Urban Studies
48(9) 1909–1924, July 2011

# An Evaluation of the Establishment of a Taxi Flat Rate from City to Airport: The Case of Seville

**José I. Castillo-Manzano and Antonio Sánchez-Braza**

*[Paper first received, April 2009; in final form, June 2010]*

**Abstract**

This paper explores the establishment of a flat rate for the taxicab service that serves routes between the city and the airport in Seville. First, the advantages and disadvantages of a flat rate are explained and why it may be an advisable measure for tourist cities with a conflictive taxicab industry. In addition, by using a methodology based on the estimation of average treatment effects, an evaluation is made of the impact of the taxi flat rate on the probability of passengers choosing a taxi for their city–airport transport needs. The main conclusion is that the increased market transparency fostered by this urban transport policy has led to an increase of almost 7 per cent in both the probability of selection and the average revenues of taxi-drivers.

## 1. Introduction

Urban transport planning policies are subject to a range of objectives, with the end goal being an efficient transport system that allows for well-organised urban development. The integrated management of mobility, transport infrastructure, urban development and environmental protection is seen as essential for achieving more sustainable development (Hull, 2005).

One crucial aspect of urban transport planning is having an approriate city–airport link. An easily accessible and appropriate city–airport connection system is important for airport management (Neufville, 2006) and may be a factor that impacts interairport competition (McLay and Reynolds-Feighan, 2006; Pels *et al*., 2003). In this respect, the availability and/or frequency, and the time and the cost in monetary terms, of airport access from the city, or vice versa, take on a fundamental significance.

It also affects the urban economy, both directly (see Gaubatz, 1999, on the importance of the taxi service in the urban development of several large Chinese cities) and indirectly (see Brueckner, 2003, on the link

**José I. Castillo-Manzano** and **Antonio Sánchez-Braza** are in the Facultad de Ciencias Económicas y Empresariales, University of Seville, Avda. Ramón y Cajal, 1, Seville, 41018, Spain.
E-mail: jignacio@us.es and asb@us.es.

0042-0980 Print/1360-063X Online

DOI: 10.1177/0042098010380957

between airport-related ground traffic and employment in US cities). In short, the evidence confirms that a good airline service and an appropriate system to connect the city and the airport are important factors in urban economic development.

In addition, airport access road planning also has a very important effect on urban structure, which is felt not just in areas of close physical proximity but defined on a metropolitan scale. This broader impact is frequently taken into account in a variety of ways depending on, amongst other factors, the modes of transport and planning strategies chosen (Freestone, 2009).

In this context, the implications of the liberalisation of air transport seen in recent years for urban transport planning must also be borne in mind. Within this liberalisation, the development of the low-cost carriers has led to secondary and regional airports strengthening their position compared with the hubs (Reynolds-Feighan, 2001). In short, this phenomenon has brought about spectacular increases in passenger traffic at regional airports that were generally underutilised. In other respects, this increase in traffic has meant that city halls, that in the case of secondary and regional airports start out with a situation where the passenger is offered few connection choices, face more complex planning for the city–airport connection.

In the case of Spanish cities with regional airports (see Dobruszkes, 2006, on the different airport categories) taxis are significant airport–city modes of transport. First, there are few efficient rail-based public transport links to regional airports, near-essential alternatives in big cities with airport hubs (Lythgoe and Wardman, 2002). Influenced by the increases in traffic, a number of medium-sized Spanish cities with regional airports have planned costly investments to develop new city–airport metro and high-speed rail connections (see Castillo-Manzano and López-Valpuesta, 2009, for an analysis of metro projects in medium-sized Spanish cities), the high cost of which is being funded with aid from the European Regional Development Fund (ERDF). However, most of these connections have been planned without taking into account that, even for the major hubs with much higher passenger traffic, rail-based public transport links between the city and the airport are not necessarily profitable in economic terms (see Winston and Maheshri, 2007, on San Francisco's BART). As a result, these new rail-based links are a serious threat to future municipal budgets.

Secondly, alternative airport–city transport models, such as a 'shuttle service', have been tested in other countries but not in Spain (see Zhao and Dessouky, 2008; Sohail *et al.*, 2006; and Rogerson and Rogerson, 1997).[1] The shuttle minibus model allows more flexible travel from the airport to any point in the city. Additionally, a low rate that is fixed beforehand can be charged rather than the higher and unpredictable prices associated with taxi services (Zhao and Dessouky, 2008; Loo, 2007).

Thirdly, a regular bus service to the airport offers tourists little flexibility and the cost is high in terms of time (Cullinane, 2003). In addition, these bus services are often run by concessionaires—i.e. private companies that are not part of the public urban transport network. Consequently, their fares are usually higher and they do not allow free transfers onto other bus networks.

Generally, a passenger's choice of transport to or from the airport depends on a number of variables, mainly cost and time (Tsamboulas and Nikoleris, 2008). Any set of airport–city connections should therefore offer flexible, time-saving and price-competitive options (Zografos *et al.*, 2008; Quadrifoglio *et al.*, 2006).

However, a specific analysis of the taxicab industry shows this to be a clear case of an economic activity with asymmetric information (see Akerlof, 1970, on the introduction of this concept into modern economics).

When taking a taxi, the passenger may only have an approximate idea of what the trip will cost (Bonsall *et al.*, 2007). The final cost depends on various factors, such as distance, the route taken and traffic congestion. This uncertainty leads to numerous complaints, including formal complaints against the taxicab operator when passengers feel they have been overcharged (Howard, 2009). The number of complaints also increases when the city has an improper fare structure that reflects either the cost of transporting or of finding a passenger, but not both at the same time (Glazer and Hassin, 1983). It is not surprising if the passenger, especially the non-resident, is reluctant to choose this mode of transport. Even so, in some cases, such as city–airport connections at secondary airports, the taxi is virtually irreplaceable as it is the only viable alternative for a passenger with several pieces of baggage, for example.

The taxicab industry is also subject to various types of regulation, such as fare control and entry restriction (Yang *et al.*, 2010; Wong *et al.*, 2008) which affect the availability and quality of the taxicab service, driver incomes and the accountability of service providers (Schaller, 2007). One common effect of controlling both fares and capacity is the development of a monopoly market value (Gwilliam, 2008; Cairns and Liston-Heyes, 1996), with the consequent loss of economic efficiency implied in monopolies.

These features of taxi services mean they can develop into a sector where a specific collective—the licence owners—obtain large amounts of revenue from a monopoly that they are prepared to defend at any cost. This is the case in Seville. As a result, there are high rates of both industrial conflict and illegal practice by people who want to cash in.

The vital role that the taxi industry plays in keeping the tourist business running smoothly must also be highlighted. From a sectoral perspective, taxis are one element of the tourism system transport component, playing a leading role in urban tourism and often providing a means of access between major facilities, attractions and accommodation (Waryszak and King, 2000). A tourism-oriented city must offer a suitable and honest taxicab service (Eisinger, 2000). If we consider that the quality of tourism service starts when the passenger steps off the plane at the airport (Rendeiro, 2006), the taxi is the next link in this service chain. This maxim certainly applies to most cities in countries like Spain, which is currently one of the premier tourism destinations in the world (second in the world ranking of tourism destinations; WTO, 2008).[2]

Conscious of this situation, some local governments around Spanish regional airports have decided to establish a flat rate for the airport–city taxi service. This rate is independent of the exact in-city location. The aim is to increase the degree of transparency in an imperfect and asymmetric information service that rarely informs the user about possible extra fare charges, such as the airport-based service, the number of pieces of baggage or night fare supplements. Apart from which only frequent users will know the optimal route to their destination, which may vary depending on the time of day.

The aim of this paper is to evaluate whether the establishment of a taxi flat rate has any influence on passengers choosing the taxi as a city–airport transport option. We evaluate the consequences of this measure in terms of urban transport policy. This article goes beyond a purely descriptive view of this urban transport measure by trying to evaluate its quantitative contribution to urban transport flows. We examine the specific case of the effect of the taxi flat rate on the city–airport service in Seville, the most highly populated city in Spain to implement the measure to date.

This paper is contextualised by literature that focuses on the evaluation of different regulations and rules for taxicab services. Examples of this type of work are those that estimate the effects of taxicab industry deregulation (Schaller, 2007; Gaunt, 1996) and those that explore the consequences of certain measures implemented in this industry (see Yang *et al.*, 2005 and Yang *et al.*, 2002, on entry regulations and price controls). However, even though this measure has been progressively taken up in a number of other cities, both in Spain and other countries such as the US, there is a lack of literature on the subject; this article intends to fill the gap regarding both an analysis and an evaluation of the effects of flat rate implementation for city–airport connections.

The paper is organised as follows. Section 2 explains the characteristics of this policy in the city of Seville. Sections 3 and 4 describe the data and the econometric methodology (average treatment effects). Section 4 presents our empirical outcomes, while section 5 offers conclusions.

## 2. The Flat Rate for Taxis in Seville

At the beginning of the past decade, there was a great deal of conflict surrounding the less-than-transparent airport–city taxi service in Seville. Formal complaints were frequent, not only from users, mainly tourists, but also from taxi-drivers themselves who filed formal complaints against each other for illegal monopolistic airport access practices.[3] They also confronted the public bus operators that connect the airport with the city, incomprehensibly accusing them of anti-competitive practices. At the height of the conflict, there were recurrent vandalism (puncturing of taxi, airport workers' car and public bus tyres, destruction of street furniture and damage to airport facilities) and illegal strikes with no guaranteed minimum services.

Consumer associations and the Seville hospitality industry complained about these conflicts. Fearing that this situation could damage the city's tourist image, Seville City Council decided to establish a flat rate for this service, a first in Spain. The new rate was approved in 2003, after years of negotiation and studies to make sure that the pricing was appropriate in the context of the average cost of airport–city transfers.[4]

Today, many see this measure as a lesser evil and not an optimal solution, considering that passengers with permanent or temporary residences in districts close to the airport pay more than they would under a variable-fare system. Opponents also argue that the measure is a concealed subsidy to the tourism industry, since most hotels are located in the city centre, one of the most distant and less accessible districts from the airport (narrow streets, high traffic congestion and a high number of pedestrian zones).

To a great extent, the success of this policy relies on advertising. In particular, notices in Spanish and English were installed in the arrival and baggage-claim halls at Seville airport calling attention to the existence of the flat rate and explaining how it works. In addition, all taxis must display information about the flat rate on the inside of their windows. Hotel reception desks play an informal but relevant role in conveying information about the flat rate to tourists and the main tourist guides also provide information about its existence.

## 3. Data

In order to evaluate the effects of the taxi flat rate in Seville, the required control group was taken from Spanish cities where the airport–city dyad presents similar characteristics to those of Seville, namely

- —The airport is not connected to the city by train or subway (unlike the hubs in Madrid and Barcelona).
- —It is not a secondary airport that supports a hub airport (unlike those of Valladolid and Reus). If it were, many of the taxi rides would be to the hub-airport city (from

Valladolid airport to Madrid and from Reus airport to Barcelona) and not to the associated city itself.
—The airport is not close to any of the main sun-and-sand tourist areas (Malaga, Valencia or Alicante airports). Most of the rides would then be to the beach-front hotels, often on hotel or tour-operator courtesy shuttles and not to the city associated with the airport.
—The airport is a similar distance from the city (approximately 10 km).
—The size of the metropolitan area around the airport is similar to that of Seville.
—There is no taxi flat rate (unlike in Alicante or Santiago de Compostela).

The need to meet all these characteristics significantly restricted possible cities with Bilbao and Zaragoza and their airports most approximating to Seville.

The size of the selected sample, both for the treatment group (Seville) and for the control group (Bilbao and Zaragoza), was greater than the average for other studies using the same methodology, albeit these were in other areas of study quite unlike the taxicab industry (see Dehejia and Wahba, 2002, in the context of training programme evaluation; and Hirano and Imbens, 2001, in the context of medical science). The database was made up of 7341 passengers (3632 participants and 3709 in the control group) interviewed in the departure lounges of the three Spanish airports (Seville, Bilbao and Zaragoza). The database was collated from a set of survey campaigns conducted by AENA, the Spanish Public Airport Authority, from 2006 to 2007 (see Castillo-Manzano and López-Valpuesta, 2010, for a full explanation of this database). The main characteristics are listed in Table 1.

The original size of the sample was reduced to eliminate passengers who took connecting flights at the airport—that is, those passengers who had not travelled to the airport from the city. As with similar databases, each observation was weighted according to the total number of passengers on the flight so that the sample could be expanded to the total population (see Dresner, 2006, for a full explanation of weighting methodology).

## 4. Methodology

Studies of the effects of certain measures or policies on the taxicab industry have traditionally taken a variety of forms, from mere descriptive analyses (Schaller, 2007; Gaunt, 1996) to more analytical approaches. The latter also range from generic choice models of the taxi compared with other transport alternatives (Martínez *et al.*, 2009; Cantillo *et al.*, 2006; Bolduc, 1999) to analyses of how accessibility can impact this transport option (see the studies by Brons *et al.*, 2009, and Givoni and Rietveld, 2007, about access to railway stations).

Following the line of other papers evaluating public policies in broad terms (Blundell *et al.*, 2004; Dehejia and Wahba, 2002; Card and Krueger, 2000), the methodology proposed here is framed by statistical causal inference. It is thus based on the estimation of the causal effect (Pearl, 2000; Holland, 1986) that a specific measure can have on one or more relevant variables (Dawid, 2000; Cox, 1992). In contrast with traditional analyses, this methodology allows consistent estimators of the effects of the evaluated measure to be obtained (Rotnitzky and Robins, 1995) by determining and isolating the possible impact of additional contaminating variables.

Causal inference techniques have been used in a wide range of scientific disciplines, such as statistics (Rubin, 2008; Rosenbaum, 2002), medicine (Christakis and Iwashyna, 2003; Hirano and Imbens, 2001), epidemiology (Oakes and Church, 2007), sociology (Morgan and Harding, 2006; Smith, 1997), political science (Duch and Stevenson, 2006; Imai, 2005) and education (Blundell *et al.*, 2004; Card and Krueger, 2000).

**Table 1.** Survey of technical data

| | *Seville* | *Bilbao* | *Zaragoza* |
|---|---|---|---|
| Airport traffic in 2008 | 4 391 794 | 4 172 901 | 594 952 |
| *Information gathering* | | | |
| Questionnaire | Available in 12 languages | | Available in four languages |
| General | Departing passengers > 15 years of age | | |
| *Sampling* | | | |
| Sample size (before weighting) | 4 140 | 3 182 | 1 137 |
| Passengers not on connecting flights (before weighting) | 3 632 | 2 713 | 996 |
| Sampling method | Stratified by traffic segments<br>A number of flights were selected for each route with passengers selected by systematic sampling | | |
| Sampling error (percentage)[a] | ±1.5 | ±1.7 | ± 2.5 |
| Number of waves | | 1 | |
| *Field work* | | | |
| Time-period | 6–12 June | 4–10 May | 15–21 June |
| Location | Departure lounges | | |
| Timetable | Monday–Sunday<br>Shifts from 6am to 10pm, extended during periods of high traffic | | |
| Year | 2006 | 2007 | 2006 |

[a] Note that

$$\pm Error = k\sqrt{(N-n)/(N-1)}\sqrt{pq/n}$$

where, $N$ is the population size; $n$ is the sample size; $p = q = 0.5$ are the complementary probabilities of an event at the point of greatest indeterminacy; $k$ is an event parameter, where $k = 2$ for a 95.45 per cent confidence level.

The theoretical framework for the development of these evaluation methods began with Rubin's causal model (RCM; see Rubin, 1974, and Rubin, 1978, on this line of research) in the context of potential outcome models (POM). The relevant variables in this model are compared according to observations of individuals participating (treated) or not participating (non-treated controls) in the evaluated measure.

Starting with an $N$-size random sample, we defined the binary variable $D$ (flat rate) that indicates if the observation corresponds to a city with ($D_i = 1$) or without ($D_i = 0$) a taxi flat rate. Thus, our $N$ observations were divided into $n_1$ and $n_0$ observations (with vs without the flat rate). In our case, $n_1$ represented the observations in Seville and $n_0$ were the observations in Bilbao and Zaragoza.

We defined the outcome variable $Y$ as the decision to take a taxi for the city–airport transfer. Using the potential-outcome notation of the RCM (Rubin, 1974), we denoted the response variable as $Y_i(1)$ when $i$ stood for a city that had a flat rate, and as $Y_i(0)$ when $i$ corresponded to a city without a flat rate. Hence, $Y_i$ was equal to

$$Y_i = D_i Y_i(1) + (1 - D_i)\, Y_i(0)$$

We also defined a $K$-dimensional vector of observed covariates as $X$. In this way, a triad was observed for each individual ($D_i$, $Y$, $X_i$).

Under the unconfoundness condition and the overlap condition (see Appendix, section 1) we were able to obtain the average effect of the flat rate fare by comparing participating and control individuals with the same value in the context of the vector $X$ of observed covariates. In practice, when the number of observed covariates included in vector $X$ is sufficient, it is no longer appropriate to compare individuals by searching for exact equal values for each covariate. In our case, the evaluation procedure would be developed in several stages (Hahn *et al.*, 2010; Heckman and Vytlacil, 2005; Heckman *et al.*, 1998).

As a result, this methodology proposes a three-step procedure for estimating the average effect that a flat rate has on choosing a taxi for the city–airport connection. First, the propensity score is estimated. This allows individuals in the treatment group (Seville) and the control group (Bilbao and Zaragoza) to be homogenised on the basis of the covariates for the two groups being compared. Secondly, a comparison is made between the two groups to estimate the average effect that a flat rate has on the likelihood of them taking a taxi. Finally, the marginal effect is calculated for a better interpretation of the results. This gives a direct indication of the average increased likelihood of choosing the taxi as a means of airport access due to the flat rate. These three steps are next explained in detail.

### 4.1 Estimation of the Propensity Score

First, we noted whether the observations corresponded to a city with or without a flat rate. We then estimated the propensity score, defined by Rosenbaum and Rubin (1983) as the conditional likelihood of 'participating in the evaluated measure', given a vector $X$ of observed covariates (see the considerations regarding the propensity score as discussed in Abadie and Imbens, 2006; Imbens, 2004; Hirano *et al.*, 2003). We denoted this as $\varepsilon(X)$

$\varepsilon(X) = P(D=1 | X=x) = E[D | X=x]$, assuming that 0 and< $\varepsilon(X)$ and< 1

If the variables $D$ and $Y$ are mutually independent when conditional upon the covariates $X$ vector, then the same will also be true when they are conditional upon $\varepsilon(X)$. This ensures that the effects introduced by the covariates in vector $X$ are controlled (Smith and Todd, 2005; Hahn, 1998) and, with these effects therefore controlled, the individuals in the two groups can be compared.

Different binary response models can be used to estimate the propensity score depending on the configuration of the F distribution function in the chosen hypothesis. We used the binary response model (logit or probit) that maximised the log pseudo-likelihood.

$$\varepsilon(X) = P(D=1 | X) = F(\beta X)$$

where, $\beta$ is the vector of parameters associated with $X$. In our case, $X$ was composed of 18 covariates, presented in Table 2 together with their descriptive statistics.

### 4.2 Estimation of the Causal Effect

In a second phase, we calculated the causal effect of the measure on the response variable—in our case, the likelihood of choosing a taxi. The average effect on the selected sample was estimated from the equality

$$\alpha = E[\alpha(X)]$$

Following Imbens and Wooldridge (2009), the estimator of $\alpha$ ($\hat{\alpha}$) was obtained from a regression adjustment (see the Appendix, section 2). In our case, $Y$ being a binary

**Table 2.** Covariates and their descriptive statistics

| *Variable* | *Description* | *Observations* | *Mean* | *Maximum* | *Minimum* | *Median* | *S.D.* |
|---|---|---|---|---|---|---|---|
| *Socio-demographic factors and employment status (base category includes unemployed)* | | | | | | | |
| Sex | 1 = male; 0 = female | 4759 | 0.563 | 1 | 0 | 1 | 0.496 |
| Age | 1 = under 30; 2 = 31–49; 3 = 50–64; 4 = 65 or older | – | 1.977 | 4 | 1 | 2 | 0.785 |
| Nationality | 1 = passenger is non-Spanish; 0 = otherwise | 1740 | 0.206 | 1 | 0 | 0 | 0.404 |
| Currency | 1 = currency of passenger's country is not Euro; 0 = otherwise | 717 | 0.085 | 1 | 0 | 0 | 0.279 |
| Homemaker | 1 = passenger is homemaker; 0 = otherwise | 262 | 0.031 | 1 | 0 | 0 | 0.173 |
| Self-employed | 1 = passenger is non-salaried, generally self-employed; 0 = otherwise | 1370 | 0.162 | 1 | 0 | 0 | 0.368 |
| Salaried worker | 1 = passenger is salaried worker; 0 = otherwise | 5344 | 0.632 | 1 | 0 | 1 | 0.482 |
| Retired | 1 = passenger is retired; 0 = otherwise | 572 | 0.068 | 1 | 0 | 0 | 0.251 |
| Student | 1 = passenger is student; 0 = otherwise | 668 | 0.079 | 1 | 0 | 0 | 0.270 |
| *Trip category (base category includes passengers visiting friends and relatives (VFR) on a domestic flight)* | | | | | | | |
| Low-cost carrier | 1 = passenger is flying a low-cost carrier (LCC); 0 = otherwise | 2019 | 0.239 | 1 | 0 | 0 | 0.426 |
| Vacation | 1 = vacation trip; 0 = otherwise | 2849 | 0.337 | 1 | 0 | 0 | 0.473 |
| Business | 1 = business trip; 0 = otherwise | 3482 | 0.412 | 1 | 0 | 0 | 0.492 |
| *Social interaction (base category includes passengers travelling without children and not accompanied to/from the airport)* | | | | | | | |
| Group size | 1 = travelling alone; 2 = 2 people; 3 = 3 people or more. | - | 1.608 | 3 | 1 | 1 | 0.715 |
| Children | 1 = travelling with children; 0 = otherwise. | 495 | 0.059 | 1 | 0 | 0 | 0.235 |
| Seen off | 1 = passenger seen off at airport; 0 = otherwise. | 2264 | 0.268 | 1 | 0 | 0 | 0.443 |

*(Continued)*

**Table 2.** (Continued)

| *Variable* | *Description* | *Observations* | *Mean* | *Maximum* | *Minimum* | *Median* | *S.D.* |
|---|---|---|---|---|---|---|---|
| *Environment (base category includes passengers travelling on workdays and from own primary or secondary home)* | | | | | | | |
| Weekend | 1 = survey was taken on Saturday or Sunday, when taxi rate—either flat or regular—is higher; 0 = otherwise | 1987 | 0.235 | 1 | 0 | 0 | 0.424 |
| Friends or family | 1 = passenger departs from home of friends or relatives; 0 = otherwise | 1876 | 0.222 | 1 | 0 | 0 | 0.415 |
| Hotel | 1 = passenger departs from hotel, boarding house or other paid accommodation; 0 = otherwise | 4532 | 0.536 | 1 | 0 | 1 | 0.499 |

variable (to take a taxi or not), we used another binary response model to estimate this equation. Again, we chose the model that maximised the log pseudo-likelihood.

### 4.3 Calculating the Marginal Effect of $\hat{\alpha}$

Only the sign of coefficients can be interpreted in binary models, which is why marginal effects are usually calculated (see the Appendix, section 3). In our case, the marginal effect of $\hat{\alpha}$ measured the average increase in the likelihood of a passenger choosing a taxi to make the transfer to the airport as a consequence of the taxi flat rate.

## 5. Results

### 5.1 Estimation of the Propensity Score

In Table 3, we summarise the results of estimating the propensity score in the context of the 18 covariates from Table 2. We opted for a logit specification, since it maximised the log pseudo-likelihood (–468 309.81) compared with a probit (–468 330.86).

The resulting coefficients indicate the degree to which each of the 18 considered covariates contribute to the propensity score. The significance of each of the individual covariates is of no importance for our analysis, however. As explained earlier, the purpose of the propensity score is simply to make the individuals from the treatment group (Seville) and the control group (Bilbao and Zaragoza) as homogeneous as possible as far as the 18 covariates are concerned.

### 5.2 Estimation of the Causal Effects

In order to estimate the causal effects, we used a probit specification that maximised the log pseudo-likelihood (–4 567 785.7) in contrast to the logit estimation (–4 567 933.7). The results are shown in Table 4.

The results show the high statistical significance of the different coefficients, including the variables $\left(\hat{\varepsilon}(x_i)\right)^2$ and $\left(\hat{\varepsilon}(x_i) - E\left[\hat{\varepsilon}(x)\right]\right)^2 D_i$, which justifies the use of a non-linear approach instead of a linear model, even under a hypothesis as restrictive as the estimation of a cluster-robust

**Table 3.** Logit estimation of the propensity score ($N$ = 7341)

| *Covariate* | *Coefficient* |
|---|---|
| Sex | 0.014 (0.027) |
| Age | –0.021 (0.023) |
| Nationality | 0.298 (0.179) |
| Currency | 0.340*** (0.046) |
| Homemaker | –0.555*** (0.034) |
| Self-employed | –0.457*** (0.053) |
| Salaried worker | –0.411*** (0.019) |
| Retired | –0.575*** (0.070) |
| Student | –0.698*** (0.134) |
| LCC | –0.025 (0.283) |
| Vacation | –0.295 (0.236) |
| Business | –0.518*** (0.105) |
| Group size | 0.217** (0.096) |
| Children | 0.242 (0.143) |
| Seen off | –0.161*** (0.031) |
| Weekend | 0.111** (0.044) |
| Friends or family | 0.985** (0.457) |
| Hotel | –0.040 (0.411) |
| Constant | 0.091 (1.719) |
| Log pseudo-likelihood | –468 309.81 |
| Pseudo $R^2$ | 0.058 |
| Wald $\chi^2$ without clusters (*p-value*) | 50169.65 (0.000) |

*Notes*: In the coefficient column, standard errors robust to heteroscedasticity and clustered by airport of origin are given in parentheses. ** and *** indicate coefficient significance at the 5 per cent and 1 per cent levels respectively.

**Table 4.** Probit estimation of relevant causal effects ($N$ = 7341)

| *Variable* | *Coefficient* |
|---|---|
| Flat rate ($D_i$) | 0.181*** (0.038) |
| $\hat{\varepsilon}(x_i)$ | –5.391*** (0.042) |
| $\left[\hat{\varepsilon}(x_i)\right]^2$ | 5.788*** (0.004) |
| Constant | 0.712*** (0.019) |
| $\left\{\hat{\varepsilon}(x_i) - E\left[\hat{\varepsilon}(x)\right]\right\} D_i$ | 1.456*** (0.038) |
| $\left\{\hat{\varepsilon}(x_i) - E\left[\hat{\varepsilon}(x)\right]\right\}^2 D_i$ | –2.834*** (0.004) |
| Log pseudo-likelihood | –4567785.7 |
| Pseudo $R^2$ | 0.0303 |
| Wald $\chi^2$ without clusters (*p-value*) | 285065.55 (0.000) |

*Notes*: In the coefficient column, standard errors robust to heteroscedasticity and clustered by airport of origin are given in parentheses. ** and *** indicate coefficient significance at the 5 per cent and 1 per cent levels respectively.

variance matrix. The positive coefficient sign and the significance of variable $D$ at the 1 per cent level allow us to conclude that the flat rate has a positive average effect on the likelihood of a taxi being chosen for the city–airport connection. This will be quantified in the following section.

### 5.3 Obtaining the Marginal Effect of $\hat{\alpha}$

Finally, Table 5 shows the estimation of the marginal effect of $\hat{\alpha}$ at the mean. The

**Table 5.** Marginal effect of $\hat{\alpha}$

| *Variable* | *Coefficient* | *S.E.* | $\bar{D}_i$ |
|---|---|---|---|
| $\partial p/\partial \bar{D}_i$ | 0.068 = 6.804 per cent*** | 0.014 | 0.473 |

estimation of the marginal effect coefficient is statistically significant at the 1 per cent level and shows there is a 6.804 per cent greater likelihood on average of a passenger choosing a taxi to make the transfer to the airport as a result of the city's taxi flat rate.

## 6. Conclusions

A good city–airport connection system is a priority for urban transport policy-makers, but they should be mindful of the fact that the planning of the various airport access options also has a very important influence on urban structure, as analysed in the introduction. The taxi plays a significant role in this city–airport connection system and is essential in cities that lack a rail-based public transport connection or a shuttle service to the airport. This includes tourist cities located in and around many regional Spanish airports, where taxis are one of the elements of the urban tourism system transport component.

Furthermore, the major development of the low-cost carriers in recent decades after the liberalisation of the air sector has considerably increased numbers of passengers, especially at regional airports, but it has also led to an increased need for developing more efficient city–airport connection options with a more competitive price point in the cities to which they fly (see Castillo-Manzano, 2010).

With no economic logic to support them, many of these cities have opted for developing enormously costly new rail-based city links. As an example, in the Andalusia region alone, high-speed rail links of this type have been planned for Jerez, Malaga and Seville airports and construction work has already been started in the last of these. Many of these projects are not viable for medium-sized Spanish cities, given the current economic circumstances, with Spain having lost funding from the ERDF (see Lima and Cardenete, 2008, on the important role played by these funds in the development of investments in transport both in Andalusia and in Spain as as whole) due to the enlargement of the EU eastwards and the obligations that Ecofin has imposed on the Spanish state to reduce the public deficit which will mean a *de facto* freeze on a large number of planned investments and even on some that have already begun.

Compared with these huge investments, in some places, including Seville, the city council has also established a flat rate for the city–airport taxi ride. One of principal advantages of this policy is transparency as it draws a line under the previous asymmetric information situation in the taxicab service. Assuming there is appropriate advertising, the result could be a perfect information scenario providing protection for users, especially non-residents and tourists. The measure proves especially useful in cities like Seville, where historically tourists have frequently been victims of fraud. Other Spanish cities, such as Alicante, Santiago de Compostela and Valladolid have adopted a similar measure in the wake of the Seville experience.

One of the disadvantages of the taxi flat rate is the high cost of negotiating with taxi-driver associations before the measure can be implemented (in Seville, several years of studies and bargaining were required). Evaluations like those of this study are therefore necessary as they might facilitate negotiations.

In other respects, the high cost of negotiating with taxi-driver associations and the pressures they exert witihin cities have thus far led to city halls being reluctant further to broaden the liberalisation of the sector in favour of the shuttle services that are frequently called for by associations in the hospitality industry, for example.

In addition, this measure requires a solid marketing policy so that a lack of knowledge does not result in passengers being charged extras by taxi-drivers, something that was frequent in Seville during the first years after the introduction of the flat rate. Although this problem seemed to have been solved, a recent study by the Spanish consumer association/lobby has stated one result of the flat-rate marketing policies being relaxed in recent years is that foreign passengers are being defrauded more and more frequently. Consequently, although with the passing of time the people targeted by these campaigns have been reduced to non-city residents, and independently of whether it already appears in the tourist guides or not, the rate should continue to be publicly displayed in airport arrival areas, especially around the baggage-claim areas, and taxi-drivers should continue to be made to affix an explanation of the rate in their rear right-hand windows, in both Spanish and English.

Meanwhile, our results establish that a flat rate offers an additional advantage for taxi-drivers as it increases the likelihood that passengers will choose a taxicab as their means of transport to and from the airport by almost 7 per cent on average. If the cost of the flat rate reflects the average cost of the airport–city ride, this would mean just under a 7 per cent increase in taxi-drivers' revenues.[5]

This outcome is congruent with the general microeconomics hypothesis that, as information in dark markets or information-asymmetric markets increases, bringing imperfect competition situations closer to perfect competition, the number of products and services exchanged will also increase. In our particular case, there will be an increase in the taxis' market share of the city–airport connection. The fact that taxi-driver associations recognise this advantage should be sufficient to mitigate some of the measure's disadvantages, such as the cost of the negotiations. Sabotage efforts should be similarly discouraged.

In conclusion, in terms of implications for urban transport planning, the taxi flat rate is confirmed as a useful measure in cities with major problems in their taxicab industries. However, neither passengers needing a ride to or from districts close to the airport nor taxi-drivers who specialise in exploiting tourists will consider it an optimal solution.

## Notes

1. The 'shuttle service' transport model refers to a system of shared minibuses that function like taxis and serve to connect the airport with a hotel or the city centre.
2. According to the WTO (2008), France is the first country considering the number of tourist arrivals, followed by Spain and the US. According to international tourism revenues, the same three countries are in the first three places, but with the US in first place, France in third and Spain still in second place.
3. Even today, taxi-drivers continue to operate using mafia-like practices in the city, grouping together in gangs that control the most lucrative services (airport and high-speed train station routes). According to non-public data, the Airport Authority estimates that, even though all taxi-drivers can take passengers to the airport, only 20 per cent can collect at the airport.
4. There are two fares in the new system: fare 1 (20.70 € in 2009) is for the day schedule on weekdays, while fare 2 (23.08 € in 2009) is for the night and weekend schedules.
5. The longer the negotiation preceding the establishment of the flat rate and the greater the power of the taxi-driver associations, the higher the probability of the rate being above average. Seville is a good example: in September 2003, after a long negotiation process, fare 1 was fixed at 15 € and fare 2 at 18 €. The average fare at that time was approximately 16.5 €, while *The Rough Guide to Andalusia* (4th edition, June 2003) specified that a taxi from the airport to the centre (one of the most distant areas from the airport) costs around 14 €.

## Acknowledgements

The authors would like to express their gratitude to the Spanish Public Airport Authority (AENA) for the data made available for this study to be conducted. The authors are also grateful to Kenneth Gibb, the four anonymous reviewers and Guido Imbens for very helpful comments.

## References

Abadie, A. and Imbens, G. W. (2006) Large sample properties of matching estimators for average treatment effects, *Econometrica*, 74(1), pp. 235–267.

Akerlof, G. (1970) The market for 'lemons': quality uncertainty and the market mechanism, *Quarterly Journal of Economics*, 84(3), pp. 488–500.

Barnow, B., Cain, G. and Goldberger, A. (1980) Issues in the analysis of selectivity bias, in: E. W. Stromsdorfer and G. Farkas (Eds) *Evaluation Studies Review Annual*, 5, pp. 43–59. Beverly Hills, CA: Sage Publications.

Blundell, R., Costa, M., Meghir, C. and Reenen, J. van (2004) Evaluating the employment impact of a mandatory job search program, *Journal of the European Economic Association*, 2(4), pp. 569–606.

Bolduc, D. (1999) A practical technique to estimate multinomial probit models in transportation, *Transportation Research Part B*, 33(1), pp. 63–79.

Bonsall, P., Shires, J., Maule, J. *et al.* (2007) Responses to complex pricing signals: theory, evidence and implications for road pricing, *Transportation Research Part A*, 41(7), pp. 672–683.

Brons, M., Givoni, M. and Rietveld, P. (2009) Access to railway stations and its potential in increasing rail use, *Transportation Research Part A*, 43(2), pp. 136–149.

Brueckner, J. K. (2003) Airline traffic and urban economic development, *Urban Studies*, 40(8), pp. 1455–1469.

Cairns, R. D. and Liston-Heyes, C. (1996) Competition and regulation in the taxi industry, *Journal of Public Economics*, 59(1), pp. 1–15.

Cameron, A. C. and Trivedi, P. K. (2005) *Microeconometric Methods and Applications*. New York: Cambridge University Press.

Cameron, A. C. and Trivedi, P. K. (2009) *Microeconometrics Using Stata*. College Station, TX: Stata Press.

Cantillo, F., Heydecker, B. and Ortúzar, J. (2006) A discrete choice model incorporating thresholds for perception in attribute values, *Transportation Research Part B*, 40(9), pp. 807–825.

Card, D. and Krueger, A. (2000) Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania: reply, *American Economic Review*, 90(5), pp. 1397–1420.

Castillo-Manzano, J. I. (2010) The city-airport connection in the low-cost carrier era: implications for urban transport planning, *Journal of Air Transport Management*, 16(6), pp. 295–298.

Castillo-Manzano, J. I. and López-Valpuesta, L. (2009) Urban retail fabric and the metro: a complex relationship: lessons from middle-sized Spanish cities, *Cities*, 26(3), pp. 141–147.

Castillo-Manzano, J. I. and López-Valpuesta, L. (2010) The decline of the traditional travel agent model, *Transportation Research Part E*, 46(5), pp. 639–649.

Christakis, N. A. and Iwashyna, T. I. (2003) The health impact of health care on families: a matched cohort study of hospice use by decedents and mortality outcomes in surviving, widowed spouses, *Social Science & Medicine*, 57(3), pp. 465–475.

Cox, D. R. (1992) Causality: some statistical aspects, *Journal of the Royal Statistical Society, Series A*, 155(2), pp. 291–301.

Cullinane, S. (2003) Hong Kong's low car dependence: lessons and prospects, *Journal of Transport Geography*, 11(1), pp. 25–35.

Dawid, A. P. (1979) Conditional independence in statistical theory, *Journal of the Royal Statistical Society, Series B*, 41(1), pp. 1–31.

Dawid, A. P. (2000) Causal inference without counterfactuals (with discussion), *Journal of the American Statistical Association*, 95(450), pp. 407–448.

Dehejia, R. H. and Wahba, S. (2002) Propensity score matching methods for non-experimental causal studies, *Review of Economics and Statistics*, 84(1), pp. 151–161.

Dobruszkes, F. (2006) An analysis of European low-cost airlines and their networks, *Journal of Transport Geography*, 14(1), pp. 249–264.

Dresner, M. (2006) Leisure versus business passengers: similarities, differences and implications, *Journal of Air Transport Management*, 12(1), pp. 28–32.

Duch, R. M. and Stevenson, R. (2006) Assesing the magnitude of the economic vote over time and across nations, *Electoral Studies*, 25(3), pp. 528–547.

Eisinger, P. (2000) The politics of bread and circuses: building the city for the visitor class, *Urban Affairs Review*, 35(3), pp. 316–333.

Freestone, R. (2009) Planning, sustainability and airport-led urban development, *International Planning Studies*, 14(2), pp. 161–176.

Gaubatz, P. (1999) China's urban transformation: patterns and processes of morphological change in Beijing, Shangai and Guangzhou, *Urban Studies*, 36(9), pp. 1495–1521.

Gaunt, C. (1996) The impact of taxi deregulation on small urban areas: some New Zealand evidence, *Transport Policy*, 2(4), pp. 257–262.

Givoni, M. and Rietveld, P. (2007) The access journey to the railway station and its role in passengers' satisfaction with rail travel, *Transport Policy*, 14(5), pp. 357–365.

Glazer, A. and Hassin, R. (1983) The economics of cheating in the taxi market, *Transportation Research Part A*, 17(1), pp. 25–31.

Gwilliam, K. (2008) A review of issues in transit economics, *Research in Transportation Economics*, 23 (1), pp. 4–22.

Hahn, J. (1998) On the role of the propensity score in efficient semiparametric estimation of average treatment effects, *Econometrica*, 66(2), pp. 315–331.

Hahn, J., Hirano, K. and Karlan, D. (2010) Adaptative experimental design using the propensity score, *Journal of Business and Economic Statistics* (online).

Heckman, J. J. and Vytlacil, E. (2005) Structural equations, treatment effects, and econometric policy evaluation, *Econometrica*, 73(3), pp. 669–738.

Heckman, J. J., Ichimura, H. and Todd, P. E. (1998) Matching as an econometric evaluation estimator, *Review of Economic Studies*, 65(2), pp. 261–294.

Hirano, K. and Imbens, G. (2001) Estimation of causal effects using propensity score weighting: an application to data on right heart catheterization, *Health Services & Outcomes Research Methodology*, 2(3/4), pp. 259–278.

Hirano, K., Imbens, G. and Ridder, G. (2003) Efficient estimation of average treatment effects using the estimated propensity score, *Econometrica*, 71(4), pp. 1161–1189.

Holland, P. W. (1986) Statistics and causal inference (with discussion), *Journal of the American Statistical Association*, 81(396), pp. 945–970.

Hotz, V. J., Imbens, G. W. and Mortimer, J. H. (2005) Predicting the efficacy of future training programs using past experiences at other locations, *Journal of Econometrics and Statistics*, 125(1/2), pp. 241–270.

Howard, R. W. (2009) Risky business? Asking tourists what hazards they actually encountered in Thailand, *Tourism Management*, 30(3), pp. 359–365.

Hull, A. (2005) Integrated transport planning in the UK: from concept to reality, *Journal of Transport Geography*, 13(4), pp. 318–328.

Imai, K. (2005) Do get-out-the-vote calls reduce turnout? The importance of statistical methods for field experiments, *American Political Science Review*, 99(2), pp. 283–300.

Imbens, G. W. (2004) Nonparametric estimation of average treatment effects under exogeneity: a review, *Review of Economics and Statistics*, 86(1), pp. 4–29.

Imbens, G. W. and Wooldridge, J. M. (2009) Recent developments in the econometrics of program evaluation, *Journal of Economic Literature*, 47(1), pp. 5–86.

Lechner, M. (1999) Earnings and employment effects of continuous off-the-job training in east Germany after unification, *Journal of Business and Economic Statistics*, 17(1), pp. 74–90.

Lima, M. C. and Cardenete, M. A. (2008) The impact of European structural funds in the south of Spain, *European Planning Studies*, 16(10), pp. 1445–1457.

Loo, B. P. Y. (2007) The role of paratransit: some reflections based on the experience of residents' coach services in Hong Kong, *Transportation*, 34(4), pp. 471–486.

Lythgoe, W. F. and Wardman, M. (2002) Demand for rail travel to and from airports, *Transportation*, 29(2), pp. 125–143.

Martínez, F., Aguila, F. and Hurtubia, R. (2009) The constrained multinomial logit: a semi-compensatory choice model, *Transportation Research Part B*, 43(3), pp. 365–377.

McLay, P. and Reynolds-Feighan, A. (2006) Competition between airport terminals: the issues facing Dublin airport, *Transportation Research Part A*, 40(2), pp. 181–203.

Morgan, S. L. and. Harding, D. J. (2006) Matching estimators of causal effects: prospects and pitfalls in theory and practice, *Sociological Methods & Research*, 35(1), pp. 3–60.

Neufville, R. (2006) Planning airport access in an era of low-cost airlines, *Journal of the American Planning Association*, 72(3), pp. 347–356.

Oakes, J. M. and Church, T. R. (2007) Invited commentary: advancing propensity score methods in epidemiology, *American Journal of Epidemiology*, 65(10), pp. 1119–1121.

Pearl, J. (2000) *Causality: Models, Reasoning and Inference Applications*. New York: Cambridge University Press.

Pels, E., Nijkamp, P. and Rietveld, P. (2003) Access to and competition between airports: a case study for the San Francisco Bay, *Transportation Research Part A*, 37(1), pp. 71–83.

Quadrifoglio, L., Hall., R. W. and Dessouky, M. M. (2006) Performance and design of mobility allowance shuttle transit services: bounds on the maximum longitudinal velocity, *Transportation Science*, 40(3), pp. 351–363.

Rendeiro, R. (2006) Tourism service quality begins at the airport, *Tourism Management*, 27(5), pp. 874–877.

Reynolds-Feighan, A. (2001) Traffic distribution in low-cost and full-service carrier networks in the US air transportation market, *Journal of Air Transport Management*, 7(5), pp. 265–275.

Rogerson, C. M. and Rogerson, J. M. (1997) The changing post-*apartheid* city: emergent Black-owned small enterprises in Johannesburg, *Urban Studies*, 34(1), pp. 85–103.

Rosenbaum, P. R. (2002) Covariance adjustment in randomized experiments and observational studies, *Statistical Science*, 17(3), pp. 286–327.

Rosenbaum, P. R. and Rubin, D. B. (1983) The central role of the propensity score in observational studies for causal effects, *Biometrika*, 70(1), pp. 41–55.

Rotnitzky, A. and Robins, J. (1995) Semiparametric regression estimation in the presence of dependent censoring, *Biometrika*, 82(4), pp. 805–820.

Rubin, D. B. (1974) Estimating causal effects of treatments in randomized and non-randomized studies, *Journal of Educational Psychology*, 66(5), pp. 688–701.

Rubin, D. B. (1978) Bayesian inference for causal effects: the role of randomization, *Annals of Statistics*, 6(1), pp. 34–58.

Rubin, D. B. (2008) For objective causal inference, design trumps analysis, *Annals of Applied Statistics*, 2(3), pp. 808–840.

Schaller, B. (2007) Entry controls in taxi regulation: implications of US and Canadian experience for taxi regulation and deregulation, *Transport Policy*, 14(6), pp. 490–506.

Smith, H. L. (1997) Matching with multiple controls to estimate treatment effects in observational studies, *Sociological Methodology*, 27, pp. 325–353.

Smith, J. A. and Todd, P. E. (2005) Does matching overcome Lalonde's critique of nonexperimental estimators?, *Journal of Econometrics*, 125(1/2), pp. 305–353.

Sohail, M., Maunder, D. A. C. and Cavill, S. (2006) Effective regulation for sustainable public transport in developing countries, *Transport Policy*, 13(3), pp. 177–190.

Tsamboulas, D. A. and Nikoleris, A. (2008) Passengers' willingness to pay for airport ground access time savings, *Transportation Research Part A*, 42(10), pp. 1274–1282.

Waryszak, R. and King, B. (2000) Tourists and taxis: an examination of the tourism transport interface, *Journal of Vacation Marketing*, 6(4), pp. 318–328.

Winston, C. and Maheshri, W. (2007) On the social desirability of urban rail transit systems, *Journal of Urban Economics*, 62(2), pp. 362–382.

Wong, K. I., Wong, S. C., Yang, H. and Wu, J. H. (2008) Modeling urban taxi services with multiple user classes and vehicle modes, *Transportation Research Part B*, 42(10), pp. 985–1007.

WTO (World Trade Organisation) (2008) UNWTO World Tourism Barometer (June–08), 6(2).

Yang, H., Wong, S. C. and Wong, K. I. (2002) Demand–supply equilibrium of taxi services in a network under competition and regulation, *Transportation Research Part B*, 36(9), pp. 799–819.

Yang, H., Leung, C., Wong, S. C. and Bell, M. (2010) Equilibria of bilateral taxi–customer searching and meeting on networks, *Transportation Research Part B*, 44(8–9), pp. 1067–1083.

Yang, H., Ye, M., Tang, W. H. and Wong, S. C. (2005) Regulating taxi services in the presence of congestion externality, *Transportation Research Part A*, 39(1), pp. 17–40.

Zhao, J. and Dessouky, M. (2008) Service capacity design problems for mobility allowance shuttle transit systems, *Transportation Research Part B*, 42(2), pp. 135–146.

Zografos, K. G., Androutsopoulos, K. N. and Sihvola, T. (2008) A methodological approach for developing and assessing business models for flexible transport systems, *Transportation*, 35(6), pp. 777–795.

## Appendix

### 1. The Unconfoundness Condition and the Overlap Condition

In the evaluation process, we aimed to guarantee the unconfoundness condition (Lechner, 1999; Rosenbaum and Rubin, 1983; Barnow *et al.*, 1980), also known as the conditional independence assumption. This assumption states that, conditional on *X*, the results are treatment-independent and denoted as (Dawid, 1979)

$$D \perp \left[ Y(1), Y(0) \right] \mid X$$

The unconfoundness condition implies (Cameron and Trivedi, 2005)

$$\mathrm{F}(Y_j \mid X, D=1) = \mathrm{F}(Y_j \mid X, D=0) = \mathrm{F}(Y_j \mid X), j=0,1$$

$$\mathrm{F}(u_j \mid X, D=1) = \mathrm{F}(u_j \mid X, D=0) = \mathrm{F}(u_j \mid X), j=0,1$$

Note that F is the distribution function and *u* the regression model error.

In addition, we required compliance with the overlap condition, according to which every value of vector *X* is associated with a positive probability that allows $D = 0$ (Cameron and Trivedi, 2005; Hotz *et al.*, 2005). This assumption ensures that there are both treated (participant) and non-treated (control) cases associated with each *X* value. Each treated individual matches another individual who has a similar *X* value. The condition can be read as follows

$$0 < P\left(D=1 \mid X\right) < 1$$

### 2. Regression Adjustment to Estimate the Causal Effect

The estimator of $\alpha$ was usually obtained from the following regression by using ordinal least squares (OLS) on the complete sample (Hirano and Imbens, 2001)

$$Y_i = \tau_0 + \alpha D_i + \tau_1' \hat{\varepsilon}(x_i) + \tau_2' \left\{ \hat{\varepsilon}(x_i) - E\left[ \hat{\varepsilon}(x) \right] \right\} D_i + u_i$$

The model linearity hypothesis was surmounted by considering quadratic functions of the terms that include the propensity score (Imbens and Wooldridge, 2009)

$$Y_i = \tau_0 + \alpha D_i + \tau_1' \hat{\varepsilon}(x_i) + \tau_2' \left\{ \hat{\varepsilon}(x_i) - E\left[ \hat{\varepsilon}(x) \right] \right\} D_i + \tau_3' \left[ \hat{\varepsilon}(x_i) \right]^2 + \tau_4' \left\{ \hat{\varepsilon}(x_i) - E\left[ \hat{\varepsilon}(x) \right] \right\}^2 D_i + u_i$$

In our case, *Y* being a binary variable (to take a taxi or not), we used another binary response model to estimate this equation. Again, we chose the model that maximized the log pseudo-likelihood.

In this way, the estimated model was expressed as follows

$$P(Y=1 \mid) = F(\mathrm{x'}\tau) = \left. \frac{e^{\mathrm{x'}\tau}}{1 + e^{\mathrm{x'}\tau}} \right|_{\mathrm{logit}} = \left. \int_{-\infty}^{\mathrm{x'}\tau} \phi(z)\, dz \right|_{probit}$$

Note that $\phi(z)$ is the standard normal density function for $-\infty < z < \infty$ and

$$\mathrm{x'}\tau = \tau_0 + \alpha D_i + \tau_1' \hat{\varepsilon}(x_i) + \tau_2' \left\{ \hat{\varepsilon}(x_i) - E\left[ \hat{\varepsilon}(x) \right] \right\} D_i + \tau_3' \left[ \hat{\varepsilon}(x_i) \right]^2 + \tau_4' \left\{ \hat{\varepsilon}(x_i) - E\left[ \hat{\varepsilon}(x) \right] \right\}^2 D_i + u_i$$

### 3. The Marginal Effect

The expression of the marginal effect for binary models is (Cameron and Trivedi, 2009)

$$\frac{\partial p}{\partial D_i} = \Lambda(\mathrm{x'}\tau)\{1 - \Lambda(\mathrm{x'}\tau)\} D_i \Big|_{\mathrm{logit}} = \varphi(\mathrm{x'}\tau) D_i \Big|_{probit}$$