Draft

Effectiveness of Industrial Effluent Standards Enforcement in Montevideo, Uruguay

Marcelo Caffera[1]

Facultad de Ciencias Empresariales y Economía, Universidad de Montevideo, Prudencio de Pena 2440, CP 11600, Montevideo, Uruguay

Abstract

Unfortunately, the empirical literature on the enforcement of industrial emissions standards refer to case studies in the developed world, mostly the U.S. and Canada. There does not exist any example of this type of empirical work for Latin America. In fact, Dasgupta, et al. (2001) is the only example for the case of a less developed country. This constitutes a very important shortcoming because Latin America has a long tradition in water pollution control laws, but both public opinion and papers that have analyzed environmental policy in the region have regarded them as poorly enforced. Furthermore, many resources are being devoted to developing new regulations and instruments, but no effort is being made to assess the effectiveness of the existing ones. This paper contributes to fill this gap by empirically testing the effect of (a) plant-level economic characteristics, and (b) inspections and enforcement actions of the municipal and state governments on industrial plants' emissions of BOD5 in Montevideo, Uruguay, using monthly data of seventy four industrial plants during the period July 1997 – October 2001.

1. INTRODUCTION

The theoretical literature on enforcing pollution regulations is now quite extensive. (See, for example, Polinsky and Shavell, 2000; Heyes, 2000; Cohen, 1997) In contrast, the empirical literature is fairly recent. This literature basically deals with two issues. First, the effect of inspections, fines and other enforcement actions (letters, phone calls and enforcement orders) on pollution levels, non-compliance and self-reporting [Magat and Viscusi (1990), Laplante and Rilston (1996), Gray and Deily (1996), Nadeau (1997); Helland (1998), Gray and Shadbegian (2002)]. Invariably, these papers have shown that

[1] E-mail address: marcaffera@um.edu.uy

firms do react to enforcement actions by reducing levels of emissions, increasing reports and reducing the length of time they spend in violation. A second issue is the determinants of the allocation of enforcement actions among the regulated plants. With respect to it, the hypotheses tested have been mainly two: the existence of certain types of targeting on the part of regulators [Gray and Deily (1996); Helland (1998); Gray and Shadbegian (2002)], and the role played by political considerations, such as the possibility that the plant will be forced to shut down, the per capita income and the level of pollution in the surrounding community [Helland (1998)]; the probability of closing and impact on employment rate in their communities [Gray and Deily (1991)]; the plant's share of the employment in the local labor market and the level of unemployment in a region [Dion, et al. (1998)]. These papers showed that regulators are sensitive to economic, social and political considerations when allocating enforcement actions among regulated plants.

Unfortunately, all of the above cited papers refer to case studies of the U.S. and Canada (basically, BOD_5 emissions of the US pulp and paper industry, BOD_5 and TSS emissions of the Quebec pulp and paper industry and air pollution from the US steel industry). There does not exist any example of this type of empirical work for Latin America.[2] In fact, Dasgupta, et al. (2001) is the only example for the case of a less developed country.[3, 4] This constitutes a very important shortcoming because Latin

---

[2] Dasgupta, et al. (2000) conducted a statistical analysis on determinants of "environmental performance" in Mexico. However their work is not comparable with the previously cited papers. First, their data resulted from a survey of 236 plants. Plant managers/owners self assessed the compliance status of their plants in a five possibilities scale, and a plant was classified as compliant if it was "always" or "almost always" in compliance. The questionnaire was not designed to obtain information on the level of emissions. Second, the survey asked for the overall "environmental performance" of the plant. Consequently, answers referred to either water, air, toxic or non-toxic pollution.

[3] Their study is also the first to include levies as determinants of levels of emissions.

[4] There are a few other examples of empirical analyses of pollution regulation in LDCs. Wang, et al. (2002) used the same database as Dasgupta, et al. (2001) to test for the determinants of the enforcement activities. They found that private firms had less bargaining power, measured as the percentage of levy actually paid relative to what they should have paid, and those plants with higher expenditures on pollution abatement paid lower levies, suggesting that regulators may had been compensating firms for such

America has a long tradition in water pollution control laws, but both public opinion and papers that have analyzed environmental policy in the region (Russell and Powell, 1996; Eskeland and Jimenez, 1992; O'Connor, 1998; Tietenberg; 1996) have regarded them as poorly enforced. Furthermore, many resources are being devoted to developing new regulations and instruments, but no effort is being made to assess the effectiveness of existing regulations.

The objective of this paper is to contribute to fill this gap by empirically testing the effect of (a) plant-level economic characteristics, and (b) monitoring and enforcement actions of the municipal and state governments on industrial plants' emissions of BOD5 in Montevideo, Uruguay.

Through this estimation I address the following questions: (1) How effective have inspections and the different enforcement actions of both municipal and state governments been in terms of reducing BOD emissions? (2) Could this effectiveness be improved by a reallocation of enforcement actions among different type of plants? (3) Could this effectiveness be improved by substitutions among different monitoring and enforcement actions (sampling inspections, non-sampling inspections, compliance orders, fines)? This question is relevant since inspections and orders were almost the only actions used by regulators. Fines were rarely levied (See Table 1 below). If this is the expression of a strategy such as the one suggested by Garvie and Keeler (1994) in the presence of institutional and political "constraints", then a study like the one proposed here could estimate the effects of such a strategy in terms of pollution abatement (tons of BOD emissions).

---

investments. Pargal, Mani and Huq (1997) estimated the impact of inspections and community characteristics (acting as proxies for political power) on water pollution in eight states of India. Their sample included 250 industrial plants surveyed in 1996. "Inspections" were the total number of inspections that the plant had been subject to in the period 1990 – 1994. They found that BOD emissions "are unaffected by inspections", The authors recognized that this result is conditional on the nature of their database, which did not allow them to analyze the impact of lagged inspections. They also found little evidence of informal enforcement (as measured by the community characteristics proxy). Gupta and Saksena (2002) estimated a relationship between inspections and compliance in the State of Punjab, India. However, their results may have been affected by the poor quality of their database: …"there is no comprehensive database" in India according to the authors.

A fourth objective would be to test for the presence and extent of under-reporting. Magat and Viscusi (1990) considered the issue was not a serious problem. Helland (1998) assumed its existence and constructed a model to correct for undetected unreported violations. Laplante and Rilstone (1996) were the only ones that attempted to test falsification of results by firms conducting a paired difference of mean test between the levels of BOD and TSS self-reported by the firms and those obtained through sampling inspections. These authors did not found statistical evidence of falsification of results, although the test was performed only with 54 observations. Shimshack and Ward (2002) opted for another approach, given their data. They tested whether current inspections, after correcting for inspection targeting, had any effect on reported emissions. They did not find evidence of inaccurate self-reporting.

A unique feature of the proposed research is the availability of a third source of information on emissions, apart from the plants themselves and the regulators. During part of the analyzed period the municipal government of Montevideo implemented a Monitoring Program in the framework of the IADB-financed Sewage Plan for the city. (See Section 2 below). An independent consulting firm that conducted its own samples of emissions ran the program. As a result, information on emissions is available for thirty-eight (38) plants for six (6) out of the thirteen (13) reporting periods studied. I will try to use this information to test for the extent of under-reporting and the effects of enforcement actions. Nevertheless, results may need to be interpreted with caution because during this period industrial plants knew they were been sampled and therefore may have changed their reporting strategy.

A fifth issue considered is the effect of inspections and enforcement actions on the compliance status of firms.

Finally, the present paper also differs from the existing literature because of the set of variables included in the estimation. This will be the first paper to include the three main types of monitoring and enforcement actions (inspections, orders and fines). In addition, previous analyses did not include production and input consumption variables. Helland (1998) was the only study in which level of production was included. No study included levels of input consumption apart from number of employees. Also, I have information on both the quantities of sampling and non-sampling inspections, the results of the sampling inspections, and quantities and values of fines collected by both the municipal and state government.

2. INDUSTRIAL WATER POLLUTION CONTROL IN MONTEVIDEO

Montevideo, capital of Uruguay, has a population of approximately one million and a half, almost half of the population of the country. The main water streams that run across the city are the Miguelete, Pantanoso and Carrasco. These, jointly with the Río de la Plata, receive the domestic and industrial effluents of the city, along with solid wastes of squatters at their margins (I.M.M., 2001). The estimated present contributions of each source of organic pollution are: (1) Industrial Effluents: between 16 and 20 tons of DBO_5 per day; (2) Sanitary system: 50 tons of DBO_5 per day, and (3) Squatters: approximately 120 tons of DBO_5 per day (originating from 300 tons of solid wastes, near 25% of the total generated by the city). It must be said though that the increase in squatters is a rather new phenomenon in Montevideo (see Amarante and Caffera, 2002) and the problem with households organic pollution through the sanitary system is being addressed with the construction of a new system that aims to dispose the effluents directly into the Río de la Plata. Furthermore, organic pollution is the main quantitative type of industrial pollution.

Water pollution control policy in Montevideo is in the hands of both the city and state government. More specifically, the Industrial Effluents Unit (Unidad de Efluentes Industriales, UEI) of the Municipal Government of Montevideo (Intendencia Municipal de Montevideo, IMM) and the Environmental Control Division (División de Control Ambiental (DCA) of the Ministry of the Environment (Ministerio de Vivienda, Ordenamiento Territorial y Medio Ambiente, MVOTMA). Firms report emissions and other economic variables to the Industrial Effluents Unit (UEI) of the Municipal Government of Montevideo (IMM), which is also in charge of regular monitoring inspections and enforcement activities. The Environmental Control Division (DCA) of the Ministry of the Environment is in charge of giving the authorization for industrial discharge (Autorización de Desagüe Industrial, ADI). Some coordination existed between this two offices between 1996 and 2001, during which an informal agreement conveyed to the city government unit (UEI) the exclusivity to make regular inspections in Montevideo, allowing the state government division (DCA) to concentrate its monitoring and enforcement efforts in the rest of the country. But this coordination began to deteriorate in the year 2002.

Instruments used are emissions standards, defined in terms of concentration of a given pollutant per liter of effluent discharged, and not in terms of quantities of pollutants discharged.

Through 1997 to 1999 the Municipal Government (IMM) established a Plan (Plan de Reducción de la Contaminación Industrial) that relaxed emissions standards for almost all the pollutants and all industrial sectors. The plan, approved in 1996 as Resolution 761/96 of the Department of Environmental Development of the IMM, was intended to bring plants into compliance by giving them a period of adjustment in order to start investing in abatement technology. The application of the plan started in March 1st 1997 and ended in December 31st 1999, day when the standards converged again to their initial values. Leather tanners and wool processors had laxer standards, converging to levels above the initial ones. As stated in the Resolution, this plan was inspired by the recognition of the difficult economic situation of the industrial sector, from the part of the municipal government (IMM).

As said, every four months plants report to the municipal government unit (UEI) monthly quantities of the following variables: (1) the level of production for each good produced, (2) the level of water consumption, including underground water, (3) the level of electric energy consumed, (4) the level fuel and/or firewood consumed, (5) the number of employees and days worked. Plants are also required to report samples of their discharges with the following information: total effluent flow and its concentration of pollutants. Plants with an effluent flow larger than 50m3/day are required to take samples every two weeks instead of monthly.

The city government unit (UEI) conducts regular monitoring inspections. There are two types of inspections: sampling and non-sampling inspections. Sampling inspections are inspections in which the inspectors take samples from the plant‘s effluents for latter analysis. These inspections always include inspections of treatment plant performance, the overall treatment process performance, as well as general questions regarding the economic situation of the firm, including changes in levels of production, or special events that could affect the effectiveness of treatment processes. Non-sampling inspections include the latter but not a sample of the plant effluents. Reasons for not sampling may be several: the plant is not working at the time of the inspection, or the plant is not discharging effluents at the time of the inspection.[5]

3. Data Set

I have three sources of information: the Municipal Government of Montevideo (Intendencia Municipal de Montevideo; I.M.M.), the Ministry of the Environment and a private consulting consortium, (MULTISERVICE-SEINCO-TAHAL; SEINCO). The core information comes from the IMM. As explained above, industrial plants in Montevideo report economic and pollution variables levels to the municipal government unit (UEI) on a four-month basis. From these reports I obtained monthly information on a set of variables that I divide into three categories (1) Pollution Variables, (2) Production Variables, and (3) Input Variables. The first category, Pollution Variables, is composed of Biological Oxygen Demand (BOD5) concentration of the industrial plants effluent discharges measured in mg/l, the average monthly flow of discharges (FLOW) measured in cubic meters (m3)/day, and the total volume of effluents discharged in the month (VOLUMES), in m3 per day. The second category, Production Variables, is composed of monthly levels of production for each good produced. The third category, Input Variables, is composed of a list of key inputs that the plants are required to report to the UEI. These are: tap water consumed per month (TAP) in m3, underground water consumed per month (UW) in m3, total water consumed per month (W) in m3, electric energy consumed per month (EL) in Kwh, fuel consumption per month (FUEL) in tons, firewood consumption per month (WOOD) in tons, gas consumption per month (GAS) in tons, number of employees per month (EMPLOY) and number of days worked per month

---

[5] Some plants discharge discontinuously and/or at given hours. This represents a problem for the DCA inspectors, who usually reserve specific days for inspections in Montevideo since they also have to inspect firms in the rest of the country.

(WD). Finally, I also gathered information from the IMM records on inspections and fines. The information on inspections is composed by the number of inspections done per month per plant, and the result of the sample in terms of mg/l of BOD_5 in those cases where a sample was taken. The information on fines levied by the UEI is composed of the number of fines levied on each industrial plant per month and their amounts.

My second source of information is the Environmental Control Division (Division de Control Ambiental, DCA) of the Ministry of the Environment. This information includes number of inspections, enforcement orders, postponements, fine threats, and fines per plant per month. In the case of inspections, results of samples in terms of BOD_5 effluent concentration (in mg/l) were also computed. In the case of fines, amounts levied are also available.

Finally, my third source of information is SEINCO, the name chosen for the private partnership MULTISERVICE-SEINCO-TAHAL that was in charge of the Monitoring Program that the IMM implemented in 1998 as part of the Third Stage of the Urban Sanitary Plan (Plan de Saneamiento Urbano – Tercera Etapa, PSUIII, financed by the IADB). The main objective of the Monitoring Program was to design, implement and execute a watercourses and industrial effluents monitoring scheme for the control of industrial pollution (Multiservice-Seinco-Tahal, 2001).

My database includes seventy-four (74) industrial plants located in Montevideo. The selection of these 74 plants was not random. First, there are all privately owned plants. Second, they were selected from a list of industrial plants that were being sampled by SEINCO during the years 2000 and 2001. Most of these plants also were the ones that were regularly being inspected by the UEI. The list included the most important industrial polluters in the city.[6] Of the eighty-seven plants in the list I excluded twelve (12) plants that reported less than six (6) times during the 13 four-month periods in my sample,

[6] It included a maximum of eighty-seven (87) industrial plants in November 2000 – February 2001. The number of plants in the list did not remain fixed during the consulting period of SEINCO. For example, in March- June 2001 there were seventy-eight (78) firms in the list. The reasons, according to SEINCO employees interviewed, were that some plants closed and others were inactive during some periods. In these cases, the "next plants in the list" of the most important polluters of the city (made from a previously performed census) were included and regularly inspected.

although they were active throughout the 13 periods. From the remaining 75 I had to exclude one more because it was not reporting BOD_5 emissions; it reported only metals emissions. Consequently, conclusions from my analysis must be interpreted according to this sample selection bias. It can be said though that this bias is intrinsic to this type of empirical analysis.

In order to conclude this subsection I present in Table 1 descriptive statistics for Input, Pollution and Enforcement Variables (descriptive statistics for Production variables are not presented).[7]

Table 1: Descriptive Statistics for Input and Pollution Variables in Database (Sample July 1997 – October 2001)
Total Potential Observations: 3,848

Variable	Mean	Std. Dev.	Missing Values
BOD_5 (mg/l)	1031	2334	952
FLOW (m3/day)	203	453	1034
TAP (m3/month)	3848	8271	638
UW (m3/month)	2792	4873	1279
EL (Kwh/month)	179409	278828	449
FUEL (m3/month)	34	50	862
WD (per month)	22	4	594
EMPLOY (number of employees)	145	610	342

[7] These number does not include the missing values for VOLUMES, GAS and FIREWOOD. GAS could not be used in the analysis because the IMM did not ask for it before 2001, and in 2001 only one plant reported gas consumption in two reporting periods. The problem with FIREWOOD is that not all of the industrial plants in the sample used firewood as an input and not all of those who did not use it reported zero consumption. Instead, a value was missing in the respective cell. Possibly this reporting defect is the consequence of rotation in the professional in charge of preparing and submitting the reports. Thirteen plants did not report firewood consumption for the entire sample period, and 32 plants alternated non-reports of firewood consumption with zero consumption, suggesting that in fact they were not using firewood as an input. Given these, I discarded these two variables from the analysis.

Table 2: Descriptive Statistics for Monitoring and Enforcement Variables – IMM and DCA
(Sample July 1996 – October 2001)
Total Observations: 4,736

	Mean	Std. Dev.	Maximum	Sum
INSPWSIMM	0.085	0.286	3	401
INSPWOSIMM	0.031	0.212	6	148
FINEIMM	0.003	0.052	1	11
FINEIMM (UR)	0.258	6.04	200	1030
INSPWSDCA	0.026	0.158	1	122
INSPWOSDCA	0.019	0.137	2	89
ORDERDCA	0.024	0.155	2	112
POSTDCA	0.013	0.123	2	60
FTDCA	0.015	0.126	2	72
FINEDCA	0.0008	0.029	1	4
FINEDCA (UR)	0.190	6.657	300	900

Note: Observations for FINEIMM were available from May 1997 (3,996 observations)

Every variable ends with the name of the institution to which the information belongs: DCA and IMM. “ORDER” stands for the total number of administrative or compliance orders issued. Types of these include: an order to present the “Industrial Discharge Application” (Solicitud de Autorizacion de Desagüe Industrial, SADI) form, an order to present periodic reports of the treatment plant (TP) performance; an order to finish the construction of the TP; an order to present the “Start of Operation Report” (Informe de Puesta en Operacion, IPO); an order to designate a competent professional as responsible for the TP operation; an order to present a “Technical Report” (Informe Tecnico, IT); and an order to present modifications to the TP. “POST” stands for postponements. The DCA sometimes deferred the due dates set in the process of application for the SADI and for their orders. Consequently, actions postponed include the same list of actions as executive orders. “INSPWSDCA” and “INSPWOSDCA” correspond to inspections with and without sample conducted by the DCA in a given period. “FTDCA” stands for fine threats from the DCA. Passed the due date, the DCA issued a note communicating the plant that it was potentially subject to a fine due to non-compliance of the previous order. “FINEDCA” is equal to one if the plant was inspected in that period and zero if not. “FINEDCA(UR)” is the total amount of the fines levied in that period. "UR" stands for "Unidad Reajustable", a monetary unit indexed by wages. Its value was approximately US$ 15 in October 2001. Similar definitions apply for the "IMM" variables.

4. Missing values

As evidenced by Table 1, I have missing values (MV) in my panel. Observations are missing either because the plant did not report in a given period, in which case I have a missing value for the entire set of variables for that period ("unit non-report"), or because the plant did report but the report had missing values for one or a subset of variables ("item non-report").

There were four main reasons for unit non-report. First, the plant simply failed to submit a report. Second, the plant went out of business. Third, the plant reported no activity in that period.[8] And fourth, the plant had not yet started business in that period.

In summary (Table 3) there were a total of sixty (61) non-reports over a potential 962 observations (74 plants times 13 periods). Of these, two corresponds to two plants that shut down in period 13. Twelve corresponds to "no-activity" periods of five different plants. Thirteen (13) corresponds to two plant that started business in periods five and eight, respectively. The remaining thirty-four (33) corresponds to "random" non-reports.

[8] I treated these as missing values because in some cases the firms indicated (usually in a letter to the UEI Director) that they were producing "very low" quantities of products, and therefore it was not worth it to report emissions. Even more, this letter was frequently followed by non-reports in following periods without any clear information regarding the exact point in time in which production re-started. These last cases involved a total of five plants and twelve reporting periods.

Table 3: Distribution of Reporting Failures by Reason

Reason	Number of Non-Reports
"Went Out of Business"	2
"No Activity"	12
"Not in Business Yet"	13
"Random"	34
Total	**61**

The following two tables further break down the distribution of these non-reports.

Table 4: Distribution of Reporting Failures by Number of Industrial Plant

	Number of Non-Reports	Number of Plants
	8	1
	6	3
	5	2
	4	1
	3	4
	1	9
	0	54
Total	**60**	**74**

Table 5: Distribution of Reporting Failures by Period

Period	Number of Non – Reports	Period	Number of Non – Reports
1	9	8	4
2	5	9	2
3	4	10	4
4	4	11	7
5	4	12	6
6	3	13	6
7	3	**Total**	**61**

Item non-reports have also several reasons. Some firms never reported a specific variable. Others reported a specific variable unsystematically. For example, in the case of underground water consumption some firms reported zero consumption in some periods and did not report in others. Finally, other values appear to be randomly missing.

Taking into consideration item and unit non-reports there were a total of 4,777 observations missing for the inputs and pollution variables described in Table 1 plus the production variables reported by the industrial plants, out of a total of 40,144 possible observations. In other words, 11.9% of the data set was missing.[9]

4. Dealing with the Missing Observations

The problem with MV is that an estimation based only on the complete observations (those having no MV) may bias parameter estimates.

Several methods are used in the applied literature and others are proposed in a more recent theoretical literature to deal with MV. The issue when selecting a method to deal with MV is that some of them (for example, imputing means) may reduce the efficiency of the final estimators. A review of these methods, along with a discussion of their properties, can be found in Little and Rubin (1987) and Little (1992). For the case of panel data, a review of the literature of incomplete panels and selection bias can be found in Verbeek and Nijman (1992b). It is not the purpose of this section to review these methods, but to inform the reader about how I dealt with the missing observations.

4.1. "Missing at Random" and "Ignorability"

First, one should distinguish between the concepts of "missing completely at random" (MCAR) and "ignorability"(Little and Rubin, 1987). Called Z the complete data set. Z is an $n*(k+1)$ matrix, where n is the number of observations and k is the number of independent variables, excluding the intercept. Now, $Z = Z_{obs} + Z_{mis}$, where Z_{obs} and Z_{mis} are the subsets of observed and missing values, respectively. Define a "response indicator" matrix R, such that $r_{i,j} = 1$ if $z_{i,j}$ is observed and zero otherwise. Then Z_{mis} is

[9] An Appendix describing the distribution of missing values per variable by industrial plant is available from the author.

MCAR if $f(R/Z,q) = f(R,q)$ for all Z, where q is a scalar or vector that indexes the density function. That is, data is MCAR if the "missing-ness" is independent of the particular realization of the data at hand. The probability distribution of the missing observations does not depend on the particular sample at hand. Similarly, Z_{mis} is "missing at random" (MAR) if $f(R/Z,q) = f(R/Z_{obs},q)$ for all Z_{mis}. In other words, data is MAR if observations for one or more variables are missing when certain values are realized for other observed variables. Finally, data is not MAR if the missing observations depend on the values of the unobserved variables for that cases. In other words, you do not observe a certain variable or the whole set of variables when the value of some variable is larger or smaller than a specific amount. It is important to note that MCAR implies MAR, but the reverse is not true.

Practical estimation procedures use the concept of "ignorability" instead of the concept of MCAR. Ignorability is a weaker concept than MCAR. A missing data mechanism is said to be ignorable for both sampling-based and likelihood-based inferences when the data is MCAR, but it is also ignorable (only for likelihood-based inferences) when the data is MAR, although not MCAR. Finally, it is non-ignorable when the data is not MAR (Little and Rubin, 1987). Therefore a missing data mechanism can be ignorable for inferences purposes even if missing values are not MCAR.

Verbeek and Nijman (1992a) proposed formal tests for ignorability in panel data. These test are worth performing because of the complexities involve in estimating a panel incorporating the selection rule. The advantage of the tests proposed are its simplicity and the fact that they take into accounts both wave (unit) and item non-response (although the authors refer to the latter as when information on the endogenous variable is missing, and they restricted their attention to linear regression models).

The general idea of the tests is to compare the estimates obtained by using only the available observations with the estimates obtained using only the complete observations. Using the available observations for each unit (plant) leads to an unbalanced panel. Using only complete observations, that is, only those units observed during all periods for all variables leads to a balanced panel. Differences between estimators from the balanced and unbalanced panels can be used to construct a "simple (quasi-) Hausman test for selectivity bias" (p. 683). The test can be performed for both the fixed effects (FE) and the random effects (RE) models. The test is based on the idea that were the selection rule is ignorable, there would be no reason why the estimates obtained using the balanced panel should differ from those obtained using the unbalanced panel, since there is no reason why the inconsistencies of the estimators from the balanced panel would coincide with the inconsistency of the estimator of the unbalanced panel.

I faced three problems that prevented me from using these tests. First, these tests do not seem to take care of item non-responses for the cases of right hand side variables, only in the case of the dependent variable. Second, Verbeek and Nijman tests assumed exogeneity of right hand side variables. Third, and most determinant, I have zero valid observations for my balanced sub-panel (I have no month with observations for all variables for all 74 plants).[10]

Given these reasons, I did not perform the tests. The consequences may not be serious for several reasons. First, I think that it is fairly simple to conclude that there exists selection bias in my data set. I have twelve (12) observations missing as a consequence that the plants informed "no activity" or "very low" activity. Missing-ness is clearly related to the level of production in those cases. In other words, the selection rule is not independent, among other possible things, of the overall economic situation of firms. These twelve cases make my selection rule not ignorable, in spite of the fact that apart from them I have another thirty-three (33) non-report cases whose causes are not as clear. Second, even when the missing data mechanism (also called the selection rule) is "ignorable" and the estimators obtained using only the balanced sub-panel are consistent, it will pay to use all the information available in the original (unbalanced) panel, since this would produce more efficient estimators. The latter is particularly true in those cases when many individuals are incompletely observed, as it is my case. Consequently, I proceed assuming that my missing data mechanism is non ignorable and an unbalanced panel.

[10] This was also true after imputing for item non-responses as explained in the following section.

4.1. Imputing item non-responses

In spite of the fact that I proceed with an unbalanced panel, I impute for the item non responses before estimating my panel. The reason was that item non-responses count for 40.9% of the total 4,777 observation missing.

According to the literature on missing values, there are basically two ways to proceed when imputing values for item non-reports: conditional mean imputation or multiple imputation (Little, 1992).

Conditional mean imputation methods are based on Buck (1960), Dagenais (1973) and Beale and Little (1975). The basic idea is to use the information on the observed X´s or on the observed X´s and Y´s to fill in missing values, correcting for the variances and covariances. LS on the filled-in data produces consistent estimates assuming MCAR.

Multiple imputation (Rubin, 1987) is proposed as a way to handle the problem that whatever the conditional mean imputation procedure, "estimated standard errors of the regression coefficients from OLS or WLS in the filled-in data would tend to be too small, because imputation error is not taken into account." (Little, 1992, p. 1232). By multiple imputation (Rubin, 1987), basically, one imputes $m \geq 2$ values for each missing observation to obtain *m* different data sets. With each data set one obtain the desired estimates, and "average" them to obtain a final parameter estimate and variance estimate that "correct" for the underestimation of variances produced by filling in missing observations.

Both type of methods were developed and applied for cases of cross section data and therefore share a problem when applied to panel data. First, it makes little sense to fill in item non-responses in one plant conditional on information observed for the rest of the plants, with different technologies and output. This critique is valid also for multiple imputation if I perform it based on the entire panel. In both methods I am "averaging" across units and time.

A possible solution to this problem is to perform the imputations within units (plants). By this way I not only preserve between plants variability, minimizing bias and variance problems for the final estimates, but I also use plant-specific information about the missing values.[11] But again if I perform multiple imputation within units, in order to preserve variability, it will not be very helpful since it would produce *m* data sets for each different unit.

Consequently, I used an iterated Buck procedure within plant to impute for item non-reports, in the spirit of the suggestion made by Beale and Little. I present this iteration briefly below.

Assume there is a data set consisting of *N* observations and *k* variables, but one or more of the *k* variables are not observed in some of the *N* observations. Define the following variables:

$\tilde{x}_j = \sum_{i \in C} x_{ij}$; where $C<N$ is the subset of complete observations. $\tilde{x}_j$ is then the average of the variable x_j over the complete observations.

$\hat{x}_{ij}$ is the filled-in data where $\hat{x}_{ij} = x_{ij}$ (the observed value) if the variable j is observed in the observation i or $\hat{x}_{ij} = \tilde{x}_j + \sum_{l \in p} b_{jl}(x_{il} - \tilde{x}_l)$ i.e.: the fitted value of a linear regression on the p observed variables for that observation.

$\bar{x}_j = \sum_{i=1}^{N} \hat{x}_{ij} / N$; the mean of variable j over the filled-in data.

$a_{jk} = \sum_i (\hat{x}_{ij} - \bar{x}_j)(\hat{x}_{ik} - \bar{x}_k) + c_{ijk}$; the jk^{th} element of the corrected matrix of sums of squares and products, where c_{ijk} is the corrected term. c_{ijk} equals the residual variance computed form the regression of x_j on the observed variables in that observation i over the complete cases, if only x_j is missing in observation i, or the residual covariance computed from the regression of x_j and x_k on the rest of the observed variables in that observation if both x_j and x_k are missing in that observation, always regressing over the

[11] Such as "Montlhy volumes of effluents discharged" divided by "Days Worked in the month" to inpute FLOW.

complete cases. In mathematical notation, call v_{jk} the covariance of $(x_j - \sum_p b_{jp} x_p)$ and $(x_k - \sum_p b_{kp} x_p)$ where p is the subset of observed variables in the observation in question. Then, $c_{ijk} = v_{jk}$ if x_j and x_k are both unknown and 0 otherwise.

The steps of the version of the iterated Buck´s procedure proposed by Beale and Little are:

1) Fit all the missing items as suggested by Buck and compute a_{jk}.
2) Calculate $\bar{x}_j$ and substitute it for $\tilde{x}_j$ in $\hat{x}_{ij} = \tilde{x}_j + \sum_{l \in p} b_{jl}(x_{il} - \tilde{x}_l)$
3) Repeat until $\bar{x}_j$ and a_{jk} have no further significant changes.

To perform this procedure I constructed the following variables for each plant: (1) *WATER = TAP + UW*: Total water consumption in m3/month; (2) *ENERGY = EL*3.6 + FUEL*43,752.06*: Total energy consumption in mega joules (MJ); (4) *LABOR* = *WD*EMPLOY*: Total days-employee worked; (5) *POLLUTION = FLOW*BOD5*1000*: Total organic pollution discharged in (mg/day); (6) *PRODUCTION* = Quantity of good(s) produced by month. The original variables were fitted after fitting these constructed variables. I estimated the linear auxiliary regressions with the variables in natural logarithms forms. These did not necessarily always provided better fits than linear auxiliary regressions with variables in original form, but they are closer to "the spirit" of a Cobb-Douglas type of production function.[12]

Finally, I do not use the monitoring and enforcement variables in this imputations. This is good for two reasons: first, I conserve degrees of freedom in the auxiliary regressions within firms, and second, it would be like cheating to use these variables to impute for the MV and then use the resulting data to test for the effect of them in pollution.

[12] An Annex in which I discuss in detail the processes followed to impute for item non-responses in each plant and the corresponding iteration procedures are available from the author upon request.

5. The Model and Estimation issues

In the textbook static case, a polluting firm is assumed to be a risk neutral profit maximizing unit. In such a case, and with the information at hand, the profit function in a given month t for a given plant i would be the following:

$$
\begin{aligned}
E(p_{i,t}) = P_{q_i,t} * Q_{i,t} - w_{i,t} * (Labor_{i,t}) - P_{W,t} * (TAP_{i,t}) - P_{UW,t} * (UW_{i,t}) \\
- P_{EL,t} * (EL_{i,t}) - P_{Fuel,t} * (FUEL_{i,t}) \\
- E(Insp_{i,t}) * FINE_{i,t}(BOD5_{i,t} - \overline{BOD5})
\end{aligned}
$$

i =1,…, 74, and t = July 1997, …., October 2001.

where $E(p_{i,t})$ = expected profit, $P_{q_i,t}$ = price of the good produced, $Q_{i,t}$ = quantity produced, $w_{i,t}$ = wage, $Labor_{i,t}$ = total number of employee-days worked, $P_{w,t}$ = price of tap water, $TAP_{i,t}$ = tap water consumed, in m^3, $P_{UW,t}$ = cost of underground water, $UW_{i,t}$ = underground water consumed, in m3, $P_{EL,t}$ = price of Kw/h, $EL_{i,t}$ = electric energy consumed, in Kw/h, $P_{Fuel,t}$ = price of fuel, $FUEL_{i,t}$ = fuel consumption, $E(Insp_{i,t})$ = expected probability of inspection, $BOD5_{i,t}$ = Biochemical Oxygen Demand concentration of discharges, in mg/l, $\overline{\mathrm{BOD5}}$ = maximum concentration level of BOD_5 per litter allowed by legislation, $FINE_{i,t}(BOD5_{i,t} - \overline{\mathrm{BOD5}})$ = fine corresponding to the level of violation.[13]

In real life firms maximize, in a given month, the present value of future expected profits. To make things simpler I suppose that prices and quantities are known with certain for every period. This is not the case for inspections and fines. There exists a probability of being inspected in a given month and also fines are not applied instantly and automatically as stated in the legislation. This is not only because in real life the

---

[13] Fines are not the only penalty for not complying. Plants can also be temporarily closed. Nevertheless, neither the municipal nor the national government have records of these measures. It can be said though that these types of measures were more rare than fines.

process of fining a plant takes time, but also it may be the consequence of several issues ranging from the present economic situation of the firm, as perceived by the regulators; the ability of the firm to fight the penalty in the judicial system and the willingness of this system to decide against firms (all of what Garvie and Keeler (1994) called the "socio-legal" institutional environment). The point is that future fines are also not certain for the firm, either in time or amount. So when deciding how much to emit in a given period the firm must assign a present value to the future fines that may derive from the present level of chosen emissions. In such a case the expected profit function would look like:

$$PVE(p_{i,t}) = \sum_{s=0}^{s=n} \frac{1}{(1+r)^s} \begin{bmatrix} P_{q_i,t+s} * Q_{i,t+s} - w_{i,t+s} * (Labor_{i,t+s}) - P_{W,t+s} * (TAP_{i,t+s}) \\ - P_{UW,t+s} * (UW_{i,t+s}) - P_{EL,t+s} * (EL_{i,t+s}) \\ - P_{FUEL,t+s} * (FUEL_{i,t+s}) \\ - E(Insp_{i,t+s}) * PVE\left(FINE_{i,t+s}(BOD5_{i,t+s} - \overline{BOD5})\right) \end{bmatrix}$$

where $PVE(p_{i,t})$ = present value of expected profit of plant i in month t, n = relevant time horizon for plant manager, r = discount rate for plant manager, and $PVE\left[FINE_{i,t+s}(BOD5_{i,t+s} - \overline{BOD5})\right]$ = present value of expected fines in month $t+s$. Given this profit function, the plant chooses an optimal level of emissions path in month t. The correct way to proceed would be to solve formally for BOD_5 from the Kuhn-Tucker conditions of the stochastic dynamic programming problem of the plant manager, to obtain the optimal emissions path. This way to proceed turned out to be not trivial and I am presently working on it. Therefore what I present below are the preliminary results of the estimation of a linear equation in the spirit of the previous literature (Magat and Viscusi, 1990; Laplante and Rilstone, 1996; Dasgupta, et al., 2001). Such an equation would be:[14]

---

[14] I should point out that this linear equation differ with the one I obtain with the formal derivation of the dynamic programming problem in that the coefficients of the enforcement variables enter the equation in a linear fashion, and the fact that the future is not taken into account.

$$
\begin{gathered}
\ln\left(BOD5_{i,t}\right) = l_0 + l_1 * \ln\left(P_{q,t}\right) + l_2 * \ln\left(Labor_{i,t}\right) + l_3 * \ln\left(Water_{i,t}\right) \\
+ l_4 * \ln\left(Energy_{i,t}\right) + l_5 * \ln(Flow_{i,t}) + l_6 * E\left[Insp_{i,t}\right] + l_7 * INSPCUM_{i,t} \\
+ l_8 * ORDERCUM_{i,t} + l_9 * FTCUM_{i,t} + l_{10} * FINECUM_{I,T} \\
+ m_i + u_{i,t}
\end{gathered}
$$

$i = 1, \ldots, 74$; t = July 1997, …, October 2001.

with variables as defined above.

BOD_5 in a given month is a function of the cumulative number of: inspections $\left[INSPCUM_{i,t} = \sum_{s=1}^{12}(Insp_{i,t-s})\right]$, enforcement orders $\left[ORDERCUM_{i,t} = \sum_{s=1}^{12}\left[Order_{i,t-s}\right]\right]$, fine threats $\left[FTCUM_{i,t} = \sum_{s=1}^{12}\left[FT_{i,t-s}\right]\right]$, and fines $\left[FINECUM_{i,t} = \sum_{s=1}^{12}\left[Fine_{i,t-s}\right]\right]$ received in the last year.

Finally, m_i is plant specific effect. I chose a fixed effect model, as oppose to a random effects, given that I am basing my inference on this 74 specific plants, which were not randomly selected from a large population. $u_{i,t}$ is the remainder stochastic disturbance assumed IID$(0, s_u^2)$.[15]

The previously cited literature include the contemporaneous number of inspections or a dummy indicating whether the plant was inspected or not in that month. I do not consider this a possibility because the fact that a plant is inspected in a given month cannot have an effect on the average level of pollution in that month, given that this depends more on decisions already taken at that moment, regarding production,

---

[15] I am currently working on the construction of a very important variable missing from the equation. This would capture the effect caused by modifications to the treatment plant. Some plants, ordered by regulators, modify their treatment plants during the period because the original treatment plants were not able to make them comply with the emissions standards. Information regarding the exact date in which each plant put into operation its modified treatment plant is difficult to obtain.

technology, etc. Furthermore, the sample from which the reported concentration of BOD_5 comes from, could have well been taken before the inspection took place. What I include in the equation is the expected probability of inspection. I calculate this expected probability of being inspected as the fitted value of a probit inspection equation.

It can be said the inspection strategy of regulators obeys four rules.[16] The first one would be a "sample without replacement" rule. The time that takes the regulator to "sample" all plants is six months. During this length of time the regulator tries to visit two times Priority 1 plants and one time Priority 2 plants. Priority 1 plants (25 of the 74 plants in my sample) are the most heavy polluters in terms of organic pollution and metals. They count for 80% of this pollution. Second, plants with worse compliance history and those showing less "cooperation" with regulators (they do not take the promised measures to abate emissions or delay the modifications to their treatment plant) are inspected more often. Third, citizens complaints about unusual emissions episodes also trigger inspections. Finally, unusual levels of reported pollution and the failure to report in subsequent periods may trigger inspections. As a result, inspections can be modeled as a function of: (1) the number of inspections performed in the plant during the last twelve months, (2) the priority group to which the plant belongs, (3) the number of detected violations, compliance orders (and postponements) issued to the plant in the last twelve months, (4) the number of reporting failures in the previous two reporting periods:

$$\begin{aligned} Insp_{i,t} = g_1 * \left[\sum_{s=1}^{12} (Insp_{i,t-s})\right] + g_2 * \left[\sum\nolimits_{s=1}^{12} \left(* DV_{i,t-s}\right)\right] + g_3 * \left[\sum\nolimits_{s=1}^{12} \left(* Order_{i,t-s}\right)\right] \\ + g_4 * \left[\sum\nolimits_{s=1}^{12} \left(Post_{i,t-s}\right)\right] \\ + g_5 * Vol_t + g_6 RF_{i,t} + g_7 Pty_i + h_{i,t} \end{aligned}$$

[16] The discussed strategy is that of the municipal government unit (UEI). As mentioned, also the national government office (DCA) conducted inspections in the period. Nevertheless, as also mentioned, a previous arrangement between these two offices had left main regular monitoring activity in the hands of the UEI.

where *DV* is a dummy variable equal to one if the plant was inspected and found out of compliance with the emissions standards in that month.[17] *Order* is another dummy variable equal to one if a compliance or other type of order was issued to the plant in that month. *Post* is also a dummy variable equal to one if the DINAMA gave more time to the plant to comply with a previous order. These three variables are included as proxies of the level of cooperation commented above. The more of these recent records, the less the cooperation of the plant.[18] It must be said though that this level of cooperation perceived by regulators is not only a function of the recent formal history of the plant. It depends also on incommensurable facts in which inspectors also base their decisions.[19] $RF_{i,t}$ (= 0, 1, or 2) is the number of reporting failures in the previous two reporting periods. In the first reporting period I set the reporting failure history of every plant equal to zero as an indicator the a new enforcement period has began. Pty_i is a dummy variable equal to 1 f the plant is a Priority 1 plant and $h_{i,t}$ is the error term, assumed to be identically and independently distributed normal variables with zero mean.

---

[17] In the case of water pollution in Uruguay the law does not punish non-compliance with emission standards but punishes actions related to the maintenance and operation of the treatment plant (which supposedly result in emissions' concentration levels above the standards). In the legislation, fines are set as an increasing function of the number of past offences of this type. In spite of this, DV are defined in this way because non-compliance with emissions standards continue to be the most important indicator that the treatment plant is not being operated properly. Nevertheless, it is not a perfect proxy if we take into account the cases when a new or modified treatment plant is put in operation and subsequent samples are taken to confirm that it is effective in making the industrial plant to comply with the standards.

[18] I do not have the number of compliance orders issued by the municipal government of Montevideo, just those issued by the national government office, DINAMA.

[19] An example is the following: sometimes inspectors are kept waiting at the plant entrance for the length of time needed to make some quick cleanings and other measures (like diluting) to comply with the emissions standards (this is more typical of small plants, with lesser time of effluents retention). Another example is the quickness to response to suggested changes. It is worth noting that this makes the effectiveness of water pollution control very dependable on those specific inspectors with long experience in the job. In other words, a lot of the compliance history of plants is lost when an inspector retires or is appointed to another office.

The plant manager updates its expected probability of being inspected and fined with information that starts in July 1996. This of course obeys to the period covered by the sample but it also has real life sense. In the first six months of 1997 the UEI implemented a new enforcement strategy. It issued a fax to every plant in its data base explaining the new 4-month Reporting Form and communicating the plants that the municipal government was undertaking new efforts in pollution control. Therefore, in July 1997 plant managers had to learn the new rules of the game.

The Uruguayan industrial sector was going through an important contraction process during part of the analyzed period. Particularly, the industry production volume index dropped 8.6% on average in 1999 and 7.2% in 2001 (during 2000 it experienced a positive 2% change). The contraction was larger as measured by the industry real GDP variation: 23% between 1996 and 2001, with an average drop of 4% in the period 1997 – 2001 and 8% in the period 1999 – 2001. Although not recognized by authorities, as a consequence of this contraction of the industrial sector inspectors may have eased or loosened their enforcement pressure on plants. I include the monthly level of the industry production volume index (*Vol*) to capture this possible effect.

Results of this estimation are presented below:

Table
Inspections Equation

Dependent Variable: INSP				
Included observations: 3848				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
INSPCUM	0.088	0.012	7.47	0.000
DVCUM	-0.033	0.032	-1.044	0.296
ORDERCUM	0.009	0.043	0.214	0.831
POSTCUM	0.072	0.049	1.478	0.139
VOL	-0.012	0.0003	-31.63	0.000
RF	0.02	0.07	0.285	0.775
PTY	0.119	0.052	2.269	0.023

The Wald statistic for the overall goodness of fit of the model was 1750.125. Although a preliminary regression, results show that INSPCUM is statistically significant and its coefficient would indicate that regulators do target plants in the sense that they seem to concentrate inspections on plants that are being inspected more regularly. The coefficient of DVCUM, although not statistically significant, is difficult to interpret. It may tell that regulators do not take into consideration violations to emissions standards in the compliance history of the plant as they take the level of cooperation from the part of the plant manager, as captured by ORDERCUM and POSTCUM (although both variables are not statistically significant). VOL is the other statistically significant variable. The objective behind including this variable was testing the idea that regulators had eased enforcement efforts due to the contraction process suffered by the industrial sector. Results are suggesting that there is no evidence that inspectors have eased their monitoring (inspection) efforts. This result is not entirely against the original idea. It may well be possible that inspectors increased their monitoring (inspection) effort decreasing their enforcement efforts (fines) for the same reason. Finally reporting failures (RF) and Priority (Pty) have the expected signs, although RF is not statistically significant.

Using this equation I obtained the fitted values (expected probabilities of inspection) to use in the pollution equation. The result for this estimation is presented below (coefficients for plants fixed effects are not presented).

Table
Pollution Equation

Dependent Variable: LOG(BOD5)				
Method: GLS (Cross Section Weights)				
Total panel (unbalanced) observations: 3536				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(PQ)	-0.375	0.111	-3.376	0.0007
LOG(LABOR)	0.480	0.033	14.67	0.0000
LOG(WATER)	0.171	0.023	7.551	0.0000
LOG(ENERGY)	0.308	0.024	12.61	0.0000
LOG(FLOW)	-0.276	0.019	-14.178	0.0000
INSPF	-1.197	0.346	-3.460	0.0005
INSPCUM	0.050	0.010	5.000	0.0000
ORDERCUM	-0.060	0.021	-2.807	0.0050
FTCUM	-0.061	0.027	-2.240	0.0251
FINECUM	0.001	0.000	1.721	0.0853
Weighted R^2	0.966			
Unweighted R^2	0.730			

All the production function variables (LABOR, WATER, ENERGY) are statistically significant and enter the equation with the correct sign.[20] The sign of the coefficient of FLOW and the fact that it is also statistically significant may be telling that plants are diluting its effluents as a compliance strategy. The expected probability of inspection has an important deterrent effect. With respect to INSPCUM, this variable seems to be telling that plants tend to increase its pollution levels once they have been inspected, apparently aware of the fact that regulators employ a “sample without replacement” strategy. In other words, if plants have already being inspected they know that they are not going to be inspected again in the short run so they increase the level of emissions. Orders and fine threats has the expected coefficient signs, although FTCUM is not statistically significant. Finally, fines are both not very important as deterrent mechanisms and statistically insignificant. Of course, this may be the consequence of the fact that only fifteen fines were levied during the whole period.

6. Preliminary Conclusions

I have just presented preliminary results on the effect of different monitoring and enforcement actions on the level of BOD_5 concentration of industrial effluents in Montevideo, Uruguay. I did not emphasise the level of the coefficients estimated, just the sign, precisely because the preliminary nature of these results. Taking this into considerations, some general conclusions can be driven. First, results of the inspection equation seem to be telling that regulators did target inspections and increased the inspection rate in periods of industrial contraction, probably in substitution of harder enforcement mechanisms such as orders and fines, which are more costly to industries. In other words, in difficult economics times regulators opted to closely monitor plants as a softer mechanism of enforcing emissions standards through personal negotiation with plant managers. Effectively, only fifteen fines were levied in the period to this seventy-

[20] This significance could be distorted by the fact that 5% of the data set were missing item non-responses that were imputed using these variables.

four plants (eleven by the municipal governments and four by the national government). Second, with respect to the effect on pollution levels, the clearer results seem to be that the probability of being inspected has an important deterrent effect, and that plants seem to be aware of the "sample without replacement" strategy increasing their levels of pollution with the number of inspections received in the last twelve months. Also, compliance orders has its own deterrent effect, and so does fine threats, although this last one is not statistically significant. Fines are insignificant both statistically and in terms of the deterrent effect on BOD_5.

7. Next steps

Probably the most important step to take is to try to correct for sample selection. The natural thing to do would be to test first given that correcting for sample selection is not simple in the cases of panel data. Nevertheless, I cannot perform the test for ignorability proposed by Verbeek and Nijman (1992a) because I do not have enough data to estimate a balanced panel with these 74 plants in my original data set.

According to Verbeek and Nijman (1992b), a first way to obtain consistent estimators of the parameters in the cases of one way error correction models when the selection rule is non-ignorable is by a generalization for the case of panel data of the two-step Heckman procedures for selectivity bias in cross sectional data sets.

Consider the following one-way error component linear regression model,

$$y_{it} = \underline{x}_{it}^{'} \underline{b} + m_i + v_{it}; \qquad i = 1,..., N; t = 1,...T,$$

whith i denoting units (industrial plants in my case) and t denoting time. $\underline{x}_{it}^{'}$ is the $1*k$ vector of k explanatory variables and $\underline{b}$ is the corresponding $k*1$ parameter vector. m_i denotes the unobservable individual specific effect and v_{it} denotes a usual disturbance term. It is assumed that the errors terms m_i and v_{it} are independent of the explanatory variables. It is also assumed that m_i and v_{it} are mutually independent with $E(m_i) = E(v_{it}) = 0$, $E(m_i m_j) = d_{ij} s_m^2$ and $E(v_{it} v_{js}) = d_{it} d_{js} s_v^2$, where $d_{kl} = 1$ when $k=l$ and 0 otherwise and $d_{it} d_{js} = 1$ when $i=j$ and $t=s$, and zero otherwise. As before, we define

$r_{it}=1$ if the unit *i* is observed in period *t* and $r_{it}=0$ otherwise. Finally let $c_i=\prod_{t=1}^{T} r_{it}$, so that $c_i=1$ if and only if the unit is observed for all *t*.

I assume that the selection rule (the missing data mechanism) is given by

$$r_{it}^{*}=\underline{z}_{it}^{'}\underline{g}+x_i+h_{it}$$

where r_{it}^{*} is a latent variable such that when r_{it}^{*} is greater or equal to a threshold level $r_{it}=1$ (it is usually assumed that this threshold level is zero for simplicity). $\underline{z}_{it}^{'}$ is a vector of variables, usually containing a subset of the variables in $\underline{x}_{it}^{'}$. $\underline{g}$ is the corresponding vector of parameters, and x_i accounts for the unobserved individual specific effect in the selection process. Finally, h_{it} is the error term. For simplicity we assume normality of the error terms and independence of $\underline{z}_{it}^{'}$ and $\underline{x}_{it}^{'}$. More specifically,

$$\begin{pmatrix}\underline{v}_i\\ \underline{h}_i\\ m_i\\ x_i\end{pmatrix}\sim N\left(0,\begin{pmatrix}s_v^2 I_T & & & \\ s_{vh}^2 I_T & s_h^2 I_T & & \\ 0 & 0 & s_m^2 & \\ 0 & 0 & s_{mx}^2 & s_x^2\end{pmatrix}\right)$$

where $\underline{v}_i=\left(\underline{v}_{i1},\dots,\underline{v}_{iT}\right)'$ and $\underline{h}_i=\left(\underline{h}_{i1},\dots,\underline{h}_{iT}\right)'$.

Two correction terms are needed in the case of panel data and not just one (known as the standard Heckman correction term). This is because know there are two error components both in the equation of interest and in the selection mechanism equation. Nevertheless, the idea remains similar to the original cross section Heckman's case in the sense that these terms are the conditional expectations of m_i and v_{it} given $\underline{x}_{it}^{'}$ and the selection rule. From Verbeek and Nijman (1992b), these conditional expectations are $E\{m_i/\underline{r}_i\}=s_{mx}A_{1i}$ and $E\{v_{it}/\underline{r}_i\}=s_{vh}A_{2it}$, where

$$A_{1i}=\frac{1}{s_h^2+Ts_x^2}\sum_{s=1}^{T}E\{x_i+h_{is}/\underline{r}_i\}$$

and

$$A_{2i} = \frac{1}{s_h^2}\left[E\{x_i + h_{it} / \underline{r}_i\} - \frac{s_x^2}{s_h^2 + T s_x^2} \sum_{s=1}^{T} E\{x_i + h_{is} / \underline{r}_i\} \right]$$

Not surprisingly, the authors concluded that this solution is "still computationally unattractive" (Verbeek and Nijman, 1990, p. 692). Given this problem, Verbeek and Nijman conclude, "it may be worthwhile to have some simple variables that can be used instead to approximate the true correction term to check for the selectivity of non-response" (p. 692). Examples of such variables presented by these authors are: (1) the number of waves the plant participate, (2) a dummy variable equal to one if the plant is observed in all periods, and (3) a dummy variable indicating whether the plant is observed in the previous period. These type of variables would be more helpful in the case of the RE models, because in the case of the FE model, the selection rule bias is captured entirely by the individual effect term and it would not be possible to identify the parameters of the proposed correction terms from the individual effect parameter ("the fixed effect estimator is more robust for selectivity bias than the random effects estimator", p. 682). At this point of my research I have not decide yet on what correction terms to use.

Another important future steps are estimating the model with different dependent variables. One of these could be the extent of the violation (or a dummy variable equal to 1 if the plant is in violation of the concentration standard.) to test whether the Plan elaborated by the municipal government in1997 was successful in its objective of increasing compliance level. Another interesting dependent variable would be the total organic load (FLOW*BOD_5) of effluents.

Finally, I want to briefly discuss the issue of testing for under-reporting. One easy way to accomplish this is to conduct a difference of means test using the mean of BOD_5 reported and the mean of BOD_5 measured by the IMM and DCA in inspections, or BOD_5 measured by SEINCO, or the three of them assuming they are technically equivalent measures. Another more ambitious objective is to estimate the effectiveness of monitoring and enforcement actions on the extent of under-reporting. Here I have the problem of sample size since I have regular four-month information for 38 plants during periods 6, 7, 8, 9, 11 and 12. In period 10 SEINCO sampled only 17 plants, apparently because most plants began to be sampled every six months, and period 10 fell between two sample campaigns. Consequently I am left with only 228 observations in order to do this estimation, which may not be a sufficient number.

APPENDIX

Table A.1.: Distribution of Missing Values per Industrial Plant

Inputs and Pollution Variables

PLANT	BOD5	ELECTR.	EMPLOY	FLOW	FUEL	OSE	PERFOR	VOL.	W.D.	% over Potential
1	22	0	0	24	4	0	44	4	0	**20,9**
2	0	3	0	28	7	6	31	12	0	**18,6**
3	5	6	0	2	1	8	32	8	0	**13,2**
4	12	1	0	17	1	0	32	27	13	**22,0**
5	4	0	0	0	8	1	0	0	0	**2,8**
6	0	2	0	0	0	2	36	0	0	**8,5**
7	17	4	5	24	1	3	30	21	12	**25,0**
8	10	4	0	6	7	9	12	17	9	**15,8**
9	2	6	1	2	7	11	10	1	1	**8,8**
10	11	5	0	6	20	2	2	2	2	**10,7**
11	6	1	1	38	32	3	16	38	1	**29,1**
12	4	0	0	12	0	7	33	18	16	**19,2**
13	18	0	0	23	32	4	23	0	0	**21,4**
14	12	9	0	27	15	7	16	0	0	**18,4**
15	1	5	0	1	0	36	0	1	0	**9,4**
16	0	1	0	3	0	1	14	16	12	**10,0**
17	8	10	17	52	6	7	44	52	32	**48,7**
18	10	10	16	52	5	6	44	52	32	**48,5**
19	5	0	0	9	0	0	0	16	16	**9,8**
20	19	8	8	9	12	12	22	11	0	**21,6**
21	6	2	0	6	0	23	0	0	0	**7,9**
22	14	1	0	23	5	11	24	5	0	**17,7**
23	1	4	4	0	1	7	6	0	0	**4,9**
24	9	0	4	2	14	0	0	4	0	**7,1**
25	20	1	0	18	0	8	36	0	0	**17,7**
26	30	1	0	29	0	1	4	4	0	**14,7**
27	1	2	4	0	3	0	0	0	0	**2,1**
28	2	0	0	1	0	1	0	0	0	**0,9**
29	39	0	0	0	0	0	0	0	0	**8,3**
30	7	2	1	8	1	4	29	8	8	**14,5**
31	7	0	0	3	4	0	0	3	0	**3,6**
32	3	11	5	9	10	10	9	2	0	**12,6**
33	0	6	4	0	4	0	16	12	0	**9,0**
34	0	4	0	4	3	4	4	12	16	**10,0**
35	8	2	0	2	52	14	31	1	0	**23,5**
36	42	4	4	11	36	10	21	8	0	**29,1**

37	36	4	1	8	9	6	25	7	0	**20,5**
38	14	0	0	0	0	4	0	0	0	**3,8**
39	11	20	4	7	8	4	4	6	4	**14,5**
40	37	3	0	8	0	4	1	8	0	**13,0**
41	32	0	0	3	0	1	1	1	0	**8,1**
42	11	2	0	2	48	5	1	8	8	**18,2**
43	11	2	0	48	1	4	1	52	0	**25,4**
44	1	0	0	1	0	0	0	0	0	**0,4**
45	3	0	0	2	0	0	0	18	20	**9,2**
46	12	0	4	4	8	0	0	0	1	**6,2**
47	0	1	0	51	0	2	1	52	44	**32,3**
48	23	2	0	23	4	7	45	22	0	**26,9**
49	3	2	0	0	0	0	32	8	8	**11,3**
50	18	5	0	7	5	2	33	7	0	**16,5**
51	14	0	0	8	1	1	0	1	0	**5,3**
52	0	3	0	0	3	4	7	0	0	**3,6**
53	19	6	3	31	4	7	10	10	1	**19,4**
54	2	0	0	3	1	0	0	40	0	**9,8**
55	3	1	0	8	0	4	40	6	0	**13,2**
56	3	7	0	1	48	2	0	0	0	**13,0**
57	3	4	4	40	44	4	44	44	44	**49,4**
58	1	0	0	8	0	0	10	4	0	**4,9**
59	2	1	0	0	0	0	36	1	0	**8,5**
60	12	1	0	0	0	3	2	0	0	**3,8**
61	0	1	0	0	1	5	0	0	0	**1,5**
62	3	3	0	0	5	8	17	0	0	**7,7**
63	3	0	0	0	0	1	0	20	12	**7,7**
64	3	0	4	4	4	9	0	0	0	**5,1**
65	1	4	0	8	0	15	4	19	0	**10,9**
66	24	3	0	19	8	4	0	17	12	**18,6**
67	1	9	4	6	0	9	2	15	8	**11,5**
68	0	0	0	2	42	0	0	12	12	**14,5**
69	2	0	0	7	0	2	3	0	0	**3,0**
70	3	0	0	0	0	0	48	0	0	**10,9**
71	10	0	0	4	0	0	0	0	0	**3,0**
72	0	2	0	24	52	2	32	25	0	**29,3**
73	7	0	0	25	0	48	0	22	0	**21,8**
74	25	4	1	0	52	10	6	0	0	**20,9**
% over Potential	**18,4**	**5,3**	**2,6**	**21,1**	**16,6**	**10,3**	**26,7**	**20,3**	**8,9**	

Table A.2.: Names, Definitions and Units of Measure of Production Variables

	Name	Description	Units of Measure		Name	Description	Units of Measure
1	ALCOHOLBEB	Alcohol Beverage	(Ton.)	34	HYPOCLORITE	Sodium hypo chlorite	(m3)
2	ALUMINUM		(Ton.)	35	ICECREAM		(Kg/month)
3	BEER		(hl)	36	JELLY		(Kgs.)
4	BEERROOT	Root beer	(hl)	37	JUICE	Fruit and tomato	l/month
5	BICYCLE	Bicycles	Q x month	38	MILKCREAM	Milk and cream	l/month
6	BOVINES		Q x month	39	OFFALBOV	Bovine offal	(Kgs.)
7	BUTTER		(Kgs./month)	40	OFFALOV	Ovine offal	(Kgs.)
8	BUTTEROIL		(Kgs./month)	41	OFFALSETC	Offal, fat, eggs	(Kgs.)
9	CARAMEL		(Kgs./month)	42	OILFISH	Fish Oil	(Ton.)
10	CARDBOARD		(Ton.)	43	OILRAW	Raw Oil	(m3)
11	CASEIN		(Kgs./month)	44	OILREF	Refined Oil	(m3)
12	CHICKEN	Chickens processed	(Ton.)	45	OVINES		Q/month
13	CIDERBOT		(Bot.)	46	PAINTS		(m3/month)
14	CRAB		(Ton.)	47	PAINTS2		(m3/month)
15	CS		(Kgs./month)	48	PAPER		(Ton.)
16	DETERDESOD	Detergents/deodorants	(Ton.)	49	PELLETS		(Ton.)
17	FABRICASH	Fabrics		50	PNF		(Kgs./month)
18	FABRICM	Fabrics	(m)	51	PORKS		Q/month
19	FABRICSYNTHET	Synthetic fabrics	(Kgs.)	52	POTATO	Potato chips	(Ton.)
20	FAT		(Kgs.)	53	RENAULT	Units processed	Units
21	FILLET	Fillet fish	(Ton.)	54	SALTS		(Ton.)
22	FISHWHOLE	Whole fish processed	(Ton.)	55	SAUSAGES		(Kgs.)
23	FLOURFISH	Flour produced	(Ton.)	56	SNACKS		(Ton.)
24	HAM	HG Fish processed	(Ton.)	57	SOAPS		(Ton.)
25	HG		(Kgs.)	58	SODA		(Ton.)
26	HIDECASTR	Hides castrated ram	(Q x month)	59	SODABOT		(Bot.)
27	HIDEDYED	Dyed cow hides	Q x month	60	SQUID		(Ton.)
28	HIDEFINISH	Finished cow hides	(m2)	61	SUPERGAS		(Ton.)
29	HIDELAMB	Lamb hides	Units	62	TOPS		(Ton.)
30	HIDESEMIFINISH	Semi finished cow hides	Units	63	WATERSODA		(hl)
31	HIDESHEEP	Sheep hides	Units	64	WETBLUE		Units
32	HIDETANNED	Tanned cow hides	Units	65	WOOLCLEAN	Cleaned wool	(Ton.)
33	HIDEUNFURRED	Unfured cow hides	Units	66	WOOLDIRTY	Dirty wool	(Kgs.)
				67	YOGURT		l/month

Table A.3.: Percentage of Missing Values of Production Variables

Number	Product Variable		Number	Product Variable	
1	**ALCOHOLBEB**	0,0	34	**HYPOCLORITE**	7,7
2	**ALUMINUM**	0,0	35	**ICECREAM**	11,5
3	**BEER**	1,9	36	**JELLY**	0,0
4	**BEERROOT**	8,7	37	**JUICE**	9,6
5	**BICYCLE**	1,9	38	**MILKCREAM**	9,6
6	**BOVINES**	0,5	39	**OFFALBOV**	0,0
7	**BUTTER**	11,5	40	**OFFALOV**	15,4
8	**BUTTEROIL**	25,0	41	**OFFALSETC**	7,7
9	**CARAMEL**	11,5	42	**OILFISH**	7,7
10	**CARDBOARD**	0,0	43	**OILRAW**	0,0
11	**CASEIN**	84,6	44	**OILREF**	1,9
12	**CHICKEN**	0,0	45	**OVINES**	10,1
13	**CIDERBOT**	0,0	46	**PAINTS**	9,6
14	**CRAB**	7,7	47	**PAINTS2**	0,0
15	**CS**	0,0	48	**PAPER**	0,0
16	**DETERDESOD**	0,0	49	**PELLETS**	0,0
17	**FABRICASH**	1,0	50	**PNF**	3,8
18	**FABRICM**	3,8	51	**PORKS**	5,8
19	**FABRICSYNTHET**	0,0	52	**POTATO**	0,0
20	**FAT**	3,8	53	**RENAULT**	14,4
21	**FILLET**	14,4	54	**SALTS**	0,0
22	**FISHWHOLE**	10,3	55	**SAUSAGES**	15,4
23	**FLOURFISH**	7,7	56	**SNACKS**	0,0
24	**HAM**	0,0	57	**SOAPS**	7,7
25	**HG**	27,9	58	**SODA**	3,8
26	**HIDECASTR**	1,9	59	**SODABOT**	1,9
27	**HIDEDYED**	44,7	60	**SQUID**	7,7
28	**HIDEFINISH**	23,1	61	**SUPERGAS**	2,9
29	**HIDELAMB**	0,0	62	**TOPS**	0,0
30	**HIDESEMIFINISH**	46,2	63	**WATERSODA**	17,3
31	**HIDESHEEP**	14,1	64	**WETBLUE**	0,0
32	**HIDETANNED**	2,9	65	**WOOLCLEAN**	1,9
33	**HIDEUNFURRED**	19,6	66	**WOOLDIRTY**	0,0
			67	**YOGURT**	11,5
				TOTAL	8,7

BIBLIOGRAPHY

Amarante, V. and M. Caffera, Los Factores Determinantes de la formación de asentamientos irregulares. Un análisis económico. Documento de trabajo Nº 01/01, Facultad de Ciencias Empresariales y Economía, Universidad de Montevideo.

Beale E. M. L. and R. J. A. Little, Missing Values in Multivariate Analysis, *Journal of the Royal Statistical Society,* Ser. B, **37**, 129-145 (1975)

Bohm, P. and C. S. Russell, Comparative analysis of alternative policy instruments, *in* "Handbook of Natural Resource and Energy Economics", Vol. I, A.V. Kneese and J.L. Sweeney. (Eds.), Elsevier, (1985),

Buck, S. F., A Method of Estimation of Missing Values in Multivariate Data suitable for use with an Electronic Computer, *Journal of the Royal Statistical Society,* Ser. B, **22**, 302-306 (1960)

Caffera, M., "Análisis Económico de la Política de Control de la Contaminación Hídrica de Origen Industrial en el Departamento de Montevideo", Programa de Apoyo a la Iniciación de la Investigación, Comisión Sectorial de Investigación Científica, Universidad de la República, Junio (2002).

CEPAL (Comisión Económica para América Latina y el Caribe), "Instrumentos económicos para el control de la contaminación del agua: condiciones y casos de aplicación" (Versión preliminar), LC/IN. 137, 28 de diciembre, 2000

CEPAL,"Aplicacion de Instrumentos Economicos en la Gestion Ambiental en America latina y el Caribe: desafios y factores condicionantes", Serie Medio Ambiente y Desarrollo **31**, , Santiago de Chile, Enero (2001).

Cohen, M., "Monitoring and Enforcement of Environmental Policy", Working Paper, Owen School of Management, Vanderbilt University, (1998).

Cousillas, M. y M. Castaño, "Fundamentos de Derecho Ambiental Uruguayo" CEJU – IFCA. Programa de Fortalecimiento del Rol de la Justicia en la Protección del Medio Ambiente y la Biodiversidad. Montevideo (1996).

Dagenais, M. G., The use of incomplete observations in multiple regression analysis, *Journal of Econometrics,* 1 317-328 (1973).

Dasgupta, P. , P. Hammond and E. Maskin, On Imperfect Information and Optimal Pollution Control, *Review of Economic Studies,* **XLVII** , 857-860 (1980).

Dasgupta, S., H. Hettige and David Wheeler, What Improves Environmental Compliance? Evidence from Mexican Industry, *Journal of Environmental*

Economics and Management, **39**, 39 – 66 (2000).

Dasgupta, S., B. Laplante, N. Mamingi and H. Wang, Inspections, Pollution Prices and environmental performance: evidence from China, *Ecological Economics*, 36(3), 487-498 (2001).

Deily, M. E. and W. B. Gray, Enforcing of Pollution Regulations in a Declining Industry, *Journal of Environmental Economics and Management* **21** 260-274 (1991)

Dion, C., P. Lanoie and B. Laplante, Monitoring of Pollution regulation: Do Local Conditions Matter?, *Journal of Regulatory Economics* **13** , 5-18 (1998)

Eskeland, G. S., A Presumptive Pigouvian Tax: Complementing Regulation to Mimic an Emission Fee, *The World Bank Economic Review*, **8** (3), 373 – 394, (1994).

Eskeland, G. S. and E. Jimenez. Policy Instruments for Pollution Control in Developing Countries, *The World Bank Research Observer*, **7** (2), p. 145-169, (1992).

Eskeland, G. S. and S. Devarajan, Taxing bads by taxing goods: toward efficient pollution control with presumptive charges, *in*, "Public Economics and the Environment in an Imperfect World" (Bovenberg, L. and S. Cnossen, Eds.), Kluwer (1995)

Garvie, D. y A. Keeler, Incomplete enforcement with endogenous regulatory choice, *Journal of Public Economics*, **55**, 141-162 (1994).

Gray, W. B. and M. E. Deily, Compliance and Enforcement: Air Pollution Regulation in the U.S. Steel Industry, *Journal of Environmental Economics and Management* **31**, 96 - 111 (1995).

Gray, W. B. and Shadbegian, When and Why do Plants Comply? Paper Mills in the 1980s Draft, October (2002).

Gupta, S. and S. Saksena, Enforcement of Pollution Control Laws and Firm Level Compliance: a study of Punjab, India, Draft presented at the 2nd World Congress of Environmental Economists, July (2002). (DO NOT CITE WITHOUT PERMISSION)

Guzmán, M., Un Instrumento Económico para el Control de la Contaminación Hídrica en Costa Rica: El Canon por Vertidos, *Instrumentos Económicos y Medio* Ambiente, **2** (3), Octubre, Centro Andino para la Economía en el Medio Ambiente (2001).

Harrington, W., Enforcement Leverage when Penalties are Restricted, *Journal of Public Economics*, **37**, 29-53 (1988).

Heckman, Sample Selection Bias as a Specification Error, *Econometrica*, **47** (1), 153-161 (1979).

Helland, E., The Enforcement of Pollution Control Laws: Inspections, Violations and Self-Reporting, *The Review of Economics and Statistics*, 80 (1), 141-153 (1998).

Heyes, A.G., Implementing Environmental regulation: Enforcement and Compliance, *Journal of regulatory Economics*, 17(2), 107-129, (2000).

I.M.M., "Primer Informe Ambiental, Montevideo XXI" (http://www.imm.gub.uy/ambiente/documentos/infoamb.htm) (2001).

I.M.M., "Documento Base. Elaboración Agenda 2002", Taller de Recursos Hídricos, Grupo Ambiental de Montevideo, (2002).

Laplante, B. and P. Rilstone, Environmental Inspections and Emissions of the Pulp and paper Industry in Quebec, *Journal of Environmental Economics and Management* **31**, 19 – 36 (1996).

Little R. J. A., Regression with Missing X´s: A Review, *Journal of the American Statistical Association,* **87** (420), 1227-1237 (1992)

Little, R. J. A. and D. B. Rubin, "Statistical Analysis with Missing Data", Wiley, New York (1987) .

Magat, W. A. and W. K. Viscusi, Effectiveness of the EPA's Regulatory Enforcement: The case of Industrial Effluent Standards, *Journal of Law and Economics* **33**, 331 - 360, (1990).

Mc Closkey, D. N. and S. T. Ziliak, The Standard Error of Regressions", *Journal of Economic Literature*, **34**, 97-114 (1996).

Multiservice – Seinco – Tahal, Presentación de Resultados del "Programa de Monitoreos de Industrias y Cuerpos de Agua", Octubre (2001).

Nadeau, L. W., EPA Effectiveness at Reducing the Duration of Plant-Level Noncompliance, *Journal of Environmental Economics and Management* **34**, 54 - 78 (1997).

O'Connor, D., Applying economic instruments in developing countries: from theory to implementation, *Environment and Development Economics*, **4**, 91 – 100 (1998).

Pargal, S. and D. Wheeler, Informal regulation of industrial pollution in developing countries: evidence from Indonesia, *Journal of Political* Economy, **6** (104), 1314 -

1327 (1996).

Pargal, S., M. Mani and M. Huq, Regulatory Inspections, Informal Pressure and Water Pollution. A Survey of Industrial Plants in India, The World Bank Policy Research Department Working Paper, November 4 (1997).

Polinsky, A. M. and S. Shavell, The Economic Theory of Public Enforcement of Law, *Journal of Economic Literature,* **38**, 45-76 (2000).

Rubin, D.B., "Multiple imputation for Non-response in Surveys", Wiley, New York (1987)

Russell, C. S. and P. T. Powell, Choosing Environmental Policy Tools, Theoretical Cautions and Practical Considerations, IADB, Washington D.C., June 1996 - No. ENV-102

Russell, C. S., "Applying Economics to the Environment", Oxford University Press, (2001).

Seroa da Motta, R., R. M. Huber and H. J. Ruitenbeek, Market based instruments for environmental policy making in Latin America and the Caribbean: lessons from eleven countries", *Environment and Development Economics,* 4; 177-201 (1999).

Shimshack, J. P. and M. B. Ward, The Impact of Fines, Enforcement Actions and Inspections on Environmental Compliance. A Statistical Analysis of the Pulp and Paper Industry", *mimeo* (2002).

SOGREAH-SEURECA-GKW-CSI, *Estudio de Impacto Ambiental PSU III,* Informe de Solicitud de Préstamo, Marzo (1995).

Tietenberg, T., Private Enforcement of Environmental Regulations in Latin America and the Caribbean. An Effective Instrument for Environmental Management? No. ENV – 101, IADB, Washington, D.C., June (1996).

Verbeek, M. and T. Nijman, Testing for selectivity bias in panel data models, *International Economic Review,* **33** (3), 681-703 (1992a).

Verbeek, M. and T. Nijman, Incomplete panels and selection bias, *in* "The Econometrics of Panel Data" (L. Mátyás and P. Sevestre, Eds.), Kluwer Academic Publishers (1992b).

Wang, H., N. Mamingi, B. Laplante and S. Dasgupta, Incomplete Enforcement of Pollution Regulation: Bargaining Power of Chinese Factories, Development Research Group, The World Bank, Washington D.C., April (2002)

World Bank, "Greening Industry: New Roles for Communities, Markets and Governments", Washington D. C. (1999).