1. INTRODUCTION

It is commonly accepted that the outdoor mail delivery process exhibits strong scale economies, even if models differ on the precise measurement of this phenomena. Engineering models based on analytical analyses of the delivery process estimate scale factors in excess of 4 or, equivalently, to a scale elasticity smaller than 0.25, so that an increase of 1% of the quantity of mail would imply a cost variation of only 0.25% (see Roy 1999 and Bernard et al 2002). Econometric models have also provided estimates of significant scale factors (around 2 or 3) (see Cazals et al 1997 and 2001). The econometric models are based on the estimation of a relation between the cost and the quantity of mail. Such a relation is usually estimated using micro data measured at the level of delivery offices. Then the cost function should incorporate variables describing the heterogeneity between offices. This diversity is introduced through observable explanatory variables and, if the sampling scheme permits, using unobservable heterogeneity components. One of the strongest conclusions of ten years of estimation of delivery cost models is that the more precise the treatment of heterogeneity between delivery offices, the lower the estimated scale elasticity. In particular the use
of panel data and of individual unobservable effects leads to estimations of scale elasticity around 3 if the model also incorporates some observable heterogeneity components; the same argument applies when using a non-parametric specification, which allows very different cost structures between “small” and “large” offices.

Using a new data set describing the cost of outdoor delivery for French offices in 2001 we estimate a very simple constant elasticity model where the cost is modeled as a function of the total traffic and of two heterogeneity components, the surface\(^1\) and the population (measured by the 1999 census; hereafter referred to as population\(^99\)). As we note below, this model is also equivalent to representing delivery cost as a function of the traffic, the population density and the traffic per capita. Section 2 presents the main result of such an elementary econometric model: The elasticity of the cost with respect to the traffic is estimated to be 0.28. Furthermore, if the sample is sub-divided according to the traffic it appears that the elasticity is almost zero for the very small offices (costs are essentially fixed) and jumps to 0.7 for very large offices.

The models presented in section 2 are more “descriptive” than “structural” because they do not incorporate technology choices or design of delivery offices, which introduce selection bias or endogeneity bias in the estimation. In this paper we carry out two steps in the direction of the specification of a more structural model. Actually, we consider two (partial) phenomena.

Section 3 presents the first one. We show that two technologies may be used for the delivery of packets: The packets may be delivered jointly with the rest of the mail or by specific motorized rounds. We consider a model which incorporates this choice as an endogenous selection mechanism.

Secondly, the model of section 4 shows that the surface of the office is an important explanatory variable. However the strategy of the design of the offices is conducted in order to reduce the cost. The decomposition of the territory into delivery offices defines units of observation with a surface depending on unknown heterogeneity components of the costs. Then the surface should be treated as endogenous: This is the goal of the last section of this paper. The model we present is not fully satisfactory because the process of construction or regrouping offices is not explicitly modeled and in particular the number of offices\(^2\) is kept exogenous.

---

1 Surface: size of the geographical postal area covered by a delivery office (denoted for simplicity the surface in the paper).
2 The sample size is also kept exogenous.
2. AN ELEMENTARY COST MODEL OF OUTDOOR DELIVERY OFFICES

Econometric costs models are specified by a relation between the total cost of the production unit as a function of the level of the output, input prices and some other explanatory variables. We estimate the model using a single cross section. The input prices are then constant and their effect is not identified.

This model is based on the empirical analysis of elementary Cobb Douglas models estimated in this section by ordinary least squares method. These models have the form:

\[ \ln C = \alpha \ln T + \sum_{j=1}^{K} \beta_j \ln Z_j + \gamma + u \]  

where \( C \) is the total cost of outdoor delivery, \( T \) the total traffic and \( Z_1, ..., Z_K \) several cofactors describing the demographical/geographical structure of the offices. The parameter \( \gamma \) is a constant and \( u \) is a random unobservable component.

This type of model is, inevitably, a simplification and is certainly an imperfect summary of the reality. However it provides the best approximation of the real complex cost function by a constant elasticity function. The value of \( \alpha \) (or its inverse \( 1/\alpha \) defining the scale economy factor) represents the relations between traffic and cost.

Our data set comprises 4739 observations, each one corresponding to a delivery office in France; for each office we observe the total outdoor cost, the total number of delivered items, and several heterogeneity variables. A selection of most relevant explanatory variables based on usual t-statistics and on economic interpretation of the results suggests two leading observable heterogeneity components, namely the population and the surface. Let us recall that each office is attached to a well defined geographical area and its population was measured by the 1999 census. The surface is defined by the number of square meters of this area. In comparison with previous studies, we have less observations and this new data set

---

3 Total outdoor cost: mixed rounds (on foot +bicycle) charges + motorized mixed rounds charges +dedicated rounds charges
4 Some of the delivery offices consist of a main office and satellite offices (“échelon de distribution”). In previous studies these satellite offices were considered
concerns a unique year only. The introduction of an unobservable heterogeneity component is then impossible.

The model may be written in two equivalent ways

\[ \ln C = 0.28 \ln T + 0.48 \ln P + 0.19 \ln S + 0.64 \]  
\[ R^2 = 0.88 \]  
\( (26.3) \quad (40.8) \quad (50.3) \quad (5.9) \)  

Where \( C, T, S \) and \( P \) represent, respectively, cost, traffic, surface and population; the numbers in parentheses are the student-t values.

An equivalent form of the previous equation is:

\[ \ln \left( \frac{C}{T} \right) = -0.05 \ln T - 0.67 \ln P + 0.19 \ln S + 0.64 \]  
\[ R^2 = 0.88 \]  
\( (-7.8) \quad (-56.5) \quad (-50.3) \quad (5.9) \)  

In this relation the explanatory variable is the unitary cost \((C/T)\) defined by the ratio of the total cost divided by the quantity of mail.

The relations between the coefficients of the two models are obvious consequences of elementary computations. For example, the elasticity of the total cost to the traffic equal to 0.28 is decomposed in 1-0.05-0.67 in the second model: The elasticity of unitary cost to the traffic (with a traffic per capita constant) is -0.05; the elasticity to the traffic per capita (with a traffic constant) is 0.67. The difference between the cost elasticity and the unitary cost elasticity is one, by definition of the elasticity. The estimate of the elasticity of the total cost to the traffic of 0.28 corresponds to a scale factor greater than 3.5.

Model (2) shows that the main unitary cost drivers are the traffic per capita and the density (a smaller effect of the traffic). A change in the traffic has more of an impact on the cost than does the direct influence of the traffic per capita.5

In order to capture a more complex relation between cost and cost drivers, we have divided the sample into sub groups.

i) We have first stratified the sample into three groups of equal size function of the traffic per capita

---

5 Traffic per capita: traffic divided by population.

99.
The elasticity of the cost to the total traffic of the three groups is respectively 0.04, 0.36 and 0.58; elasticity increases with the traffic per capita.

ii) A second stratification of the sample is based on the level of the density.

Low \( \frac{P}{S} \) (<42 hab/km

\[
\ln \frac{C}{T} = -0.12 \ln T - 0.69 \frac{T}{P} - 0.29 \ln \frac{P}{S} + 1.38 \quad R^2 = 0.55
\]  

Intermediate \( \frac{P}{S} \) (42 118 hab/Km

\[
\ln \frac{C}{T} = 0.008 \ln T - 0.89 \frac{T}{P} - 0.26 \ln \frac{P}{S} + 0.48 \quad R^2 = 0.7
\]

High \( \frac{P}{S} \) (>118 hab/Km

\[
\ln \frac{C}{T} = 0.001 \ln T - 0.69 \frac{T}{P} - 0.098 \ln \frac{P}{S} + 0.7 \quad R^2 = 0.63
\]  

The elasticity of the cost to the total traffic of the three groups is 0.19, 0.03 and 0.3. This U shape of the elasticity function of the density is one of the main results of this empirical study. In intermediate density, the cost is
almost completely independent from a small variation in the traffic. In low or high density regions the cost reacts to a variation of traffic and this variation is more important in the high density areas.

Finally the joint effect of stratification by traffic per capita and by density is analyzed by the estimation of nine models. The empirical evidence is summarized by graph (figure 1) representing total cost elasticity to traffic or to unit cost to traffic per capita.

The U shape of the relation between elasticity and density is preserved and this curve is shifted up when the traffic per capita increases.

The graph also gives the relative mean unit cost by group of delivery offices. The higher elasticities correspond to the lower mean unit cost but the mean unit cost increases when the density or the traffic per capita increases (numbers in brackets are mean unit cost)

The mean unit cost (where, by convention, the mean unit cost of high traffic and high density offices is 1) is computed from the data and not from the estimated model.

**Table 1: Mean unit costs**

<table>
<thead>
<tr>
<th>Density</th>
<th>Traffic per Capita</th>
<th>Low</th>
<th>Intermediate</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td></td>
<td>4.25  (7.25)</td>
<td>3.125 (10.6)</td>
<td>3 (13)</td>
</tr>
<tr>
<td>Intermediate</td>
<td></td>
<td>3 (1)</td>
<td>1.75 (0.37)</td>
<td>1.25 (0.37)</td>
</tr>
<tr>
<td>High</td>
<td></td>
<td>2.5 (1)</td>
<td>1.375 (0.37)</td>
<td>1 (0.25)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.25 (4.25)</td>
<td>2.25 (7.125)</td>
<td>1.5 (6.37)</td>
</tr>
</tbody>
</table>

*(Numbers in brackets are the standard deviation)*
These different values of cost elasticities and the particular shape of the function need an explanation, drawn from the economics of delivery, as described in Cohen and Chu (1997), Roy (1999), and Bernard et al (2002).

To make things clearer, we can briefly recall that in the delivery process, the fixed costs are accounted for by the need for the carrier to move from one stop to another, whereas the variable costs consist in loading the mailboxes. The stop time needs particular attention. Indeed, the delivery model is made more complex by the fact that the proportion between fixed and marginal costs is not constant. According to the level of traffic, the stop time can be either considered as fixed or variable. For very low levels of traffic, a single letter requires its own stop. But for very high levels of traffic, a letter just adds to an existing stop as described in Jasinski and Steggles (1977) through
a Poisson process. Therefore, the marginal cost (i.e., the derivative of the cost function) is not constant and depends on the level of traffic.

These effects exist in the econometrics presented here, while we examine the range of elasticities obtained over the different values of traffic. As the level of traffic increases, the scale economies are exhausted, because the relative weight of loading time increases whereas the route time and stop time remain fixed. The average cost goes down as the marginal cost has become constant. Therefore, the value of cost elasticity with respect to traffic increases, although the value of average cost decreases. This explains why costs elasticity obtained with higher levels of traffic are above the ones obtained with lower levels.

The singular result is the U shape, which leads to an analysis of the cost elasticities with respect to traffic over the different values of density. For a very high value of density, it is easy to understand that the relative amount of fixed costs is lower than in medium density: The distance between two stops is lower, and the stops are shared with more addresses.

The quite interesting case occurs for very low levels of density. In these areas, we should expect the value of cost elasticity to be even lower than in medium density, since the distance between two stops is even greater. But if we look at it closer, one stop may correspond to one addressee only.

For an equal level of traffic, the probability that an extra letter generates an extra stop is higher in low density areas than in medium density areas, in which one stop can serve several addresses. So, as described above, the stop-part of the costs can be considered more as a variable cost in very rural areas.

It also can be mentioned that in very low density areas, an extra letter can lead to a detour (increasing the route costs), or conversely a decrease of volumes may allow short cuts in the delivery process, because some parts of the route can be avoided. This is very accurate when the levels of traffic are quite low. Again, this explains why we can find higher cost elasticities in areas with very low density.

3. AN ENDOGENOUS SELECTION MODEL FOR OPTIMAL DELIVERY OF PACKETS

As described in Roy (1999) and Bernard et al (2002), in France, the optimal mode for delivering letter mail is often by foot or bicycle. But the existence of parcels renders quite complex the choice of the mode of delivery in these
areas. If we assume that it is impossible to deliver parcels in something other than a van (ignoring the special cases of very small or very few parcels), then it is quite difficult to determine whether routes must be dedicated for parcels or integrated with letter mail.

In the case where the optimal mode for delivery for letter mail is the van, the question is straightforward. In the case where the optimal mode for delivery is foot or bicycle, the trade-off is the following: Is the loss made by a van constraint (sub-optimal) lower than the gains made by sharing the fixed costs of the route?

In very high-density areas, using a van for delivering letter mail can be significantly more expensive. As the distance between two stops becomes smaller, the speed of traveling becomes very slow, perhaps even slower than a pedestrian. This is amplified by the necessity of parking the van at each stop (or so), which creates prohibitive extra costs. For a dedicated parcel delivery, however, the stop coverage is quite low, so the van does not have to stop at every point. It is quite natural therefore to organize dedicated routes in very dense areas.

So the key question is that of the “grey zone”: The zones with medium density, where the problem of organizing dedicated routes or not becomes relevant. In La Poste, an engineering model (taking its roots in Roy 1999), has assisted the decision making process. It is interesting to study now ex post how econometrics can treat the results.

Two categories of delivery offices appear in the sample:

- in most of the offices (3937 of 4739\textsuperscript{6}) packets are delivered jointly with the other mail (\(D = 0\))
- in 802 offices (17%) packets are delivered by specific motorized rounds (\(D = 1\))

This decision is considered endogenous and the model incorporates this mechanism by an endogenous switching model (see Gourieroux 2000), which has the following structure:

- the counterfactual model: Each delivery office has two cost functions depending on the choice between \(D\) equal to either 0 or 1
- the assignment equation which explains how the decision between 0 and 1 is taken.

\textsuperscript{6} Due to a few missing data, the sample size has been modified a little depending on the variables.
The model assumes that for any office only one of the two regimes is observed and that the decision rule is not independent of the potential outcome, even conditionally on the explanatory variables.

The estimation is conducted using Heckman’s two-step method and gives the following results:

Cost function without dedicated rounds:

\[
\ln C = 0.18 \ln T + 0.43 \ln P + 0.23 \ln S + 1.66
\]

(9) (21.4) (39.0) (8.2)

or equivalently

\[
\ln \frac{C}{T} = -0.16 \ln T - 0.66 \ln \frac{T}{P} - 0.23 \ln \frac{P}{S} + 1.66
\]

(10)

Cost function with dedicated rounds

\[
\ln C = 0.39 \ln T + 0.91 \ln P + 0.11 \ln S - 4.46
\]

(7.0) (10.2) (11.4) (-2.9)

or equivalently

\[
\ln \frac{C}{T} = 0.41 \ln T - 1.02 \ln \frac{T}{P} - 0.11 \ln \frac{P}{S} - 4.46
\]

(12)

Decision equation (Probit model):

\[
\text{Prob}(D = 1) = \Phi \left( 0.36 \ln T + 0.51 \ln P + 0.18 \ln \text{Revp} - 13.4 \right)
\]

(5.4) (3.9) (1.9) (-25.2)

where \( \Phi \) is the cumulative function of the normal distribution and \( \text{Revp} \) is the average household revenue.

The two lambda coefficients of the Mills Ratio are significant and confirm that the selection between the two regimes is not merely random.
As the decision whether or not to use dedicated rounds is highly correlated to the size of the office, it is not surprising to verify that the size of the elasticity is smaller for offices without dedicated rounds.

In the linear case under consideration, identification of the model with an endogenous selection rule does not require specific instrumental variables in the selection equation. However we have selected the number of firms as an instrument in order to increase the robustness of the procedure.

An interesting validity test of this model is to compare actual decisions and optimal ones. Indeed for each office the models (9) and (11) permit a computation of the (latent) costs in the two regimes given T, P and S. We can then compute the difference\(^7\) \(\Delta\) and the office is classified as either one, which should use dedicated rounds or one which should deliver all mail jointly. The results are in the following table:

### Table 2: Actual decisions and optimal ones

<table>
<thead>
<tr>
<th>Number of offices which should</th>
<th>use dedicated rounds</th>
<th>not use dedicated rounds</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Used dedicated rounds</strong></td>
<td>488</td>
<td>314</td>
</tr>
<tr>
<td><strong>Did not use dedicated rounds</strong></td>
<td>168</td>
<td>3 769</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>656</td>
<td>4 083</td>
</tr>
</tbody>
</table>

The number of mismatches between prediction and reality is 482 (10% of the offices).

4. ENDOGENEITY OF DELIVERY OFFICES’ DESIGN

Let us consider a territory with a particular density of inhabitants and mail distribution per capita. The postal operator will divide this territory into \(n\) units with a particular spatial shape. This division minimizes the total delivery cost. In such a model the number and the shape (in particular the surface area) of the delivery offices become endogenous. Indeed the cost structure depends on local heterogeneity factors known by the operator and then incorporated into the optimal partition of the territory.

\(^7\) \(\Delta\) is the exact average treatment effect of the use of dedicated rounds.
Even if in reality, the division of France into delivery offices is largely the consequence of historical considerations, it is certainly also partially based on cost minimization considerations. It would be difficult to write a theoretical model of optimal division and to estimate such a fully structural specification. However we want to consider the surface of each delivery office as an endogenous variable in the model and to estimate the cost function using an instrumental variables method. This approach is an approximation of the implicit structural model and, in particular, the number of observations is treated as exogenous.

Then we just consider the same equation as equation (1) (in section 2) but in the case in which the surface is endogenous and the estimation is done by instrumental variables (or 2SLS in that case). The obtained estimation is:

\[
\ln C = 0.33 \ln T + 0.44 \ln P + 0.26 \ln S - 0.97
\]

or equivalently

\[
\ln \frac{C}{T} = 0.03 \ln T - 0.7 \ln P - 0.26 \ln S - 0.97
\]

A test of endogeneity has been performed. This test estimates the previous equation by OLS but introduces the estimated residuals of the instrumental equation. Endogeneity is then equivalent to the significance of the coefficient of the residuals and is not rejected in our case.

Despite this evidence in favor of endogeneity the model does not provide dramatically different results from our previous ones: In particular, the scale factor for the whole sample is around 3. The increment of the elasticity of traffic between model (1) and model (6) depends upon the surface where

---

8 This problem is strongly related to the problem of observing the integral of a continuous time process between random times. This question has been considered recently in finance. In the case of delivery offices, time processes become spatial processes and the question becomes extremely complex. To our knowledge no econometric model is based on this kind of computation.

9 Instruments selected are the household revenue index, the number of firms and the average mailboxes by block of flats.

10 See section 2.
the traffic changes. The model estimated in this section shows that if the traffic increases, the surface will decrease.

If we allow a variation of the surface to a modification of traffic, the elasticity becomes:

$$\frac{\partial \ln C}{\partial \ln T} = 0.33 + 0.26 \frac{\partial \ln S}{\partial \ln T}$$

$$= 0.33 + 0.26(-0.40) = 0.226$$

Finally, it should be noted that it is valid to consider the surface of the density as endogenous, given the hypothesis of the exogeneity of the population.

5. CONCLUDING REMARKS

This paper presents several parametric log linear models for the outdoor delivery cost functions based on a survey of all the French delivery offices in 2001. The main empirical results derived from this set of estimations are the following:

The best approximation of the model by a constant elasticity model gives an elasticity of the cost to the traffic equal to 0.28, confirming the existence of a strong scale economy in the delivery process. A relevant selection of cost drivers for the unit cost of a delivery office uses the traffic per capita, the density and, at a lower level, the total traffic. The total cost elasticity is actually not constant: It increases with the traffic per capita and is a U shape function of the density.

This U shape, highlighted econometrically for the first time, is one of the main contributions of this paper: It is very well explained by the Economics of delivery and it reconciles different anterior results upon costs elasticities.

The design of the delivery process should be treated endogenously. In most of the previous models of the cost of delivery, the design of the process, i.e. the division of the territory into delivery area and the organization of rounds, is considered as given. We have shown that it is useful to treat it as an endogenous component and to explain this by the model. In this paper the design was catch by the use of rounds not dedicated and by the surface of area of each office. In future research we plan to consider jointly endogeneity of the design of the delivery offices and non constant elasticity models.
APPENDIX: The Data

Definition of the variables

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>4764</th>
</tr>
</thead>
<tbody>
<tr>
<td>Metropolitan French Post Delivery Offices (2001)</td>
<td></td>
</tr>
<tr>
<td>25 Delivery Offices with missing data</td>
<td></td>
</tr>
</tbody>
</table>

| Lcost (Cost) | Mixed rounds (on foot+bicycle) charges + motorized mixed rounds charges + motorized dedicated rounds charges |
| LTraffic (Traffic) | Mixed rounds (on foot+bicycle) traffic + motorized mixed rounds traffic + motorized dedicated rounds traffic |
| Lpop99 (Population 99) | INSEE\(^{11}\) definition |
| Lsurface (Surface) | Surface covered by each Delivery Office |
| Lnb_etab (Number of firms) | INSEE definition |
| Ltraf_tete (Unit cost) | Traffic per capita |
| Ldensity (population density) | Population99 divided by surface |
| LIRE (average mailboxes index) | Number of mailboxes divided by block of flats |
| Lrev_moy_men (average Households revenue) | = Number of Households x Households Delivery Offices revenue index\(^{12}\) |

---

\(^{11}\) INSEE: Institut National de la Statistique et des Etudes Economiques (French census bureau).

\(^{12}\) DGI: Direction Générale des Impôts, France (French Taxes Administration).
REFERENCES:


