Politics and Regulation: The Case of Public Transit*

Isabelle Perrigne

Pennsylvania State University

Sunita Surana

University of Southern California

Preliminary Draft

Do Not Cite

Do Not Circulate

November 2004

* We thank Jean-Jacques Laffont for thoughtful discussions and Philippe Gagnepain for providing the data analyzed in this paper. The first author is grateful to Hewlett Packard Laboratories for their financial support through a grant from the Solution and Service Technologies Center.

Correspondence to: I. Perrigne, Department of Economics, Pennsylvania State University, University Park, PA 16802. e-mail address : perrigne@psu.edu.

Abstract

Political economy has long acknowledged the role played by politics in policy decisions. The purpose of this paper is to analyze the impact of politics on regulatory contracts in public transportation. Public transportation is regulated at the municipality level and hence could be influenced by the political composition of the City Council. Data from bus transit contracts in 57 French cities from 1985 to 1993 show that the price charged to consumers is significantly different between left-wing and right-wing cities, the former being lower. The election held in 1989 has shown the emergence of a new party acting as an interest group with environmental interests. Its influence has been significant irrespective of a leftist or rightist majority.

While adopting an incentive regulation framework, we develop a model to take into account such political factors by considering a nonbenevolent regulator. Relying on a structural model, we estimate the parameters of the demand and cost functions as well as the shadow cost of public funds, the density of firms' types and the weight associated to the consumer surplus for every city as a function of political factors and cities financial situation. The empirical results show that cities with a leftist majority and with environmentalists in their council tend to put a larger weight on the consumer surplus leading them to lower the price for public transportation. As expected, the level of debt per capita restrains the cities in their redistributive policies.

Key words: Regulation, Political Economy, Contracts, Public Transportation.

JEL classification: L51, L92, C50, R48, D 79

Politics and Regulation: The Case of Public Transit

I. Perrigne and S. Surana

1 Introduction

to be completed

2 Politics and Regulation: Some Evidence from Public Transit

This section presents data from two different sources. A first data set contains detailed information on regulatory contracts for 57 French cities from 1985 to 1993. To assess the role played by politics on these regulatory contracts, data on city council elections and composition of city councils have been collected.

2.1 Political Data

The 1985-1993 period has experienced one election at the city level, which took place in March 1989. Since the data covers years before 1989, we also collected data on the previous election, which took place in March 1983.

Every city council in France is renewed every six years through a democratic election. The size of the City council is proportional to the population in every city. In our data, the number of its members varies from 37 to 101 according to the population size. Several parties enter into the electoral competition while proposing a list of candidates. The election takes place into one or two rounds. The final allocation of seats in the city council among the various lists in competitionm responds to a mix of majority and proportional rules. If a list of candidates obtains the absolute majority at the first round, there is no second round for the election. The winning list will then obtain 50% of the seats, while the remaining 50% of the seats will be allocated through a proportional rule among all the lists in competition. This allocation rule gives a confortable majority to the winner of the election.[1] The new city council elects the new mayor of the city, who is in general at the very top of the winning list. Given the strong majority of the winning list, there is no conflict for the election of the new mayor.

For the election held in 1983, 37 cities out of the 57 cities in our sample have experienced a single round, while there were 27 in the 1989 election. When any of the lists obtains the absolute majority at the first round, a second round is organized a week later. The second round is organized as follows. Only the lists, which have obtained a minimum of 10% of votes at the first round, will enter into the competition for the second round. It is, however, possible for the lists which have obtained between 5% and 10% of votes at the first round to merge with a list, which has obtained more than 10% at the first round. This rule favors various political alliances between the two rounds of election giving hope to parties with a low popularity to have at least a representative in the city council. The final allocation of seats follows the rule that we have described.[2]

[1] Let consider a simple example with three lists in competition. The winning list (party A) obtains at the first round 60% of the expressed votes, while party B and party C both obtain 20%. According to the allocation rule adopted in France for 30 seats to be allocated, party A will have 50% of the 30 seats plus an additional 60% of the remaining 15 seats, which makes a total of 24 ($= 15 + 9$). Both parties B and C will obtain each 20% of the remaining 15 seats, which makes 3 seats for party B and 3 seats for party C. Party A will govern the city with 80% of the seats of the City Council. Despite having obtained each 20% of the votes, party B and party C will obtain only 10% of the seats in the city council.

[2] Let consider a simple example. In the first round of the election, several lists are in competition. Only 4 have obtained at least 10% of the votes. The results of the second round give 40% to party A, while parties B, C and D obtain 20% each. For 30 seats to be allocated, party A will obtain 21 seats ($= 15 + 6$) or 70% of the council while other parties will obtain 3 seats each or 10% of the council.

When combining both elections, we find more than 70 different list names, which is quite large, many of them participating to a small number of cities. A large majority of lists are, however, affiliated to a political party. We can roughly consider seven political tendances in France from the left to the right wing. Parties associated with a communist ideology are considered as extremist left, while parties based on socialism are considered as leftist parties. Between the left and the right wing, there are several parties in the center. Every of these political parties call themselves as left centre or right centre. Since there is not always major differences between center left and left and between center right and right and the important likelihood that these parties will sign alliances with left and right parties, we have labeled these parties as either left or right. They usually obtain a low proportion of votes. Parties with a republican or democrat ideology are considered as the right wing, while parties based on a nationalism ideology are labelled as extremist right.[3] Lastly, the past 25 years have experienced the emergence of new parties in France with environmentalist interests. They tend to act as special interest groups lobbying against polluting industries and car pollution. We consider them as a special interest group as their arguments refer less to ideology than other parties. For instance, they make a great use of scientific studies on pollution and planet warming and their consequences on health. They favor free public transportation to induce more car drivers to switch to a more environmentally friendly mode of transportation.[4]

In some cities, some of their representatives have signed agreements with the left wing after the first round. After their relative success at the first round of the 1989 election, they have maintained their lists at the second round.[5] The following table provides the results at both elections for extreme leftist, leftist, rightist, extreme rightist and environmentalist parties. These figures provide the average percent of votes at the final round of the election

[3]The diversity of political parties in France is larger than in the US. It may look surprising to some readers that both republican and democrat are considered as the right wing. Various right wing parties in France use either the term republican or democrat in their affiliation.

[4]In the same spirit, they favor the development of bike paths and pedestrian areas in cities.

[5]Data provide detailed information on the political affiliation of elected council members.

as well as the percent of seats obtained in the city council.[6]

Table 1: 1983 and 1989 Election Results

	1983		1989	
Party	votes	city council	votes	city council
Extreme Left	14.33	10.21	10.41	8.71
Left	30.32	36.37	39.48	35.86
Right	53.91	52.50	45.11	45.13
Extreme Right	0.81	0.11	2.63	5.15
Environmentalist	0.36	0.60	2.23	4.51

Between 1983 and 1989, we observe some changes for the leftist and the rightist parties. The percent of seats obtained on average by the left wing was stable though the 1989 election was more favorable in terms of vote proportion. The right wing experienced a significant decrease in popularity both in terms of average vote proportion and representativity in city councils. In 1983, the right wing won 32 cities, while the left wing and the extremist left won 20 and 5 cities, respectively. In 1989, the right wing won 26 cities, while the left wing and the extremist left won 28 and 3 cities, respectively. Fourteen cities have experienced a major switch in their majority, 4 cities switched from a leftist to a rightist majority, while 10 cities have switched from a rightist to a leftist majority. The few cities with a extreme leftist majority ar emainly located in North of France, where traditional industries such as steel and mining have declined in the seventies. Consequently, these cities are characterized by an important unemployment rate and exerience a difficult economic transition.

Table 1 shows that the representativity of extreme leftist and extreme rightist parties are not negligible, the former has known a decline in popularity while the latter has experienced a increase in popularity between 1983 and 1989. In 1983, 13 cities in our sample had some council members affiliated with extreme rightist party and 56 cities

[6]The sum of the figures in Table 1 does not sum up to 100% because of some independent lists.

had some extrme leftist members. In 1989, there were 29 cities with extreme rightist council members and 45 cities with extreme left council members. The most striking change between these two elections is the emergence of the environmentalists.[7] In 1983, only 8 cities in our sample had at least one environmentalist in their city council, while there were 27 after the 1989 election. We note that the environmentalists have usually maintained their position while gaining seats in other cities.[8] We also note that the environementalists are more likely to obtain seats in cities with a leftist majority than in cities with a rightist majority. In 1989, only 7 cities out of the 26 rightist cities had at least one environmentalist in their city councils while there were 19 out of the 28 leftist cities. In 1983, the results were quite different with 6 cities out of the 32 rightist cities and 2 cities out of the 20 leftist cities. Their representativity is quite small as they obtain between 1 to 5 seats but their influence has been important in policy decisions, especially in terms of public transportation as shown in the next subsection.

To conclude this presentation, we discuss the two actors in these elections: the politicians and the voters, who have different objectives in these elections. First, we discuss the characteristics of the electors for the various political tendances. Given the well known split between the left and right wings, the left being in favor of redistributive policies while the right is in favor of more liberal policies, we have computed the corelation coefficient between the percent of votes and the unemployment rate for every city. The

[7] We consider it as the major change for the following reasons. It was the first election in which the environmentalists became important actors in French politics and this trend has been confirmed by the following elections. Moreover, it is not unusual in France that election results fluctuate so much for parties within a short period of time. Such a fluctuation happened a few years ago between the presidential election and the congress election within a short period of time. The extreme rightist candidate obtained about 20% of the votes at the first round of the presidential election, while very few representatives of his party were elected at the congress by an obvious lack of popular support.

[8] An interesting question is to understand whether this vote has been motivated by high levels of pollution. A first look at cities with environmentalist in their city councils suggests that there is no particular pattern between pollution or the presence of polluting indsutries and the environmentalist vote. We have not, however, data on pollution to assess empirically such a statement.

unemployment rate in our sample varies from 7% to 23% with an average value at 15%. The results are given in the table below when pooling the data from both elections.

Table 2: Unemployment Rate and Election Results

Party	Correlation Coefficient
Extreme Left	0.1926
Left	-0.1051
Right	-0.1088
Extreme Right	0.0950
Green	-0.1383

Not surprisingly, there is a strong positive correlation between unemployment rate and the extreme leftist vote. This tendance is well implemented in north and east of France, an area which has experienced a decline of traditional industries. In these industries, an important proportion of the labor force is unionized and unions in France are very close to extreme leftist parties. Regarding the exteme rightist vote, it seems that there is a correlation with immigration such as in South of France though the extreme rightist wing is well implemented in areas such as Brittany where the immigration has always been low. The vote for extremist parties are usually based on strong ideological factors. The correlation is negative for the other parties. Regarding other population characteristics, we do not have more data to draw some conclusions and have to rely on other studies. It seems that young generations are more liklely to vote for the left wing, while older generations are more conservative with a rightist vote. Urban population is more likely to vote for the leftist wing, while rural population is more likely to vote for the rightist wing. Individuals' wealth and social status is usually correlated with political choice such as low revenue and working class voters value the leftist wing, while high revenue and professional class voters value the rightist wing.[9]

[9] It is a very rough representation of the electorate in France and of their political preferences. We can find many examples contradicting this tendancy. In general, leftist parties favor redistribution favoring low income households, while rightist parties favor tax cuts favoring high income households.

For each city at both elections, data provide the total number of registered voters and the number of actual voters. We can use this information to compute the participation rate at these elections. In 1983, the participation rate was on average 72.55%, with values ranging from 60% to 81%, while in 1989 it decreased to an average of 64.83% with values ranging from 53% to 78%. The following table provides the correlation coefficients between the participation rate and the percent of votes obtained. On the one hand, we observe a strong correlation for extreme leftist parties indicating a strong mobilization of voters. On the other hand, the correlation is significantly negative for extreme rightist parties and environmentalists. This result is somewhat difficult to interpret suggesting swing or opportunist voters who can change their votes at the next election.[10]

Table 3: Participation Rate and Election Results

Party	Correlation Coefficient
Extreme Left	0.19
Left	-0.04
Right	0.13
Extreme Right	-0.36
Green	-0.42

Lastly, it would be interesting to analyze the motivation of politicians. There are two assumptions in the political economy literature. Namely, politicians are either rent seekers or partisans. The former refers to politicians who are self interested. Politicians can be interested by some rents associated to the position or by just maintaining their positions as long as possible. We call these politicians rent seekers or office seekers. A good indicator of self interested politicians could the proportion of cities for which the same party with the same leader is reelected in 1989. Out of the 57 cities in our sample, 36 mayors elected in 1983 were reelected in 1989, which represents about 63%. For the

[10]It may also indicate disillusioned voters by the standard leftist and rightist parties who have preferred either to vote for something different such as environementalist or to abstain.

remaining 21 cities, we find that, for 12 of them, the mayor in 1983 run for the 1989 election at the head of a list and, for 5 of them, the mayor elected at the 1989 election, was participating at the 1983 election at the head of a list. Therefore, for 84% of the cities in our sample, we find that the same political leaders have participated at both elections. For the remaining 16%, it is, however, possible that the political leader at one election was participating at the other election taking the second or third position on the list. The data provide the identity of the list head only. Many of the mayors elected in large cities have later hold offices in national elections and in the government suggesting again their office-seeking motivation.[11] Nonetheless, the empirical analysis of the following section will show that their behavior responds to both ideological and office-seeking motivations.

2.2 Public Transportation Data

The data provide detailed information on the operating costs including labor, energy and material costs, the number of buses as well as the number of employees. Fixed costs such as capital are not included in the data as these costs are not born by the firm. In particular, buses are provided by the city. Detailed information on production can be found such as the network size, the number of seat kilometers, the average speed and the number of passengers. In addition, the revenues from the bus fares and subsidies are also provided. All the data have been deflated and expressed in constant 1985 FF. See Perrigne (2002) for a detailed description of the data. In this paper, we are interested on whether the political composition of the city council and the political affiliation of the mayor have influenced the regulatory contract. Since the contract is characterized by the pair price and transfer, we assess any relationship between the price and the political composition of the city council. Moreover, the politicial composition of the city council may have also influenced the quality of the service offered to passengers.

[11] The data provide some interesting examples of political dynasties such as in Toulouse, where two generations of the Baudis family have been mayors holding the city hall for more than 40 years. Another example includes Bordeaux with Chaban-Delmas, who has ruled the city hall for more than 30 years.

Since public transportation is a highly subsidized activity, the fare covers on average only 48% of operating costs, public transportation can be considered as a redistributive toll among the population. In this sense, we can expect that cities with a leftist majority will subsidize more public transportation than cities with a rightist majority by offering a lower price to commuters. In the same spirit, we can also expect that cities with a leftist majority will offer a better service with an extended coverage to render public transportation accessible to all. Given the strong arguments used by environmentalists during their campaign on free transportation and pollution, we can expect them to lobby on these issues by reducing the price of public transportation. Given that their representativity in city councils is small (see table 1), it is unclear whether they can really influence the cities policies in terms of public transportation. To analyze the right/left opposition, we have spilt the data set into two data sets according to the affiliation of the majority in the city council. To simplify the analysis and because of the low number of cities with an extreme leftist majority, we have considered these cities as leftist wing.[12] Cities with a leftist majority have an average fare equal to 2.25 French Francs (FF), while cities with a rightist majority have an average fare equal to 2.44FF, confirming our expectation. Figure 1 displays both price densities. Figure 1 shows a larger range of values for right wing cities with a larger mode than for left wing cities.[13]

Since some cities have experienced a political change between the two elections, we have drawn similar graphs while comparing cities with the same majority in 1983 and 1989 and cities experiencing a change in the majority between 1983 and 1989. Though we have a small number of observations for assessing these changes, we observe a clear change in pricing policies. For instance, we compare price densities for cities with a leftist majority at both elections with cities switching from a leftist to a rightist majority. See

[12]The results are not dramatically different.

[13]The data do not provide detailed information on the pricing policy. About 80% of cities in France offer special fares to some groups of the population such as for young, eldery, unemployed and low revenue people. Leftist and rightist cities may differ in their pricing policies through these discounts given to subgroups of the population for redistribution purposes.

Figure 2. The density for the former cities is unimodal with a large variance, while the density for the latter is bimodal with a small variance. The former is slightly at the left of the latter. The comparison of these densities suggests that only one city experiencing a switch has maintained its low prices while the other cities have (gradually) changed their prices for public transportation while increasing them. Figure 3 displays the price density for cities with a rightist majority at both elections and for cities with a rightist majority in 1983 switching to a leftist majority in 1989. The contrast is not as striking in this case but we can observe a smaller mode for the latter than for the former with a second mode at a larger value indicating than cities experiencing a change in majority have slightly and gradually lowered the price for public transportation. We have performed a similar analysis for various quality indicators such as the network coverage measured by the ratio network size by the area of the city and the number of seats offered divided by the size of the population. The observed differences are minimal with a slightly larger mode for the left wing cities for the network coverage. Similarly, the number of available seats per capita is slightly larger for cities with a leftist majority. These results suggest that the quality of service does not differ much according to the left/right majority of the city council.[14]

As mentioned earlier, the environmentlaists are strong advocate of public transportation in the hope that its development will reduce the level of car pollution in cities. It would be interesting to assess whether the presence of their representatives in city councils has an impact on the policy for public transportation.[15] For this purpose, we conduct a similar study relying on price densities as well as for other various quality indicators. When considering the full sample of 57 cities, we find that the average price of public transportation for cities with no environmentalist in their city council is equal to 2.41FF,

[14]The graphs are available upon request to the authors.

[15]Following the 1983 election, we have 8 cities with a green representative in the city council, while there are 27 cities with a green representative following the 1989 election. This provides enough observations to assess such a change. Since the election is held in Spring 1989, we consider that some changes, if any, occured only starting in 1990.

while this average price drops to 2.15FF when there is at least one environmentalist in the city council. This implies that on average passengers pay about 12% more for public transportation in cities with no environmentailst in the city council. Figure 4 displays both price densities.

To analyze this effect further, we consider separately cities with a leftist or rightist majority. For leftist cities, the average bus fare is equal to 2.32FF when there is no environmentalist council member, while it is equal to 2.12FF when there are some. We have observed some alliances between the environmentalists and the left wing. Not suprisingly, the leftist politicians have then to reconsider their environmental policies when governing the cities. This fact could partly explain this observed difference. For cities with a rightist majority, we observe a more striking difference with an average bus fare equal to 2.51FF when there is no environmentalist in the city council and equal to to 2.19FF when there is at least one environmentalist in the city council. This represents a difference of about 15%. Figure 5 and 6 displays the price densities for leftist and rightist cities, respectively. The range of values is clearly larger for the rightist cities with no environmentalist as displayed by Figrue 6. Figure 5 displays a larger mode for the leftist cities with no environmentalist with a larger range of values as well.

To our knowledge, the right wing has never had any particular agreement with the environmentalists between the two rounds of election. Therefore, the above argument of political alliance is not valid. We have then to find other explanations to rationalize such differences. The French election system offers a strong majority to the winning list to govern the city. In this respect, the winning list has free hands to govern according to their own aspirations. Empirical evidence would suggest that politicians' main motives are to hold office and to be reelected. Politicians have then to maximize the probability of being reelected in a dynamic perspective. When the electors express some concerns for environment through their votes to environmentalist, it is in the politician's best interest not to ignore this signal and to apply some environmental policies to capture these votes at the next election. In the political economy literature, these voters are called the swing voters as they are mobile across parties because of their low concern for ideology. If we view the environmentalists as a special interest group, their voters will vote for the party representing their interests. By applying a enviromental policy such as offering public transportation at a lower price, leftists or rightist politicians may hope to gain some of these voters at the next election.[16]

To conclude this section, we have performed a hedonic price analysis for bus fare, while relating bus fare or price to a set of various quality indicators and political factors given the previous evidence. We also introduce some indicators for the cities' financial constraints. As discussed previously, politicians are self-interested and would like to please their electorate in the objective of being reelected. They can do so by reducing the price for public transportation. When considering politicans motivated by ideology, they could reduce the price for public transportation while considering it as a redistibutive tool. Politicians are, however, refrained in their policies by financial constraints. Though all the cities in our sample have some debts, cities cannot run into systematic deficits every year. We can expect that cities with already important debts will be constrained in their redistributive objectives or policies in general. Moreover, a tax on firms has been gradually implemented in France starting in 1971 to partly subsidize public transportation. When the economic activity is low, the city will have to pay a larger proportion of subsidies to public transportation since the taxes levied on firms will be less. Therefore, the city's debt and economic activity could constrain the politicians in their public transportation pricing policies.

[16]When looking at quality indicators, we observe that the extent of the network is significantly larger for leftist cities with environementalists, namely 247 versus 172. For rightist cities, we observe the opposite with a larger extent of the network when there is at least one environmentalist, namely 166 versus 140. It is then unclear how the environmentalists have influenced the quality of public transit.

Table 4: Hedonic Price Model for Public Transportation

Variable		Fixed Effects	Random Effects
Constant	-1.9018	—	-1.9017
	(0.000)		*(0.028)*
Network Extent	0.1179	0.2691	0.1978
	(0.000)	*(0.000)*	*(0.000)*
Area	0.0929	—	0.784
	(0.000)		*(0.072)*
Seats per Capita	0.4109	0.2645	0.3479
	(0.000)	*(0.037)*	*(0.000)*
Debt per Capita	0.0970	—	0.0849
	(0.000)		*(0.157)*
Unemployment Rate	0.3111	—	0.3144
	(0.000)		*(0.034)*
Leftist Majority	-0.0767	-0.0392	-0.0396
	(0.001)	*(0.186)*	*(0.145)*
Environmentalist	-1.4880	-0.9448	-1.2362
	(0.002)	*(0.045)*	*(0.007)*
Time Trend	0.0254	0.0174	0.0277
	(0.000)	*(0.000)*	*(0.000)*
R^2	0.344	0.309	0.772

We regress the logarithm of prices on a constant, the logarithm of the city area, the logarithm of the network size divided by the area of the city, the logarithm of the number of seats available by the population size or number of seats per capita, the logarithm of the amount of city debt divided by the population size or debt per capita, the logarithm of the unemployment rate, a dummy for a leftist majority, the proportion of seats hold by environmentalists in the city council as well as a time trend.[17] All the results are

[17]Note that the data on the debt per capita are unavailalbe for the 1985-1993 period. We then collect

displayed in Table 4, while the p-values are given between parentheses.

All the coefficients in the model that does not take into account the panel structure of the data are significant and meet our expectations. The quality of the service offered to commuters as measured by the network extent and the number of seats per capita contribute to increase the price of public transportation. An increase in the quality would imply to operate more buses and therefore to hire more drivers and to consume more energy leading to an increase in operating expenses, which is reflected in the increasing price. We note, however, that an increase in 1% of either the extent of the network or the number of seats per capita increases the price by less than 1%, namely 0.12% and 0.41%, respectively. A larger city area is associated to a higher price. It is well known that sprawl cities are on average more costly to operate than compact cities. This cost increase is again reflected in the price. We have introduced two variables in the hedonic price equation to assess the financial constraints of the city. Cities with an important debt per capita have to be cautious in their expenses to control further development of debt. Thus these cities are not able to subsidize as much their public transit. A 1% increase in the debt per capita increases the price by 0.1%. This result suggests that cities put an important priority on their public transit increasing by little the bus fare when the level of debt increases. The unemployment rate has a more dramatic effect probably because of the tax levied on firms to subsidize public transit. A city experiencing a large unemployment rate is likely to receive less taxes from firms, implying a larger proportion of the transportation costs to be subsidized by the city leading the city to increase bus fares. An increase by 1% of the unemployment rate increases the bus fare by 0.32%. Given that there is a positive correlation between unemployment rate and the debt per capita, we can expect this effect to be even stronger for cities experiencing both, namely a large debt and an important unemployment rate. The two political variables are singificant. A

data for 1994 and 1995 and average over the two years. Note that these observations are related to the post 1989 election. We have then checked whether the debt is larger for leftist cities after 1989 than for rightist cities as we could expect. We have not found any particular pattern.

dummy has been considered for the left versus right majority instead of the proportion of the council seats held by leftist politicans. In cities with a rightist majority, 77.92% of council seats are held by rightist politicians, while in cities with a leftist majority 60.85% of council seats are held by leftist politicians. In view of these figures, it is not as much the proportion of seats held by leftist politicans, which matters in the policy decisions taken by the city council but rather the nature of the majority in the city council. In contrast, the proportion of seats held by environmentalist is a variable of interest as it represents their popular support. A larger proportion of environmentalists may have a greater influence on the policy decisions because it represents a larger proportion of swing voters that the majority could attract at the next election.[18] Everything else being equal, a switch from a rightist to a leftist majority would decrease the price by about 7.3%. Similarly, the coming of environmentalists in the city council for an average value of 4.80% would decrease the price by 6.89% with the same majority. The time trend indicates a price increasing over time due to other factors.

To take into account the panel structure of the data, we have also considered a model with fixed effects and a model with random effects. The magnitude of the coefficients is somewaht different but the coefficients remain significant except for the dummy for the left majority. This could be explained by the fact that a large proportion of cities, about 74%, have not experienced a change in the majority over the period. We also note that the debt variable has lost some significancy in the random effects model. Since the data provide a single observation for each city in the sample, it may have been captured in the random effect. For the fixed effect model, we have tested the equality of the fixed effects. This hypothesis is strongly rejected with a p-value equal to 0.000. It is unclear whether the city effects are correlated with the other exogenous variables in the model such as the quality variables and the political variables though our intution will favor correlation. We have then performed a Hausman test. The p-value is equal to 0.0715, which does

[18]Conditional on the presence of environmentalist in the city council, they have obtained 2.63% of council seats after the 1983 election while this number increases up to 4.80% after the 1989 election.

not provide a clear response since we can consider the fixed effect model at 10% and the random effect model at 5%. Since the 5% level is the most widely used, we consider that there is no correlation between the city effects and the exogenous variables of the model given in Table 4. The variable debt per capita could raise the possibility of a simultaneity problem in the model. We note that the amount to be subsidized computed as the total operating costs minus the revenue from bus fares is still a small proportion of the total budget of the city, namely 6.65% on average, excluding a simultaneity problem.[19] Morever, in view of regulatory models, the error term in the hedonic price equation includes the unobserved firm's type, which can be potentially correlated with the exogenous variables in the hedonic price model. This term of unobserved heterogeneity is captured by the city effect. Since the random effect model is preferred over the fixed effect model, such correlation is negligible. We can then consider that the variables in the hedonic price model do not suffer from endogeneity.

This section has provided strong empirical evidence that the political composition of the city council influences the price of public transit and therefore its regulation. The next sections present a model incorporating this political dimension in the regulatory contract as well as its estimation on the data.

3 The Economic Model and the Estimation Method

We need to consider an incentive regulatory contract model, while incorporating a political dimension in agreement with the empirical evidence displayed previously. Moreover, this model has to be structurally "estimable." We exclude regulatory capture models as in Laffont and Tirole (1991) since the environementalists directly participate to the election. In addition, these models involve three parties instead of two (three-tier models) and have been derived for two firms' types of adverse selection. A such, their estimation would

[19]Note that a tax on firms is used to subsidize public transit. We can then expect that public transit represents less than 6.65% of the cities' budget.

require additional data and the strong assumption that firms reduce to good (efficient) and bad (inefficient) firms.[20] We first discuss other possibilities that we found in the literature.

3.1 Discussion of Various Approaches

Laffont (1996) and Laffont (2000) propose some extensions of the Laffont and Tirole (1986) model while taking into account a political dimension to industrial policy. Two types of consumers are considered in the population, namely type 1 and type 2. A proportion $1-\alpha$ of type 2 consumers enjoy more the good than the proportion α of type 1 consumers. Type 2 consumers could be considered as those using public transit and type 1 consumers those who do not use public transit. Note that, for the exception of Paris which is not included in the data, $1-\alpha$ is quite large in the case of public transit.[21] When considering a simple random majority model, type 1 and type 2 consumers, whoever gets the majority at the election, will delegate to a politician the design of the regulatory contract. Majority 1 representing type 1 consumers will maximize the social welfare for the proportion of type 1 consumers only ignoring the remaining $1-\alpha$ proportion of the population.[22] Similarly, majority 2 representing type 2 consumers will maximize the social welfare for the proportion of type 2 consumers only, while taking into account that these consumers enjoy more the good by considering a weight β larger than one for the consumer surplus. To simplify, we consider the case of public ownership of the firm. The results are as follows. When majority 1, (say) right wing, wins the election, the solution to the problem gives an equation similar to the Ramsey pricing, while the contract induces less effort from

[20] The extension of such models to a continuum of types seems to be extremely difficult to conduct and is beyond the scope of this paper.

[21] We have not found exact data on the proportion of the urban population using public transit. At least one person is using public transit in 60% of French households.

[22] While not taking the bus, type 1 consumers still derive some utility from public transit but not as much as type 2 consumers. We can consider that they benefit from public transit as it decreases the level of traffic congestion.

the firm relatively to the Laffont and Tirole (1986) model. An important assumption in the model is that the proportion α needs to be larger than $1/(1+\lambda)$, λ representing the shadow cost of public funds. In developing countries this cost is estimated at 0.3. This would require α to be larger than 0.77, which is a very strong majority. Similarly, when majority 2, (say) left wing, wins the election, the solution to the problem gives a distorted level of effort, which is in this case similar to the one in the Laffont and Tirole model (1986), while the price tends to be lower than the standard Ramsey pricing because of the larger weight β on the consumer surplus. Here again, an important assumption is to have $1-\alpha$ larger than $1/(1+\lambda)$, implying a strong majority at the election. Relatively to the social optimum, majority 1 leads to a smaller amount of public transit than socially optimal, while majority 2 leads to a larger amount of public transit than socially optimal. This makes sense since politicians in this model care for their electorate only, which is supposed to enjoy more or less public transit according to their types 2 or 1, respectively.[23] The model can be further extended while considering private ownership, while assuming that the firm belongs (say) to the right (majority 1). As expected, the provision of public transit is smaller than under public ownership because the firm's rent is undervalued by majority 1.

Despite that this model provides an interesting framework taking into account the political dimension, it does not fit to the analysis of our data for the following reasons. First, it is not so simple to divide the population into 2 types of consumers and to assume that these consumers have a uniform vote. As mentioned previously, the proportion of transit commuters is probably less than 50% of the population. Moreover, to our knowledge, transit commuters are not well organized to constitute a special interest group with

[23] Laffont (2000) considers an incomplete contract framework, in which the constitution has the choice to delegate or not the design of the regulatory contract to the politicians based on the unknown value for the cost of public funds. If the economic conditions are unstable or the variance of λ is large, it is better to delegate to the politician. The problem is different in our data since the design of the contract is delegated to a transportation authority under the control of the politicians.

representatives.[24] Moreover, transit commuters represent a very heterogenous population. Though we have not found a statistical study on who are the transit commuters in France, various documents suggest the following users. Given the quasi absence of school buses in French cities, many middle school and high school students rely on public transit. According to data provided by the *Groupement des Autorités Responsables des Transports* (GART), there are 1.9 millions of students taking the bus every day in France. The car onwership rate in France is about 79.4% and only 28.5% of French households have 2 cars and more, leaving a non negligible proportion of the population relying on other transportation modes than the car. Moreover, only 77% of the population over 18 has a driving license, namely 89% of men and 67% of women. This would suggest that the population with a low revenue such as immigrants and working class households and with mobility problems such as elderly are the main users of public transit in addition to students.[25]

Given the heterogeneity of public transit consumers, it seems heroic to assume a unique vote. Students and immigrants do not vote in general, while working class households are usually leftist partisans. In contrast, elderly people tend to be more conservative with a rightist vote. Though we could roughly predict that transit users usually favor a leftist ideology, the split is not as sharp. Second, these models require a very strong majority for the politicians to ignore a part of the population. Our data show a single city with such a strong majority (at least 77% of the votes for a λ equal to 0.3).[26] Parties govern with a confortable majority given the mixed majority/proportional electoral rule in France. For cities with a rightist majority, the average proportion of seats in the city council is 77.92%. Thus, not all the cities in our sample would satisfy the assumption of having the majority

[24] A typical example of this lack of organization occurs when public transit is on strike. Transit commuters may loose days of work because of these strikes organized by a small group of unionized workers. Transit commuters have never lobbyed to restrict laws on strikes for public transit bearing many of the consequences due to the strike. This absence of lobbying is probably due to their dispersion.

[25] A study by the GART also shows that two third of the public transit users are women.

[26] Cannes in South of France had 81.26% of the votes obtained by the right wing at the 1989 election.

with a least 77% of the seats in the city council. Third, the data show strong evidence for the influence of environmentalists in the choice of the price for public transportation. The models that we have briefly discussed ignore this possibility as the politicians only care for their majority. Lastly, the ownership of firms operating public transit cannot be easily classified as private versus public. Three companies, namely Keolis, Connex and Transdev, shared 74% of the networks in France in 2001, representing 90% of the market share in terms of passengers. Two of these companies are semiprivate and are affiliated with public firms. In addition to the difficulty of defining the public versus private ownership, the data do not provide information on the identity of the operator. Moreover, it seems difficult to consider that these companies are belonging to the left or right majority. While finding inspiration from this literature, we propose an alternative model.

3.2 The Model

We adopt a Principal-Agent framework to model the incentive regulatory contracts between the regulator and the operator with a political economy dimension. As discussed in Perrigne (2002), a model of incentive regulation with expost observability of the costs as first defined by Laffont and Tirole (1986) seems to be a good approximation of the reality. Our data suggest that contracts are nonetheless influenced by the political composition of the city council. Therefore, we need to leave the assumption of a benevolent utilitarian maximizer as widely accepted in the political economy literature. We will come back on this assumption later while writing the model.

The basic idea is to assume that the firm in charge of operating the bus service has hidden information on its efficiency and takes hidden reducing cost actions called effort. These lead to the socalled problems of adverse selection and moral hazard, respectively. Both affect the firm's cost in the sense that a higher efficiency and effort level tend to decrease the costs. Both are assumed to be unknown to the regulator. The regulator does not know the firm's efficiency and effort level in the sense that even when observing expost the firm's cost, he cannot disentangle these effects from the random shock in the cost. Through the contract proposed by the regulator to the firm, which consists in a price and a cost reimbursement rule based on ex post observed costs, the firm will be induced to exert appropriate effort level without lying about its efficiency. This leads to the well known efficiency/rent trade-off, where the regulator has to give enough incentives to the firm in terms of profit to reach efficiency while extracting some firm's rent.

The firm knows its efficiency or type denoted θ, which is assumed to be drawn from a distribution $F(\cdot)$ defined on $[\underline{\theta}, \overline{\theta}]$ with a density $f(\cdot) > 0$, where $\underline{\theta}$ denotes the most efficient firm and $\overline{\theta}$ the least efficient one. Thus, the regulatory authority offers a contract to the firm based on the expected demand and cost, which are both subject to some external shocks.[27] This contract is defined as the pair (p, t), where p denotes the price that the firm will be authorized to charge to consumers and t the transfer to the firm. This transfer is a cost reimbursement rule as a function of the expost observed realized cost. After the demand and cost are realized, the transfer is paid to the firm based on its actual (observed) cost.

Because public transportation is a private good, the firm faces a demand denoted by $y(p)$ subject to some random shock denoted ϵ_d. To simplify we adopt a multiplicative random shock and denote $\overline{y}(p)$ the deterministic part of the demand, namely the expectation of the demand with respect to ϵ_d is equal to $\overline{y}(p)$. The revenue generated from the bus fares is equal to $R(p) = py(p)$. This demand generates a gross consumer surplus denoted $S(p)$, while $S(p) - R(p)$ denotes the net consumer surplus and can be computed as $\int_p^\infty y(\tilde{p})d\tilde{p}$ following Assumption 1 in Perrigne (2002).

The firm's effort is denoted by e. Effort is costly to the firm leading to an effort

[27] The Laffont and Tirole (1986) model assumes a constant marginal cost with an additive random shock, while the demand is known with certainty. A fixed demand is unrealistic in the case of public transit, whose demand can fluctuate from one year to another. French data show that the demand for public transportation has increased over the period of study, while public transit demand has experienced a slight decrease over the recent years. It seems difficult to predict with certainty these fluctuations justifying a random demand.

disutility $\psi(e)$. Usual assumptions are made such as $\psi'(\cdot) > 0$, $\psi''(\cdot) > 0, \forall e > 0$ and $\psi(0) = 0$.[28] The firm's cost denoted by $C(y, \theta - e)$ is subject to a multiplicative random shock ϵ_c. The cost is increasing in θ and decreasing in e. The firm's utility is defined as $U = t - \psi(e)$, where t is the net transfer to the firm, which is a function of the firm's type θ and the realized cost $C(y, \theta - e)$, namely $t(\theta, C(y, \theta - e))$. The gross transfer is by definition equal to $t + C(y, \theta - e)$. Lastly, giving a subsidy to the firm implies that additional taxes need to be raised introducing a distortion to the economy. Thus the shadow cost of public funds denoted by λ needs to be taken into consideration. Though we will estimate λ, in western countries where the tax collection system is efficient, λ is estimated at 0.3.[29]

The firm's optimization problem is assumed to be independent of political considerations as, to our knowledge, the firms operating public transportation do not change over time according to the political majority. Though some firms are semiprivate, we still consider that their optimization problem is to maximize their utility $U(\theta, e) = t(y(\theta), \theta - e) - \psi(e)$ with respect to the two decision variables (θ, e). Note that y becomes a function of θ through the price decided by the regulator. The firm may lie about its efficiency parameter θ. Let denote by θ its true efficiency and by $\tilde{\theta}$ its announced efficiency. Thus, a firm with a θ efficiency announcing a $\tilde{\theta}$ efficiency has a utility $U(\tilde{\theta}, \theta, e) = t(\tilde{\theta}, C(y(\tilde{\theta}, \theta - e))) - \psi(e)$. Because the demand and the cost are subject to some random shocks unknown at the time of the design of the contract, the firm maximizes the expected value of its utility, namely $\mathrm{E}[U(\tilde{\theta}, \theta, e)]$, where $\mathrm{E}[\cdot]$ denotes the expectation of the term between brackets with respect to the random shocks (ϵ_d, ϵ_c). Expectation has to be taken with respect to both random shocks since ϵ_d affects y.

Maximizing $\mathrm{E}[U(\tilde{\theta}, \theta, e)]$ with respect to e for a fixed value of $\tilde{\theta}$ gives an effort as a function of the annouced $\tilde{\theta}$ conditional on the true θ, i.e. $e(\tilde{\theta}|\theta)$. Thus the firm's

[28] We could add $\psi'''(\cdot) > 0$ to avoid stochastic mechanisms.

[29] This value is widely accepted among economists. It is expected to be larger in developing economies. See Ballard, Shoven and Whalley (1985) for an estimate of λ from US macroeconomic data.

maximization problem can be written as $\max_{\tilde{\theta}} \mathrm{E}[U(\tilde{\theta}, \theta)]$, where $U(\tilde{\theta}, \theta) = t(\tilde{\theta}, C(y(\tilde{\theta}), \theta - e(\tilde{\theta}|\theta))) - \psi(e(\tilde{\theta}|\theta))$. The solution of this program gives $\tilde{\theta} = \tilde{\theta}(\theta)$. Using truth-telling, $U(\tilde{\theta}, \theta) = U(\theta, \theta)$. To simplify notations, we denote the latter by $U(\theta)$. Maximizing $\mathrm{E}[U(\theta)]$ gives the following first-order condition after some algebra

$$\mathrm{E}[U'(\theta)] = -\psi'(e). \tag{1}$$

Equation (1) provides the incentive compatibility (IC) constraint to the regulator's maximization problem.

We need to pay particular attention to the regulator's maximization problem. First, it is more complex than for the case of a public good as considered usually in the theoretical literature. As a matter of fact, it has to take into account various factors such as the net consumer surplus $S(p) - py(p)$, the revenue evaluated at the cost of public funds $(1 + \lambda)py(p)$, the cost for the gross transfer to the firm evaluated at the cost of public funds $(1 + \lambda)(t + C(y, \theta - e))$ in addition to the firm's profit or utility U. We have discussed in details that politicans are self-interested in the sense that they maximize the probability of being reelected to maintain their position. Thus they have to please their electorate while trying to convince more electors to vote in their favor at the next election. Moreover, they can be motivated by partisan ideologies, which can be dictated by their parties such as redistributive policies for leftist parties. For both reasons, the politician is not a benevolent utilitarian maximizer. Given the heterogeneity among bus commuters, it seems, however, difficult to target successfully a particular group of the population to gain their votes. This problem has been acknowledged in the political economy literature for the provision of public goods. Thus, the politicians cannot consider only a part of the population in their maximization problem. We propose the following solution to the Principal's maximization program. The politician will maximize the various components of the social welfare, while putting a different weight on the consumer surplus according to his/her political affiliation and composition of the city council. In particular, we expect that cities with a leftist majority and cities with environmentalists in their city council will put a larger weight on consumer surplus than other cities. The impact on the contract will be studied further.

We consider the following "social welfare" for self-interested politicians as

$$SW = \beta(S(p) - py(p)) + (1+\lambda)py(p) - (1+\lambda)(t + C(y, \theta - e)) + U,$$

where β is a weight on the consumer surplus. Note that this weight can take a large range of values depending on how much politicians care about consumers. Using $t = U + \psi(e)$, it can be written equivalently as

$$SW = \beta S(p) + (1 + \lambda - \beta)py(p) - (1+\lambda)[\psi(e) + C(y, \theta - e)] - \lambda U, \tag{2}$$

where $U = U(\theta)$.

The regulator's objective is to maximize this quantity under the firm's incentive compatibility and individual rationality constraints. The latter states that the firm should expect a positive profit to accept the contract. Because the regulator decides about the contract to offer to the firm at the beginning of the period and the demand and the cost are both subject to random shocks unknown to both parties exante, everything need to be taken in expectation with respect to these two random shocks. This gives the following maximization problem for the regulator or politician

$$\begin{aligned} &\max_{p,e,U} \int_{\underline{\theta}}^{\overline{\theta}} \mathrm{E}[SW] f(\theta) d\theta, \qquad (3)\\ &\text{Subject to } \ \mathrm{E}[U'(\theta)] = -\psi'(e) \quad (IC),\\ &\qquad\qquad \mathrm{E}[U(\theta)] \geq 0 \qquad\qquad (IR), \end{aligned}$$

where SW is given by (2). Because the firm's efficiency or type θ and therefore its effort level $e(\theta)$ are unobserved, the regulator maximizes an expected social welfare with respect to θ. Because $U(\cdot)$ is decreasing in θ, the individual rationality constraint is satisfied if and only if $\mathrm{E}[U(\overline{\theta})] = 0$. It is then commonly replaced by $\mathrm{E}[U(\overline{\theta})] = 0$, i.e. the least efficient firm will have a zero profit, since SW is decreasing in U. The firm will then reveal the truth about its efficiency level θ and will exert appropriate level of effort.

Such a maximization problem can be solved using the Pontryagin principle. Let $E_{\epsilon_d}[y(p)] = \overline{y}(p)$, the Hamiltonian is

$$\begin{aligned} H &= \Big\{(1+\lambda)p\overline{y}(p)+\beta\int_p^\infty \overline{y}(\tilde{p})d\tilde{p}-(1+\lambda)\psi(e)-(1+\lambda)\mathrm{E}[C(y,\theta-e)]-\lambda\mathrm{E}[U(\theta)]\Big\}f(\theta) \\ &\quad +\mu(\theta)(-\psi'(e)), \end{aligned}$$

where $\mu(\cdot)$ is the co-state variable and $\mathrm{E}[U(\cdot)]$ is the state variable. After some basic algebra, the first-order conditions are as follows

$$H_p = \Big\{(1+\lambda)p\overline{y}'(p)+(1+\lambda-\beta)\overline{y}(p)-(1+\lambda)\mathrm{E}\left[\frac{\partial C(y,\theta-e)}{\partial y}\frac{\partial y(p)}{\partial p}\right]\Big\}f(\theta)=0, \quad (4)$$

$$H_e = \Big\{-(1+\lambda)\psi'(e)-(1+\lambda)\frac{\partial \mathrm{E}[C(y,\theta-e)]}{\partial e}\Big\}f(\theta)-\mu(\theta)\psi''(e)=0, \quad (5)$$

$$-H_{\mathrm{E}[U(\theta)]} = \lambda f(\theta) = \mu'(\theta), \quad (6)$$

where H_X denotes partial differentiation of the Hamiltonian H with respect to X and $\overline{y}'(p)$ denotes the expectation of the derivative of $y(p)$ with respect to p.

By integrating (6) and using the transversality condition $\mu(\underline{\theta}) = 0$, it gives $\mu(\theta) = \lambda F(\theta)$. Replacing $\mu(\theta)$ by $\lambda F(\theta)$ and rearranging terms, (4) and (5) become

$$\frac{p-\mathrm{E}\left[\frac{\partial C(y,\theta-e)}{\partial y}\frac{\partial y(p)}{\partial p}\right]/\overline{y}'(p)}{p} = \frac{-(1+\lambda-\beta)}{1+\lambda}\frac{\overline{y}(p)}{p}\frac{1}{\overline{y}'(p)}, \quad (7)$$

$$\psi'(e) = -\frac{\partial \mathrm{E}[C(y,\theta-e)]}{\partial e} - \frac{\lambda}{1+\lambda}\frac{F(\theta)}{f(\theta)}\psi''(e). \quad (8)$$

Equation (7) provides the pricing rule adopted by the regulator. When considering a cost function with a constant marginal cost and an additive random shock such as $C(y,\theta-e) = (\theta-e)y+\epsilon_c$, (7) gives the Ramsey pricing rule for $\beta = 1$, where the price depends on the cost of public funds λ, the marginal cost and the demand elasticity. Such a specification is in general chosen by theorists giving the numerator of the left-hand side of (7) as $p - \partial C/\partial y$ (see for instance Laffont and Tirole (1986)). Because we consider a more general functional form, (7) does not have the same interpretation. Moreover, because of β, the price distortion takes a different form. While taking the simple example of a constant marginal cost, we observe that the price will be closer to the marginal cost for $\beta > 1$ since $(1+\lambda-\beta)/(1+\lambda) < \lambda/(1+\lambda)$ assuming a constant price elasticity. In contrast, when $\beta < 1$, the price will be larger than the Ramsey price.

Equation (8) provides the effort level. The optimal effort level is distorted relatively to the first best as in the Laffont and Tirole (1986) model providing similar incentives to the firm. Equations (7) and (8) will be the basis of the econometric model to estimate incentive regulatory contracts.

3.3 Estimation Method

The econometric model is defined from (7) and (8). Since the estimation problem is similar to the one studied in Perrigne (2002), we will rely on this method to estimate the model. As a matter of fact, (7) differs by the introduction of the consumer weight β. To define the econometric model, we need first to parameterize the demand, the cost and the disutility of effort.

The demand function for public transportation takes the simple following form

$$y(p) = \exp(d_0) Z_d^{d_1} p^{d_2} \exp(\epsilon_d), \tag{9}$$

where Z_d denotes a vector of exogenous variables such as the city's characteristics and indicators for quality, ϵ_d is a random shock. Using this specification, d_2 ($d_2 < 0$) is interpreted as the price elasticity. As in Wolak (1994), we can assume that this demand is defined only for $p \leq p_{\max}$ and takes a zero value for any price above this maximum price.[30]

Given the complexity of the model, we choose a rather simple representation of the technology using a Cobb-Douglas production function. Under this assumption, the variable cost function is expressed as

$$C(y, \theta - e) = \exp(\beta_0) \exp(\beta_L(\theta - e)) p_L^{\beta_L} p_M^{\beta_M} y(p)^{\beta_y} \exp(\epsilon_c), \tag{10}$$

[30] Assuming the existence of a maximum price allows a well-defined net consumer surplus for this constant elasticity demand function for any $d_2 < 0$.

where p_L and p_M denote the price for labor and material, respectively, and ϵ_c is a random shock. We assume that the firm's type and effort affect labor efficiency. As in Wolak (1994), capital costs are excluded and C represents the operating costs.[31] We impose homogeneity of degree one in price by setting $\beta_L + \beta_M = 1$. We could also introduce in (10) some exogeneous variables Z_c associated with a vector of coefficients β_Z. Regarding the disutility of effort $\psi(e)$, we choose the following exponential form

$$\psi(e) = \exp(\alpha e) - 1, \tag{11}$$

where α is a parameter to be estimated. To satisfy the assumption of the model, α needs to be strictly positive. The random shocks ϵ_d and ϵ_c are unobserved exante by both the regulator and the firm. These random shocks will be interpreted as the error terms in the econometric model that we will define in the next section. We make the following assumption on ϵ_d and ϵ_c.

Assumption A: *The random shocks ϵ_d and ϵ_c satisfy* $\mathrm{E}[\exp(\epsilon_d)|Z] = 1$ *and* $\mathrm{E}[\exp(\epsilon_c)|Z] = 1$*, where Z is a vector of exogenous variables. Moreover, the firm's type θ is independent of ϵ_d conditionally to Z.*

The first part of assumption A is quite natural following the multiplicative random shocks in (9) and (10).[32] Note that we do not make any other assumptions on the distribution of the random shocks beyond their first conditional moments. Moreover, we do not make any independence assumption between these two random shocks as they can be correlated.[33] The second part of Assumption A is in agreement with the definition of the firm's type

[31] Working with operating costs can be easily justified. The main capital (buses) is provided by the City or the local transportation authority and does not show up in the firm's accounting report. Moreover, an accurate measure of capital costs is usually difficult to obtain. When adding the labor, energy, maintenance and other material costs, we obtain about 90% of the costs reported by the firms in the sample.

[32] The expected net consumer surplus $\mathrm{E}[S(p) - py(p)]$ becomes $\int_p^\infty \overline{y}(\tilde{p})d\tilde{p}$.

[33] This assumption precludes maximum likelihood estimation, which requires a full parametric specification for the joint density of the random shocks as in Wolak (1994). We could have considered a log-log form of (9) with an expected random shock ϵ_d equal to zero. In this case, $\overline{y}(p)$ would be interpreted as

as made by theorists in the sense that θ is idiosyncratic to the firm and known before any realization of the demand. Moreover, it is crucial to identify the model as discussed later. Assumption A will be used later in the estimation method to specify moments defining GMM estimators.

It remains to discuss the parameterization of the firms' type density $f(\cdot)$. Because of its flexibility, we consider a Gamma density, i.e. $f(\theta; r, \gamma) = \gamma(\gamma\theta)^{r-1}\exp(-\gamma\theta)/\Gamma(r)$, where $r \in \mathbb{N}^+$ and $\Gamma(r) = \int_0^\infty x^{r-1}\exp(-x)dx$.[34]

We observe $N = 57 \times 9 = 513$ contracts and index by i the ith contract between a firm and its corresponding city. Rewriting the price equation (7) using the above parameterization and the logarithm, the econometric model is defined by the following three equations

$$
\begin{aligned}
y_i &= \exp(d_0) Z_{di}^{d_1} p_i^{d_2} \exp(\epsilon_{di}), && (D)\\
C_i &= \exp(\beta_0) p_{Li}^{\beta_L} p_{Mi}^{1-\beta_L} y_i^{\beta_y} \exp(\beta_L(\theta_i - e(\theta_i))) \exp(\epsilon_{ci}), && (C)\\
\log p_i &= \beta_0 + \log E_c + \log \beta_y - \log\left(1 + \frac{1+\lambda-\beta_i}{1+\lambda}\frac{1}{d_2}\right) \\
&\quad + \beta_L \log p_{Li} + (1-\beta_L)\log p_{Mi} + (\beta_y - 1)\log \overline{y}(p_i) + \beta_L(\theta_i - e(\theta_i)), && (P)
\end{aligned}
$$

where β_i is the weight for the consumer surplus and $e(\theta_i)$ is the solution of the system of equations (7) and (8) at a given value of θ_i defining the optimal effort level, for $i = 1, \ldots, N$. The term E_c arises because the cost needs to be considered in expectation with respect to ϵ_d and ϵ_c, namely $E_c = \mathrm{E}[\exp(\beta_y \epsilon_d + \epsilon_c)]$. In two of these equations, the effort $e(\theta_i)$ appears. Both $e(\theta_i)$ and θ_i are unobserved. Another equation could be used for the price. In particular, solving (7) and (8) in (p, e) given θ_i gives the regulator's price

the expected logarithm of the demand, which would complicate (7) and (8) and therefore the derivation of the econometric modelling.

[34]Given that we have no information on the type density, it would have been interesting to consider a model where $f(\cdot)$ is left unspecified or non parameterized. This would avoid any misspecification issue and would reinforce any policy analysis. This would lead to a semiparametric model. Additional information would be needed to identify nonparametrically the type distribution. Such an issue is left for future research.

schedule $p(\theta_i)$ and the firm's optimal effort $e(\theta_i)$. Elementary algebra gives

$$\begin{aligned}
\log p_i &= \frac{1}{\alpha+\beta_L+d_2(-\alpha\beta_y+\alpha+\beta_L)}\Bigg\{K-(\alpha+\beta_L)\log\left(1+\frac{1+\lambda-\beta_i}{1+\lambda}\frac{1}{d_2}\right) \\
&\quad +\alpha\beta_L\log\frac{p_{Li}}{p_{Mi}}+\alpha\log p_{Mi}+(\alpha(\beta_y-1)-\beta_L)(d_0+d_1\log Z_{di}) \\
&\quad +\beta_L\left(\alpha\theta_i+\log\left(1+\frac{\alpha\lambda}{1+\lambda}\frac{F(\theta_i;r,\gamma)}{f(\theta_i;r,\gamma)}\right)\right)\Bigg\}, \qquad (P') \\
e_i &= \frac{1}{\alpha+\beta_L+d_2(-\alpha\beta_y+\alpha+\beta_L)}\Bigg\{K'-\beta_y d_2\log\left(1+\frac{1+\lambda-\beta_i}{1+\lambda}\frac{1}{d_2}\right) \\
&\quad +\beta_L(1+d_2)\log\frac{p_{Li}}{p_{Mi}}+(1+d_2)\log p_{Mi}+\beta_y(d_0+d_1\log Z_{di}) \\
&\quad +\beta_L(1+d_2)\theta_i-(1-d_2(\beta_y-1))\log\left(1+\frac{\alpha\lambda}{1+\lambda}\frac{F(\theta_i;r,\gamma)}{f(\theta_i;r,\gamma)}\right)\Bigg\}, \qquad (E)
\end{aligned}$$

for $i=1,\ldots,N$, where K and K' are constant terms. Namely, $K=\alpha(\beta_0+\log E_c)+(\alpha+\beta_L)\log\beta_y+\beta_L(\log\alpha-\log\beta_L)$ and $K'=(1+d_2)\beta_0+\log\beta_L(1-d_2(\beta_y-1))+(1+d_2)\log E_c+\log\alpha(-1+d_2(\beta_y-1))+\beta_y d_2\log\beta_y$.

From an econometric point of view, the random shock ϵ_d in (D) is interpreted as an error term. Similarly, the term $\theta_i-e(\theta_i)$ in (P) and the last term in (P') as a function of θ_i can be interpreted as an error term with a nonzero mean. The term $\theta_i-e(\theta_i)$ in (C) is part of the error term $\exp(\beta_L(\theta_i-e(\theta_i)))\exp(\epsilon_{ci})$. As a matter of fact, this term can be interpreted as a term of unobserved firm's heterogeneity in (C). The difficulty in estimating such a model arises from the fact that we observe neither the θ_i, nor the effort level $e(\theta_i)$ and that we cannot disentangle the part arising from the firm's type θ_i and the one from the random shock ϵ_{ci} from the error term $\exp(\beta_L(\theta_i-e(\theta_i)))\exp(\epsilon_{ci})$. Moreover, the density $f(\cdot;r,\gamma)$ of the firms' types needs to be estimated whose parameters (r,γ) appear in $e(\theta_i)$ and in $p(\theta_i)$ as in (P').[35]

[35] Wolak (1994) introduces an additional "econometrician" error term in the price equation. By parameterizing the joint distribution of all the error terms including θ, the model is identified. Such an "econometrician" error term is not necessary. Moreover, our method does not require to parameterize the distribution of the error terms ϵ_d and ϵ_c as only a first moment restriction is used to identify and estimate the model.

In addition to the parameters $(d_0, d_1, d_2, \beta_0, \beta_L, \beta_y, \lambda, \alpha, r, \gamma)$, we need to estimate the weight β_i for each observation in our sample. The reduced form analysis that we have previously conducted suggests that bus fares are a function of political factors. These political factors are taken into account in the model through the weight β. We then propose to parameterize the weights β_i as a linear function of exogenous variables, namely

$$\beta_i = \eta_0 + \eta_1' Z_{Pi}, \tag{12}$$

where Z_{Pi} is a vector of variables reflecting the political composition of the city council and the city's financial constraints. Thus the parameters to be estimated are $(d_0, d_1, d_2, \beta_0, \beta_L, \beta_y, \eta_0, \eta_1, \lambda, \alpha, r, \gamma)$, while the observables are $\{C_i, p_{Li}, p_{Mi}, y_i, p_i, Z_{di}, Z_{Pi}, i = 1, \ldots, N\}$. We briefly discuss the identification of the model as the estimation procedure closely follows the identification.

Using Assumption A providing a first conditional moment for $\exp(\epsilon_{di})$, the parameters d_0, d_1 and d_2 are identified from (D) as well as the error terms ϵ_{di}, $i = 1, \ldots, N$. Instead of using (P) to identify the parameters $(\beta_L, \beta_y, \eta_0, \eta_1)$, we propose to use (P') because (P') will also allow us to identify α. In particular, using Assumption A, (P') allows us to identify the parameters α, β_L and β_y. Note that (P') (as well as (P)) allows us to identify only $\eta_0/(1+\lambda)$ and $\eta_1/(1+\lambda)$.[36] As β_L, β_y, d_2, $\eta_0/(1+\lambda)$ and $\eta_1/(1+\lambda)$ are identified, (P) allows to recover the error terms up to a constant, namely $\beta_0 + \log E_c + \beta_L(\theta_i - e(\theta_i))$, $i = 1, \ldots, N$. Using these recovered error terms and Assumption A providing a first conditional moment for $\exp(\epsilon_{ci})$, (C) allows to identify the error terms ϵ_{ci}, $i = 1, \ldots, N$. The recovered error terms $(\epsilon_{di}, \epsilon_{ci}, i = 1, \ldots, N)$ and β_y allow to compute E_c.[37] It remains to identify β_0 and λ as well as the parameters (r, γ) of the density $f(\cdot)$. Several identifying strategies can be used. A first strategy will consist to use the recovered values for $\beta_0 + \beta_L(\theta_i - e(\theta_i))$ from (P), which is also equal to a function of parameters $\beta_0, \lambda, r, \gamma$ using (E). As the distribution of the left-hand side is known (from (P)), the latter parameters

[36] Note that the constant term in (P') is not identified because we use an orthogonality condition.

[37] Note that we could have used this step to identify the parameters $\eta_0/(1+\lambda)$ and $\eta_1/(1+\lambda)$ instead of identifying the latter from (P').

are identified. Similarly, a second strategy could consist in using (P'), which provides a function of parameters $\beta_0, \lambda, r, \gamma$. Using a similar argument, the parameters are identified.

The estimation method closely follows the identification discussed previously. It is a multi-step procedure, where equations are estimated in a specific order allowing at each step to recover information, which will be used in the following step.[38] The method is, however, simple to implement despite the complexity of the model.

The first step consists in estimating the demand equation defined by (D) using Assumption A. In particular, the conditional moment $\mathrm{E}[\exp(\epsilon_{di})|Z_i] = 1$ is used to define the unconditional moment $\mathrm{E}[\Phi(Z_i)(\exp(\epsilon_{di}) - 1)] = 0$, where $\Phi(\cdot)$ is a vector function of exogenous variables or instruments Z. This defines a Nonlinear GMM estimator based on the following moment

$$\mathrm{E}\Big[\ \Phi(Z_i)\{y_i \exp(-d_0) Z_{di}^{-d_1} p_i^{-d_2} - 1\}\Big] = 0. \tag{13}$$

The vector of instruments $\Phi(Z_i)$ could be chosen optimally following Chamberlain (1987). In the application, we choose an identity function for $\Phi(\cdot)$. We perform a two-step nonlinear GMM, where the matrix of weights is computed optimally to reduce the variance of the estimated parameters. See Hansen (1982) and Hayashi (2000) for GMM estimation. We then obtain estimated parameters $(\hat{d}_0, \hat{d}_1, \hat{d}_2)$ and estimated residuals $\hat{\epsilon}_{di}$, $i = 1, \ldots, N$.

The second step consists in estimating the price equation (P'), whose error term (say) η_i is a function of θ_i, namely $\eta_i = \beta_L(\alpha\theta_i + \xi(\theta_i))/(\alpha + \beta_L + d_2(-\alpha\beta_y + \alpha + \beta_L))$. The moments of η_i are unknown. We are using instead Assumption A, which implies that this error term is independent of ϵ_{di} conditionally on Z_i. This independence can be expressed as a conditional covariance equal to zero. This defines a Nonlinear GMM estimator based

[38] This multi-step procedure is certainly not the most efficient as estimating the full model jointly will bring some efficiency gains. Nonetheless, estimating the full model could introduce inconsistency of parameter estimates due to a misspecification of $f(\cdot)$. As the estimation of $f(\cdot)$ is performed in the final step involving (β_0, r, γ), the estimates for $(d_0, d_1, d_2, \alpha, \beta_L, \beta_y)$ do not suffer from such a potential inconsistency.

on the following moment

$$\begin{aligned}
\mathrm{E}\Bigg[\Phi(Z_i)\Bigg\{ & \log p_i - \frac{\alpha\beta_L}{\alpha+\beta_L+d_2(-\alpha\beta_y+\alpha+\beta_L)} \log \frac{p_{Li}}{p_{Mi}} \\
& + \frac{\alpha}{\alpha+\beta_L+d_2(-\alpha\beta_y+\alpha+\beta_L)} \log p_{Mi} \\
& \frac{\alpha+\beta_L}{\alpha+\beta_L+d_2(-\alpha\beta_y+\alpha+\beta_L)} \log\left(1+\frac{1}{d_2}-\frac{\eta_0}{d_2(1+\lambda)}-\frac{\eta_1}{d_2(1+\lambda)}Z_{pi}\right) \\
& + \frac{\alpha\beta_y-\alpha-\beta_L}{\alpha+\beta_L+d_2(-\alpha\beta_y+\alpha+\beta_L)}(d_0+d_1\log Z_{di})\Bigg\}(\exp(\epsilon_{di})-1)\Bigg] = 0, \qquad (14)
\end{aligned}$$

where $d_0 + d_1 \log Z_{di}$, d_2 and ϵ_{di} can be replaced by their estimates obtained from the first step, i.e. $\hat{d}_0 + \hat{d}_1 \log Z_{di}$, $\hat{d}_2$ and $\hat{\epsilon}_{di}$, respectively.[39] The estimated coefficients give an unique solution for the estimated coefficients $(\hat{\alpha}, \hat{\beta}_L, \hat{\beta}_y, \eta_0/\widehat{(1+\lambda)}, \eta_1/\widehat{(1+\lambda)})$.

The third step consists in estimating the cost equation (C), which involves $(\theta_i - e(\theta_i))$. Note that this term of unknown value appears in (P) as well. Using the previous estimates, (P) provides an estimate for this term up to a constant, namely $\beta_0 + \log E_c + \beta_L(\theta_i - e(\theta_i))$, $i = 1, \ldots, N$. From Assumption A, we can exploit the first moment condition on the error term ϵ_{ci}, namely $\mathrm{E}[\exp(\epsilon_{ci})|Z_i] = 1$. Using the estimates from the previous two steps and the terms $\theta_i - e(\theta_i)$ estimated up to a constant, the first moment condition on $\exp(\epsilon_{ci})$ provides an estimate for the error terms ϵ_{ci}, $i = 1, \ldots, N$. Note that we can test whether the random shocks on demand and cost are independent. For instance, we can compute a simple correlation coefficient using the $\hat{\epsilon}_{d_i}$s and $\hat{\epsilon}_{c_i}$s to assess the degree of correlation between the two. The term E_c is by definition $\mathrm{E}[\exp(\beta_y \epsilon_{di} + \epsilon_{ci})]$. A simple method for estimating E_c consists in averaging the exponential of the estimated terms $\hat{\beta}_y \hat{\epsilon}_{di} + \hat{\epsilon}_{ci}$. The estimated E_c is noted $\hat{E}_c$.

It remains to estimate the constant term of the cost function β_0, the cost of public funds λ and the parameters of the firms' type density (r, γ). This is the purpose of the fourth and final step of our estimation method. These parameters appear in (P'). The

[39] As for the first step, we choose the identity function for the vector of instruments and perform a two-step nonlinear GMM estimator using an optimal weight matrix.

three previous steps provide estimates for $d_0, d_1, d_2, \alpha, \beta_L, \beta_y, E_c$, as well as $\eta_0/(1+\lambda)$ and $\eta_1/(1+\lambda)$, which could be used in (P'). Thus, we obtain $\log p_i = \Psi(\theta_i; \beta_0, \lambda, r, \gamma)$, $i = 1, \ldots, N$. As $r \in \mathbb{N}^+$, we can develop for each value of $r = 1, 2, \ldots$ a Method of Moments estimator based on the first three moments for estimating $(\beta_0, \lambda, \gamma)$. This gives

$$\begin{aligned} \mathrm{E}[\log p_i] &= \mathrm{E}[\Psi(\theta_i; \beta_0, \lambda, r, \gamma)], \\ \mathrm{E}[\log p_i^2] &= \mathrm{E}[\Psi^2(\theta_i; \beta_0, \lambda, r, \gamma)], \\ \mathrm{E}[\log p_i^3] &= \mathrm{E}[\Psi^3(\theta_i; \beta_0, \lambda, r, \gamma)]. \end{aligned}$$

The function $\Psi(\cdot; \beta_0, \lambda, r, \gamma)$ does not lead, however, to some moments that are computationally tractable. We propose instead to estimate the first moment $\int \Psi(\tilde{\theta}; \beta_0, \lambda, r, \gamma) f(\tilde{\theta}; r, \gamma) d\tilde{\theta}$, the second moment $\int \Psi^2(\tilde{\theta}; \beta_0, \lambda, r, \gamma) f(\tilde{\theta}; r, \gamma) d\tilde{\theta}$ and the third moment $\int \Psi(\tilde{\theta}; \beta_0, \lambda, r, \gamma) f(\tilde{\theta}; r, \gamma) d\tilde{\theta}$ by simulated moments.

Using the importance sampling method, the first moment will be estimated by

$$\frac{1}{N} \sum_{i=1}^{N} \frac{1}{S} \sum_{s=1}^{S} \frac{\Psi(\theta_{is}; \beta_0, \lambda, r, \gamma)}{g(\theta_{is})} \times f(\theta_{is}; r, \gamma), \tag{15}$$

where θ_{is} is a simulated value for θ_i drawn from a density $g(\cdot)$. Likewise, the second and third moments can be estimated by their simulated counterpart. Without loss of generality, we can choose $g(\cdot)$ to be an exponential density with parameter equal to 1. This defines a Method of Simulated Moments (MSM) estimator. See Gourieroux and Monfort (1996) for the asymptotic properties of MSM estimators. In particular, the estimator provides consistent estimates for S fixed. We can perform a MSM estimator for each value of $r = 1, 2, \ldots$. For each of r, we obtain a triplet of estimates $(\hat{\beta}_0, \hat{\lambda}, \hat{\gamma})$. We can find different criteria to choose among these different triplets. We can look at the best adjustement for higher moments or compare the value function.

4 Empirical Results

As detailed in the previous section, the model consists in estimating the parameters of the demand, the cost, the weights for consumer surplus, the disutility of effort and the firms' types density as well as the cost of public funds. The first step consists in estimating the demand equation (D) involving the variables y, the price p and a vector of exogenous variables Z_d. The demand y is defined as the number of passengers. The number of passengers offers an important variability with a variation coefficient larger than 1 and is strongly correlated with the size of the population as expected. Thus, the size of the population is a natural choice as an exogenous variable.[40] The demand for public transit can be also affected by the quality of the service offered. The quality of public transit is multidimensional and many aspects are difficult to measure such as the clealiness of buses . Following our previous reduced form analysis, we consider two indices measuring the extent of the network and the number of seats offered per capita. Both variables should have a positive impact on demand as a higher value make public transit more attractive to commuters. Both variables offer enough variability to be entertained in the estimation with coefficients of variation equal to 1.3 and 0.41, respectively. Many empirical studies on demand for public transportation also include the speed, measured as the average speed of buses. Despite that this variable offers a small variability in our data, we have included it in the model. The results are given in Table 5.[41]

The price variable is expected to be endogenous in a demand equation. In our model, the price is determined as a function of the firm's type, i.e. $p_i = p(\theta_i)$. Following Assumption A, the firm's type is independent of the error term ϵ_d. As such, the price varialbe should not be endogenous. For security, we perform the estimation with various instruments for the price variable. In particular, we choose three instruments, which are

[40] We could expect that the demand for public transit will vary with the age distribution of the population and with the average revenue. Data are either incomplete or offering not enough variability across cities to be included in our empirical analysis.

[41] Table 5 does not include a time trend as its coefficient is insignificant with a negative sign.

the unemployment rate, the city area and the debt per capita. The Hansen test statistic for the validity of the instruments is equal to 5.439 distributed as a Chi square with 2 degrees of freedom, which leads to consider these instruments as acceptable.

Table 5: Demand Parameters

Variable	Coefficient	t-value
Constant (d_0)	0.939	19.70
Population	1.094	61.39
Network Extent	0.109	7.65
Seat per Capita	0.767	14.57
Speed	-0.165	-2.85
Price (d_2)	-0.359	-3.621

The price elasticity is found to be equal to -0.359 resulting in a relatively inelastic demand to price variation. The magnitude of this coefficient is in the $[-0.2, -0.5]$ range as surveyed by Oum, Waters and Young (1992). The coefficient for population is positive and strongly significant. We can interpret this coefficient as follows. If we divide the demand equation by the population, we would find a coefficient equal to 0.094. When the population increases by 1%, the demand per capita would increase by about 0.09%. An increase in traffic congestion due to an increase in the population size would increase the demand for public transportation per capita, inducing more people to choose the public transit as their transportation mode. The extent of the network and the number of seats per capita have both significantly positive coefficients. The demand is more responsive to an increase in the offered capacity per capita than for an increase in the extent of the network. By increasing the extent of the network, the transportation authority probably offers bus service to remote areas or suburbs, where, given the distance involved, commuters tend to prefer to use their own car explaining this relatively low demand elasticity with respect to network extent. Regarding the speed variable, a negative value may be related to the nature of the city. A low average speed may indicate a compact city, where the demand for public transit is usually high. In contrast, a large average speed may indicate a sprawl city, where the demand for public transit is usually large. Such an argument could explain the obtained negative coefficient for the speed variable.

The second and third steps involve the estimation of the price and cost equations. These steps allows to recover important coefficients such as the parameters of the cost function β_y and β_L, the parameter for effort disutility α as well as the parameters η to recover the weights β_i. The following table provides the former estimates.[42]

Table 6: Cost and Effort Disutility Parameters

Variable	Coefficient
β_y	0.917
β_L	0.752
α	0.496

Detailed information on the price for labor and material inputs can be found in Perrigne (2002). The coefficient β_y is slightly smaller than 1 indicating slighlty decreasing returns to scale. Given that y measures the number of passengers and not the offered capacity, we could have expected a smaller value as an increase in 1% in the number of passengers is expected to increase the cost by a smaller amount than 1%.[43] The estimated coefficient β_L implies a coefficient for material price equal to 0.248 as homogenity of degree one in input prices has been imposed. The noticeable result of Table 6 is the parameter for the disutility of effort. This parameter has been estimated without imposing any constraint. Its positive value is in agreement with theory. The third step allows us to recover the error terms ϵ_c in the cost equation. Using the recovered error terms ϵ_d from the demand equation, we can test whether these error terms are independent. We

[42]Note that we do not provide any t-values for the estimated coefficients in Tables 6 and following. As discussed previously, the multi-step estimator is not the most efficient one and corrections for standard errors should be computed. An alternative method to compute standard errors would be to conduct bootstrap.

[43]A similar result is obtained in Perrigne (2002). This result can be explained by the strong correlation between the number of passengers and the capacity offered.

find a correlation coefficient equal to 0.1497 indicating a low correlation. We can think about weather conditions, which can affect both demand and cost.[44]

The second step also provides estimates for the parameters of the weight function.[45] In view of the reduced form analysis conducted in the data section, we have chosen three variables, which will determine the weight on the consumer surplus, namely a dummy for a leftist majority, the proprotion of environmentalists in the city council and the amount of debt per capita. The results are given in Table 7.[46]

Table 7: Parameters for the Weight Function

Variable	Coefficient
Constant Term (η_0)	0.9792
Leftist Majority	0.3439
Proportion of Environmentalists	1.2913
Debt per Capita	-0.0747

The constant term tells us that a city with a rightist majority, no environmentalist in its city council and no debt will consider a weight for the consumer surplus slightly less than one, which almost corresponds to the case of the benevolent utilitarian maximizer. If this rightist city experiences some debt, its weight value will decrease with the amount of debt per capita to attain values smaller than one. Similarly, a city with a leftist majority, no environmentalist and no debt will consider a weight equal to 1.3231. This weight will increase with environmentalists. These results are interesting as we could have expected more partisan views from politicians. For instance, the difference between a leftist adn a

[44] Since error terms capture ommitted variables, it is likely that these variables are not the same for the demand and cost.

[45] The second step allows us to recover the values of the coefficients η_0 and η_1 up to a multiplicative term, which is a function of the cost of public funds λ. The results displayed in Table 7 have been obtained by dividing the estimated coefficients by $1 + lam\hat{b}da$ with $\hat{\lambda}$ has been obtained in the final and fourth step of our estimation.

[46] We use the proportion of environementalists, i.e. taking a value between 0 and 1. For the debt per capita, we consider a value expressed in 10,000FF.

rightist majority is just equal to 0.3439, which may look still quite small, everything else being equal. Subsidizing public transportation can be viewed as redistributing revenue among the population. Though a rightist majority does not care as much, both parties care about redistribution and consumers when deciding the pricing for public transit. The coefficient for environmentalists confirms their successful influence on politicians when deciding the pricing of public transit as displayed in the reduced form analysis. The city financial constraint is a non negligible factor in determining the weight. The amount of debt per capita is on average 4,632FF. For a city with an average debt per capita, the weight will decrease by the amount of 0.0346 everything else being equal, suggesting a relatively low impact of financial constraints in their decisions of subsidizing public transit. The following table displayed some summary statistics for the weights.

Table 8: Summary Statistics for Consumer Surplus Weights

	Mean	STD	Minimum	Maximum	N
Weight (Full Sample)	1.1300	0.1825	0.8862	1.4216	514
Left with No Green	1.2903	0.0135	1.2448	1.3175	163
Left with Green	1.3507	0.0352	1.2790	1.4216	90
Right with No Green	0.9428	0.0180	0.8862	0.9734	202
Right with Green	0.9891	0.0278	0.9555	1.0451	58

The cities with a leftist majority tend to put more weight on the consumer surplus than the rightist cities. This tendance is accentuated with the presence of environmentalists. Though the values are relatively close to one, we observe some variability across cities with a maximum value equal to 1.4216, which is about 60% more than for the city with the minimum value at 0.8862. It would be interesting to conduct some counterfactual simulations. One of the major political change between the 1983 and the 1989 elections has been the access of environmentalists in city councils. The data show that 11 cities with a leftist majority experienced such a change.[47] Everything else being equal, the

[47]Note that 24 cities out of 57 did not experience any political change regarding the nature of the council majority and the presence of environmentalists in the city council.

estimation results of the model predict that these cities would decrease their price for public transit on average by 20.36% for an average value of debt per capita and for 6.2% of environmentalists in the city council at the new election.[48] We observe a decrease by less than 10% in the data. Note that our simulation considers everything else being equal while the quality of the public transit has improved over the period explaining this difference.

We have also simulated the impact of an increase by 1,000FF of the debt per capita. While considering a leftist city with no environmentalist and an average debt per capita, we find that the price for public transit would increase by 2.37% everything else being equal. If the city has (say) 5% of environmentalists in the city council, the increase would be only by 1.43%. Conducting a similar exercise for a city with a rightist majority and no environmentalist, the model predicts an increase by 15.04% of public transit price, while the price increase wouldbe only 7.51% with (say) 5% of environmentalists in the city council. Such a difference can be explained by the different weight on consumer surplus according to these different scenario.

The final step of our estimator provides an estimate for the cost of public funds λ, the constant term in the cost function β_0 and the parameters of firms' type density r and *gamma*. The results are given in the following table.

Table 9: Cost of Public Funds and Efficiency Parameters

Parameter	Estimate
Constant term (β_0)	-2.000
Cost of Public Funds	0.360
r	1
γ	1.130

We find that $r = 1$ gives the best adjustement for the final step of our estimator.[49] This

[48]We observe that these 11 cities had on average 6.2% of environementalists in their city council after the 1989 election.

[49]This value has been chosen as it provides the lowest value for the value function.

corresponds to a density with an important concentration of relatively good types. The estimated value of γ gives a expected value for firms' types equal to 0.885 and a variance equal to 0.783. This gives an expected value for $\theta - e$ of the order of 0.5, indicating that firms' labor is on average half efficient since we have assumed that firms' effort and type affect labor efficiency. When considering a simple example of an employee working on average 40 hours per week, his efficient labor is equivalent on average to only 20 hours of work.

The cost of public funds is slightly larger than the value estimated for developing countries. Given that public transit is subsidized through a tax to firms, we can expect that raising additional taxes to subsdize public transit is quite costly to the local economy. In this respect, a value larger than 0.3 is not really surprising.

to be completed

5 Conclusion

to be completed

References

to be completed