were not successfully contradicting the implication of the tournament model presented that the winner prize should increase at an increasing rate with the number of competitors.

Harbring and Irlenbusch (2004) investigate the tournament incentive schemes offer payments dependent on relative performance. Thus they create an experimental tournament setting conducted in the laboratories for experimental research at the universities of Bonn and Erfurt with 240 students of different disciplines involved in the experiment – 192 take part as agents and 48 as principals to observe the Interplay

between fairness considerations towards the principal and the competition among agents. The results reveal that in environments with higher competition agent may also be tempted to destroy the production of his competitors in order to improve the own relative position but also retaliation toward the principal if wages are perceived to be too low.

CHAPTER 2 – DATA

2.1. Approach and Data

The empirical literature on Tournaments has been grown. However, many studies use no appropriate data from different contexts than employees in firms, such as baseball, golf, auto racing, agricultural production, and lab experiments. Then I consider a special case surrounding the incentives for workers in public sector.

Briefly, tournament theory shows that a firm may motivate employees by running competitors for rewards (e.g. bonus, promotions, etc.) either for a group or individualistic schemes. Most of the cases, accurately measuring individual performance are costly to the firm may economize by measuring only rank ordering of performances.

The dataset is drawn from the Coordenação de Fiscalização (i.e., the Inspections Division) of the Secretaria da Receita Federal (SRF). This data were monthly collected by the SRF during the Brazilian bonus program for Tax collection, from August 1989 to April 1993. The panel is unbalanced in the sense that we have more observations on some tax agencies than on others, and because these observations correspond to different points in historical time. This record contains information on the bonus program created by the Brazilian government to compensate tax officials for their efforts in collecting taxes and uncovering tax violations. The Brazilian reform provides monetary compensation to tax collectors based on their individual and group performance in finding and collecting taxes from tax evaders. Bonuses are paid with revenues raised in the collection of fines – for every dollar of fines collected, 68 cents are allocated to pay bonuses. So, the tax agency performance should be affected positively. An executive committee, called Comissão de

Administração da Retribuição Adicional Variável (CRAV) supervises and evaluates the performance of each tax agency. There are 111 tax agencies¹ in the country: one central agency, ten regional agencies and 100 local agencies. See Table 1 below for a list of all tax agencies included in this study.

Table 1: Group of Tax agencies

Groups	Tax Agencies
North	Boa Vista, Belém, Macapá, Manaus, Monte Dourado, Porto Belém, Porto Manaus, Porto Velho, Rio Branco, Santarem
Northeast	ALF/Fortaleza, Aracaju, Araçatuba, Caruaru, Feira de Santana, Fortaleza, Ilheus, Joao Pessoa, Juazeiro, Maceió, Natal, Porto Recife, Porto Salvador, Recife, Salvador, São Luiz, Teresina, Vitoria da Conquista
South	Cascavel, Caxias do Sul, Curitiba, Florianopolis, Foz do Iguaçu, Salgado Filho Airport, Imperatriz, IRF/Chuí, IRF/Porto Alegre, Joaçaba, Joinville, Londrina, Maringá, Novo Hamburgo, Santana do Livramento, Paranaguá, Passo Fundo, Pelotas, Ponta Grossa, Porto Alegre, Rio Grande, Santa Maria, Santo Angelo, São Sebastião, Uruguiana
Southeast	Angra dos Reis, Rio de Janeiro Airport, Bauru, Belo Horizonte, Campinas, Campos, Contagem, Curvelo, Divinópolis, Governador Valadares, Guarulhos, Guarulhos Airport, Juiz de Fora, Limeira, Monte Claros, Niterói, Nova Iguaçu, Osasco, Porto Rio de Janeiro, Presidente Prudente, Tancredo Neves Airport, Rio de Janeiro, Rio Preto, Santo André, Santos, São Jose do Rio Preto, São Paulo, Sorocaba, Taubaté, Viracopos Airport, Uberaba, Uberlândia, Varginha, Vitória, Volta Redonda
Midwest	Brasilia Airport, Brasília, Campo Grande, Corumbá, Cuiabá, Goiania, Itajaí, Mundo Novo, Órgão Central, Ponta Porã, Remessas Internacionais

The 111 agencies were ranked according to their performance each month, inducing managers to allocate resources efficiently within the tax agencies in light of

¹ Our study considered tax agencies as the participants within the tournament. The tax agencies were considered to be competing for a prize of the employees.

the inter-agency competition. In the tournament theory the principal designs tournaments and prizes to motivate agents to exert effort. The prize is the main factor to motivate the competitors (Harbring & Irlenbusch, 2003). Several author also indicated the size of the tournament in terms of number of firms competing in the market, or in terms of total number of employees competing in the market.

It is important to note that this paper does not attempt to answer the question under what circumstances tournament incentive systems might dominate other forms of incentive systems.

2.1.1. Variables in the core model

In order to test the implications of the tournament theory, it is necessary to find particular moments when the economic environment has radically challenged the nature and quality of competition due the structure of rewards. The Brazilian government created a bonus program to compensate tax officials for their efforts in collecting taxes and uncovering tax violations provides an inedited dataset to test some tournaments predictions.

The quantitative variables of interest in this paper are: (i) ranking; (ii) number of high level auditors (AFTNs); (iii) number of administrative bureaucrats (TTNs); (iii) fines collected; and (iv) gross product domestic (GDP) by state.

Getting accurate data on performance (ranking) in many jobs is quite difficult and the reason is that it is high costly to the firm obtain accurately measuring effort of performance². Ranking variable used here is a privileged source of information of the study to analyze our understanding of compensation and incentive systems. This is an ordinal variable (ranking) and three factors were used to measure the tax agencies

² The group reward is calculated according to the relative efficiency of the tax agency vis-à-vis other tax agencies in the country.

efficiency rank: (i) the amount of fines collected; (ii) the relative performance in reaching pre-established goals; and (iii) the size of the agency (in terms of the number of officials). The first two items (i) and (ii) are used to determine the agency's total due compensation, because they are directly related to the group reward. It is very important to understand carefully item (ii): The pre-established goals was adjusted, based on the percentage of the total amount of fines collected by the tax agency relative to overall amount collected based on the agency average of the last three months. The use of the agency quarterly average is to avoid potential problems (i.e., disturbances or seasonality), that might distort the rewards. Thus, is natural to expect outstanding performance, $(ranking_{t-1}, ..., ranking_{t-3})$ carried forward into period t . Second, at agency level, what can be seen as manager incentive to reallocate their staffs, determining strategies, in ways which increase group performance. This enables us to see the time adjustments within tax agency; item (iii) is indirectly related to the group reward. Thus, each agency is ranked based on their relative performance on their rounded result with the most successful agency being ranked highest (score 1), and so on. Another reason we could want to include a lagged dependent variable is that it may help *proxy* for slower moving omitted variables that are not captured by the controls included (i.e. types of taxpayers residing in the region).

The other data set contains information on fines collected for each of the tax collection agencies. In the tournament theory the principal designs tournaments and prizes to motivate agents to exert effort. The prize is the motivating factor for the competitors³ (Harbring, & Irlenbusch, 2003). This variable will capture the agency effort in collecting underreported taxes as well as the effectiveness of the reform. The

³ The bonus paid to an official is composed of two types of rewards: An individual reward and a group reward. For this reason, both types of rewards increase with the amount of fines collected, consequently both the individual and the group have incentives to increase their productivities.

analysis of fines is complicated both by the fact that the sample period was a time of hyperinflation and the currency was converted on four occasions.

The number of employees was a more traditional *proxy* for the number of participants in the tournament found in many studies of tournament theory. In this study there are two types' variables that convey information about the number of workers inside the SRF: auditors and high level supervisors assigned (AFTNs), and the administrative bureaucrats denoted (TTNs). The number of AFTNs was a *proxy* for the number of participants in the tournament. AFTNs are highly skilled workers and their duties include field or external investigations, examinations of tax returns, customs' inspections, collection of overdue taxes and fines, and supervision of tax agencies. Tournament theory predicted larger prizes (Bognanno, 1994; O'Shaughnessy, 1998) as the number of contents increased.

The number of Administrative bureaucrats (TTNs) *proxy* the size of the tax agencies. Campbell (1993) used the firm's employment as a *proxy* for firm size. TTNs usually play a more passive role in the investigation and collection procedures. Because of the discrepancies in skills and duties, the rules governing the bonus program establish that the reward paid to a low-level bureaucrat cannot exceed 30% or that paid to the AFTNs. Moreover, the tournament predicts that wage differentials directly compensate workers for higher effort at higher levels of the hierarchy, implying that differentials arise to maintain incentives to lower rungs of the tax agencies.

A natural question to raised is whether a disproportionate number of tournaments were won by the same tax agencies is because those agencies were truly better than others, or because they were (for whatever reason – Tax code, price level, nominal GDP, wages) given superior quality inputs. A common control variable used here and in many other research was the gross domestic product (GDP) by state. Note also that the GDP by act as a fairly good *proxy* for most aspects of development. For

instance, it can be argued that rising income levels ultimately and inevitably translate into better economic conditions by looking at a cross section of regions. The GDP by state were converted to U.S. dollars on the basis of the real exchange rates. The conversion factors were constructed on the basis of information from the Banco Central do Brasil (i.e., Federal Bank) databases. The real GDP by state was calculated on the basis of data expressed in 2000 prices (in local currency). Again, the analysis of GDP by state is complicated both by the fact that the sample period was a time of hyperinflation and the currency was converted on four occasions.

In Table 2 below we give summary statistics on the ranking and nine explanatory quantitative variables. These statistics are based on monthly observations.

Table 2: Descriptive Statistics for Dependent and Explanatory Variables

Variables	Observations	Mean	Std. Dev.	Minimum	Maximum
Ranking	4671	53.7101	30.59653	1	110
Ranking_{t-1}	4671	53.7101	30.59653	1	110
Ranking_{t-2}	4561	53.66893	30.56853	1	110
Ranking_{t-3}	4451	53.62076	30.53442	1	110
Fines	3201	456093.6	1054789	0	1.11e+07
Fines_{t-1}	3104	457521.4	1058211	0	1.11e+07
AFTN	4581	53.08404	85.86319	1	618
TTN	4581	49.77101	60.43109	1	570
GDP	4725	28730.65	47337.08	48.95	349735.3

There are large variations across agencies in ranking as well as fines collected, number of employees (AFTNs and TTNs) and gross domestic product (GDP) by state. It is worth noting that the GDP by state variation suggests that income is distributed unequally within all country.

As discussed earlier we expect that the current *ranking_t* should be negatively related with number of the high level auditors (AFTNs) and positively related to the number of bureaucratic administrators (TTNs). Using annual data at a suitable level of regions could capture the economic conditions. Thus, we expect that either an expansion in the GDP by state or a higher GDP by state for a particular region could improve the tax agencies performance ranking in collecting fines, consequently inversely correlated with ranking.

Table 3 presents detailed descriptive statistics for our dependent variable.

Table 3: Detailed Descriptive Statistics for Tax Agencies Ranking

Table 1: Summary of the 1990-1992 Survey Data by Region																
		North			Northeast			South			Southeast			Midwest		
Ranking	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD	Obs	Mean	SD	
1990	132	45.09	27.21	216	67.60	27.96	252	47.11	27.67	492	57.67	30.86	108	34.59	28.39	
1991	132	57.81	29.98	216	64.62	28.28	259	45.56	27.80	485	53.20	33.24	108	43.38	19.92	
1992	132	71.20	31.05	216	78.58	23.45	270	47.30	26.47	492	40.63	29.20	108	52.74	23.09	

Notice that the number of observation for regions overtime is the same except for the South which slightly increase. Looking at the number of observations for the Southeast region decrease slightly from the first to the second year period and returned for the initial number of observation in the third year period. In the 1990s cross-section the best mean value of the ranking was for the Midwest region. The average was 34.59 with a standard deviation of 28 followed by the regions North (45.09), South (47.11) both with a standard deviation of 27, Southeast (57.67 with a standard deviation of 30.86) and Northeast (67.60 with a standard deviation of 27.96).

We observe a similar variance across the five regions. This suggests that the within 5 regions' performance in terms of ranking might be relatively homogeneous in what regards the bonus program for tax collection. In the 1991 the best and the worse ranking average again was for the regions Midwest (43.38 with standard deviation of 19.92), and Northeast (64.62 with a standard deviation of 28.28), respectively. The second best ranking was the region South (45.56 with a standard deviation of 27.80), followed by Southeast (53.20 with a standard error of 33.24), and North (57.81 with a standard deviation of 29.98). Notice that the average ranking for the Midwest region was lower for the two uninterrupted period, on the other hand the Northeast was the worse one. In 1992 the higher ranking means was for the Southeast (40.63 with a standard deviation of the 29.20), followed by South (47.30 with a standard deviation of 26.47), Midwest (52.74 with a standard deviation of 23.09), North (71.20 with a standard deviation of 31.05), and Northeast (78.58 with a standard deviation of 23.45).

In the 1990-1992 panels, the distribution for the region Southeast drifts upwards over time, reflecting growth in the average ranking and downwards over time for the regions Northeast, Midwest and North. Our table also shows an interrupted upward drift for the South region.

Table 4 below provides us information about the correlation coefficients between pairs of quantitative variables. The first column informs us about the correlation coefficient between the dependent variable and each of the explanatory variables. The highest correlation coefficients between the dependent variable and the three lags used here presented in the first column are 0.7926, 0.7220 and 0.6780, respectively, which suggests that ranking is strongly characterized by inertia.

Table 4: Correlation Coefficients for Quantitative Variables

	ranking	L1rank	L2rank	L3rank	finest	L1finest	AFTN	TTN	GDP	North	Northeast	South	Southeast	Midwest
ranking	1.0000													
L1rank	0.7926	1.0000												
L2rank	0.7220	0.8004	1.0000											
L3rank	0.6786	0.7469	0.8288	1.0000										
finest	-0.1700	-0.1606	-0.1599	-0.1702	1.0000									
L1finest	-0.1654	-0.1760	-0.1650	-0.1727	0.9428	1.0000								
AFTN	-0.1181	-0.1235	-0.1243	-0.1298	0.4516	0.4518	1.0000							
TTN	-0.0781	-0.0830	-0.0846	-0.0935	0.3681	0.3700	0.8782	1.0000						
GDP	-0.1475	-0.1622	-0.1676	-0.1991	0.1622	0.1568	0.1351	0.0677	1.0000					
North	0.0604	0.0562	0.0497	0.0528	-0.1006	-0.1004	-0.1295	-0.1438	-0.1821	1.0000				
Northeast	0.2810	0.2771	0.2696	0.2780	-0.1086	-0.1083	-0.0885	-0.0844	-0.2168	-0.1532	1.0000			
South	-0.0892	-0.0891	-0.0856	-0.0946	-0.0889	-0.0895	-0.1121	-0.0719	-0.1252	-0.1896	-0.2546	1.0000		
Southeast	-0.1283	-0.1194	-0.1098	-0.1110	0.2551	0.2521	0.2093	0.0965	0.4819	-0.2753	-0.3698	-0.4576	1.0000	
Midwest	-0.0815	-0.0874	-0.0924	-0.0911	-0.0551	-0.0497	0.0631	0.2073	-0.1617	-0.1050	-0.1410	-0.1745	-0.2535	1.0000

The dependent variable ranking appeared to be not highly correlated with any of the remaining independent variables. The dependent variable ranking is negatively correlated with fines collected, number of AFTNs, TTNs, GDP and all dummies variable except for North and Northeast dummies. The sign of the coefficient the number of bureaucratic administrators (TTNs) was opposite to expectations; the TTNs was hypothesized to positively affect ranking. Second, shows that the highest correlation coefficients between the lagged dependent and two lagged dependent variables are positive and strongly correlated (0.8004 and 0.7469), indicating high inertia of this variable. Excluding this both high correlations, the other correlation coefficients in the second column are negative and quite low, except for North and Northeast region dummies. Third, the strength of the correlation between two lagged ranking variable, and the third lagged ranking was high positive (0.8288). The remaining coefficients is negative and small, except for both, north and northeast

region that are positive low correlated with the two lagged dependent variable. Fourth, the correlation coefficients for the third lagged dependent and other independent variables are negative and low, except for the regions north (0.05) and northeast (0.278). For the latter it inform us that that the tax agency grouping the northeast region is the worse ranking. Fifth, the three highest correlation coefficients for the fines collected with the lagged fines collected (0.9428), which is strongly characterized by inertia, the number of AFTNs (0.4516), and the number of TTNs (0.3681), as expected. The remaining other independent variables are also negatively and low correlated except for GDP and southeast dummy. Seventh, we note that the correlation coefficient between AFTNs and TTNs are very high (0.8782). This correlation coefficient raises our attention for the potential problem of multicollinearity. Ninth, shows that GDP is highly correlated with southeast, as expected. The other variables correlations in this column are negatively and low.

CHAPTER 3 – EMPIRICAL METHODOLOGY

3.1. GMM Estimators

Working with panel data sets possess several major advantages over conventional cross-sectional or time-series data sets. Combining the time-series with the cross-tax agency give us a large number of data points, increasing the degrees of freedom and reducing the collinearity among explanatory variables permit richer model moreover improving the efficiency of econometric estimates.

In order to examine the tournaments predictions we use a dynamic panel data econometric model. For instance, let's assume no unobserved region-specific effects. The following moment condition is specified as:

$$y_{i,t}^* = \alpha y_{i,t-1}^* + \beta' X_{i,t} + \mu_{i,t} \quad (1)$$

Where the subscript it denotes tax agency i at time t , respectively; $y_{i,t}^*$ is the “true” tax agency ranked according to their relative performance, X is a set of explanatory variables and $\mu_{i,t}$ is the residual term representing omitted factors that determine $y_{i,t}^*$. Since we estimate a dynamic model, this will induce measurement errors in the dependent variable namely lagged ranking. For this section, assume that measurement error is only standard random noise. Then, we have:

$$y_{i,t} = y_{i,t}^* + v_{i,t} \quad (2)$$

Where, $v_{i,t}$ is *i.i.d.* and $y_{i,t}$ represents the measured tax agency ranking. Substituting equation (2) into equation (1) we have:

$$y_{i,t} = \alpha y_{i,t-1} + \beta' X_{i,t} + \varepsilon_{i,t} \quad (3)$$

Where, $\varepsilon_{i,t} = \mu_{i,t} + v_{i,t} - \alpha v_{i,t-1}$. Thus, equation (3) is our population regression model.

One immediate problem in applying OLS to this empirical problem and to (3) in general, is that by construction $y_{i,t-1}$ is endogenous and potentially correlated with $v_{i,t-1}$, and $X_{i,t}$ is possibly correlated with $\mu_{i,t}$, which gives rise to “dynamic panel bias”, and then OLS estimator is inconsistent. Whatever the source of the correlation, if there is a valid instrumental variable, then the consistent estimation can be found. One solution to the econometric problems presented above would be to employ the Generalized-Method of Moments⁴ (GMM) estimators that use the dynamic properties of the data to generate proper instrumental variables introduced by Arellano and Bond (1991), and Arellano and Bover (1995). So far we are assuming that there is no tax agency region-specific fixed effect, we base our estimates on the so-called *levels* GMM estimator. The GMM technique allows the use of instruments to deal with the random measurement error in the lagged dependent variable and the possible endogeneity of outstanding independent variables, $X_{i,t}$. Credibility depends on the instrument relevance. Assuming the assumption of weak exogeneity instead of strict exogeneity, would mean, for instance that the explanatory variables might be affected by past and present realizations conditional on the dependent variable, but that they must be uncorrelated by its future innovation. Based on the assumptions that the error term, ε_{it} is not serially correlated, and the explanatory variables are weakly exogenous, in both cases, the following moment conditions can be written as:

$$E[y_{i,t-s} \cdot \varepsilon_{i,t}] = 0 \text{ for } s \geq 2 \quad (4)$$

⁴ The Generalized Method of Moments was introduced by L. Hansen in 1982. There are other texts that cover GMM, such as Hayashi (2000), Hansen (2000), Wooldridge (2002), Davidson and MacKinnon (1993), and Greene (2000).

$$E[X_{i,t-s} \cdot \varepsilon_{i,t}] = 0 \text{ for } s \geq 1 \quad (5)$$

It is appropriate to discuss the basic framework of GMM systems estimators allowing and controlling for unobserved Tax Agency-specific effect and time-specific effect. Consider the model:

$$y_{i,t} = \alpha y_{i,t-1} + \beta' X_{i,t} + \eta_i + \delta_t + \mu_{it} \quad (6)$$

where the subscript it denotes tax agency i at time t , respectively; η_i represents the tax agency-specific effect, reflect the differences in the initial level of efficiency, whilst δ_t is a time-specific effect, capture ranking changes that are common to all agencies. Tax agency effects and time effects may also reflect tax region-specific and time-specific components of measurement errors. To deal with the fact that the measurement errors are likely to be determined not only by random errors but by specific and persistent characteristics of each region, we use equation (6) to estimate the so-called system GMM estimator that combines into a single system the regression equation in both differences and levels, each with its specific set of instrumental variables. Taking the first differences of our regression equation (6) to eliminate the tax agency-specific effect, we obtain,

$$y_{i,t} - y_{i,t-1} = \alpha(y_{i,t-1} - y_{i,t-2}) + \beta'(X_{i,t} - X_{i,t-1}) + (\delta_t - \delta_{t-1}) + (\mu_{i,t} - \mu_{i,t-1}) \quad (7)$$

Since $\Delta X_{i,t} = X_{i,t} - X_{i,t-1}$ is likely to contains endogenous variables, which by design are correlated with the residual, the use of the instruments is required to obtain unbiased and consistent estimates. Arellano and Bond (1991) developed a GMM difference estimator which uses lagged levels of $X_{i,t}$ as instruments for $\Delta X_{i,t}$. By construction, is generated between the new error term, $(\varepsilon_{i,t} - \varepsilon_{i,t-1})$, and the differenced

lagged dependent variable, $(y_{i,t} - y_{i,t-1})$. Once again, we assume two conditions. The first condition is that the lagged levels of $X_{i,t}$ are uncorrelated with future errors, called weak exogeneity. The second condition is that the error term is not serially correlated. The following moment conditions can be used to obtain a GMM difference estimator:

$$E[X_{i,t-s}(\varepsilon_{i,t} - \varepsilon_{i,t-1})] = 0 \text{ for } s \geq 2 \quad (8)$$

$$E[y_{i,t-s}(\varepsilon_{i,t} - \varepsilon_{i,t-1})] = 0 \text{ for } s \geq 3 \quad (9)$$

Although asymptotically consistent, the use of the GMM estimator comes with a drawback that when the lagged levels may not be good instruments for first differences when there is persistence over time in the dependent and explanatory variables. Weak instruments⁵ has repercussion on both the asymptotic and small-sample performance of the differences estimator that may yield imprecise estimates and finite sample biases based on studies by Alonso-Borrego and Arellano (1996) and Blundell and Bond (1997). Another concern with the GMM difference estimator is that it directly eliminated the cross agency region-specific effect. To overcome this problem, Arellano and Bover (1995) and Blundell and Bond (1995) proposed to complement the difference equations with the original equation in levels and estimate the system as a whole. For the regression in levels, the agency region-specific effect is not eliminated, but it is potentially correlated with some independent variables, then we need to control for it using instrumental variables. Therefore, the instruments used are the lagged *differences* of the explanatory variables in the levels equations. It is worth noticing that the lagged differences are valid instruments under the assumption that

⁵ This estimator has low asymptotic precision and large biases in small samples, which leads to the need to complement it with regression levels equation.

there is no correlation between the agency region-specific effect, and explanatory variables are constant over time. This assumption results from the following stationarity property,

$$E[y_{i,t+p} \eta_i] = E[y_{i,t+q} \eta_i], \text{ for all } p \text{ and } q. \quad (10)$$

$$E[X_{i,t+p} \eta_i] = E[X_{i,t+q} \eta_i], \text{ for all } p \text{ and } q. \quad (11)$$

Moreover, the lagged differences are used as valid instruments in the levels-specifications. Other lagged differences would result in redundant moment conditions (Arellano and Bover 1995). The moment conditions for the levels equations are:

$$E[(y_{i,t-s} - y_{i,t-s-1})(\eta_i + \varepsilon_{i,t})] = 0 \text{ for } s = 2 \quad (12)$$

$$E[(X_{i,t-s} - X_{i,t-s-1})(\eta_i + \varepsilon_{i,t})] = 0 \text{ for } s = 1 \quad (13)$$

Using the set of moment conditions given by equations (4) and (5), and (8) to (11), and following Arellano and Bond (1991) and Arellano and Bover (1995), we utilize the GMM *systems* estimators procedure to improve precision estimates of the parameters of interest and reduce the potential biases in finite samples as shown by Blundell and Bond (1995) using Monte Carlo simulations. In practice, we calculate the GMM in two-step procedure. First, assume that the residuals, $\varepsilon_{i,t}$, are independent and homoskedastic over time and across agencies. This first step estimation assumption corresponds to a specific weighting matrix that is used to produce first-step coefficient estimates. The GMM estimator is not yet feasible estimator, because the covariance matrix of the moment condition is not known. In the second step, construct a consistent

estimate of the variance-covariance matrix of the moment conditions from the residuals obtained in the first step, and use this matrix to re-estimate the parameters of interest. The second step estimates relax the assumptions of independence and homoskedasticity. Asymptotically, the second-step estimates are superior to the first-step ones in so far as efficiency is concerned.

The validity of the *system* GMM as a consistent estimator can be ascertained by showing that the error term is not serially correlated, and the instrumentals used are adequate ones to our regression. The literature addresses this issue by considering two specification tests gauged by Arellano and Bond (1991). The first is a Sargan test⁶ of over-identifying restrictions, under the null hypothesis⁷ that the full set of instruments are not correlated with the residuals by the moment conditions used in the estimation. Failure to reject the null hypothesis gives support to the model. Second, a serial correlation test, which determines whether the error term $\varepsilon_{i,t}$ of the difference equation has second order. The null hypothesis is that the differenced error has no second-order serial correlation.

3.2. Hypotheses

We can state our hypothesis more specifically by rewriting the general equation (6):

$$\begin{aligned} \text{Ranking}_{it} = & \alpha_1 \text{Ranking}_{i,t-1} + \alpha_2 \text{Ranking}_{i,t-2} + \alpha_3 \text{Ranking}_{i,t-3} + \alpha_4 \text{Fines}_t + \alpha_5 \text{Fines}_{i,t-1} \\ & + \alpha_6 \text{AFTN} + \alpha_7 \text{TTN} + \alpha_8 \text{GDP} + \eta_i + \delta_t + \varepsilon_{it} \end{aligned} \quad (14)$$

⁶ It is often called Hansen's overidentification statistic or Hansen's J statistic. Davidson and MacKinnon (2004) prefer to call it the Hansen-Sargan statistic.

⁷ Davidson and MacKinnon (2004) indicate that the Sargan test may reject the null hypothesis for more than one reason. Perhaps one or more of the instruments are invalid. Perhaps the model is misspecified, either because one or more of the instruments should have been included among the regressors. Or perhaps the finite sample distribution of the test statistics differs substantially from its asymptotic distribution. Even if we do not know how to interpret a significance value of the overidentification test statistic, it is always a good idea to compute it.

Where the subscript it denotes tax agency i at time t ; *ranking* is represented here as the concept of the prize within the Brazilian bonus program. Lower ranking indicates better performance consequently larger rewards to pay the tax officials, for their individual effort; *Fines* represents the amount of fines collected by each of the tax collection agencies, and will capture the level of effort of the agents; *AFTNs* is the number of auditors and high level supervisors; *TTNs* represents the number of administrative bureaucrats; *GDP* is the gross domestic product by state. Finally, η_i , δ_i , and $\mu_{i,t}$, which are defined in equation (6) which are used to adjusted estimate for the ranking.

Consistent with our discussion in section 2, our hypothesis is to expect outstanding performance carried forward to period t , due time adjustment within tax agency, realized by the manager (learning by doing), either reallocating their staffs, determining better strategies, or both in order to improve tax agency performance (ranking), thereby resulting in a higher performance inertia ($\alpha_1 > 0$), ($\alpha_2 > 0$), and ($\alpha_3 > 0$). Regarding to amount of fines collected, our conjecture is that ($\alpha_4 < 0$) a larger amount of fines collected is associated with the level of effort of the group in finding and collecting taxes and therefore, lower ranking (better ranking); we also expect that the amount of fines collected in the previous period by the tax agency is negatively related to its current ranking ($\alpha_5 < 0$); we expect that the number of auditors and high level supervisors, *AFTNs*, have negatively relationship with the ranking, demonstrating support for the hypothesis that larger prizes as the number of contents increase ($\alpha_6 < 0$). The administrative bureaucrats play a more passive role in the investigation and collection procedures. Because of the discrepancies in skills and duties, the reward paid to a low level bureaucrat bonus cannot exceed 30% paid

of that paid to $AFTNs$. Thus, is natural to expect that the coefficient of the number of administrative bureaucrats positively related with the ranking ($\alpha_7 > 0$). Thus, this result will find out evidences to support the hypothesis that wage and bonus differentials between workers directly compensates workers for higher levels of the hierarchy, directly affecting workers incentive; Regarding the control variable GDP by state, our conjecture is that $\alpha_5 < 0$, rising income region levels translate into better economic conditions (inputs) for tax agency collect fines, therefore, a lower ranking.

CHAPTER 4 – RESULTS

This section presents the regression on ranking. Our equation includes eight explanatory variables: the lagged dependent variable, the tax agencies fines collected, the lagged tax agencies fines collected, the number of auditors and high level supervisors (AFTNs), the number of administrative bureaucrats (TTNs), and the gross domestic product (GDP) by state. As mentioned in the previous chapter, our main econometric methodology is the GMM-*system* estimator. The OLS, within groups, 2SLS and GMM-*differences* is added in this work only for comparative purposes. Recall that the OLS does not control for region-specific effects and the estimate is likely to be biased upwards. Although the Within transformation induces a non-negligible correlation between the transformed lagged dependent variable and the transformed error. This correlation does not vanish as the number of individuals in the sample increases, with the direction of the inconsistency estimate is likely to be downwards. The GMM first-differenced, provides a convenient framework for obtaining asymptotically efficient estimators for the AR(1) panel data model. Although asymptotically consistent, the use of the GMM differences estimator comes with a drawback when the correlation between lagged levels of the series $\Delta y_{i,t-1}$ if either the true value of the parameter α approaches unity, or if the ratio $\frac{\text{var}(\eta_i)}{\text{var}(\varepsilon_{it})}$ becomes large.

In this case, instruments available for the equations in differences are likely to be weak when the series have near unit root properties. Thus, we are going to ground our analyses on GMM *system* estimation. Accordingly to Bond (2002) the estimator *system* GMM extends straightforwardly to autoregressive-distributed lag models, and also has

smaller finite sample bias and greater precision estimating autoregressive parameters using persistent series⁸.

First, the initial proposed model, only the dependent variable is included with one lag, was estimated. However, in practice the model was estimated with included lags for all variables from four to zero. Models with three and two lags are accepted by the Sargan test and the correlation test. The number of lags⁹ used here is determined as the smallest model that is accepted by the Sargan statistic and the correlation statistic. Bond (2002, p.21) argues that "... *system* GMM estimator also extends straightforwardly to autoregressive-distributed lag models like (6). In this context Blundell and Bond (2000) show that a sufficient condition for the validity of additional moment conditions like (12) is that the $(y_{i,t}, x_{i,t})$ series each satisfy a mean stationarity assumption".

Table 5 provides the regression results that allow a comparison of the OLS, within groups, 2SLS, GMM-differences and GMM-*system* estimator. As expected in presence of tax agencies region-specific effect, OLS appears to give an upwards-biased estimate of the coefficient on the lagged dependent variable, whilst Within Groups appears to give a downwards-biased estimate of this coefficient. In this study we are using results for the two-step estimator rather than the one-step estimator. A large number of applied works using GMM estimator shows efficiency gains from using the two-step¹⁰ rather than the one-step estimator.

⁸ See tables A.1 and A.2 in Appendix A for careful consideration on the series being used here.

⁹ Insignificant variables, at the 10% level, were then deleted from the initial model.

¹⁰ A finite-sample correction for the asymptotic variance proposed for Windmeijer is used here for the two-step estimates. On the contrary, asymptotic standard errors tend to be quite small; as a result t statistics tend to be too large.

Table 5: Regression results

	OLS LEVELS	WITHIN GROUPS	2SLS DIF	GMM DIF t-2	GMM DIF t-3	GMM SYS t-2	GMM SYS t-3
Ranking_(t-1)	0.5770* (0.02416)	0.4587* (0.0257)	1.4994* (0.0278)	0.7043*** (0.1920)	0.6596*** (0.1995)	0.4442* (0.1596)	0.4823* (0.1766)
Ranking_(t-2)	0.1769* (0.0259)	0.1149* (0.0226)	0.3426* (0.0174)	0.3296 (0.3069)	0.3472 (0.3042)	0.3384 (0.2098)	0.3114 (0.2051)
Ranking_(t-3)	0.0985* (0.02423)	-0.0150 (0.0260)	0.3311* (0.0139)	-0.4293** (0.2080)	-0.4020*** (0.1594)	0.1138 (0.1325)	0.1057 (0.1099)
Fines_(t)	-0.00000428* (8.66e-07)	-0.00000697* (1.32e-06)	-9.53e-07 (8.62 e-07)	-0.0000071* (3.15e-06)	-0.00000705* (2.65e-06)	-0.00000316*** (1.62e-06)	-0.00000417*** (1.59e-06)
Fines_(t-1)	0.00000351* (7.66e-07)	-0.000000189 (1.37e-06)	-2.77e-06 (1.36 e-06)	0.00000345 (2.20e-06)	0.00000354*** (1.58e-06)	0.00000294 (1.49e-06)	0.00000394** (1.56e-06)
AFTN_(t)	-0.00443 (0.00664)	0.1507** (0.0508)	-0.036 (0.0393)	-0.01312 (0.1512)	-0.1798 (0.1567)	-0.0375*** (0.0228)	-0.0158 (0.019)
TTN_(t)	0.0068 (0.0087)	0.0282 (0.0485)	0.03126 (0.05)	-0.1589 (0.1270)	-0.1256 (0.0960)	0.05613 (0.0370)	0.02152 (0.0319)
PIB_(t)	0.000000881 (-0.00000812)	0.0000257** (0.0000117)	-0.000094* (0.00003)	0.0000367*** (0.0000207)	0.0000302** (0.0000204)	0.00000433 (0.0000112)	-0.0000114 (0.0000102)
lmonth_13	-4.6668** (2.1392)	-3.2744*** (1.799)	1.6331 (1.8404)			-2.7915*** (1.644)	-1.3566 (1.573)
lmonth_38	-2.5340 (1.7132)	-1.9585 (1.7692)	-1.9288 (1.8997)	-8.5467 (2.4562)	-8.1219 (2.2297)	-6.012* (1.9368)	-6.016* (1.8439)
Constant	8.2055* (1.2380)	16.418* (3.9916)	57.100* (3.1653)			5.1973* (1.882)	5.834* (1.673)
AR(1)	0.0413		0.0000	0.040	0.025	0.05	0.055
AR(2)	0.0000		0.0000	0.387	0.446	0.321	0.705
Sargan Test				0.223	0.298	0.482	0.447

Notes: Time dummies are included as instruments in all specifications. Standard errors reported in parentheses. AR (1) and AR (2) are tests for first-order and second order serial correlation, asymptotically N (0, 1). These test the levels residuals for OLS levels, and the first-differenced residuals in all columns. GMM results are two-step estimates. Sargan is a test of the overidentifying restrictions for the GMM, asymptotically χ^2 . P-value is reported. This test uses the minimized value of the corresponding two-step GMM estimators. ***, ** and * indicate coefficient at the 1%, 5% and 10% level, respectively. Estimations performed using GMM-system procedure combining transformed and level instruments. Variables instrumented: Ranking, Energy and GDP.

For the first regression in the column 1 first, second and third lagged ranking have significant coefficient. The variables fines and lagged fines have a quite small significant coefficient. Only the first binary explanatory variable (*_lmonth_13*) is significant. Nevertheless, this specification is rejected by the test of the first-order and second-order serial correlation, which point toward either omitted variables with high over-time persistence or ignored dynamic effects coming from the lagged dependent

variable. In the second column, we have the within groups estimation and indicate that the ranking, lagged ranking, fines, the number of AFTNs, PIB and the event dummy (*_imonth_13*) are significant. Again fines have a small significant effect on ranking. Accordingly with Bond (2002) since the estimator is not highly persistent ($\alpha \rightarrow 0$) these estimates of the first-differenced residuals are likely to be biased, unless the time period is too large.

A particular way of dealing with the problem of serial autocorrelation in the residuals is to control for the presence of tax agencies region-specific effects. The fixed-effect is the possibility that there is permanent region-specific effect on fines collections per inspection, possibly arising from differences in management style within tax agencies or from types of taxpayers residing in the region.

The third column reports the 2SLS estimator for the equation in first-differences, using $t - 2$ as instrumental variable. However, this specification is rejected by the test of the first-order and second-order serial correlation.

In the fourth and fifth columns we presented regression in differences, the region-specific effect is directly eliminated however must be controlled by the use of instruments. The validity of lagged levels dated $t - 2$ as instruments in the first-differenced equation is not rejected by the Sargan test of overidentifying restrictions. In this specification, only the first, and third lagged ranking, fines, and the GDP by state are statistically significant and with the same sign as before except for the number of administrative bureaucratic (TTNs). Although fines collected and GDP by state are statistically significant there was an insignificant support for the relationship between ranking and fines and GDP. Column (5) presents instruments dated $t - 3$ are not rejected, in which the statistically significant regressors were the same as dated $t - 2$, except for lagged fines, however a barely lower and higher estimated coefficients.

Column (6) and (7) present our main econometric methodology the GMM-*system*¹¹. Accordingly to Bond (2002) the estimator *system* GMM extends straightforwardly to autoregressive-distributed lag models, and also has smaller finite sample bias and greater precision estimating autoregressive parameters using persistent series¹². These results are supported by the specification tests, and the test of second-order serial correlation. Our results suggest that in the model that the significant determinants of the ranking are $ranking_{t-1}$, $fines_t$, $fines_{t-1}$, AFTNs and both dummy variables. Note that the second and third lag of the dependent variable and the lagged fines are insignificants, but they are included because otherwise the second order serial correlation test indicates serial correlation problem.

According to column (6) our results indicate that ranking in the first lag is helpful for forecasting current ranking. As explained earlier, ranking is also a function of the pre-established goals. The agency pre-established goals was adjusted based on the agency quarterly average directly related to the group reward. Thus, the outstanding performance ($ranking_{t-1}$) was carried forward into period t could be caused by the program rules. Another interpretation for the performance (ranking) inertia, is for the time adjustments within tax agency, realized by the manager (learning by doing), either reallocating their staffs or possibly determining strategies, or both, to improve the tax agency performance on the sources most valuable for collection and it will take effect in the next periods.

The negative effect of $fines_t$ collected on the ranking can be interpreted as level of effort of the group in finding and collecting taxes from tax evaders. The lagged

¹¹ Investigating the time series properties of the individual series is strongly recommended when using these GMM estimators for dynamic panel data model, to assure the validity of the additional moment conditions.

¹² Tables A.1 and A.2 in Appendix reports simple AR (1) specifications for the four series, ranking, fines, TTNs, AFTNs and GDP. All five series being used here are found to be highly persistent, although even using OLS estimates none is found have exact unit root.

amount of fines collected (fines_{t-1}) is also and can be seen as evidence of the learning by doing effect, but it was not significant. Interestingly, current fines collected by each tax agency have a quite small effect on ranking. As mentioned before, the analysis using this variable is complicated both by the fact that the sample period was a time of hyperinflation and the currency was converted on four occasions. Moreover, the fine collected is linked to group effort rather than individual effort, particularly in large agency. At this time we look at the tax agencies as the participants within the tournament competing for a prize of the employees. Thus, this result demonstrated evidence to support the hypothesis that prizes motivate agents to exert effort.

The numbers of auditors and high level supervisors denoted as AFTNs have negatively relationship with the ranking, as expected demonstrating support for the hypothesis that larger prizes as the number of contents increased. We are using ranking¹³ as a *proxy* for larger prize, for this reason the most successful tax agency being ranked highest get score 1, and so on. This suggests that the number of AFTNs (contents) were important to stimulate the tax agency ranking. In summary, both the efficiency and tournament models were indicative that the AFTNs as a whole acted upon ranking pressuring them downwards as the number of AFTNs increased and as the number of tax agencies employing them increased.

The administrative bureaucrats assigned (TTNs) is intended to capture the size of the Tax agency. As explained earlier, TTNs usually play a more passive role in the investigation and collection procedures. Because of the discrepancies in skills and duties, the rules governing the bonus program establish that the reward paid to a low-level bureaucrat cannot exceed 30% or that paid to an auditor. However, our results

¹³ . Ranking is represented here as the concept of the prize within the Brazilian bonus program. Lower ranking indicates better performance consequently larger rewards to pay the tax officials, for their individual effort.

observed the estimated coefficient of the number of administrative bureaucrats is statistically insignificant.

A natural question to raised is whether a disproportionate number of tournaments were won by the same tax agencies is because those agencies were truly better than others, or because they were (for whatever reason – Tax code, price level, nominal GDP, wages) given superior quality inputs. A common control variable used here was the gross domestic product (GDP). The results shows that the relationship between ranking and GDP by state was not statistically significant within the bonus program. Although insignificant, GDP by state was included because otherwise the second order serial correlation test indicates serial correlation problem.

Additional control variables are included here to isolate certain periods that may be systematically different from other periods covered in our dataset: two binary explanatory variables in our model also called event study. Both of them are statistically significant and capture seasonality in the series.

Column (7) is added only for comparative purposes. Although, this GMM estimation with endogenous variable in $t-3$ is valid instruments, accepted by the Sargan-statistic and the correlation test, we decided to use the preliminary model, which are instruments dated $t-2$. Furthermore, one very important coefficient, AFTNs, is statistically not significant.

CHAPTER 5 - CONCLUSIONS

This paper provides the first empirical evidence on two key assumptions in these models using a special case surrounding the incentives for workers in public sector. First, we verify if the principal designs tournaments and prizes to motivate agents to exert effort. Second, the study sought to understand if the number of participants increased as the size of the prize increase. Finally, this study also pondered the possible inertia for time adjustments within tax agency, possibly in determining strategies to improve the tax agency performance on the sources most valuable for collection.

In view of the potential problems found in the literature with testing the theory of tournament using natural data, some successfully works have been compelled to test the theory in an experimental setting. We illustrated our estimation procedures using a rich data base obtained from the Coordenacao de Fiscalizacao (i.e., the Inspections Division) of the Secretaria da Receita Federal (SRF) during the major incentive reform instituted in 1989 to improve tax enforcement. We constructed a larger unbalanced panel data containing information upon 110 tax agencies distributed between 10 regions and 45 time period by month, which allowed us to support the predictions raised during this work. These contributions were new within the context of the tournament model.

The tournaments theory model was evidenced within the bonus program context. As emphasized in section 2, the ranking was measured upon three factors. This enable us to pondered the possible inertia caused by time adjustments within tax agency, possibly in determining strategies to improve the tax agency performance on the sources most valuable for collection. This result can also be seen as learning by

doing, where the manager reallocate their staffs and refocus the attention on an explicit goal of the agency to improve your relative performance (ranking).

Ranking represented the concept of the prize within the Brazilian bonus program context. Lower ranking indicates better performance consequently larger prizes (bonus). The results demonstrated evidence to support the hypothesis that prizes motivate agents to exert effort. The tournament theory also concept that prize size must increase as the number of competitor increase in order to motivate participation. This suggests that the number of AFTNs (contents) were important to stimulate the tax agency ranking. In summary, both the efficiency and tournament models were indicative that the AFTNs as a whole acted upon ranking pressuring them downwards as the number of AFTNs increased and as the number of tax agencies employing them increased. The study also examined this conception and results confirmed this hypothesis.

A natural question to raised is whether a disproportionate number of tournaments were won by the same tax agencies is because those agencies were truly better than others, or because they were (for whatever reason – Tax code, price level, nominal GDP, wages) given superior quality inputs. A common control variable used here and in many other research was the gross domestic product (GDP) by state. However, the results shows that the relationship between ranking and GDP by state was not statistically significant within the bonus program. Although insignificant, GDP by state was included because otherwise the second order serial correlation test indicates serial correlation problem.

A potential limitation of this work is that the payment during the period studied is linked to group rather than individual performance, dampening individual incentives, particularly in large tax agencies. Another concern is that the bonus program was

during the period of high inflation rates over 1,000% that have motivated a number of actions by the federal government avoid corrosion of tax revenue.

Finally, a possible extension of this research, therefore, would be examine the effects of pay spread on the cooperation behavior of employees in multitasks setting. Furthermore, the pervert effects generated by this theory would be investigated in the RAV program, since this successfully government bonus program was finished for reasons unknowledgeable. For this reason, a richer prediction could be tested here, such as the potential for sabotage and pay compression. Such prediction might require new approaches, and requiring detailed information within-firm data. Furthermore, Tournament theory usually assumes two identical contents. Employees also differ in wage growth, performance ratings, and other observable characteristics. Tournaments with heterogeneous contestants have not been extensively analyzed.

APPENDIX A

Table A.1: Alternative Estimates of the AR (1) Specification

$AR(1)$	OLS Levels	Within Groups	2SLS DIF	GMM DIF	GMM DIF
$Ranking_{t-1}$	0.816 ^{***} (0.016)	0.596 ^{***} (0.020)	-4.124 ^{***} (0.477)	0.045 (0.045)	0.084 ^{***} (0.040)
AR(1)	-7.28		-0.58	-5.63	-6.12
AR(2)	1.90		-1.83	-0.88	-0.21
Sargan				0.071	0.897
Instruments			$Ranking_{t-2}$	$Ranking_{t-2}$ $Ranking_{t-3}$	$Ranking_{t-2}$ $Ranking_{t-3}$ $Ranking_{t-4}$

Notes: Quarters dummies included in all models. Standard errors reported in parentheses. AR (1) and AR (2) are tests for first-order and second order serial correlation, asymptotically $N(0, 1)$. These test the levels residuals for OLS levels, and the first-differenced residuals in all columns. GMM results are two-step estimates. Sargan is a test of the overidentifying restrictions for the GMM, asymptotically χ^2 . P-value is reported. This test uses the minimized value of the corresponding two-step GMM estimators. ***, ** and * indicate coefficient at the 1%, 5% and 10% level, respectively

Table A.2: AR (1) Model Estimates

$Fines_t$	OLS Levels	Within Groups	GMM DIF t-3	GMM Sys t-3
$Fines_{t-1}$	0.928*** (0.026)	0.382*** (0.040)	0.218*** (0.049)	0.726*** (0.040)
AR(1)	-2.59		-2.54	-2.50
AR(2)	-1.56		-1.02	-0.71
Sargan			0.006	0.402
$AFTN_t$	OLS Levels	Within Groups	GMM DIF t-3	GMM Sys t-3
$AFTN_{t-1}$	0.992*** (0.001)	0.931*** (0.016)	0.609*** (0.072)	0.986*** (0.002)
AR(1)	-1.61		-3.07	-3.23
AR(2)	0.55		2.25	2.27
Sargan			0.031	0.036
TTN_t	OLS Levels	Within Groups	GMM DIF t-3	GMM Sys t-3
TTN_{t-1}	0.997*** (0.001)	0.895*** (0.019)	0.608*** (0.051)	0.973*** (0.007)
AR(1)	1.79		-3.61	-3.86
AR(2)	1.26		0.83	-1.79
Sargan			0.025	0.999
GDP_t	OLS Levels	Within Groups	GMM DIF t-3	GMM Sys t-3
GDP_{t-1}	0.724* (0.006)	0.490* (0.0007)	0.487* (0.0004)	0.3835* (0.008)
AR(1)	-6.17		-5.13	-5.11
AR(2)	-5.89		-5.10	-5.06
Sargan			96.98	102.57

Notes: Quarters dummies included in all models (but not reported) Quarters dummies variables are also included as instruments in GMM-diff. specifications. Standard errors reported in parentheses. AR (1) and AR (2) are tests for first-order and second order serial correlation, asymptotically $N(0, 1)$. These test the levels residuals for OLS levels, and the first-differenced residuals in all columns. GMM results are two-step estimates. Sargan is a test of the overidentifying restrictions for the GMM, asymptotically χ^2 . P-value is reported. This test uses the minimized value of the corresponding two-step GMM estimators. ***, ** and * indicate coefficient at the 1%, 5% and 10% level, respectively.

REFERENCES

- Arelano, M. and S.R. Bond (1991), Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, *Review of Economic Studies*, 58, 277-297.
- Arelano, M. and O. Bover, (1995), Another Look at the Instrumental Variable Estimation of Error Component Models. *Journal of Econometrics* 68, 29-51.
- Baum, C. F. (2006): "An Introduction to Modern Econometrics Using Stata". Stata Press. College Station, Texas.
- Bhattacharya, S. (1982), Aspects of Monetary and Banking Theory and Moral Hazard. *J. Finance* 37 (May 1982): 371-84.
- Bhattacharya, S. (1984), Tournaments and Incentives: Heterogeneity and Essentiality (mimeo).
- Blundell, R. W and S.R. Bond (1998), Initial Conditions and Moment Restrictions in Dynamic Panel Data Models, *Journal of Econometrics*, 87, 115-143.
- Bognanno, Michael L. (1994), CEO pay as a tournament prize. *Labor Law Journal*, 45, 290-315.
- Bognanno, Michael L., Corporate Tournaments. *Journal of Labor Economics*, Vol. 19, No. 2 (Apr., 2001), pp. 290-315.
- Bond, S. (2002): "Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice", Working Paper CWP09/02, London: The Institute for Fiscal Studies Department of Economics UCL.
- Bond, S.R., C. Nauges and F. Windemeijer (2002), Unit Roots and Identification in Autoregressive Panel Data Models: A Comparison of Alternative Tests, *mimeo*, *Institute for Fiscal Studies*, London.
- Bull, Clive; Schotter, Andrew; Weigelt, Keith. Tournaments and Piece Rates: An Experimental Study, *Journal of Political Economy*, 1987, vol. 95, no. 1
- Carmichael L. (1983), The Agent Agents Problem; Payment by relative Output" *Journal of Labor Economics*, 1: 37-50.
- Eriksson, Tor. Executive Compensation and Tournament Theory: Empirical Tests on Danish Data, *Journal of Economics*, Vol. 17, No. 2 (Apr., 1999), pp. 262-280

- Green, Jerry R., Stokey, Nancy L., A. (1983), Comparasion of Tournaments and Contracts." *Journal of Political Economy*. 91, 349-64
- Hansen, L.P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators", *Econometrica*, 50, 1029-1054.
- Harbring, Christine., & Irlenbusch, Bernd (2003). An experimental study on tournament design, *Labor Economics*, 10, 443-464.
- Harbring, Christine., Irlenbusch, Bernd. Incentives in Tournaments with Endogenous Prize Selection IZA DP No. 1340, discussion paper series, October 2004, German.
- Hisao, Cheng. "Analysis of Panel Data". *Econometric Society Monographs*. Second Edition.
- Holmstrom, Bengt, (1982), Moral Hazard in Teams, *Bell Journal of Economics*, Autum, 13, 324-40.
- Kahn, Charles M; Silva, Emilson C.D; Ziliak, James P. (2001), Performance-based wages in tax collection: the Brazilian tax collection reform and its effects. *The Economic Journal*, vol. 111, No. 468, pp. 188-205.
- Knoeber, Charles R.; Thurman, Walter N. (1994), Testing the Theory of Tournaments: An Empirical Analysis of Broiler Production, *Journal of Labor Economics*, Vol. 12, No. 2, pp. 155-179
- Lamber, Lambert A; Larcker, David F. and Weigelt, Keith (1981), Rank Order Tournaments as Optimum Labor Contracts, *The Journal of Political Economy*. LXXXIX, 841-64.
- Lazear, Edward P., and Sherwin Rosen, (1981), "Rank Order Tournaments as Optimum Labor Contracts", *The Journal of Political Economy*. LXXXIX 841-64.
- Main, Brian G.M., Charles A. O'REILLY III, and James Wade (1993), Top Executive pay: Tournament or Teamwork? *Journal of Labor Economics* 11(4), 606-28
- Malcomson J. (1984), Work Incentives, Hierarchy, and Internal Labor Markets, *Journal of Political Economy*. 92: 486-507.
- Malcomson J. (1986), Rank-Order Contracts for a Principal with Many Agents, *Review of Economic Studies*. LIII, 807-817.
- Nalebuff, Barry J. and Stiglitz, Joseph E. (1983), Prizes and Incentives: Toward a General Theory of Compensation and Competition, *Bell Journal of Economics*, Spring, 14, 21-23.

- O'Reilly, C. A., Main, B. G., & Crystal, G. S. (1988). CEO compensation as tournament and social comparison: A tale of two theories. *Administrative Science Quarterly*, 33, 257.
- Rosen, Sherwin. (1986): Prizes and incentives in elimination tournaments. *Am. Econ. Rev.* 76(4):701-15.
- Sargan, J. D. (1958): "The estimation of Economic Relationships Using instrumental Variables", *Econometrica*, 26, 393-415.
- Windmeijer, F. (2000): "A finite sample correction for the variance of linear two-step GMM estimators", *Institute for Fiscal Studies Working Paper Series*, N°. W00/19, London.
- Wooldridge J.M. (2002): "Econometric Analysis of Cross Section and Panel Data", *The MIT Press*. Cambridge, Massachusetts. London, England.