A Smooth Gravity Model for Origin-Destination Flows: Estimating How Local Taxes Impinge upon the Relocation of French Establishments

Autores y e-mail de la persona de contacto:
M. Manjón-Antolín[a], O. Martínez-Ibañez[a] y J.M. Arauzo-Carod[a,b]

josepmaria.arauzo@urv.cat

Departamento:
[a]QURE-CREIP Department of Economics, Universitat Rovira i Virgili
[b]Institut d'Economia de Barcelona, Universitat de Barcelona

Universidad:
Universitat Rovira i Virgili

Área Temática: Localization, clusters and industrial districts

Resumen:
We propose a new spatial gravity model for origin-destination flows that is constructed under mild distributional assumptions and uses a continuous distance-based definition of neighbourhood. We illustrate the main features of our model by estimating the effect of corporate taxes on the number of relocated establishments between French “Zones d'Emploi" (travel-to-work areas). Results indicate that, once we control for taxable agglomeration rents, local taxes deter relocations. We also find weak evidence that larger relocating establishments tend to be more affected by taxes than the smaller ones. Lastly, estimates of the pushing and pulling effects, measured by the ratio of the indirect-to-direct marginal effects, may extend up to 150 around the origin and 300 km. around the destination, respectively. For larger firms, however, these effects barely reach a hundred km.

Keywords:

Palabras Clave: gravity models, relocation, count data

Clasificación JEL: C21, C52, R12

A Smooth Gravity Model for Origin-Destination Flows: Estimating How Local Taxes Impinge upon the Relocation of French Establishments*

M. Manjón-Antolín[a], O. Martínez-Ibañez[a] and J.M. Arauzo-Carod[a,b]

[a]QURE-CREIP Department of Economics, Rovira i Virgili University

[b]Institut d'Economia de Barcelona, University of Barcelona

Abstract

We propose a new spatial gravity model for origin-destination flows that is constructed under mild distributional assumptions and uses a continuous distance-based definition of neighbourhood. We illustrate the main features of our model by estimating the effect of corporate taxes on the number of relocated establishments between French "Zones d'Emploi" (travel-to-work areas). Results indicate that, once we control for taxable agglomeration rents, local taxes deter relocations. We also find weak evidence that larger relocating establishments tend to be more affected by taxes than the smaller ones. Lastly, estimates of the pushing and pulling effects, measured by the ratio of the indirect-to-direct marginal effects, may extend up to 150 around the origin and 300 km. around the destination, respectively. For larger firms, however, these effects barely reach a hundred km.

Keywords: count data, gravity models, relocation, spatial interactions.

JEL Classification: C21, R3, R12.

*This research was partially funded by SEJ2010-19088/ECON, SEJ2010-16934/ECON, the "Xarxa de Referència d'R+D+I en Economia i Polítiques Públiques"and the SGR Program (2009-SGR-322) of the Catalan Government. We thank Magda Lleixà for excellent research assistance. The usual caveats apply.

1 Introduction

Quite often, we hear on the news that a multinational has relocated a plant to a another country. However, the fact that a substantial number of establishments and firms relocate their premises to a nearby location within the same country or region goes largely unnoticed by the media. According to Lee (2008), for example, in the U.S. manufacturing industry and over the period 1972-1992, "[f]or every 100 new entrants starting operation over a five-year period, more than 10 plants turn out to be relocated from other counties". Also, van Dijk and Pellenbarg (2000) report that the number of Dutch migrant firms has raised from $36,000$ to $68,000$ between 1987 and 1995. Similarly, the relocation of establishments in France have practically doubled between 2004 and 2010, increasing from $72,770$ to $137,859$ according to the *Institut nacional de la statistique et des études économiques* (INSEE). In particular, out of the approximately $400,000$ relocations accounted by the SIRENE database over the period 2008 to 2010, around one fourth were migrations to another travel-to-work area (*Zones d'Emploi*, ZdEs).

These figures contrast with the relatively small number of empirical studies on this issue, particularly when compared to the extensive literature on the determinants of new business creation —see e.g. Arauzo-Carod *et al.* (2010) for an overview. Furthermore, previous studies on the determinants of the relocation flows do not account for both origin and destination determinants (see, e.g., Holl 2004, Mariotti 2005 and Manjón-Antolín and Arauzo-Carod 2011). Rather, they analyse the determinants of the number of relocations to a set of destinations (regardless of the origin, which is unknown) using the destinations' features as the main explanatory variables. It is also interesting to note that these papers implicitly assume that origin-destination (OD) flows are spatially independent, a potential error of specification that may result in biased estimates of the model coefficients (Griffith 2007).[1]

We propose a new approach for the analysis of spatially-dependent OD flows when both

[1] Referring to those studies that analyse individual relocation decisions (using e.g. binary outcomes as the variable of interest), Kronenberg (2013) observes that "while numerous studies acknowledge the relevance of both the characteristics of a firm's region of origin (which might push a firm from its present location, or keep it there), and the qualities of a firm's region of destination (which might pull a firm towards this location), most analyses either focus on reasons underlying the outmigration of firms, or concentrate on regional features which attract relocating businesses". To our knowledge, none of these studies on individual relocation decisions account for spatial dependence.

origin and destination characteristics are available. In particular, we follow LeSage and Page (2008) in modelling the relocation process using a spatial gravity equation for the OD flows. We differ from them in that we use count-data rather than log-linear specifications (see also Lambert *et al.* 2010). In this respect, our model is closer to that proposed by LeSage *et al.* (2007). We also differ from these related works in the way we account for spatial dependence. Rather than using a threshold to define neighbours (e.g., contiguity or an *ad-hoc* distance), we construct the spatial weighting matrices using a smooth function of the distance between the geographical units involved (Martínez *et al.* 2013). Moreover, we do not consider a spatial autoregressive specification —as e.g. LeSage *et al.* (2007) and LeSage and Page (2008) do– but apply the spatial weighting only to the covariates —as e.g. Lambert *et al.* (2010) do.

We illustrate the main features of our model using census French data on the number of relocated establishments between the 341 mainland ("Métropole") ZdEs that existed in 2008.[2] More precisely, the SIRENE database provides information on the number of relocations that occurred within each ZdE as well as on the number of relocations that occurred between ZdEs. The question that arises then is whether the driving forces behind the flows of "within-ZdEs relocations" are the same as those behind the flows of "between-ZdEs relocations" (as e.g. the model of LeSage *et al.* 2007 assumes; see, in contrast, the evidence provided by Weterings and Knoben 2013). LeSage and Pace (2008), for example, propose using a different function for the conditional expectation of each flow. Our model uses a less restrictive assumption, namely that the conditional distribution of the "within-ZdEs relocations" may be different from that of the "between-ZdEs relocations".

As for the vector of explanatory variables, the available data includes centroid distances, corporate taxes (the "taxe professionnelle"), proxies for agglomeration economies, local labour market conditions, institutional features, transport infrastructures and educa-

[2]We use *Zones d'Emploi* as the unit of analysis (rather than *communes* or municipalities, as e.g. Houdebine and Schneider 1997 and Rathelot and Sillard 2008 do) for two reasons. First, ZdEs are defined by economic (rather than administrative) criteria. Namely, they are constructed using commuting data showing what is the geographic area where most of the active population lives. Thus, they roughly correspond to French local labour markets, which in principle makes them close to the relocating area that firms may actually consider. Second, the number of ZdEs in France makes the sample size large but still manageable, for there were 348 ZdEs in 2008 (which results in 348×348 observations, while, had we analysed municipalities, the number of observations would have been $36,695 \times 36,695$). However, we dropped the ZdEs from the region of *Corse* because of the specificities of these territories (there were for example no relocations from mainland France to these ZdEs and there are fiscal advantages in the corporate taxes that do not apply in mainland ZdEs).

tion. This specification raises a problem of identification though, since it is well known that taxes and proxies for agglomeration economies are positively correlated (Jofré-Monseny 2013, Koh *et al.* 2013, Luthi and Schmidheiny 2013). To deal with the existence of "taxable agglomeration rents" (Charlot and Paty 2007), we use total taxes to control for non-observed agglomeration economies and destination-origin differences in local taxes to assess the impact of taxes on the relocation flows. This limits the interpretation of our results, for although we can provide evidence on whether taxes deter the relocation of establishments, we cannot provide an estimate of the impact of taxes on the number of relocated establishments (*ceteris paribus*). On the other hand, this approach provides consistent estimates of the model coefficients under the assumption of independence between destination-origin differences in local taxes and the unexplained component of the agglomeration forces.[3]

We report coefficient estimates both to motivate this empirical strategy and to asses whether taxes impinge upon the relocation flows. In particular, we provide results for OD flows of establishments of different sizes (defined by intervals on the total number of employees) to asses the differential effects of taxes across the size distribution of the establishments (Charney 1983, Baldwin and Okubo 2009). In order to analyse the geographical scope of the whole set of covariates, however, we use the ratio of the indirect-to-direct marginal effects (see e.g. LeSage and Pace 2009 for definitions). Plots of this ratio versus distance nicely summarise how far can reach the effect of a unitary change in the covariates and how is the shape of this effect with respect to the distance reached (Martínez *et al.* 2013).

The rest of the paper is organised as follows. Section 2 briefly reviews the relevant literature. In Section 3 we present the model and in Section 4 the main empirical results. Section 5 concludes.

[3]The assumption made in the IV approaches followed by e.g. Rathelot and Sillard (2008) and Brülhart *et al.* (2012) is the independence between the instruments of the taxes and the unexplained component of the agglomeration forces (i.e., the exogeneity of instruments).

2 Related literature

2.1 Relocation models

Location behaviour of new firms has been widely analysed from a empirical point of view in recent years (see Arauzo-Carod et al., 2010, for a detailed review of this literature) but knowledge regarding firm relocation (a specific case inside location) is so scarce, mainly because the lack of data for firm migration behaviour. In fact, usually, data about location puts together strictly new and relocated firms, so there is no distinction between both (different) phenomena. But recently some datasets have started to distinguish between them and, consequently, to make separate analysis becomes now possible.

This lack of analysis implies also that there are no specific theories regarding location decisions of relocated firms and, as Brouwer et al. (2004, pp. 336) point out "[r]elocation theories are hardly applied and are often treated as a special case of location theories". But they are more than this, since characteristics of new and relocated firms differ in one essential feature, the availability of information about markets and institutions, which could be reasonably assumed to be larger for relocated firms. These information asymmetries imply also that location preferences could not be the same since decisions are made upon available information. Accordingly, location and relocation are not exactly the same phenomenon and must be analysed separately or, at least, in a way allowing to take into consideration such particularities (Lee, 2006; Holl, 2004; Pellenbarg et al. 2002a, 2002b).

Whilst most of contributions regarding relocation behaviour of firms have followed mainly a descriptive approach until the late 90s, as Mariotti (2005) reviews in detail and others also point to their dislocating effects for firms and workers (Carter, 1999), recent contributions try to modelise relocation decisions and to analyse which are the main determinants of such firm migrations as well as to take into account individual characteristics of firms, as Stam (2007, p. 46) points out when discussing influence of life-cycle: "Firms that decide to move out of their home region in the early phases of their existence do so for different reasons than do those that decide to do so at a later stage. The entrepreneur's personal relationships, for example, become less important as time goes by". It is worth noting that we will focus on these contributions that identify several variables that are hypothesized to influence relocation decisions of firms (Lee, 2006; Brouwer at el., 2004), like age (firm's age decreases with firm's age), size (firm's mobility decreases with firm size), performance (firm's growth increases firm's mobility) and market (firm's mobility increases with market size).

2.2 Empirical results

Van Dijk and Pellenbarg (2000) use firm data from a sample of 1,300 firms across the Netherlands and assume that firm behaviour is guided by profit maximisation strategies and that this applies for every dimension of firm activities, including decisions regarding location. Accordingly, if firms consider that profits could be higher in another site they could decide to relocate there. But decisions regarding whether to stay or move will depend on push (reasons to leave present location), pull (reasons to move to another location) and keep (reasons to stay at the same location) factors (van Dijk and Pellenbarg, 2000).

Empirical analysis of relocation determinants have mainly identified the following factors as the most important according to previous push-pull-keep pattern: lack of space, accessibility and labour market. Lee (2008) analysis relocation of manufacturing plants across the U.S. and concludes that incidence of incentive programs (tax and financial incentives) over relocation decisions is so small, that role of wage differentials is small and diminishes over time . Regarding role played by plant level characteristics, Lee (2008) shows that probability of relocating is negatively associated with age, size, capital intensity, lack of specialization, lower productivity levels and lower labour skills. At a state level, Lee (2008) also demonstrates that higher union membership rates imply lower probabilities of relocations. Duranton and Puga (2001) use the same dataset regarding French relocations (SIRENE) that in this paper and approach this phenomenon in terms of diversification vs. specialization of French employment areas (Zones d'Emploi). Concretely, they demonstrated that for 1993-1996 most of firms' departures (72.0%) were coming from specialised municipalities and were going to specialised municipalities. Their point is that, at the beginning, new firms need to be located in a diversified area in order to maximise learning process from a large set of different industries, but once firm decides a production technology and gets an standardized product, their need to be located inside such diversified and innovation oriented area decreases and, cost saving strategies appear. At this moment it is time to move to a specialised (usually smaller) area, where costs are lower and there is less industry diversity. And for the specific case of headquarters relocation, Strauss-Kahn and Vives (2009) use data from the U.S. and conclude that headquarters tend to move to bigger metropolitan areas with an important concentration of economic activities business services and airport facilities. In a similar centre-periphery pattern, Baudewyns et al. (2000) analyse the effect that better public infrastructures have on the (de)location decisions of Belgian firms from the city of Brussels and the region of Wallonia.

Even if acknowledging the importance of previous determinants over relocation decisions, it is hard to assume that they influence homogeneously several types of relocated firms. In this sense, a key distinction relies on distance covered by relocated firm, as one could argue that shorter distance relocations are not influenced in the same way or/and by the same determinants that those travelling longer distances (Weterings and Knoben, 2013). Concretely, Weterings and Knoben (2013) assume that long-distance and short-distance relocations drivers differ, being that short-distance relocations are mainly triggered by internal characteristics of the firms as lack of space due to firm's growth. As this requirement has nothing to do with external environment, it is more likely that these firms will try to look for alternative sites close to their current locations. On the contrary, long distance relocations are more likely to occur due to differences in opportunities related to markets (Weterings and Knoben, 2013; Stam, 2007), so external characteristics as level of specialization, R&D intensity or agglomeration economies matter.

2.3 Evidence from France

French case has received little attention regarding determinants of relocated concerns and most of contributions about location focus on entry determinants of new plants. Among those contributions, there are those of Rocha (2008), who carries out a very spatial detailed analysis (zip-code level) for the determinants of firm entry. In this paper zip-code level fixed effects are used and it is demonstrated that agglomeration economies really matter for firm location decisions ; Rathelot and Sillard (2008), who focus on a detailed issue like local taxes and show that local taxation clearly matters for firms' location decision; Coris (2008), who analyses software industry and how its specificities in terms of geographical proximity among agents at different stages of production process allow firms to delocalise their activities; Huiban et al. (2006), who uses the same dataset that this paper (SIRENE), check whether entries, exits and relocations depend on types of spatial areas where manufacturing firms operate (urban, suburban and rural) and, concretely, they found higher mobility rates at urban and suburban areas; in a similar paper Huiban et al. (2002) analyse firm's relocation in food industry activities and show that between 1993 and 1998 this phenomena affects about 2,17% of existing establishments. Additionally, they show that about 90% of relocations take place inside the same type of areas (namely urban or rural) and that most of relocations that imply changes in such areas consist of urban to rural movements; Autant-Bernard et al. (2008, 2006) and Autant-Bernard and Massard (2007), who concentrates on location patterns of high-tech firms and obtain results that highlight role played by extant technological infrastructures in attracting such firms; and Crozet et al. (2004), who analyse location patterns of FDI firms at a province level and conclude that agglomerations economies matter but that there are also important differences depending on each industry. Apart from previous aggregated analysis, there are other papers focusing on location dynamics inside urban areas, trying to underline whether different activities show differences in their location patterns in terms of CBD vs. peripheral manufacturing areas or even how different activities coexist at the same places. Among papers covering such issues there are those of Baumont et al. (2004) for the metropolitan area of Dijon (Communaute de l'Aggomeration Dijonnaise) and Boiteux-Orain and Guillain (2004) for 'le-de-France.

3 The model

We seek to empirically analyse the determinants of flows between geographical units. In particular, we are interested in the determinants of the number of establishments that move from location $i = 1, 2, ..., n$ to location $j = 1, 2, ..., n$, which we denote by y_{ij}. Thus, the number of observations is $n \times n = N$ and the variable of interest, y_{ij}, contains information not only about the flows occurring between different geographical units ($i \neq j$) but also about the flows occurring within the same geographical unit ($i = j$).[4]

We propose using a spatially weighted gravity model to analyse these data (see LeSage and Pace 2009). This means that our vector of explanatory variables includes characteristics of the geographical units from which the flows originate ($\mathbf{x}_i$), characteristics of the geographical units to which the flows are destined ($\mathbf{x}_j$), and the (centroid) distances between origins and destinations (d_{ij}). Also, (some of) these origin and destination covariates may be spatially weighted to account for the spatial dependence between locations.

This model specification largely follows the work of LeSage *et al.* (2007) and LeSage and Pace (2008), the former concentrating on the case in which the variable of interest is a count and the latter on the case in which (a transformation of) the variable of interest is normally-distributed. Our proposal differs from them, however, in two important features. First, the distributional assumptions we impose on the variable of interest. Second, the definition of the spatially weighting matrices and the way we account for the spatial dependence. Next we discuss these features in more detail.

3.1 Distributional assumptions

To construct the econometric model, one may consider that the behaviour of the variable of interest is the same regardless of whether the flows occur within the same location or between different locations. This is the assumption implicitly made by e.g. LeSage *et al.* (2007) in their analysis of the patent citations made by high-tech firms located in European regions. In our empirical application, this amounts to assume that the behaviour of the intra-ZdEs flows of establishments does not differ from that of the inter-ZdEs flows of establishments. However, empirical evidence suggests otherwise.

Figure 1 shows that while the number of relocated establishments between ZdEs is often zero, the number of establishments relocated in the same ZdE is always strictly positive. Since

[4]This is an important difference with respect to previous studies on the relocation of firms and establishments, such as e.g. Holl (2004) and Manjón-Antolín and Arauzo-Carod (2011). Since they do not have information about the origin of the flows, their dependent variable is the number of establishments that move to location $j = 1, 2, ..., n$ and their number of observations is consequently n.

we are dealing with count data, this typically means that a different econometric model is needed (Cameron and Trivedi 2013). Moreover, as Weterings and Knoben (2013) argue, short-distance relocations (essentially, within ZdEs) are mostly driven by growth factors, whereas long-distance relocations (essentially, between ZdEs) tend to be more related with potentially better opportunities in the new site. Thus, regardless of whether one uses the same econometric model to analyse inter- and intra-ZdEs flows, its covariates are likely to differ.[5]

[INSERT FIGURE 1 AROUND HERE]

If one is not willing to impose the assumption of common behaviour across geographical units, then "a separate model for flows from the main diagonal of the flow matrix" is needed (LeSage and Pace 2008: 960). In their analysis of the population migration flows in the US states, for example, LeSage and Pace (2008) use a different function for the conditional expectation of the variable of interest. Yet they assume the same distribution for both intraregional and interregional flows. In our empirical application, the "excess of zeros" we observe in the inter-ZdEs flows of establishments questions the validity of this assumption (see Figure 1). This lead us to consider a more general assumption, namely that the conditional distribution of the flows when $i \neq j$ may be different from that when $i = j$. In maths:

$$
\begin{aligned}
y_{ii} &\sim F_1(\mathbf{x}, \mathbf{d}, \theta_1) \\
y_{ij} &\sim F_2(\mathbf{x}, \mathbf{d}, \theta_2) \qquad \text{for } i \neq j
\end{aligned}
$$

where F_1 and F_2 are appropriate distributions functions (e.g. F_1 is Poisson and F_2 is an Inflated Poisson), $\mathbf{x}$ is an $n \times K$ matrix of covariates, $\mathbf{d}$ is the $n \times n$ matrix of distances between locations, and $\theta = (\theta_1, \theta_2)$ is a vector of parameters to be estimated. Notice that, depending on the values of θ, this parameterisation still allows for a common behaviour across geographical units as well as for different determinants for inter- and intra-ZdEs flows.

[5] Most of the within-ZdEs flows can be considered short-distance relocations and, by the same token, most of the between-ZdEs flows can be considered long-distance relocations. However, since the size of a ZdE varies widely (from 45 km^2 to 6,267 km^2, being the mean about 1,578 km^2), this is not always the case. In the Paris region, for example, one can easily find several ZdEs within a radius of 20 km. In contrast, moving 20 km from the centroid one of the large ZdEs of the South and Central France will still leave you in the same ZdE.

3.2 Spatial weighting matrices

Flow data is likely to show spatial dependence, i.e., the values of the variable of interest in a particular origin-destination are likely to be correlated with those of nearby origins and/or destinations (Griffith 2007). This is apparent in Figure 2, which plots the spatial distribution of the intra- and inter-ZdEs flows in France. Paris stands as a powerful network of spatial interactions and, in general, the larger flows are concentrated in the more dense areas (e.g. Lyon, Toulouse, Nanterre, Marseille-Aubagne and Bordeaux-Zone-Centrale). On the other hand, one observes much less activity in the ZdEs of Central France. To control for this geographical correlation, one may include among the covariates (a function of) the distance between origins and destinations. However, this may not suffice to address spatial dependence (Curry 1972). Distance may be indeed a relevant explanatory variable, but some spatial weighting is needed (Griffith and Jones 1980).

[INSERT FIGURE 2 AROUND HERE]

LeSage and Pace (2008) propose including among the covariates spatial lags of the dependent variable, computed using a contiguity spatial weight matrix. However, since we are dealing with a count variable, it is more convenient to model the interdependence among locations by means of the conditional expectation of the variable of interest (Lambert *et al.* 2010). This is also the approach followed by LeSage *et al.* (2007), who propose a Bayesian hierarchical Poisson with latent regional effects parameters that follow a spatial autoregressive structure constructed using a contiguity spatial weight matrix. Our spatial weighting scheme differs from that of LeSage *et al.* (2007) and LeSage and Pace (2008), however, in that it applies only to the covariates (as in Lambert *et al.* 2010). We also differ in the definition of the spatial weighting matrix, which in our case is a smooth function of the distance (see e.g. LeSage 2004 and McMillen and McDonald 2004). Still, it is worth noting that this approach can be seen as an approximation to the "Leontieff inverse" that arises when using the spatial autoregressive structure (Martínez *et al.* (2013); see also McMillen 2010).

Thus, let $W[i,j]$ denote the weight assigned to location j on location i. For simplicity, we assume that all covariates have the same weight. However, we assume that these weights may differ both across origins and destinations. We denote the former by $W^O[i,j]$ and the latter by $W^D[i,j]$. Also, consistent with our distributional assumptions, intra-location flows may be subject to a different weighting scheme, which we denote by $W^{OD}[i,j]$. Lastly, we construct these weighing matrices using the following exponential function:

$$W^M[i,j] = \frac{\exp\left(\frac{d_{ij}}{h^M}\right)}{\sum_{l=1}^{n} \exp\left(\frac{d_{il}}{h^M}\right)} \quad \text{for} \quad M = O, D, OD.$$

where h^M is a set of parameters to be estimated.

3.3 Model specification

Since our interest lies in analysing the determinants of the number of relocating establishments, the model specification employs commonly used distributions for count data (see e.g. Cameron and Trivedi 2013). In particular, we use the Poisson and Negative Binomial to model the intra-ZdEs flows (F_1 when $i = j$) and the Inflated versions of these models to model the inter-ZdEs flows (F_2 when $i \neq j$). This means that our model specification, in conditional expectation form, is given by:

$$E\left(y_{ij}|\mathbf{x}, d\right) = \mu_{ij}\left(1 - \varphi_{ij}\right)\mathbf{1}(i \neq j) + \mu_i \mathbf{1}(i = j)$$

with $\varphi_{ij} = \frac{\exp\left(\gamma_0 + \mathbf{x}_i\gamma^O + \mathbf{x}_j\gamma^D + \gamma^d d_{ij}\right)}{1+\exp\left(\gamma_0 + \mathbf{x}_i\gamma^O + \mathbf{x}_j\gamma^D + \gamma^d d_{ij}\right)}$ being the probability that $y_{ij} = 0$ (alternatively, a probit function may be used here), $\mathbf{1}(\cdot)$ being an indicator function and

$$\begin{aligned} \mu_{ij} = \quad & exp\left(\beta_0 + \sum_{l=1}^{n} W^O[i,l]\mathbf{x}_l\beta^O + \sum_{l=1}^{n} W^D[j,l]\mathbf{x}_l\beta^D + \beta^d d_{ij}\right) \\ \mu_i = \quad & exp\left(\beta_0^{OD} + \sum_{l=1}^{n} W^{OD}[i,l]\mathbf{x}_l\beta^{OD}\right). \end{aligned}$$

Notice that, for the sake of simplicity, we have considered the same set of covariates in all the elements of the model. However, in applications the determinants of the probability that the variable of interest is zero may well differ from those of its conditional expectation (see e.g. Manjón-Antolín and Arauzo-Carod 2011). Also, one may argue that the determinants of the intra-ZdEs flows may differ from those of the inter-ZdEs flows (see e.g. Weterings and

Knoben 2013). In any case, including these exclusion restrictions in the previous expressions would only complicate the notation without providing further insights.

In general, the coefficients of the model are $\theta_1 = \left[\beta_0^{OD}, h^{OD}, \beta^{OD}\right]$ and $\theta_2 = \left[\gamma_0, \gamma^O, \gamma^D, \gamma^d,\right.$ $\left.\beta_0, h^O, \beta^O, h^D, \beta^D, \beta^d\right]$. More specifically, γ_0, β_O and β_0^{OD} are the constant terms; γ^d and β^d are the distance parameters; the vectors γ^O and γ^D are the origin and destination parameters of the inflated part of the inter-ZdEs flows model; the vectors β^O and β^D are the origin and destination parameters of the non-inflated part of the inter-ZdEs flows model; and β^{OD} are the parameters of the intra-ZdEs flows model.

We expect the number of relocations to be decreasing in the distance between origin and destination, so that $\gamma^d \geq 0$ and $\beta^d \leq 0$. Also, β^{OD} can contain both positive and negative parameters. As for the rest of the γ and β coefficients, they can be either positive or negative. In particular, the impact on the volume of flows makes that the sign interpretation of β's and γ's is the opposite. While positive/negative β's would increase/decrease the expected flows, positive/negative γ's would decrease/increase the expected flows.

Following Griffith and Jones (1980), we expect $\beta^O \times \beta^D \leq 0$ and $\gamma^O \times \gamma^D \leq 0$ for those characteristics that make both origin and destination either more ($\beta^O \leq 0$ and $\beta^D \geq 0$; $\gamma^O \geq 0$ and $\gamma^D \leq 0$) or less ($\beta^O \geq 0$ and $\beta^D \leq 0$; $\gamma^O \leq 0$ and $\gamma^D \geq 0$) attractive for the relocating firm. Similarly, we expect $\beta^O \times \beta^D > 0$ and $\gamma^O \times \gamma^D > 0$ for those characteristics that proxy for the size of the origin and destination, as well as those that either make the origin more attractive and the destination less attractive ($\beta^O < 0$ and $\beta^D < 0$; $\gamma^O > 0$ and $\gamma^D > 0$) or the origin less attractive and the destination more attractive ($\beta^O > 0$ and $\beta^D > 0$; $\gamma^O < 0$ and $\gamma^D < 0$). Notice, however, that this interpretation does not imply that the γ and β coefficients of variables that determine both φ_{ij} and μ_{ij} must show an opposite sign. For example, a variable may act as a proxy for the size of the origin and destination in φ_{ij} ($\gamma^O \times \gamma^D > 0$) while making both origin and destination either more or less attractive in μ_{ij} ($\beta^O \times \beta^D \leq 0$).

3.4 Estimation strategy

We estimate the model coefficients by maximum likelihood. In particular, in our empirical application the vector of explanatory variables, $\mathbf{x}$, includes both corporate tax rates and proxies for agglomeration economies. However, to the extent that these proxies do not fully account for the existence of agglomeration economies, this specification raises a problem of endogeneity that may result in biased estimates of the tax coefficients. This is because taxes tend to be positively correlated with agglomeration economies (Jofre-Monseny 2013, Koh *et al.* 2013, Luthi and Schmidheiny 2013).

Figure 3 shows that corporate tax rates are unevenly distributed across France.[6] The ZdEs of the South-East, for example, have generally higher taxes than those of the North (the Paris area being an exception to this pattern). Thus, one may expect that, *ceteris paribus*, the relocation flows would typically go from South-East to North (Charney 1983). However, a comparison of Figures 2 and 3 reveals that the flows seem to go the other way round. This means that areas with higher corporate taxes actually receive important inflows of establishments from other ZdEs.

[INSERT FIGURE 3 AROUND HERE]

There is therefore a "taxable agglomeration rent" by which "[f]irms accept to bear a higher tax rate in order to benefit from agglomeration economies" (Charlot and Paty 2007: 248). As a result, one may expect to observe a positive correlation between taxes and relocation flows. However, once we control for the existence of agglomeration economies, taxes should impinge upon the relocation flows.[7] We make use of this idea to derive an empirical strategy that yields consistent estimates of the tax coefficients.

[6]We compute corporate tax rates as the sum of of the following components of the "taxe professionnelle": mean tax rate of the municipalities included in the corresponding ZdE ("taux communal"), mean tax rate of the municipalities included in the corresponding "commune" ("taux intercommunal"), mean tax rate of the municipalities included in the corresponding "zone d'activités économiques" ("taux applicables dans les ZAE"), tax rate of the province in which the ZdE is located ("taux departemental") and tax rate of the region in which the ZdE is located ("taux regional"). In essence, this is the measure used by e.g. Houdebine and Schneider (1997) and Rathelot and Sillard (2008), except that since they use municipalities as the unit of analysis, they do not use means but individual values.

[7]Brülhart *et al.* (2012) and Jofre-Monseny and Solé-Ollé (2010), for example, show that agglomeration economies reduce the sensitivity of firm location to local taxes; see also Jofre-Monseny and Solé-Ollé (2012) for evidence based on employment location.

Let us first introduce some further notation. Let $\mathcal{A}$ be the agglomeration economies that are not accounted for by our vector of explanatory variables (and are thus part of the non-observed part of the model). Also, let us assume that the error term u can be additively decomposed into the non-accounted agglomeration economies ($\lambda\mathcal{A}$, with λ being an appropriate parameter vector) and a random shock with the standard properties (ε). Lastly, let T denote the corporate taxes variable, so that $\mathbf{x} = [\mathbf{x}_1, T]$.

In this setting, it is well known that model coefficients can be consistently estimated if $E[\mathcal{A}|\mathbf{x}_1, T] = E[\mathcal{A}]$. However, if the previous condition does not hold, standard least-squares and maximum-likelihood estimators are in general inconsistent (Wansbeek and Meijer 2000). One way to obtain consistent estimates is to resort to instrumental variables. To this end, we require a set of variables $\mathbf{z}$ such that $E[T|\mathbf{z}] \neq E[T]$ (relevance condition) and $E[\mathcal{A}|\mathbf{z}] = E[\mathcal{A}]$ (exogeneity condition). This is the approach followed by e.g. Rathelot and Sillard (2008) and Brülhart *et al.* (2012), who ultimately replace the corporate taxes variable by a linear prediction of $\mathbf{z}$ and $\mathbf{x}_1$.

We put forward an alternative approach based on the existence of a variable $\mathcal{T}$ that is correlated with the unobservable $\mathcal{A}$ (the opposite of the previous exogeneity condition). This variable is included in the model as an additional explanatory variable, so that now $\mathbf{x} = [\mathbf{x}_1, T, \mathcal{T}]$ and $u = \lambda(\mathcal{A} - \mathcal{T}) + \varepsilon$. It is then easy to see that in this setting a consistent estimate of the coefficient associated with T requires that the following condition holds: $E[\mathcal{A} - \mathcal{T}|T] = E[\mathcal{A} - \mathcal{T}]$.

The question that arises, of course, is whether such a variable can be found among the set of covariates typically used in (re)location studies. In our case, the structure of the corporate taxation in France provides a natural candidate: the total tax rates (i.e., the sum of municipal, departmental and regional taxes) of the origin and the destination of the flow. Obviously, this variable is closely related to the presence of agglomeration economies. In fact, this is the origin of the endogeneity problem and the reason why the estimation of its associated coefficient is typically found to be positive (Rathelot and Sillard 2008, Jofre-Monseny and Solé-Ollé 2010). However, if this variable is actually controlling for (part of) the unobserved agglomeration economies, we need another one to pick up the effect of taxes on the relocation flows. We use the difference between destination and origin local tax rates (i.e., the sum of municipal taxes, which include the "taux communal", the "taux intercommunal" and the "taux applicables dans les ZAE") to construct the corporate taxes variable T. First, it is safe to assume that the difference between local taxes (T) and the unexplained component of the agglomeration forces ($\mathcal{A} - \mathcal{T}$) are unrelated. Second, this is still a very relevant part of the corporate taxes, for on average approximately half of the total tax rates correspond to the municipal tax rates.

3.5 Interpretation of the tax coefficients

Our estimation strategy is simple to implement. Also, it is valid under an independence assumption analogous to the one made in the IV approach. However, it also has some limitations, particularly with regard to the interpretation of the coefficients.

Notice that the coefficient of T is not identified when $i = j$, since in this case the differences in local taxes become zero. Thus, our approach can only be used to analyse inter-ZdEs flows. Still, one should be careful with the interpretation of the total tax rates coefficient in the intra-ZdEs part of the model ($\beta_{\mathcal{T}}^{OD}$). A positive tax coefficient may be indeed a by-product of the positive correlation between taxes and agglomeration economies. However, it may also reflect increases in relocation flows within the same ZdE from municipalities raising taxes to municipalities lowering taxes, as long as the average increase is lower than the average decrease. This is why we have considered the dispersion in corporate rates as an additional covariate (and its product with the total tax rate to control for cross-effects), measured by the standard deviation of the total tax rate across the municipalities of the corresponding ZdE. We expect this variable to have a positive coefficient, thus indicating that tax variations stimulate the relocation decisions.

In the inter-ZdEs part of the model, we expect taxes to deter the relocation flows (i.e., $\beta_T \leq 0$) once we control for the existence of agglomeration economies by including (a function of) $\mathcal{T}_O$ and $\mathcal{T}_D$ among the covariates. Also, since the total tax rates in the origin and destination correspond to characteristics that make both origin and destination more attractive for the relocating firm, we expect $\beta_{\mathcal{T}}^O \leq 0$ and $\beta_{\mathcal{T}}^O \geq 0$.[8] Lastly, under the *ceteris paribus* clause

[8] In the empirical application we have also included these variables in the inflated part of our models. However, it is not a priori clear what should be the signs of these coefficients ($\gamma_{\mathcal{T}}^O$ and $\gamma_{\mathcal{T}}^D 0$).

the impact of a unitary change in the difference between destination and origin local tax rates must be obtained keeping the rest of the variables constant. However, in order to maintain the total tax rates of the origin and destination unchanged the department and/or regional tax rates must vary to fully compensate the unitary change in T. Therefore, to compute the marginal effects we need to make additional assumptions about how this unitary change was originated.[9]

4 Results

4.1 The data

Our main data source is the INSEE, which provides information at the ZdE level on the number of relocating establishments (by number of employees) as well as on socio-economic characteristics of the ZdEs (agglomeration economies, local labour market conditions, institutional characteristics, transport infrastructures and education).[10] One of the reasons for using ZdEs as the unit of analysis is that, since are defined by economic criteria, they are more likely to correspond to the relocating area that firms may consider.[11] More specifically,

[9]For example, if we assume that the increase in the difference of local taxes is only due to an increase in the mean municipal taxes of the destination, we must consider an analogous reduction in the department and/or regional taxes of the destination. Notice, however, that these department and regional taxes are included in the computation of the local taxes of the nearby ZdEs. As a result, the total taxes rates in the nearby ZdEs will ultimately decrease, thus making the destination ZdE less attractive with respect to its neighbours. Consequently, we expect that (mean of the) flows to the destination will be reduced.

[10]The INSEE provides analogous information for alternative administrative (regions, provinces, and municipalities, as well as "arrondissements" and "cantons") and functional (e.g., the "zonage en aires urbaines") units. However, given our interest in controlling for agglomeration economies, previous evidence on the extent of their spatial range suggests that regions and provinces are way too large (see e.g. Rosenthal and Strange 2003). Also, French provinces are divided into 330 "arrondissements", which in turn are divided into 3,883 "cantons". These may consequently seem a better choice, but the "arrondissements" do not have the status of legal entities and the "cantons" are mainly used as constituencies for the election of the members of the General Council of the province. Lastly, municipalities are too small (and probably too many given the way flow data define the sample size), for their mean surface in mainland France is only about 15km^2 and the average number of inhabitants is about 1,700. ZdEs can thus be seen as a good comprise between regions and provinces, on the one hand, and municipalities, on the other. (In fact, there are formed by adjacent municipalities from the same region and usually the same province). Furthermore, compared to other functional units such as the "zonage en aires urbaines", which divides big cities into central-business-districts ("pôle urbain") and their suburban areas ("couronne périurbaine"), they have the advantage of covering the whole territory (the "zonage en aires urbaines" only covers three quarters of the population).

[11]Notice that if this is not the case and firms actually use a different geographical unit when deciding where to relocate we will be facing a potential bias in our estimates. However, an empirical assessment of this issue is clearly beyond the scope of this paper (see e.g. Arauzo-Carod and Manjón-Antolín 2012).

a ZdE is defined as "a geographical area within which most of the labour force lives and works, and in which establishments can find the main part of the labour force necessary to occupy the offered jobs". Thus, ZdEs are constructed using commuting data (flows of movement from residence to work of active persons observed during the 2006 census) to define their geographical limits, which, as a general rule, should not contain an active population of less than $25,000$ people.

There were 348 ZdEs in mainland France when we constructed the dataset.[12] However, after dropping those from Corse, we were left with 341 observations. This means that since our dependent variable is the number of establishments that relocated their premises between each of these ZdEs in 2008, our sample size consist of $341 \times 341 = 116,281$ observations ($115,940$ corresponding to inter-ZdE flows and 341 corresponding to intra-ZdE flows).[13] We report descriptive statistics of this variable in Table 1, distinguishing between intra- and inter-ZdEs flows.

[INSERT TABLE 1 AROUND HERE]

We also used data from the INSEE to construct the vector of explanatory variables, except for the centroid distances (in km., data from the "Service d'information aéronautique") and the corporate taxes (total tax rates and differences in local tax rates computed as described in footnote 6 using data from the "Ministère de l'économie et des finances"). In particular, the available proxies for the agglomeration economies include the number of entering establishments in all industries in 2006 (*entry*), the stock of establishments in all industries in 2006 (*stock*), the population density in 2006 (*dens*) and the job density in 2005 (*jobdens*). In addition, we included the squares of population density ($dens^2$) and job density ($jobdens^2$) to account for potential dis-economies (Henderson 1997). Moreover, we use unemployment levels in 2007 (*unem*), net average wages per hour in 2006 (*wage*), and the employed manufacturing (*empman*) and service (*empser*) workforce in 2007 to control for local labour market

[12] In June 2012 the INSEE changed the list and composition of the ZdEs (they became 322). We use the previous version, the so-called ZdE 1990. The list of the municipalities is the one datum by the geographical official code ("code officiel géographique").

[13] Out of the $123,729$ establishments that relocated their premises in mainland France in 2008, almost three out of four moved to the same ZdE. However, this rate may be reduced up to a half in the Île-de-France region (e.g. Roissy-en-France, Vitry-sur-Seine and Poissy) and reach almost 90% in areas of the South (e.g. Perpignan and Toulouse) and South-East (e.g. Digne).

conditions. As for the institutional features, we have considered whether the ZdE is the capital of a region (the dummy *capital*) and whether it is in the Paris region (the dummy *paris*). We have also used dummies indicating whether the ZdE has a TGV station (*tgv*) and an airport (*airport*) as measures of the transport infrastructures. Lastly, the following variables summarise the educational level of the individuals living in the ZdE: percentage of adult population with a university degree (*uni*), with secondary studies (*bep*) and without studies (*ili*) in 1999. We report descriptive statistics of these variables in Table 1. We also include descriptive statistics of the standard deviation of the total tax rate (employed in the inter-ZdEs part of the model) and total population (*ptot*, employed in the inflated part of the models).

4.2 Estimates

We initially report estimates of a non-spatially weighted version of our model in Table 2. These coefficients are likely to be biased because they are obtained without tackling the endogeneity problem of the taxes and/or controlling for the spatial dependence of the data. However, these results serve us to illustrative three major points.[14]

First, the differences between intra- and inter-ZdEs flows. As the histogram in Figure 1 suggests, the assumption of common behaviour across geographical units does not hold. Whereas the conditional distribution of the intra-ZdEs flows seems to fit better to a standard count data model (in particular, the significance of the overdispersion parameter supports the negative binomial specification), the conditional distribution of the inter-ZdEs flows requires an inflated version of a count data model to cope with the excess of zeros (in particular, the significance of the overdispersion parameter and the rejection of the Vuong test support the negative binomial specification).

Second, the distance between ZdEs is always statistically significant and shows the ex-

[14]It is important to stress that the results reported in Table 2 are not driven by a specific selection of covariates. We have explored alternative specifications that do not include all the proxies for agglomerations economies (e.g. dropping entry and/or population density), use other proxies for wages (e.g. distinguishing white-collard and blue-collard workers), use alternative education variables (e.g. the CAP rather than the BEP), and/or change the way we introduce taxes (e.g. with differences between taxes in destination and origin in the main equation and/or with taxes in origin and destination in the inflated part). However, in all the alternative specifications we considered we obtained essentially the same results.

pected sign. Thus, the relocation flows seem to be decreasing in the distance between origin and destination. This, together with the significance of both origin and destination variables as well as the differences between intra- and inter-ZdEs flows, support the use of a gravity model.

Third, the effects of the correlation between taxes and unobserved agglomeration economies. The signs of the tax variables indicate that the flows of firms are directed towards ZdEs with higher taxes. We have previously pointed out the difficulties in the interpretation of the tax coefficient in the intra-ZdEs model. Still, the positive sign of the tax variable is consistent with taxes (at least partially) stimulating relocation flows within the ZdEs. In the inter-Zdes model, taxes seem to make both origin and destination more attractive for the relocating firm. This in line with the findings of Rathelot and Sillard (2008) and Jofre-Monseny and Solé-Ollé (2010).

[INSERT TABLE 2 AROUND HERE]

We provide further evidence of the existence of taxable agglomeration rents in Table 3. In particular, the estimates reported in columns **(1)**, **(2)** and **(3)** show that this result does not depend on the way we introduce the taxes in the model (either as a variable in origin and another in destination or as the difference between the variable in destination and that in the origin, either in the main equation of the model or in the inflated part).[15] Next, however, we introduce the difference in local taxes as an additional covariate. Thus, estimates reported in columns **(4)**, **(5)** and **(6)** follow the empirical strategy described in the previous section.

We find now evidence that taxes deter the relocations flows, since the coefficient of $\mathcal{T}$ is negative and statistically significant. It is also interesting to note that the coefficients of T tend to increase when we introduce $\mathcal{T}$ as an additional covariate. This is consistent with the idea that the coefficients of the taxes variables in columns **(1)**, **(2)** and **(3)** are actually reflecting both the genuine effect of taxes and the indirect effect arising from the correlation with the unobservable agglomeration economies. Once we control for the former, the latter becomes stronger.

[15]In fact, introducing taxes in the inflated part barely changes the results since these variables are barely statistically significant.

Lastly, given that taxes are barely significant in the inflated part, we decided to drop them to see whether these results still hold. As shown in column **(7)**, it turns out that these estimates not only corroborate our previous finding but provide a better fit in terms of a penalized likelihood criterion (AIC). We have consequently decided to use this specification in the last step of our empirical strategy. Namely, the estimation of the model accounting for spatial dependence. Results are reported in Table 4.

[INSERT TABLE 3 AROUND HERE]

In particular, results were obtained by spatially weighting the following variables: the stock of establishments (agglomeration economies), the unemployment rate (local labour market conditions), the existence of a TGV station (transport infrastructures) and the percentage of people with a university degree (education). The reason for this was to reduce the computational burden and facilitate the convergence of the gauss-newton algorithm (with analytical first derivatives) that we used. Notice, however, that since variables proxying for the same determinant are likely to be highly correlated, it may suffice to account for the spatial dependence of one of them. As for the rest of the covariates it was either judged unnecessary (distance and institutional features) or complex to implement (since e.g. taxes are defined as a difference between destination and origin).

Moreover, we report results for all the establishments and for the following categories of establishment size: zero employees (self-employment), 1-2 employees (micro firms), 3 to 9 employees (small firms), 10 to 49 employees (medium-sized firms) and more than 50 employees (large firms). In this vein, we seek to investigate whether the impact of taxes on relocations is different for establishments of different size. We find weak evidence that larger relocating establishments tend to be more affected by taxes than the smaller ones. Although the coefficient estimates of $\mathcal{T}$ are in accordance with what the theory predicts (Charney 1983, Baldwin and Okubo 2009), they are only significant for the self-employees.

As for the rest of the coefficients, we find differences both across the size distribution as well as in the origin and destination of the flow. The flows of larger establishments, for example, are barely affected by the characteristics of the origin and destination. In contrast, the presence of agglomeration economies, the labour market and the education appear to be important for the flows of smaller establishments. Also, Paris stands for the expulsion of small and medium-sized establishments while being attractive for the self-employees. Lastly, most of the statistically significant variables show the same sign in both the origin and the destination. Thus, most of our determinants are either characteristics that proxy for the size of the origin and destination (e.g., entry and stock) or make the origin more attractive and the destination less attractive (e.g. the percentage of adult population without studies and the employed manufacturing and service workforce).

However, the use of a smooth continuous function to construct the spatial weighting matrices provides an additional set of parameters (h^O and h^D) that estimate the geographical scope of these determinants. More precisely, the larger the parameters h^O and h^D are the larger the geographical scope of the determinants, i.e., the further will reach the indirect marginal effects.[16] Thus, our results indicate that the larger the establishment is the smaller is the geographical scope of the determinants of its flows. Also, the destination characteristics have a larger geographical impact than the origin characteristics.

[INSERT TABLE 4 AROUND HERE]

To conclude, we provide further evidence on the geographical dimension of the pushing and pulling effects. Following Martínez *et al.* (2013), we use the ratio between the indirect marginal effect and the direct marginal effect for this purpose. This measure circumvents the difficulties of analysing marginal effects when the number of covariates is large (as it is in our case) and, since only depends on the distance between geographical units, a simple plot allows to analyse how far reaches the effect of a unitary change in the covariates and how is the shape of this effect with respect to the distance reached. In particular, in Figure 4 we report results for the different establishment sizes based on the estimates reported in Table 4.

[INSERT FIGURE 4 AROUND HERE]

There are substantial differences in the geographical scope of the covariates across the size distribution. While the relative effect in large firms practically disappears after 40 km., for

[16] In spatial models, the indirect marginal effect is the impact on the dependent variable in a geographical unit of a unitary change in a covariate in a different geographical unit (see e.g. LeSage and Pace 2009)

example, it may reach nearly 400 km. for self-employees. There also notable differences in the decreasing path, which is almost constant in large firms and becomes much more smooth the smaller the establishment is. It is also interesting to note that firms seem to consider origin and destination characteristics asymmetrically, in the sense that the characteristics that tend to push firms out of a ZdE do not have the same geographical extent than those that tend to pull firms towards another ZdE. However, this seems to be case mostly for the smaller firms. In large and medium-sized concerns, differences in the geographical scope of the origin and destination determinants are not appreciable.

5 Conclusions

Where to locate a new venture is a critical strategic decision. Yet we observe a large number of establishments and firms changing the location of their premises. These relocations may occur between countries (foreign direct investments), but also within the same country or region (what we may call "local relocations"). In any case, it is important for both managers and policy makers to understand why firms and establishments are relocated.

There is indeed an extensive literature on the determinants of f.d.i. and a growing body of research on the determinants of local relocation decisions at the firm level. However, little is known about what drives the flows of business concerns between nearby geographical units. Previous evidence on this topic has concentrated on the characteristics of the destination (since the origin of the flows is typically unknown) and has been obtained assuming that origin-destination flows are spatially independent (a potential specification error). This paper presents a new gravity model to analyse spatially-dependent origin-destination flows when both origin and destination characteristics are available. In particular, we account for the geographical interdependence using spatial weighting matrices constructed as a smooth function of the distance between the geographical units involved.

We illustrate the capabilities of our model by estimating the effect of corporate taxes on the number of relocated establishments between French "Zones d'Emploi" (travel-to-work areas). To this end, we propose an estimation strategy that deals with the endogeneity problem raised by the existence of taxable agglomeration rents. We show that the maximum likelihood estimator is consistent under the assumption of independence between destination-origin differences in local taxes and the part of the agglomeration forces that is unexplained by the model.

Our main finding is that, once we control for taxable agglomeration rents, local taxes deter the relocation flows. Also, consistent with the theoretical predictions, taxes seem to impinge more on the larger relocating concerns (although the statistical significance of the associated coefficients is marginal). In general, we find that the determinants of the origin-destination flows vary across the size distribution. In this respect, estimates of the pushing and pulling effects, measured by the ratio of the indirect-to-direct marginal effects, show that the impact of a change in the covariates may extend up to 150 km. around the origin and 300 km. around the destination for the smaller concerns, while it barely reaches 100 km. for the larger ones. This suggests that the geographical scope of the relocation determinants is asymmetrical, in the sense that firms seem to be concerned by larger areas in the destination than in the origin. However, this asymmetry is mostly a small-concern phenomenon.

References

Arauzo-Carod, Josep Maria, Daniel Liviano-Solís and Miguel Manjón-Antolín (2010), 'Empirical studies in industrial location: An assessment of their methods and results', *Journal of Regional Science* **50**(3), 685–711.

Arauzo-Carod, Josep Maria and Miguel Manjón-Antolín (2012), '(Optimal) spatial aggregation in the determinants of industrial location', *Small Business Economics* **39**, 645–658.

Baldwin, Richard and Toshihiro Okubo (2009), 'Tax reform, delocation, and heterogeneous firms', *Scandinavian Journal of Economics* **111**(4), 741–764.

Brülhart, Marius, Mario Jametti and Kurt Schmidheiny (2012), 'Do agglomeration economies reduce the sensitivity of firm location to tax differentials?', *Economic Journal* **122**(September), 1069–1093.

Cameron, A. Colin and Pravin K. Trivedi (2013), *Regression Analysis for Count Data*, Cambridge University Press (2^{nd} ed.).

Charlot, Sylvie and Sonia Paty (2007), 'Market access effect and local tax setting: evidence from French panel data', *Journal of Economic Geography* **7**, 247–263.

Charney, Alberta H. (1983), 'Intraurban manufacturing location decisions and local tax differentials', *Journal of Urban Economics* **14**, 184–205.

Curry, Leslie (1972), 'A spatial analysis of gravity flows', *Regional Studies* **6**(2), S9–S25.

Griffith, Daniel A. (2007), 'Spatial structure and spatial interaction: 25 years later', *The Review of Regional Studies* **37**(1), 28–38.

Griffith, Daniel A. and Kelvyn G. Jones (1980), 'Explorations into the relationships between spatial structure and spatial interaction', *Environment and Planning A* **12**, 187–201.

Henderson, J. Vernon (1997), 'Medium size cities', *Regional Science and Urban Economics* **27**, 583–612.

Holl, Adelheid (2004), 'Transport infrastructure, agglomeration economies, and firm birth. Empirical evidence from Portugal', *Journal of Regional Science* **44**(4), 693–712.

Houdebine, Michel and Jean-Luc Schneider (1997), 'Mesurer l'influence de la fiscalité sur la localisation des entreprises', *Économie et Prevision* **131**, 47–64.

Jofre-Monseny, Jordi (2013), 'Is agglomeration taxable?', *Journal of Economic Geography* **13**(1), 177–201.

Jofre-Monseny, Jordi and Albert Solé-Ollé (2010), 'Tax differentials in intraregional firm location: Evidence from new manufacturing establishments in Spanish municipalities', *Regional Studies* **44**(6), 663–677.

Koh, Hyun-Ju, Nadine Riedel and Tobias Böhm (2013), 'Do governments tax agglomeration rents?', *Journal of Urban Economics* **75**, 92–106.

Kronenberg, Kristin (2013), 'Firm relocations in the Netherlands: Why do firms move, and where do they go?', *Papers in Regional Science* **forthcoming**.

Lambert, Dayton M., James P. Brown and R.J.G.M. Florax (2010), 'A two-step estimator for a spatial lag model of counts: Theory, small sample performance and an application', *Regional Science and Urban Economics* **40**(10), 241–252.

Lee, Yoonsoo (2008), 'Geographic redistribution of US manufacturing and the role of state development policy', *Journal of Urban Economics* **64**, 436–450.

LeSage, James P. (2004), A family of geographically weighted regression models, *in* L.Anselin, R.Florax and S.Rey, eds, 'Advances in Spatial Econometrics: Methodology, Tools and Applications', Springer, pp. 241–262.

LeSage, James P. and Manfred M. Fischerand Thomas Scherngell (2007), 'Knowledge spillovers across Europe: Evidence from a Poisson spatial interaction model with spatial effects', *Papers in Regional Science* **86**(3), 393–422.

LeSage, James P. and R. Kelley Pace (2008), 'Spatial econometric modeling of origin-destination flows', *Journal of Regional Science* **48**(5), 941–967.

LeSage, James P. and Robert Kelley Pace (2009), *Introduction to spatial econometrics*, CRC Press.

Luthi, Eva and Kurt Schmidheiny (2013), 'The effect of agglomeration size on local taxes', *Journal of Economic Geography* **forthcoming**.

Manjón-Antolín, Miguel and Josep Maria Arauzo-Carod (2011), 'Locations and relocations: determinants, modelling, and interrelations', *Annals of Regional Science* **47**(1), 131–146.

Mariotti, I. (2005), *Firm relocation and regional policy*, Department of Spatial Sciences (University of Groningen), Netherlands Geographical Studies 331.

Martínez-Ibañez, Oscar, Miguel Manjón-Antolín and Josep Maria Arauzo-Carod (2013), 'The geographical scope of industrial location determinants: An alternative approach', *Journal of Economic and Social Geography* **104**(2), 194–214.

McMillen, Daniel P. (2010), 'Issues in spatial data analysis', *Journal of Regional Science* **50**(1), 119–141.

McMillen, Daniel P. and John F. McDonald (2004), Locally weighted maximum-likelihood estimation: Monte carlo evidence and an application, *in* L.Anselin, R.Florax and S.Rey, eds, 'Advances in Spatial Econometrics: Methodology, Tools and Applications', Springer, pp. 225–239.

Rathelot, Roland and Patrick Sillard (2008), 'The importance of local corporate taxes in business location decisions: Evidence from French micro data', *Economic Journal* **118**(March), 499–514.

Rosenthal, Stuart S. and William C. Strange (2003), 'Geography, industrial organization and agglomeration', *The Review of Economics and Statistics* **85**(2), 377–393.

van Dijk, J. and P.H. Pellenbarg (2000), 'Firm relocation decisions in the Netherlands: An ordered logit approach', *Papers in Regional Science* **79**, 191–219.

Wansbeek, Tom and Erik Meijer (2000), *Measurement Error and Latent Variables in Econometrics*, Elsevier.

Weterings, Anet and Joris Knoben (2013), 'Footloose: An analysis of the drivers of firm relocations over different distances', *Papers in Regional Science* **forthcoming**.

Figure 1: **Histogram of the dependent variable.**

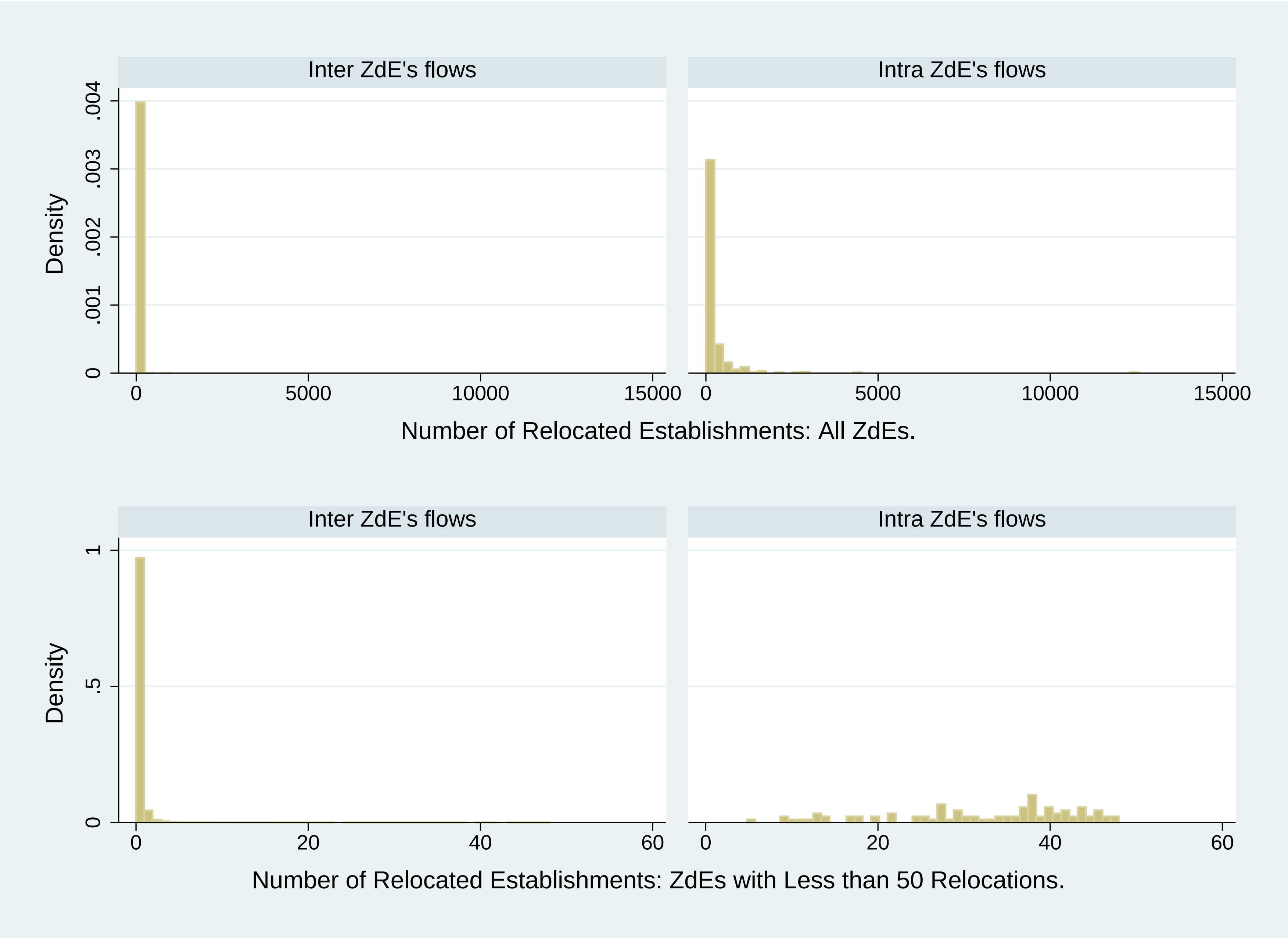

Figure 2a: **Intra ZdEs flows.**

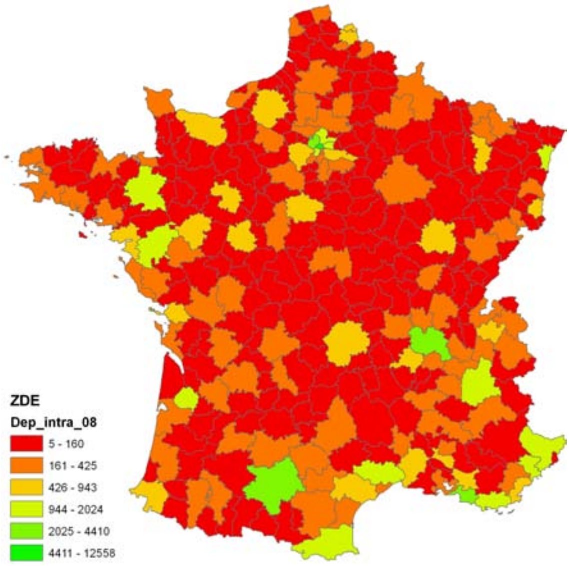

Figure 2b: **Inter ZdEs flows (by origin).**

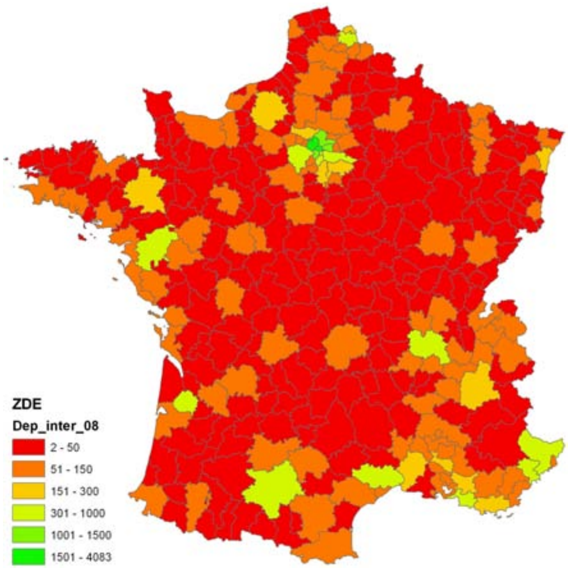

Figure 2c: **Inter ZdEs flows (by destination).**

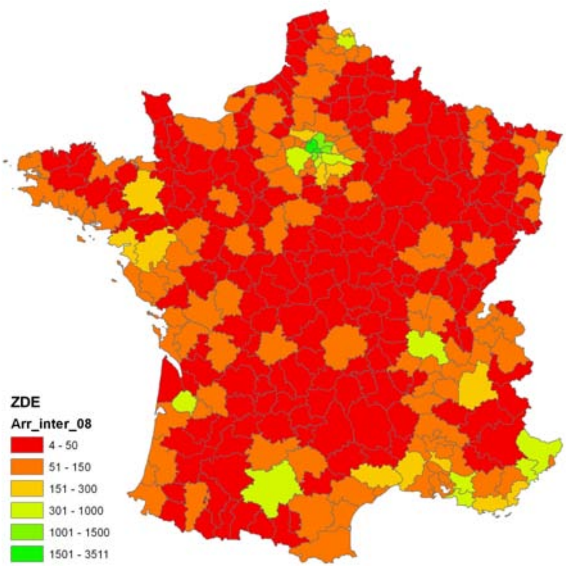

Figure 3: **Corporate Taxes (ZdE mean).**

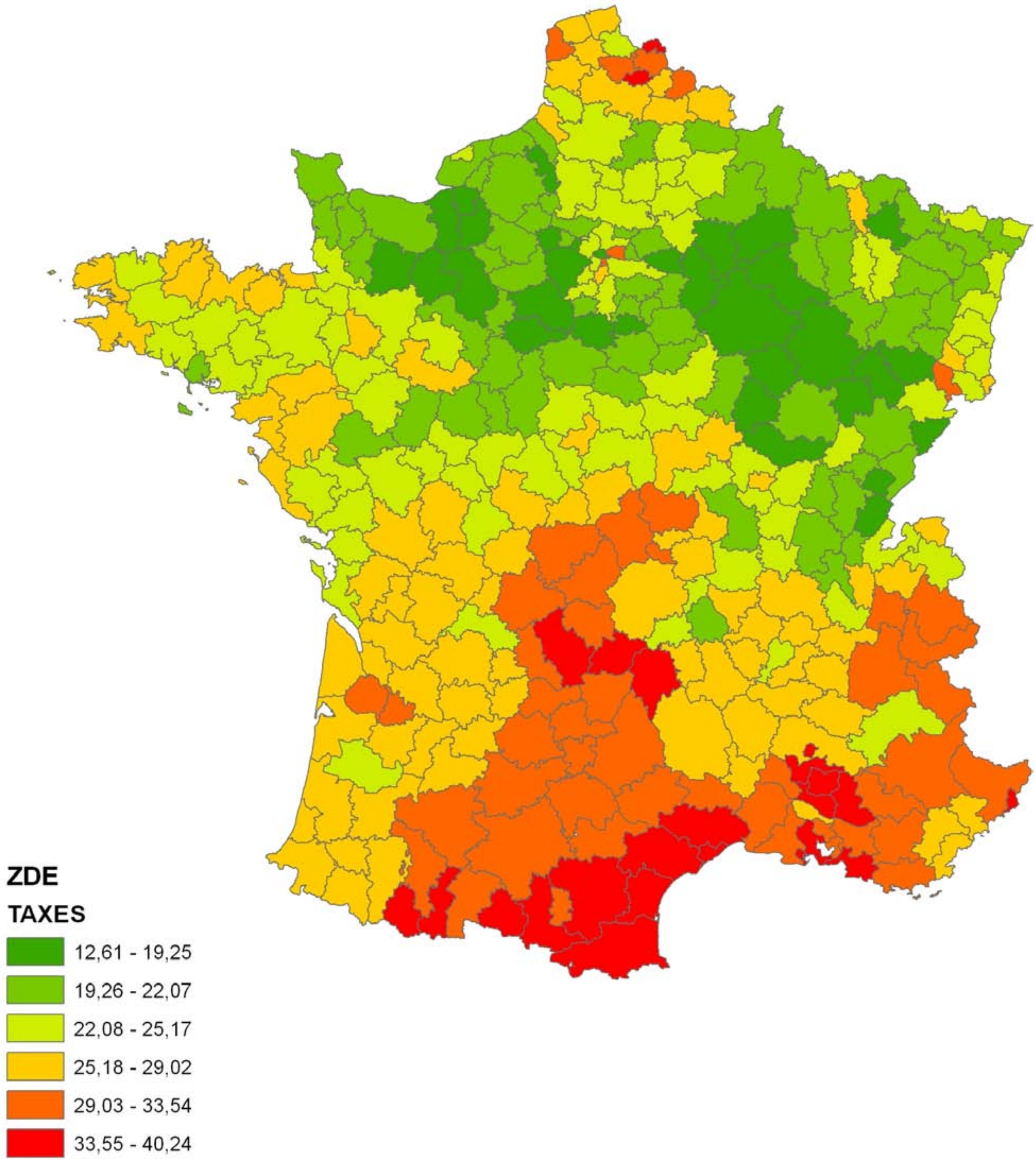

Figure 4: **Relative Marginal Effects (by Establishment Size, Zero-Inflated Negative Binomial Estimates).**

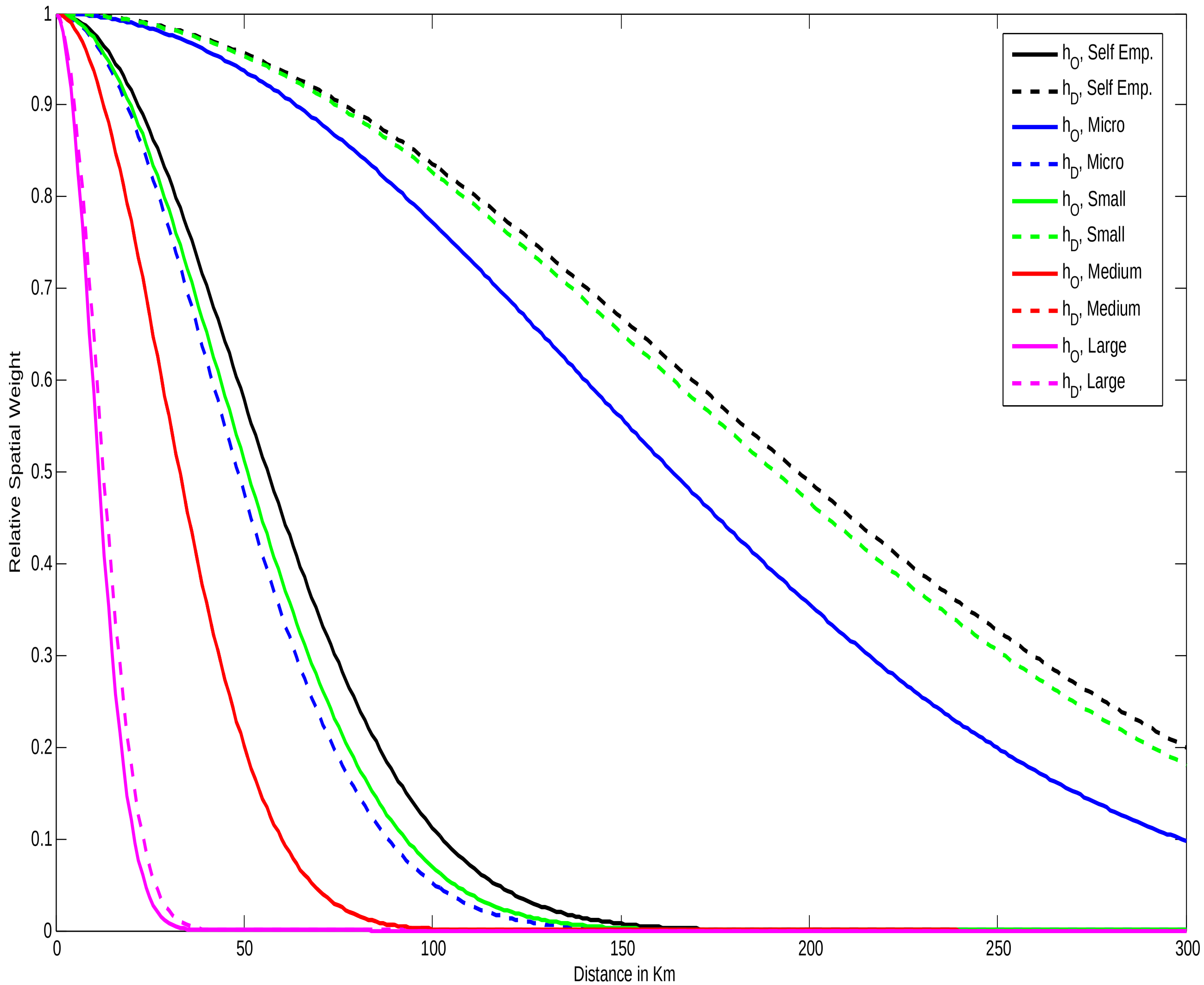

Table 1: Summary statistics.

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**
y_{ij}	1.06	45.84	0	12558
$y_{ij}(i = j)$	273.52	796.49	5	12558
$y_{ij}(i \neq j)$	0.26	5.27	0	830
entry	0.95	3.41	0.06	58.61
stock	9.76	33.95	0.66	590.84
dens	0.27	1.25	0.01	20.17
jobdens	0.11	0.51	0	7.84
unem	8.67	2.11	4.5	15.5
wage	11.08	1.17	9.5	18.4
empind	0.2	0.08	0.04	0.44
empser	0.57	0.08	0.36	0.85
capital	0.26	0.44	0	1
paris	0.08	0.27	0	1
tgv	0.41	0.49	0	1
airport	0.5	0.5	0	1
ili	0.27	0.05	0.12	0.38
bep	0.08	0.02	0.03	0.13
uni	0.06	0.05	0.02	0.31
tax (T)	25.27	4.91	14.07	40.24
dif. loc. tax ($\mathcal{T}$)	0	4.63	-18.31	18.31
distance	395.2	192.98	0	1049.27
std. dev. tax	4.01	1.25	0	10.91
ptot	180.73	318.36	14.64	4251.7

Table 2: Estimates of the Determinants of the Relocation Flows.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Model:	Poisson	Neg. bin.	Inf. Poisson	Inf. Neg. bin.	Poisson	Neg. bin.	Inf. Poisson	Inf. Neg. bin.
Flows:	All ZdEs	All ZdEs	All ZdEs	All ZdEs	Intra-ZdEs	Intra-ZdEs	Inter-ZdEs	Inter-ZdEs
Main equation								
Oentry	0.4688***	0.3894***	0.4731***	0.1344	0.5201***	-0.2151	0.9090***	0.6369***
	(0.0234)	(0.1361)	(0.0230)	(0.1380)	(0.0238)	(0.2364)	(0.0438)	(0.1042)
Ostock	-0.0359***	-0.0077	-0.0368***	0.0001	-0.0342***	0.0671**	-0.0803***	-0.0530***
	(0.0025)	(0.0146)	(0.0024)	(0.0147)	(0.0025)	(0.0264)	(0.0046)	(0.0110)
Odens	-0.4207***	0.2440	-0.3771***	-0.0951	-0.7824***	-0.1179	-0.6301***	-0.2961*
	(0.0330)	(0.2576)	(0.0331)	(0.2411)	(0.0365)	(0.6612)	(0.0626)	(0.1647)
Odens^2	0.0405***	-0.0380	0.0333***	0.0158	0.0582***	-0.0567	0.0562***	0.0389**
	(0.0032)	(0.0246)	(0.0032)	(0.0231)	(0.0038)	(0.0615)	(0.0059)	(0.0160)
Ojobdens	1.7974***	0.3791	1.6117***	0.8490	2.8315***	1.4086	2.5388***	1.4392***
	(0.0929)	(0.7065)	(0.0932)	(0.6673)	(0.1016)	(1.7865)	(0.1759)	(0.4556)
Ojobdens^2	-0.4247***	-0.1262	-0.3630***	-0.2824*	-0.5965***	-0.1365	-0.5422***	-0.4153***
	(0.0235)	(0.1718)	(0.0236)	(0.1637)	(0.0272)	(0.4068)	(0.0428)	(0.1139)
Ounem	0.0238***	0.0653***	0.0180***	0.0793***	0.0447***	0.0453***	0.0096**	0.0558***
	(0.0026)	(0.0101)	(0.0028)	(0.0123)	(0.0025)	(0.0159)	(0.0048)	(0.0083)
Owage	0.0076	0.1762***	-0.0068	0.1598***	0.0624***	0.0859*	0.0465***	0.1657***
	(0.0055)	(0.0274)	(0.0056)	(0.0300)	(0.0062)	(0.0498)	(0.0095)	(0.0208)
Oempind	-2.7022***	-1.4336***	-1.9377***	-2.0315***	-3.6880***	-3.8008***	-3.0014***	-3.2907***
	(0.1155)	(0.4710)	(0.1196)	(0.6080)	(0.1251)	(0.7574)	(0.2387)	(0.3937)
Oempser	-2.7124***	-1.9668***	-2.0659***	-2.6916***	-4.8129***	-4.6502***	-3.3949***	-3.5952***
	(0.1190)	(0.5303)	(0.1226)	(0.6719)	(0.1380)	(0.8878)	(0.2407)	(0.4390)
Ocapital	0.2931***	0.1172**	0.2885***	0.0910*	0.2324***	0.1945**	0.0211	0.0324
	(0.0106)	(0.0490)	(0.0107)	(0.0539)	(0.0112)	(0.0820)	(0.0205)	(0.0397)
Oparis	-0.3360***	-0.1074	-0.3736***	0.0743	-0.1647***	-0.1996	-0.1040***	0.1125
	(0.0236)	(0.0960)	(0.0242)	(0.1079)	(0.0261)	(0.1673)	(0.0387)	(0.0767)
Otgv	0.3588***	0.2488***	0.3091***	0.3161***	0.4866***	0.3742***	0.4316***	0.3671***
	(0.0092)	(0.0385)	(0.0094)	(0.0491)	(0.0101)	(0.0631)	(0.0178)	(0.0316)
Oaeroport	0.1737***	0.0516	0.1360***	0.0674	0.2345***	0.1986***	0.0769***	0.0847***
	(0.0089)	(0.0375)	(0.0090)	(0.0436)	(0.0099)	(0.0602)	(0.0184)	(0.0316)
Oili	-5.3419***	-5.0734***	-4.2024***	-5.5805***	-7.1095***	-6.3355***	-6.1073***	-5.3125***
	(0.2148)	(0.8366)	(0.2249)	(0.9925)	(0.2140)	(1.4088)	(0.4323)	(0.7054)
Obep	-8.7538***	-8.0184***	-8.0486***	-13.1135***	-19.3401***	-18.0523***	-6.6032***	-13.0850***
	(0.4990)	(1.9946)	(0.5286)	(2.5704)	(0.5312)	(3.3557)	(1.0021)	(1.7076)
Ouni	-1.9195***	-2.0872**	-1.3428***	-3.7604***	-3.2767***	-4.2118***	-2.3712***	-2.9021***
	(0.1871)	(0.8540)	(0.1931)	(0.9507)	(0.1843)	(1.4669)	(0.3825)	(0.6811)
Otax (T)	-0.0131***	-0.0064	-0.0155***	-0.0094	0.0289***	0.0363*	-0.0087***	-0.0088*
	(0.0014)	(0.0049)	(0.0015)	(0.0068)	(0.0030)	(0.0210)	(0.0025)	(0.0046)
Dentry	0.5989***	1.0967***	0.5948***	0.8413***			1.2562***	0.8224***
	(0.0234)	(0.1383)	(0.0230)	(0.1355)			(0.0436)	(0.0955)
Dstock	-0.0486***	-0.0756***	-0.0483***	-0.0685***			-0.1158***	-0.0691***
	(0.0025)	(0.0147)	(0.0024)	(0.0143)			(0.0046)	(0.0101)
Ddens	-0.6361***	-1.1443***	-0.5927***	-0.6531***			-0.7169***	-0.7247***
	(0.0331)	(0.2626)	(0.0331)	(0.2384)			(0.0627)	(0.1676)
Ddens^2	0.0419***	-0.0274	0.0346***	-0.0105			0.0384***	0.0257
	(0.0032)	(0.0255)	(0.0032)	(0.0227)			(0.0059)	(0.0160)
Djobdens	2.1115***	2.8729***	1.9412***	1.4579**			2.4071***	1.9812***
	(0.0932)	(0.7239)	(0.0931)	(0.6615)			(0.1761)	(0.4639)
Djobdens^2	-0.4020***	-0.0465	-0.3444***	0.0147			-0.3645***	-0.2545**
	(0.0234)	(0.1793)	(0.0234)	(0.1617)			(0.0425)	(0.1136)
Dunem	0.0126***	0.0141	0.0108***	0.0073			-0.0117**	0.0116
	(0.0026)	(0.0101)	(0.0028)	(0.0131)			(0.0047)	(0.0092)
Dwage	0.0157***	0.0507*	0.0072	0.0847***			0.1183***	0.1037***
	(0.0054)	(0.0287)	(0.0055)	(0.0315)			(0.0093)	(0.0222)
Dempind	-3.6587***	-5.7261***	-2.9893***	-3.4981***			-4.1360***	-3.7538***
	(0.1148)	(0.4632)	(0.1192)	(0.6410)			(0.2286)	(0.4274)
Dempser	-3.4433***	-5.2895***	-2.8865***	-2.5440***			-4.1839***	-3.4447***

	(0.1174)	(0.5223)	(0.1208)	(0.7020)			(0.2255)	(0.4655)
Dcapital	0.2019***	-0.1584***	0.1765***	-0.2267***			-0.1303***	-0.1843***
	(0.0105)	(0.0487)	(0.0106)	(0.0526)			(0.0198)	(0.0391)
Dparis	-0.8777***	-0.6790***	-0.8670***	-0.8416***			-0.5970***	-0.5764***
	(0.0235)	(0.1016)	(0.0240)	(0.1117)			(0.0381)	(0.0807)
Dtgv	0.4127***	0.3689***	0.3454***	0.1641***			0.4780***	0.3131***
	(0.0091)	(0.0384)	(0.0093)	(0.0499)			(0.0171)	(0.0333)
Daeroport	0.1872***	0.1250***	0.1845***	0.0646			0.1179***	0.0979***
	(0.0087)	(0.0373)	(0.0089)	(0.0439)			(0.0175)	(0.0321)
Dili	-6.9666***	-7.2714***	-6.9212***	-6.0591***			-8.2198***	-6.8720***
	(0.2134)	(0.8262)	(0.2244)	(1.0148)			(0.4081)	(0.7239)
Dbep	-8.8299***	-16.0996***	-6.5042***	-4.2360			-9.7840***	-9.1897***
	(0.4938)	(1.9847)	(0.5238)	(2.8207)			(0.9671)	(1.8463)
Duni	-3.2602***	-5.3351***	-2.9083***	-4.2781***			-5.4201***	-4.6899***
	(0.1858)	(0.8382)	(0.1930)	(0.9688)			(0.3691)	(0.6869)
Dtax (T)	0.0073***	0.0365***	0.0084***	0.0253***			0.0207***	0.0183***
	(0.0014)	(0.0048)	(0.0015)	(0.0069)			(0.0024)	(0.0047)
distance	-0.0338***	-0.0113***	-0.0295***	-0.0093***			-0.0082***	-0.0067***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)			(0.0001)	(0.0001)
Ost.dev.tax					0.1671***	0.1423		
					(0.0188)	(0.1292)		
Otax×Ost.dev.tax					-0.0050***	-0.0057		
					(0.0007)	(0.0047)		
Inflated part								
Optot			-0.0026***	0.0002			-0.0020***	0.0009
			(0.0002)	(0.0006)			(0.0001)	(0.0006)
Dptot			-0.0018***	0.0065***			-0.0010***	0.0080***
			(0.0002)	(0.0010)			(0.0001)	(0.0012)
Ostock			0.0683***	-0.0384			-0.0022	-0.2210***
			(0.0147)	(0.0415)			(0.0084)	(0.0391)
Dstock			0.0741***	-0.1833***			0.0185**	-0.2327***
			(0.0135)	(0.0618)			(0.0084)	(0.0720)
Oent06			-0.5470***	-0.7432**			0.1252	1.2926***
			(0.1493)	(0.3756)			(0.0888)	(0.2886)
Dent06			-0.6748***	-1.9856***			-0.1495*	-1.9395***
			(0.1385)	(0.5634)			(0.0886)	(0.6488)
Dtax - Otax (T)			0.0058	0.0044			-0.0005	0.0116
			(0.0047)	(0.0079)			(0.0036)	(0.0083)
distance			-0.0052***	0.0042***			-0.0002*	0.0032***
			(0.0002)	(0.0002)			(0.0001)	(0.0002)
Observations	116,281	116,281	116,281	116,281	341	341	115,940	115,940
AIC	277506.19	78951.37	255096.87	77130.32	15358.00	3868.35	83987.84	64552.57
LR Test	1148924.19***	20943.49***	564929.57***	10015.07***	189454.72***	736.62***	69748.32***	9683.78***
α		8.67***		5.59***		0.22***		2.39***
Vuong Test			19.66***	20.50***			19.83***	14.91***

Note: 115,940 observations. Standard errors in parentheses. *** $p - value < 0.01$, ** $p - value < 0.05$, * $p < 0.10 - value$. “O” and “D” indicate origin and destination covariates, respectively. AIC stands for Akaike Information Criterion. LR Test is a joint significance test. α is the value of the overdispersion parameter.

Table 3: Taxes as Determinants of Inter-ZdEs Flows (Zero Inflated Negative Binomial Estimates).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Main equation							
Otax (T)	-0.0099**			-0.0197***	-0.0188***		
	(0.0048)			(0.0072)	(0.0072)		
Dtax (T)	0.0166***			0.0267***	0.0285***		
	(0.0050)			(0.0075)	(0.0073)		
Dtax - Otax (T)		0.0130***	0.0134***			0.0234***	0.0196***
		(0.0040)	(0.0040)			(0.0068)	(0.0062)
dif. loc. tax ($\mathcal{T}$)				-0.0158*	-0.0161*	-0.0159*	-0.0153*
				(0.0087)	(0.0087)	(0.0087)	(0.0087)
Inflated part							
Otax (T)	-0.0167*	-0.0193**		-0.0171*			
	(0.0097)	(0.0094)		(0.0096)			
Dtax (T)	0.0051	0.0013		0.0060			
	(0.0104)	(0.0099)		(0.0104)			
Dtax - Otax (T)			0.0112		0.0123	0.0118	
			(0.0082)		(0.0083)	(0.0082)	
AIC	64563.10	64552.86	64544.67	64571.47	64560.83	64553.00	64543.42
LR Test	9638.85***	9637.43***	9680.01***	9642.14***	9687.17***	9727.92***	9683.35***
α	2.38***	2.38***	2.39***	2.38***	2.39***	2.39***	2.39***
Vuong Test	14.92***	15.05***	15.01***	14.93***	14.92***	15.02***	14.95***

Note: 115,940 observations. Standard errors in parentheses. *** $p-value < 0.01$, ** $p-value < 0.05$, * $p < 0.10-value$. AIC stands for Akaike Information Criterion. LR Test is a joint significance test. α is the value of the overdispersion parameter. Besides the reported variables, the main equation of the model includes proxies for agglomeration economies (*entry*, *stock*, *dens*, $dens^2$, *jobdens* and $jobdens^2$), local labour market conditions (*unem*, *wage*, *empman* and *empser*), institutional features (*capital* and *paris*), transport infrastructures (*tgv* and *airport*), education (*uni*, *bep* and *ili*) and distance. Besides the reported variables, the inflated part of the model includes total population, stock of establishments and distance as additional covariates. "O" and "D" indicate origin and destination covariates, respectively.

Table 4: Estimates of the Determinants of the Spatially Weighted Inter-ZdE Flows.

	(1)	(2)	(3)	(4)	(5)	(6)
Establishments:	All	Self.Employ.	Micro	Small	Medium	Large
Main equation						
Oentry	0.1162^{***}	0.1121^{***}	0.1723^{***}	0.1518^{***}	0.3809^{***}	0.4172
	(0.0082)	(0.0085)	(0.0152)	(0.0286)	(0.0763)	(0.2941)
W^OOstock	0.0459^{***}	0.0467^{***}	-0.0496	0.0165	-0.0260^{***}	-0.0420
	(0.0048)	(0.0050)	(0.0372)	(0.0135)	(0.0094)	(0.0359)
Odens	-0.3999^{**}	-0.4279^{***}	-0.7007^{**}	-0.6178	-0.0764	-0.6232
	(0.1596)	(0.1639)	(0.2874)	(0.4055)	(0.7001)	(2.3514)
Odens^2	0.0534^{***}	0.0579^{***}	0.0634^{**}	0.0756^{**}	-0.0221	-0.0443
	(0.0145)	(0.0151)	(0.0259)	(0.0364)	(0.0745)	(0.2283)
Ojobdens	1.8247^{***}	1.9100^{***}	2.5921^{***}	3.0098^{***}	3.1194^{*}	5.1367
	(0.4249)	(0.4383)	(0.7611)	(1.0837)	(1.8148)	(7.1196)
Ojobdens^2	-0.5392^{***}	-0.5672^{***}	-0.6396^{***}	-0.7812^{***}	-0.4168	-0.3867
	(0.0990)	(0.1035)	(0.1766)	(0.2472)	(0.4790)	(1.6039)
W^OOunem	0.0958^{***}	0.0896^{***}	-0.0655	0.0417	0.0596	0.1332
	(0.0127)	(0.0134)	(0.0781)	(0.0413)	(0.0449)	(0.2087)
Owage	0.0175	0.0260	0.0068	-0.0955^{*}	-0.2887^{***}	0.0326
	(0.0183)	(0.0193)	(0.0381)	(0.0527)	(0.0753)	(0.2789)
Oempind	-2.7289^{***}	-2.7599^{***}	-3.2111^{***}	-1.7460	-3.2032	-2.8047
	(0.3653)	(0.3871)	(0.8844)	(1.2418)	(2.0697)	(8.8886)
Oempser	-2.8535^{***}	-2.8228^{***}	-3.7459^{***}	-3.0563^{**}	-4.8205^{**}	-4.1537
	(0.4192)	(0.4452)	(0.9618)	(1.3757)	(2.4160)	(9.9648)
Ocapital	0.2041^{***}	0.2042^{***}	0.0433	0.1129	-0.0718	0.7268
	(0.0366)	(0.0389)	(0.0793)	(0.1176)	(0.1904)	(1.2758)
Oparis	-0.2123^{**}	-0.2725^{***}	0.4343^{***}	0.5687^{**}	1.2150^{***}	0.8677
	(0.0886)	(0.0941)	(0.1475)	(0.2761)	(0.2661)	(0.9511)
W^OOtgv	1.5253^{***}	1.5119^{***}	2.1062^{**}	1.4934^{***}	0.7320^{***}	0.0111
	(0.1265)	(0.1308)	(0.8689)	(0.4160)	(0.1616)	(0.9505)
Oaeroport	0.1405^{***}	0.1334^{***}	0.0367	0.1809^{**}	0.1412	0.0247
	(0.0271)	(0.0291)	(0.0624)	(0.0877)	(0.1461)	(0.9120)
Oili	-4.2718^{***}	-4.4441^{***}	-3.5547^{***}	-4.1231^{***}	-6.7974^{*}	-17.7331
	(0.4469)	(0.4767)	(0.9977)	(1.4956)	(3.5857)	(18.3674)
Obep	-9.2272^{***}	-9.1096^{***}	0.3119	-13.9076^{***}	-24.6660^{***}	-25.2965
	(1.2618)	(1.3413)	(2.9942)	(4.1899)	(9.2507)	(41.0758)
W^OOuni	-8.3952^{***}	-9.1264^{***}	28.0547	-2.0855	-2.6136	-15.0202
	(2.2600)	(2.3461)	(19.3453)	(5.9869)	(-2.6136)	(20.9456)
Dentry	0.1686^{***}	0.1646^{***}	0.1673^{***}	0.0940^{***}	0.2027^{***}	0.6480
	(0.0091)	(0.0095)	(0.0226)	(0.0225)	(0.0352)	(0.4850)
W^DDstock	-0.1389^{***}	-0.1322^{***}	0.0333^{***}	-0.1052	0.0323^{**}	-0.0485
	(0.0219)	(0.0232)	(0.0081)	(0.0691)	(0.0162)	(0.0517)
Ddens	-0.8074^{***}	-0.7508^{***}	-0.7826^{***}	-0.3106	-0.6520	1.4787
	(0.1574)	(0.1675)	(0.2939)	(0.3721)	(0.4460)	(2.3532)
Ddens^2	0.0476^{***}	0.0393^{**}	0.0657^{**}	0.0364	0.0356	-0.3283
	(0.0145)	(0.0154)	(0.0280)	(0.0341)	(0.0407)	(0.3674)
Djobdens	2.5764^{***}	2.3258^{***}	2.7774^{***}	1.3745	2.4952^{**}	0.0535
	(0.4259)	(0.4535)	(0.8114)	(1.0069)	(1.2444)	(6.0264)
Djobdens^2	-0.4851^{***}	-0.4158^{***}	-0.6385^{***}	-0.3634	-0.4748^{*}	1.2978
	(0.0992)	(0.1059)	(0.1941)	(0.2363)	(0.2841)	(2.2098)
W^DDunem	-0.1211^{***}	-0.0724	0.0343	-0.1310	-0.0003	-0.0933
	(0.0432)	(0.0450)	(0.0257)	(0.1436)	(0.0517)	(0.30668)
Dwage	0.0424^{**}	0.0377^{*}	0.0145	0.0697	-0.0939	0.0182
	(0.0203)	(0.0219)	(0.0408)	(0.0571)	(0.0621)	(0.3425)
Dempind	-3.5703^{***}	-3.6990^{***}	-5.0449^{***}	-1.1266	-1.0976	-9.0855
	(0.3921)	(0.4253)	(0.7480)	(1.3064)	(1.9693)	(10.5442)
Dempser	-3.0289^{***}	-3.1144^{***}	-4.7109^{***}	-0.7377	-3.6249^{*}	-17.9169
	(0.4461)	(0.4779)	(0.8552)	(1.3988)	(1.8805)	(13.3953)
Dcapital	-0.0292	-0.0404	0.0697	0.0912	-0.0222	0.2099
	(0.0355)	(0.0380)	(0.0752)	(0.1145)	(0.1882)	(0.9708)

Dparis	-0.1137	-0.1207	-0.6713^{***}	-0.2554	0.0751	0.4041
	(0.0767)	(0.0821)	(0.1964)	(0.2281)	(0.4469)	(1.3894)
W^DDtgv	1.8023^{***}	1.8875^{***}	1.5232^{***}	1.4986	2.2400^{***}	1.2661
	(0.4484)	(0.4992)	(0.2399)	(1.8531)	(0.4690)	(1.2706)
Daeroport	0.1046^{***}	0.1024^{***}	0.1828^{***}	-0.0680	0.1573	0.7673
	(0.0279)	(0.0300)	(0.0556)	(0.0937)	(0.1350)	(1.0471)
Dili	-2.3705^{***}	-2.6661^{***}	-3.3781^{***}	-1.9646	-4.3657^{*}	-0.7118
	(0.4272)	(0.4590)	(0.9581)	(1.5803)	(2.3841)	(22.2232)
Dbep	-3.8995^{***}	-3.7333^{***}	-10.4369^{***}	9.6697^{**}	-11.1208	-17.9390
	(1.2981)	(1.3785)	(2.8062)	(4.7076)	(6.8199)	(43.5558)
W^DDuni	72.3966^{***}	63.9670^{***}	-2.8267	35.8196	0.3265	6.1925
	(11.2506)	(11.5315)	(3.6096)	(30.7496)	(5.2058)	(15.4927)
Dtax - Otax (T)	0.0307^{***}	0.0259^{***}	0.0047	0.0065	0.0583^{*}	0.0537
	(0.0069)	(0.0073)	(0.0148)	(0.0254)	(0.0328)	(0.2277)
dif. loc. tax ($\mathcal{T}$)	-0.0279^{***}	-0.0248^{**}	0.0211	0.0150	-0.0536	-0.0389
	(0.0093)	(0.0099)	(0.0199)	(0.0319)	(0.0422)	(0.2873)
distance	-0.0068^{***}	-0.0066^{***}	-0.0085^{***}	-0.0084^{***}	-0.0105^{***}	-0.0067^{***}
	(7.6×10^{-5})	(8.2×10^{-5})	(0.0002)	(0.0003)	(0.0003)	(0.0018)
Inflated part						
Optot	0.1387^{***}	-0.0311	-8.8139^{***}	-0.6403^{***}		
	(0.0265)	(0.0256)	(0.1231)	(0.0329)		
Dptot	6.7421^{***}	7.7289^{***}	0.4973^{***}	-0.9392^{***}		
	(0.0655)	(0.0665)	(0.0283)	(0.1247)		
Ostock	-0.1538	-0.1467	0.3841^{***}	-0.0856		
	(0.5364)	(0.5388)	(0.2979)	(1.1102)		
Dstock	-0.1745^{***}	-0.2439^{***}	0.0012^{***}	0.4680^{***}		
	(0.0002)	(0.0002)	(0.0004)	(0.0006)		
Oent06	0.8988^{*}	0.9287^{*}	-5.7332^{***}	0.9065^{***}		
	(0.4970)	(0.4847)	(2.5542)	(0.4018)		
Dent06	-2.5909^{***}	-1.9671	-0.0484^{***}	-7.1880^{***}		
	(1.2477)	(1.2569)	(0.3141)	(2.2701)		
distance	0.06033^{***}	0.0033^{***}	0.0038^{***}	0.0070		
	(1.9×10^{-5})	(2.1×10^{-5})	(1.2×10^{-5})	(9.6×10^{-5})		
AIC	64769.7584	56926.77	16279.7325	7738.90	3683.5021	666.2136
α	0.8558^{***}	0.8281^{***}	0.7735^{***}	0.3486^{*}	-0.2420	-10.5779^{***}
Vuong Test	12.33^{***}	11.42^{***}	7.19^{***}	5.34^{***}	-0.0871	-0.0414
h^O	4750^{***}	4570^{***}	38700^{**}	3740^{**}	130^{**}	189^{**}
h^D	54600^{***}	55900^{***}	3380^{***}	52600^{**}	2210^{**}	229^{**}

Note: 115,940 observations. Standard errors in parentheses. *** $p-value < 0.01$, ** $p-value < 0.05$, * $p < 0.10-value$. Zero-inflated negative binomial estimates are reported except in columnn **(5)** and **(6)**, in which negative binomial estimates are reported. "O" and "D" indicate origin and destination covariates, respectively. AIC stands for Akaike Information Criterion. LR Test is a joint significance test. α is the value of the overdispersion parameter.