Available online at www.sciencedirect.com
ScienceDirect
Transportation Research Part A 42 (2008) 673–695
TRANSPORTATION RESEARCH PART A
www.elsevier.com/locate/tra

An empirical model of low-cost carrier entry

Alessandro V.M. Oliveira *

Center for Studies of Airline Competition and Regulation (NECTAR), Aeronautical Institute of Technology, Brazil

Received 7 November 2005; accepted 1 January 2008

Abstract

Competition between low-cost carriers in rapid expansion and full-service network carriers has definitely become one of the most relevant issues of the airline industry. The present paper addresses this matter by analyzing the entry of the low-cost Gol Airlines, in the Brazilian domestic market, in 2001. A route-choice model is estimated by making use of a flexible post-entry equilibrium profits equation and accounting for endogeneity of the main variables. Results indicate the relevance of market size and rival's route presence as underlying determinants of profitability. Furthermore, it is also performed an analysis of the consistency of Gol's entry patterns with the route-choice behavior classically established by the pre-eminent US low-cost carrier Southwest Airlines – that is, a focus on short-haul and high-density markets. Evidence is found that although Gol initiated operations by reproducing the behavior of Southwest, it quickly diversified its portfolio of routes and, at the margin, became more in accordance with the pattern of entry of JetBlue Airways (another successful US low-cost carrier), focusing mainly on longer-haul markets, albeit with some relevant country-specific idiosyncrasies.

Keywords: Entry; Deregulation; Low-cost airline; Discrete-choice; Endogeneity

1. Introduction

Competition between rapidly expanding low-cost carriers (LCC's) and traditional network full-service carriers (FSC's) has recently become one of the most significant issues regarding the airline industry. Although basically a phenomenon of fully or partially liberalized markets – and thus dating back to the US deregulation process of the 1970s – it was only recently, however, that the LCC segment won recognition as a relevant and distinct business strategy as well as a profitable market niche. Following the successful paradigm of the pioneer Southwest Airlines in the United States, airlines such as Ryanair and EasyJet, in Europe, flourished in the market, and soon the concept has spread worldwide. Moreover, this segment is expected to expand considerably within the next few years, and this has undoubtedly been forcing legacy carriers to respond progressively – a movement that is shaping the frontiers of competition in the industry.

* Tel.: +55 1239476969.
E-mail address: A.V.M.Oliveira@gmail.com
URL: http://www.nectar.ita.br

0965-8564/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tra.2008.01.025

The present paper addresses this matter by examining the entry of low-cost carrier Gol Intelligent Airlines in the Brazilian domestic market in 2001. By making use of this case study, one is able to make inferences on the strategy of a successful and fast-growing newcomer LCC in an airline industry with recent liberalization. The ultimate objective here is therefore to inspect Gol's route-choice decisions in order to identify entry patterns which could be associated with notable benchmarks of the LCC niche.

Gol Airlines was not only the first scheduled LCC of Brazil, but also of Latin America. Initially with operations confined to domestic flights,[1] Gol soon represented the most effective threat to the so-called ''Big Four'' legacy majors, Varig, Vasp, Tam and Transbrasil, since the establishment of liberalization in 1992. By offering basic air transport service (without frills), and prices that were lower than the average in the market, and above all with lower costs and careful choice of routes, Gol started a successful path of growth and penetration in the domestic market. The consequence was that, according to Department of Civil Aviation, after only 2 years of operations the carrier was already Brazil's only profitable airline among major players, with 5.8% of operating margin and 12% share in the domestic segment – that is, 3.2 billion pax-km in a market with 26.8 billion pax-km in 2002.[2]

The literature on LCC was, until very recently, rather scarce and the few existing studies were usually related with the investigation of the FSC's pricing behavior in response to entry. Dresner et al. (1996) examined and found significant spillover impacts of LCC entry onto other competitive routes, as on other routes at the same airport and on routes at airports in close proximity to where entry occurred. Their analysis was performed by inspecting, among others, the entry of Southwest Airlines into Baltimore–Washington International Airport, in 1993. Windle and Dresner (1999) investigated the impact of entry by ValuJet into Delta Air Line's hub, Atlanta, and refuted the US DOT's claim that the latter increased fares on non-competitive routes to compensate for lost revenues on competitive routes. More recently, Morrison (2001) assessed the total extent of Southwest Airlines' influence on competition by investigating the impacts of its actual, adjacent and potential route presence, on other carriers' fares in 1998, and reached the conclusion that the presence of the LCC permits saving to consumers that amounted to 20% of US airline industry's domestic schedule passenger revenue for that year.

In contrast, Ito and Lee (2003) and Boguslaski et al. (2004) focus on route entry decisions and entry patterns by LCC's. Whereas the former is aimed at studying the implications for further growth of the LCC's in the US market, by considering their propensity to enter high-density routes, the objective of the latter is to determine and quantify *the market characteristics which have influenced Southwest's entry decisions*. The authors' main conclusions are that LCC's are no longer a niche segment restricted to particular geographic regions or leisure travelers and that the legacy airlines' degree of exposure to LCC competition is very likely to increase from *roughly 30% today to just under 50% in the future* (Ito and Lee, 2003). The authors also find that markets with high traffic density are becoming increasingly contestable, with relevant implications to market structure and competition. Other examples of the empirical airline literature on entry are Berry (1992), Whinston and Collins (1992) and Joskow et al. (1994).

In order to study Gol's entry decisions in the Brazilian domestic market, an empirical model of route choice was designed in the same fashion of Boguslaski et al. (2004). By considering a fairly flexible post-entry equilibrium profits equation, the model is estimated by making use of Newey (1987)'s methodology, and therefore Amemiya's Generalized Least Squares (AGLS) was employed. This approach results in consistent and asymptotically efficient estimation of the parameters of a limited-dependent variable, such as the newcomer's entry decisions, for the case of the presence of some endogenous regressors. The use of a specific estimator to account for endogeneity of regressors in an empirical entry model constitutes the main contribution of the paper.

Results indicate the relevance of market size and rival's route presence as underlying determinants of profitability. Unobservables at the airport/city levels, such as sunk costs and economies of scope, are also found to be significant. Furthermore, the consistency of Gol's entry patterns with the route-choice behavior classically established by Southwest Airlines for the LCC segment – short-haul and high-density markets – is

[1] Actually, Gol started entering international routes only recently, by serving Guarulhos Airport (São Paulo) – Ezeiza Airport (Buenos Aires).

[2] *Source*: Statistical Yearbook of Department of Civil Aviation, 2002.

investigated; evidence is found that although Gol initiated operations by reproducing standards of Southwest, the carrier quickly diversified its portfolio of routes and became more in accordance with JetBlue Airways' entry pattern, focusing mainly on longer-haul markets, albeit with some relevant country-specific idiosyncrasies.

This paper has the following structure: Section 2 portrays the background of the entry of Gol Airlines in the Brazilian airline industry, with a description of the main paradigms related to LCC entry patterns along with some facts about the deregulation process in Brazil and Gol Airlines. Section 3 presents the empirical model and the econometric issues. Section 4 reports the results and includes an analysis of Gol's entry patterns consistency. Section 5 presents concluding remarks.

2. Background: LCC niche and entry of Gol Airlines in Brazil

2.1. The LCC market niche and its paradigms

The entry of low-cost carriers providing basic air transport service with no frills and lower fares has considerably transformed competition in the airline industry. Notwithstanding a phenomenon of partially or fully liberalized airline markets dating back to the US deregulation process of the seventies, it was only recently that this *low-cost revolution* (Doganis, 2001) has resulted in the formation of a well recognized and distinct business strategy and a sustainable market niche.

The LCC niche is usually associated with the *Southwest Airlines Paradigm* (hereafter *SWP*), mainly because the US carrier pioneered this sort of operations with standards that are deliberately reproduced around the world.[3] The most widely known characteristics of this paradigm are (Silva and Espírito Santo, 2003): single fleet type; simplification or elimination of in-flight service; use of less congested secondary airports; direct sales to consumers; e-ticketing; short-haul, point-to-point flights in dense markets with no interlining or transfers, which means a simple network structure, with absent or weak feed to long-range flights; single-class cabin layout; simple or no frequent-flyer program; high level of fleet utilization; and highly motivated employees.[4] Moreover, LCC's are typically associated with a very aggressive pricing strategy, typically with the use of a simplified fare structure with few or no restrictions, and low one-way fares.[5]

The cost advantage permitted by the SWP is not merely an issue of paying lower salaries or operating at cheaper airports, and, contrary to common sense, not even due only to the lack of frills; instead it is rather a function of fundamental differences in the business model associated with it, emerging mainly from a very careful choice of markets, targeting dense, short-haul routes in order to exploit economies due to higher seating density and higher aircraft utilization, especially with non-stop service. According to Boguslaski et al. (2004), Southwest has resulted in unit costs that are 28–51% lower than the US major airlines, considering 2001 US DOT unit cost figures.[6]

Since the early nineties, and in particular very recently, a plethora of *de novo*, LCC entry, has been observed around the world. Inspired by the more than three decades of success of Southwest Airlines, and stimulated by market liberalization, airlines such as Ryanair and EasyJet in Europe, Air Asia and Virgin Blue in the South Pacific, 1Time and Kulula in Africa, and Gol and U Air in South America, flourished in the market, meaning that the LCC concept has rapidly achieved global recognition.

In parallel to the worldwide spread of the low-cost operations based on the SWP, alternative standards for the segment have been successfully implemented in the United States. The *AirTran-Frontier Paradigm* (*AFP*),

3 As the Chief Executive of Ryanair (UK) once said: *We went to look at Southwest. It was like the road to Damascus. This was the way to make Ryanair work* (Doganis, 2001).

4 This description refers to what can be considered "classic" Southwest paradigm. One has to bear in mind, as we will see below, that Southwest's actual patterns of operations has had some changes recently: *its strategy evolved during the latter half of the decade to include a much more heterogeneous mix of markets, including a number of markets which were both long-haul and surprisingly thin* (Boguslaski et al., 2004).

5 Tretheway (2004) points out that the introduction of low one way fares ultimately served to undermine the ability of the FSC's to price discriminate, and not only resulted in a considerable increase in competition but also in an exposure of the problems associated with the FSC business model.

6 Moreover, Southwest Airlines' average stay length is shorter than most other carriers, understating its unit cost advantage.

with a clear focus on the low-fare business market by making use of multi-service operations, usually with mini-hubs to provide convenient connections and more possibilities in terms of origin-and-destination markets, and with a more complex fare structure and even business class[7]; and secondly, the *JetBlue Airways Paradigm* (*JBP*), which is associated with the focus on long-haul routes (usually more than 1500 km), resulting in the highest average stage length of the LCC segments.[8]

And finally, one has the Ryanair Paradigm (RYP). Ryanair, the most successful LCC in the world in terms of profitability, is extremely focused in selecting only destinations with underutilized secondary and tertiary airports. Contrary to all other major LCC's like EasyJet or even Southwest Airlines, Ryanair has always resisted entering either primary or congested airports; this permitted her to always be in a position to negotiate airport fees (cost control) and to focus almost completely on the leisure travelers segment, which makes this paradigm completely distinct from, for example, the AFP.

It is important to emphasize two caveats on the above-mentioned paradigms. First, while newer standards of operation have clearly emerged in the segment, the essence of the SWP remains dominant for most of LCC's, namely the absence or weak presence of frills and the lower costs, typically resulting in low prices. From this point of view, the SWP is still the major benchmark for LCC's. In addition to that, it is clear that, due to the ever-changing state of the competition in deregulated airline markets, it is rather unlikely to observe the three above-mentioned paradigms in a very strict basis, but rather as a *mixture* of them. Indeed, the volatile frontier of competition along with the need of market expansion have forced LCC's to also enter atypical markets, with relevant examples being the recent entry of Southwest in the coast-to-coast markets of the United States (US Department of Transportation, 2001, 2002), and even more recently, into major carriers' hubs. This trend has resulted in LCC's serving a variety of short/medium/long-haul, business/tourism, direct/indirect routes, which has ultimately increased the exposure not only to FSC competition but also among LCC's.

Nevertheless, even with carriers having a more diversified range of routes nowadays, it is clear that, by making use of the notion of paradigms as benchmarks one has useful reference in order to analyze and pinpoint patterns of entry behavior by LCC's. Nevertheless, despite those caveats, my argument here is that the idea of paradigms is still useful for a better identification of actual LCC behavior, especially because the industry all over the world is observing a wave of small entrants, most of them claiming to be LCC's. Actually it is not always straightforward to identify if a given airline is really operating as a LCC and this causes much confusion among industry analysts. Therefore, in order to better analyze and pinpoint patterns of entry behavior by entrants that are truly consistent with LCC operations, some form of representation, such as the paradigms, is needed.

For example, one can study a carrier's marginal propensity to enter a market with respect to flight haul in order to make inferences on its conformity with either the SWP or the JBP. The analysis of preference with respect to flight haul, in this specific case, is crucial to distinguish between both paradigms; in addition, a thorough investigation of the commercial and operational practices of the newcomer (with respect to strategic choices regarding frills, network, airports, labor, advertising, etc.), can be very effective in permitting its correct classification.

Fig. 1 presents a diagram with this sort of analysis.

As Fig. 1 shows, the probability of entry of a SWP-like LCC is increasing in-flight haul but with diminishing returns, in such a way that the highest probability is associated with relatively shorter-haul markets. On the other hand, a JBP-like LCC has typically an ever-increasing entry probability with respect to flight distance, with highest levels associated with long-haul domestic flights.

Therefore, by performing a simple inspection of the marginal effects of distance on the probability to enter a market by an alleged LCC, it is possible to have a straightforward analysis of consistency with either SWP or

[7] AirTran Airways operates in the eastern United States with Atlanta as its hub, being the second-largest carrier at Hartsfield International Airport, and providing service to 45 cities within the country. Frontier Airlines operates routes linking its Denver hub to 38 cities in 22 states and Mexico.

[8] With operations started in 2000, JetBlue Airways focuses on non-stop transcontinental routes in the US. The airline serves point-to-point routes between 22 destinations in 11 states, Puerto Rico and the Dominican Republic. It is important to emphasize that both JBP and AFP are usually considered in a different category from Southwest when it comes to passenger amenities and in-flight entertainment (IFE).

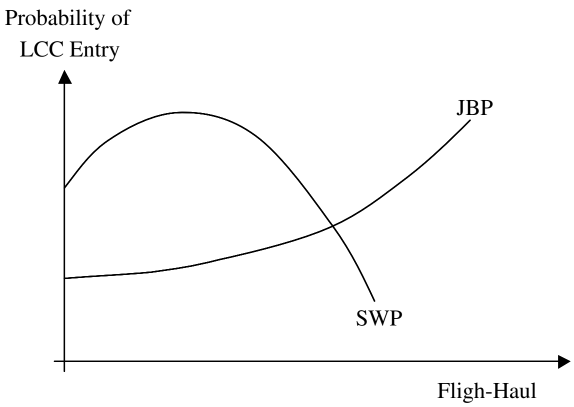

Fig. 1. SWP versus JBP: effects of flight haul on LCC entry probability.

JBP. Similarly, it would be possible to make inferences on the conformity of a given carrier with AFP by inspecting, for example, its degree of hubbing and propensity to enter business-related cities.[9]

2.2. LCC entry in Brazil

The removal of regulatory barriers in the Brazilian airline industry since the early nineties was undoubtedly a major determinant of Gol Airlines' entry in 2001, and of the unprecedented increase in competition in the market. Started at the beginning of the nineties within a broader governmental program for deregulation of country's economy, regulatory reform was then performed gradually, in three main rounds, by the Department of Civil Aviation, (DAC).

Gol Intelligent Airlines was not only the first scheduled LCC of Brazil, but also of Latin America, with operations commencing in January 2001. It is owned by Grupo Áurea – a conglomerate that owns 38 companies and a major operator of urban and long-distance coach services across Brazil.

By offering a very simple fare structure, and thus with an average yield that at the beginning was 31% below those of FSC competitors,[10] Gol started a successful path of growth and penetration in the domestic market. After only 2 years of operations, it was already Brazil's only profitable airline with operating profits of BRL 39 million[11] and an operating margin of 5.8% of total revenues. Table 1 presents some characteristics of Gol, compared with the major legacy airlines within the country in 2002; Gol's figures of 2001 are also presented to demonstrate the airline's rapid growth. One can see that Gol's unit costs and yields were roughly a third lower than its competitors and its average stage length was approximately 20% lower; also, it is possible to visualize the pace of expansion of the LCC, which, from the start-up year, 2001, to 2002, increased RPK's by 156% and RPK market share by 151%:

Some additional characteristics of the newcomer are: absence of complete food service (only snacks and cereal bars); standardized fleet (Boeing 737–700s and 800s, the largest operator of Next-Generation 737 aircraft in Latin America); availability of full e-ticketing service and heavy distribution via the internet (65% of sales, according to Silva and Espírito Santo, 2003); reservation system software acquired from JetBlue ("Open Skies"); around half of the original staff coming from outside the industry and half recruited from other airlines – especially flight crew and technical staff – although not more than 15% from any particular carrier.[12]

[9] In this case, however, an analysis of a carrier's overall service attributes is probably more useful to infer the conformity with AFP than a focus on route entry decisions. None of them are accomplished in this paper, however.

[10] Statistical Yearbook of Department of Civil Aviation, volume II, 2001.

[11] BRL means Brazilian'currency (Real). Approximately, 13 million dollars (exchange rate of 2003).

[12] According to Lima (2002), hiring personnel from other carriers was made easier due to the downsizing process taken place at Vasp and especially at the bankrupt Transbrasil (Lima, 2002). According to Silva and Espírito Santo (2003), Gol had the following internal slogan: *the youngest and most experienced airline in Brazil.*

Table 1
Comparison of Gol and incumbent FSC's (2002) – domestic market[a]

Variable	Unit	*FSC*			*LCC*		
		Tam 2002	Varig 2002	Vasp 2002	Gol 2002	Gol 2001	Growth 2001–2002 (%)
Revenue passenger kilometers (RPK)	pax * km (billion)	9.34	10.48	3.39	3.22	1.26	156
Market share RPK	fraction	0.35	0.39	0.13	0.12	0.05	151
Traffic per employee	pax * km (million)	1.23	0.75	0.70	1.56	1.08	44
Load factor	fraction	0.53	0.59	0.55	0.63	0.60	5
Unit cost	BRL/(pax * km)	0.33	0.33	0.31	0.20	0.18	10
Yield	BRL/(pax * km)	0.29	0.31	0.27	0.21	0.18	14
Operating margin	fraction	−0.12	−0.05	−0.16	0.06	0.02	153
Average stage length	km	868	1017	1016	792	772	3

Notes: (i) BRL means Brazilian currency (real, current values); (ii) pax means number of passengers traveled; (iii) market share is equal to a firm's RPK over industry's RPK; (iv) operating margin is equal to operating profits or losses over total revenues.
[a] *Source*: DAC's Statistical Yearbook, vols. I and II.

At the beginning, Gol's marketing efforts were clearly oriented to become "the people's airline", concentrating more on potential travelers with lower income than on current traveling public (Zalamea, 2001, mentions *small business officials, blue collar workers, students, farmers and others who have never flown before* as targeted segments of consumers). For example, Tarcisio Gargioni, Gol's Vice President for Marketing and Services, once revealed: *Our business plan identified that in 2000, out of the 170 million Brazilian population only 6 million flew commercial aviation. Out of the remaining 164 million, some 25 million could also become potential fliers provided fares were reduced 30%* (Lima, 2002).

Nevertheless, demand stimulation from non-traveling, lower-income consumers was eventually not enough to guarantee the expansion rate targeted by the airline; in fact Gol's rapid growth was achieved primarily at the expense of existing legacy carriers, and also enhanced by Transbrasil Airlines' exit in 2001: *We did a market survey in September [2001] and found only 4% of our passengers had never flown before* (Gargioni, as in Lima, 2002). Undoubtedly, Brazil's macroeconomic instability, lower *per capita* income and high wealth concentration can be regarded as major sources of Gol's lack of success in attracting the non-traveling public. Also, the country's high interest rates, which increase the risk of enterprise, probably forced Gol not to venture providing service to new domestic destinations where new demand could be created, but to focus only on already existing routes. Another relevant restriction is the lack of infrastructure for typical LCC operations in other cities than the major conurbations in Brazil (suitable secondary airports, for example).

This does not mean, however, that Gol's entry was totally ineffective in stimulating new demand on *existing* routes. On the contrary, if one considers the top-500 densest routes in Brazil, and by comparing traffic density of 2002 with 2000 (previous to entry), it is possible to arrive at the conclusion that routes entered by Gol observed a 13.1% average increase in traffic density (pax), against a 7.0% increase on all 500 routes. In fact, routes not entered by Gol had an 11.5 decrease in traffic density within the same period.[13]

A major issue is whether the above-mentioned difficulty in attracting new travelers and therefore not generating the amount of additional traffic expected of a LCC,[14] has ultimately forced the airline to substantially alter its initial route entry strategy in order not to affect growth. Indeed, this may be particularly true with respect to the effect on route choice of flight haul – as seen before, a crucial variable with respect to analysis of conformity with LCC paradigms. For example, it was observed that, since 2002, medium-to-long-haul routes were increasingly added to Gol's network, as one can see from the maps of Fig. 2:

At the beginning of 2001, Gol had only six 737–700's, providing service between São Paulo, Rio de Janeiro, Belo Horizonte, Florianópolis, Brasília, Porto Alegre and Salvador (i.e. the maximum haul below 1500 km).

[13] Calculations based on figures of DAC's Statistical Yearbook (Volume I). Results are consistent with findings of Dresner et al. (1996), for the US market, and the so-called "Southwest Airlines Effect".
[14] Although some increase in traffic density was actually observed after Gol's entry (see Section 4).

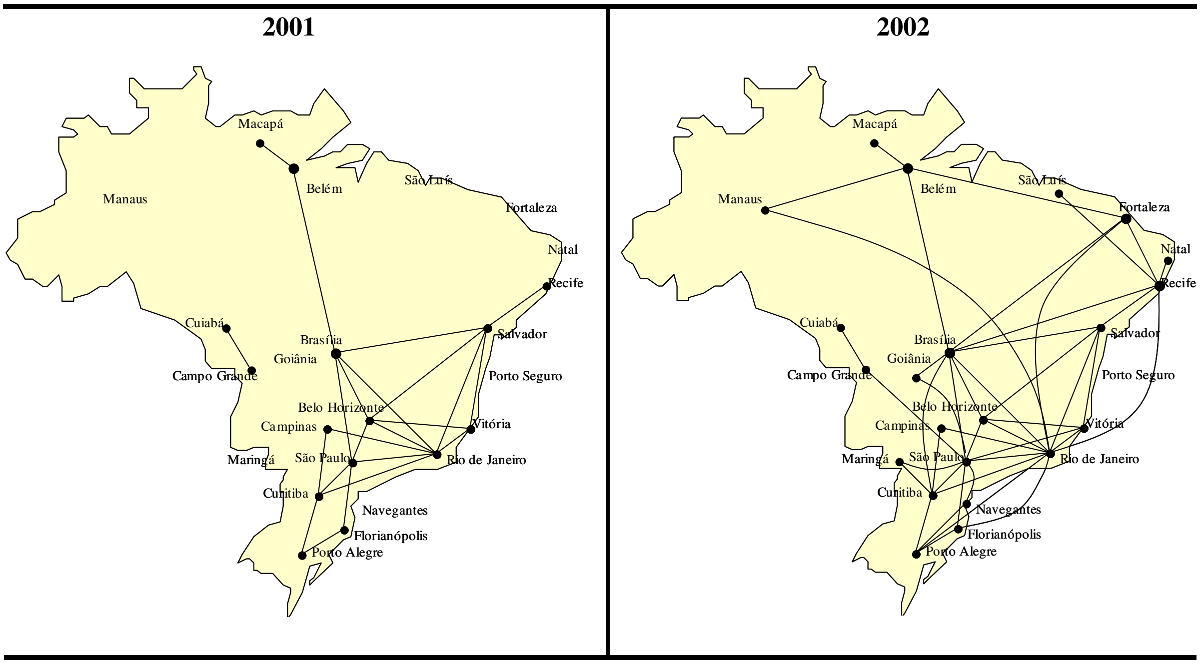

Fig. 2. Evolution of Gol's network within Brazil (*Source*: DAC's HOTRAN reports).

This straightforward link with the SWP is not surprising since, during the start-up of operations, Gol admitted it was reproducing the model of Southwest Airlines in the same way Ryanair did in the United Kingdom (Guimarães, 2002). In fact, this recipe permitted the newcomer to rapidly achieve levels of efficiency that were notably higher-than-average, with aircrafts having 10–12 flights a day and very fast ground turn-around times, within 15–30 min. In fact, by December 2001 there was only one city-pair in the entire network which could be classified as direct long-haul route: Brasília–Belém, with 1610 km.

By the end of 2002, on the other hand, the situation was clearly very different. Gol had 22 aircraft in operation, serving a much wider network with many routes with longer-than-average distance and certainly an additional target of feeding long-range flights. For example, routes like Rio de Janeiro–Manaus (2860 km), Rio de Janeiro–Recife (1863 km) and Brasília–Fortaleza (1690 km) were added to the network structure, indicating a higher propensity to enter direct, long-haul routes and rapidly increasing the possibilities of connections between extreme regions like the South and the North/Northeast.

Table 2 gives some details on the route profile of the airline with respect to flight haul, by considering entry on the top-500 densest routes:

Table 2
Direct routes served by Gol – flight-haul distribution

Flight-haul intervals (km)	#500 Top routes	Direct routes served 2001		Direct routes served 2002	
		#	%	#	%
Q_0–Q_1 less than 390	125	8	6.4	11	8.8
Q_1–Q_2 390–716	125	11	8.8	15	12.0
Q_2–Q_3 716–1466	125	9	7.2	18	14.4
Q_3–Q_4 more than 1466	125	5	4.0	10	8.0
Total	500	33	6.6	54	10.8

Notes: (i) Q_1, Q_2, Q_3 and Q_4 mean the quartiles considering a sample with the 500 densest routes; (ii) # means number of routes and % means percentage out of the top routes.

As one can observe in Table 2, Gol increased by 21 the number of direct routes served from 2001 to 2002 (54–33). Out of these 21 new routes, two-thirds (14) were medium-to-long-haul routes (that is, with flight haul above the median, 716 km). Indeed, as Gol doubled its presence on longer-haul routes in 2002 (28 routes above the median, against 14 in 2001), it consolidated a very different portfolio of routes, notably atypical for a LCC; in fact, the portfolio changed from a set in which there was a minority of long-haul routes in 2001 (14 out of 33), to a slight majority in 2002 (28 out of 54).

All these facts raise questions about the actual standard of operations undertaken by Gol in the Brazilian airline industry, especially with respect to the paradigm with which it might be consistent. One might doubt whether Gol, although claiming itself as initially inspired by Southwest Airlines (Guimarães, 2002), could resist entering a wider range of markets in order to expand or even to exploit unobservable (to the analyst) economies of scope throughout Brazil, increasing the number of actual origin-and-destination markets served. In fact, by a simple inspection on Gol's website, one can quickly arrive at the conclusion that flights with more than two stops and/or connections are much more frequently available than non-stop flights, which clearly represents a departure from the typical SWP.

The start of operations of "red-eye" flights in 2003 in order to attract more travelers from coach and to persist in expanding serves as an additional argument to the claim that the LCC's standards are probably not consistent with the SWP, but could be potentially associated with a variant of the JBP; JetBlue Airways is well known for its red-eye, transcontinental flights in the US market. In fact, it is known that, just before starting-up operations in Brazil, Gol's executives made visits to both Southwest Airlines and JetBlue Airways in order to conceive the newcomer's strategic planning.

By analyzing Gol's entry patterns, it is possible to collect further evidence on the change of directions by the LCC from 2002 and beyond and to make inferences about the determinants of entry decisions by a LCC in a recently liberalized airline market, which is the focus of Section 3.

3. Empirical modeling

In this section I present the empirical modeling of route entry decisions of Gol Intelligent Airlines. Firstly, the LCC's route entry problem is analyzed under a discrete-choice model framework; secondly, issues like sample selection, functional form and chosen empirical specification are discussed. And finally, the issue of endogeneity is examined along with a discussion of the instrumental variables and the estimator employed. Empirical results are presented in Section 4.

3.1. Discrete-choice framework

The main objective here is to develop a framework of discrete choice with random utility[15] for the analysis of the patterns of entry decisions of the newcomer Gol Airlines. It is straightforward that here we have Gol as the decision maker, and the set of decisions "to enter a route" and "not to enter a route" as the alternatives in this "route-choice problem". Consider the binary variable representative of choice, PRES_{kt}, which accounts for the presence of Gol on the kth route at time t. The probability of entry can then be regarded in the following way:

$$\Pr[\text{PRES}_{kt} = 1] = \Pr[\delta\pi^*_{kt} - SC_k > 0] \quad (1)$$

where the multiplicative term $\delta\pi^*_{kt}$ is the present value of the stream of equilibrium profits of the newcomer (δ is the discount factor) in case of entry. SC_k is the amount of sunk costs on the kth route. One can develop (1) in the following way:

[15] In the random utility approach, *the observed inconsistencies in choice behavior are taken to be a result of observational deficiencies on the part of the analyst* (Ben-Akiva and Lerman, 1985); therefore, contrary to the constant utility approach, which assume a probabilistic behavior for the decision maker, by assuming random utility I assume that the individual always select the alternative with the highest utility (profits). By doing this, here we have the standard interpretation of the error term as representing factors that are observable to the firm but not to the econometrician.

$$\Pr[\delta\pi^*_{kt} - SC_k > 0] = \Pr\left[\frac{\delta\pi^*_{kt}}{SC_k} > 1\right] = \Pr[\ln \pi^*_{kt} + \ln\delta - \ln SC_k > 0] \tag{2}$$

By introducing ε_{kt}, the disturbances associated with the choice mechanism within a random utility framework, in (2), we have the following random variable representative of equilibrium net present value profits at the route level (Π^*_{kt}):

$$\Pi^*_{kt} = \ln \pi^*_{kt} + \ln\delta - \ln SC_k + \varepsilon_{kt} \tag{3}$$

where ε_{kt} is assumed to be iid $\sim N(0, 1)$.[16]

As in a typical discrete-choice model (ex. Amemiya, 1978), we have only PRES_{kt} as an observable, whereas the other terms (δ, π^*_{kt} and SC_k) are latent. Actually, only the sign of Π^*_{kt} is observed:

$$\text{PRES}_{kt} = \tau(\Pi^*_{kt}) = \begin{cases} 1 & \text{if } \Pi^*_{kt} > 0 \\ 0 & \text{if } \Pi^*_{kt} \leqslant 0 \end{cases} \tag{4}$$

Therefore, we have PRES_{kt} assigned with one in case of entry (expectation of positive route profitability) and zero in case of no entry (no expectation of route profitability).

3.2. Sample selection, functional form and empirical specification

I now turn to the description of the sample and the empirical specification. The objective is to have a sample with a large and representative cross-section of routes, in terms of capturing a high percentage of total domestic traffic in Brazil. Fortunately, the Statistical Yearbook of the Department of Civil Aviation – volume II,[17] provides annual figures of domestic origin-and-destination traffic; thus, data were collected for the 500 densest routes in the country,[18] structured in a panel with 2 years, 2001 and 2002, and with each observation being then a route-year.

I define a "route" in the following way: routes are considered here as non-directional origin and destination markets aggregated to the city level, which means that, for instance, journeys from Rio de Janeiro to Brasília are aggregated with the journeys from Brasília to Rio de Janeiro. For example, Ito and Lee (2003) and Richard (2003) also makes use the assumption of non-directional markets; on the other hand, Berry et al. (2006), Evans and Kessides (1993) and Borenstein (1989) use directional markets. Therefore, the raw data available in the Statistical Yearbook (directional airport-pairs) were aggregated to represent non-directional city-pair markets.

Also, some additional procedures of sample selection were performed in order to reduce potential heterogeneity across routes, specially with respect to demand attributes such as the price elasticities, implicit in any specification of π^*_{kt}. More specifically, it is well known that flight distance and trip purpose are relevant sources of heterogeneity across routes. In fact, one would expect higher price elasticities of demand on routes in which there is abnormally higher competition either within modes of transportation or between scheduled and charter airlines. In the Brazilian case, one would certainly have this sort of problem with very short-haul routes – which engender lower relative disutility associated with coaches, for example – and with exceptionally highly tourism-related routes – in which there is higher availability of charter flights.[19]

[16] This is a convenient assumption, as the literature on binary probit estimation within a simultaneous equations framework is vast (examples being Amemiya, 1978; Smith and Blundell, 1986; Rivers and Vuong, 1988 and Lee, 1991), in opposition to the binary logit with endogenous variables.

[17] All information contained in the yearbooks and reports used here is monthly supplied from all scheduled airlines to the Department of Civil Aviation according to specific legislation (Instrução de Aviação Civil – 1505). One of the drawbacks of the data is that one cannot distinguish direct from indirect traffic (i.e., number of travelers with and without flight connections). These sorts of data are currently unavailable in Brazil.

[18] In order to define the 500 densest, figures of the period 1998–2002, available in the Statistical Yearbooks of the Department of Civil Aviation, were considered.

[19] Indeed, the Brazilian airline industry is characterized by a high proportion of business-related traffic, with tourism-related routes being exceptions. According to a research performed by São Paulo's aviation authorities, DAESP, in 2002, approximately 60% of the passengers in domestic trips that traveled from or to that state's airports had business-related purposes of travel.

In order to deal with this problem, I excluded routes with unusually short haul and/or a high percentage of available weekend seats – the latter considered a reasonable proxy for identifying tourism-related routes. Therefore, sample selection was conducted in the following way: first, exclusion of routes with flight haul that is lower than the 5th percentile (160 km, as measured for the top-500 densest routes sample)[20]; and second, exclusion of routes with a unusually high percentage of available seats during weekends – higher than the 95th percentile within the top-500 routes sample.[21] This resulted in a final data sample with 448 routes. With this set, one could still be able to capture the traffic of an average of 27 million passengers per year, which represents approximately 96.66% of domestic trips during that entire period.

With respect to the *functional form specification* of (3), a translogarithmic function is used. This specification has advantages and disadvantages. On the one hand, it can be regarded as flexible in such a way that it can represent any equilibrium profits function of unknown form and does not impose restriction on the substitution elasticities between the arguments (permits a full modeling of substitution and complementarity). On the other hand, however, the translogarithmic can be viewed as limited as multicolinearity may emerge among its many terms, and thus not being suitable for highly disaggregate models. As the number of second-order terms in the right-hand side increases quickly as the list of independent variables increases, there is usually a trade-off between the increased flexibility permitted by having higher order terms and the practical difficulties associated with a elevated number of parameters to be estimated; examples of flexible profit functions of this type in the empirical literature are Mullineaux (1978) and Slade (1986).

With respect to the *empirical specification* of (3), there is a large list of potential candidates for variables to be included as regressors, and many are indicated by the literature. As the major focus here was to analyze the conformity of Gol Airlines with either the SWP or the JBP – especially with respect of flight haul and route density – and, at the same time accounting for the effects of market structure at the route level (presence of the opponent FSC's), the chosen empirical specification was then:

$$\Pi^*_{kt} = \Pi[\pi^*_{kt}(\mathrm{den}_{kt}, \mathrm{km}_k, \mathrm{sdr}_{kt}), \mathrm{cpres}_{kt-1}, \mathrm{tbapres}_{kt-1}, \mathrm{DC}_l, \varepsilon_{kt}] \tag{5}$$

where den_{kt} is route density on route k and time t, km_{kt} is flight distance on route k, and sdr_{kt} is the number of seats available on direct flights of FSC rivals per passenger on route k and time t; cpres_{kt-1} and $\mathrm{tbapres}_{kt-1}$ are dummy variables that control for, the presence, at time $t-1$, of Gol at the endpoint cities, and of the bankrupt Transbrasil on route k. Finally the DC's ($l = 1, 2, \ldots, L$) are city-specific dummies. The translog representation of (5) would then be

$$\Pr[\mathrm{PRES}_{kt}=1] = \Pr[\Pi^*_{kt}>0] = \Pr\begin{bmatrix} \upsilon_0+\upsilon_1\ln\mathrm{den}_{kt}+\upsilon_2\ln\mathrm{km}_k+\upsilon_3\ln\mathrm{sdr}_{kt}+\upsilon_4\ln\mathrm{den}^2_{kt}+ \\ \upsilon_5\ln\mathrm{km}^2_k+\upsilon_6\ln\mathrm{sdr}^2_{kt}+\upsilon_7\ln\mathrm{den}_{kt}\ln\mathrm{km}_k+\upsilon_8\ln\mathrm{den}_{kt}\ln\mathrm{sdr}_{kt}+ \\ +\upsilon_9\ln\mathrm{km}_k\ln\mathrm{sdr}_{kt}+\sum_l u_l\mathrm{DC}_l+v_1\mathrm{cpres}_{kt-1}+v_2\mathrm{tbapres}_{kt-1}+\varepsilon_{kt}>0 \end{bmatrix} \tag{6}$$

where the υ's, u's and v's are parameters. Let us now present details of each of the variables present in (5) and (6):

PRES_{kt}, is a binary variable that accounts for route presence of Gol Airlines, being assigned with one if entry has occurred in any year lower or equal than t, and with zero otherwise. PRES_{kt} then means the presence of the LCC on route k in year t. Here I follow the partial equilibrium approach of Berry (1992), assuming that

[20] Availability of charter flights could not be observed and therefore it was not possible to assess the impact of charter airlines on Gol's entry decisions. Boguslaski et al. (2004) also had a procedure of sample selection regarding distance when studying the route choice of Southwest Airlines. They had a cut-off range of 100 and 3000 miles. In the present case, however, there is no route with more than 3000 miles in the initial data sample. The authors excluded markets with distance outside this interval as they are not likely to be targets for Southwest Airlines entry (the minimum and maximum distance of Southwest's markets was, respectively, 152 and 2438 miles). In the present case, the only market of Gol which was outside this range (Florianópolis-Navegantes, approximately 55 miles) has been discontinued (not available in Gol's website in February 2003).

[21] As mentioned before, this measure can be regarded as a proxy for identifying tourism-related routes. The average is 21% of total whole-week seats available and the 95th percentile is 35%. *Source*: Department of Civil Aviation's HOTRAN reports (various).

at the beginning of each period, firms take their overall network structure as given and then decide whether to operate or not in a given route.

"Route" is defined as discussed before (see Section 3.2); therefore one has to precisely define "entry". Here I define entry as Gol's presence in a given origin-to-destination (O–D) market, within the period under consideration (2001–2002); for more details, see the discussion of cpres_{kt-1} below. The information of the presence of Gol in the O–D markets was collected from Panrotas's Domestic and International Schedules and Fares Guide and from Airwise's website.[22]

The definition of PRES_{kt} in O–D markets is certainly in contrast with Boguslaski et al. (2004), which considers only non-stop markets and thus disregards routings with flight connections and stops within a given route. That procedure is certainly more reasonable for their case of Southwest Airlines, which is usually associated with non-stop and short-haul flights (the SWP, as discussed in Section 2.1). In the present case, however, one may be unconvinced whether Gol has typical SWP standards, but, on the contrary, one would believe that it has some propensity to enter a more diversified range of markets, especially longer-haul routes with stops and connections (see Section 2.2). Therefore, it would not be a reasonable procedure to include only either non-stop or direct flights in the definition of "route", as it would not be representative of Gol's operations; also, it would be impossible to investigate the conformity of Gol with either SWP or JBP (see Fig. 1). Therefore, broader definitions of both route and entry were considered in this study.

Another issue regarding the definition of entry is related to the minimum level of operations (MLO) within a year for Gol's presence to be accounted for. The previous literature usually had either absolute or relative definitions of MLO. For instance, whereas Oum et al. (1997) and Berry (1992) used MLOs of, respectively, 100 and 90 passengers per quarter in the 10% sample collected by the US DOT,[23] Evans and Kessides (1993) used a fractional definition, considering effective presence as more than 1% of total traffic on the route. The latter is certainly a more flexible filtering criterion which could be adapted for the Brazilian conditions; however, as here traffic disaggregated by airline is not observed, a proxy used was to adapt Evans and Kessides (1993)'s approach, and therefore use the minimum percentage of seats available at the endpoint cities,[24] considering then "entry" when actual figures are higher than 1%.

den_{kt} is route density of traffic (in million passengers) and was collected from the Statistical Yearbook of the Department of Civil Aviation (volume II) for the years 2001–2002. Consisting of direct flights passengers traffic figures, this variable represents total (non-airline-specific) domestic number of trips, aggregating all single-trip and round-trip traffic of revenue passengers.[25]

km_k represents route distance, that is, the one-way distance between origin-and-destination airports. This information was provided by Department of Civil Aviation's Laboratory of Simulation and was calculated by using the polar coordinates method. One important issue about km_{kt} is related to its calculation in case of an observation with more than one airport in one or both endpoint cities of the given route. In both cases the latitude and longitude of the airports closest to the city centre was employed and considered representative of the distance between cities.[26]

Another aspect of km_k it that it represents the *minimum* distance between two given airports, and therefore does not take into account neither actual airway distance nor the effect of flight connections and/or stops. In principle, one would object using this proxy for flight distance, especially for medium-to-long-haul routes because their higher availability of seats in flights with stops represents a higher actual distance flown than can be assessed by km_k.

Besides that, the lower the participation of non-stop flights on one given route the more one would underestimate the effect of actual flight distance on profits, specially because the higher distance would permit lower

---

22 Panrotas is a database that is similar to OAG's of flight schedules guide, the world's most comprehensive schedules database; Airwise's website: http//www.airwise.com.

23 US Department of Transportation (DOT) Origin and Destination Survey.

24 *Source*: Department of Civil Aviation's HOTRAN's (various).

25 Only non-connecting traffic was considered due to data availability.

26 As mentioned before, there were only three cities in this situation found in the data sample: Rio de Janeiro, Sao Paulo and Belo Horizonte. In all cases the largest city airport (in terms of figures of number of passengers and movement of aircrafts) is located closer to the city centre. *Source*: INFRAERO's website (February, 2004).

unit costs – a phenomenon known as "cost taper" in the transport literature, see Brander and Zhang, 1990. One has to be cautious with that argument, however, as more stops are also known for increasing costs – for instance, by additional landing/departure fees and higher fuel consumption; besides that, on the demand side, stops usually increase passengers' flight disutility, generating competitive disadvantage and also reducing profitability – a product differentiation effect. In spite of these arguments, we can therefore interpret km_{kt} as capturing the broad effect – the balance of the two effects mentioned above – of flight distance on the probability of entry by the LCC.[27]

sdr_{kt} is the number of total seats available per passenger on direct flights of FSC's on route k and time t. A relative measure, that is seats *per* passenger, was considered better than the absolute figure of seats available, as it avoids strong colinearity with den_{kt}. Figures of total number of flights disaggregated by airline is available in Department of Civil Aviation's HOTRAN, "Horário de Transporte", a data system that generates reports containing operational information of all scheduled flights within the country (non-published data). This information was extracted from their system in every month for the period 2001–2002, and was subsequently aggregated by year. Sdr_{kt} is then both a measure of product differentiation – that is, more seats available meaning more convenient flights and service levels generated by the FSC's – and of the degree of how well or underserved a given route actually is.

cpres_{kt-1} is a binary variable that controls for entry at any of the endpoint cities of a route in the previous year. It is assigned with one for each route in 2002 that had one of the endpoint cities entered by Gol in 2001, and zero otherwise. This variable is crucial in the present framework as it is designed to control for the effects of sunk costs at the city level, that is, once one given airport/city was entered, it is then easier for the newcomer to provide services to other routes out of it, as it had already incurred in most of the sunk costs, such as start-up costs, advertising, sales and operational structure at the airport. Therefore, it is useful to control for the fact that the probability that Gol enters a given route B–C is related to the probability that Gol has entered route A–B, for example.

Finally, cpres_{kt-1} also controls for the effects of routes already entered in the year before and thus making a distinction between "new" or "true" entry and "previous" entry; in this sense, this variable makes this study in line with Toivanen and Waterson (2005).[28]

tbapres_{kt-1} is a binary variable representative of the presence of the bankrupt (and no-longer existing) Transbrasil Airlines on route k and time $t-1$. Transbrasil started ceasing operations in 2000 and fully exited the industry in the middle of 2001. On the one hand, one would think of this exit as an opportunity to any newcomer; this would be certainly the case in city-pairs with endpoint airports subject to slots; on the other hand, however, one would think of the exit of Transbrasil as a clear signal for Gol that those routes were not profitable.

DC_l ($l = 1, \ldots, \text{L}$), which are *city-specific dummies*: assigned with 1 if the city is one of the endpoint cities of the city-pair, and 0 if not. The city dummies provide an economical way to capture and control for a large number of truly significant variables, which can be regarded as being actually city-specific, instead of route-specific; also, most of them are in fact either unobservables or potentially measured with noise by the researcher.

Below is a list of some of the potential effects that may be controlled by the city dummies:

(a) *Characteristics of travelers from/to a city*: consumers' purchasing behavior and attributes, like the percentage of the travelers which frequently makes use of the internet when searching and buying; consumers' attributes: income, niche preferences, propensity to make either tourism-related or business-related trips, etc; city's gross domestic product and wealth in general, as a factor of business-trips generation; percentage of migrants established in one city (ex: large participation of migrants from Northeast in São Paulo, a fact that is potentially trip-generation enhancing); geographic idiosyncrasies;

[27] Some colinearity with sdr_{kt} is expected *ex-ante*, however.

[28] The authors criticize the common procedure found in the entry literature, of treating "entry" of existing firms, meaning the continuation of operations, in a similar way to entry of new firms; this would be equivalent to *making the assumption that firms can in every period, without a change in costs, review their entry decision*. They state: *Given sunk costs, we believe that particular assumption to be unwarranted and we do no utilize it* (Toivanen and Waterson, 2005, p. 3).

(b) *Characteristics of the airports of a city*: sunk costs; airport accessibility and costs of the access to the airport (price of taxi, distance from the zones-of-trip-generation, etc.), the size of the zone of influence of the city's airport(s) in terms of trip generation (nearby cities); operational costs and expenses related to a particular city (airport fees, cost of hiring personnel, cost of contracts in general, etc.); presence of airports with constrained capacity (slots) or with spare capacity; subsides and incentives given by authorities to operations in one given city; presence of airports owned by the public enterprise Infraero; vacant slots or frequencies left by the bankrupt Transbrasil; airport/airway infrastructure: size of the runway, air traffic control capacity;
(c) *Characteristics of network out of a city*: size of the airlines' network (unobservable degree of product differentiation, economies of scope, etc. at the airport level); presence of hub or mini-hub in a city; frequent flyer effect: number of possible destinations out of a city;
(d) *Characteristics of the competition in a city*: quality/availability of motorways out of a city; tolls; presence of charter; number of travel agents; commission fees to travel agents of a city; airport dominance by particular airlines; number of airlines operating out of a city and concentration levels; levels of advertising and forms of effective media in one city; number of flights and excess capacity out of a city.

Most of these effects are expected to generate persistent heterogeneity in the error-term structure *across cities* (but not routes), which can be controlled via the city-specific dummies, DC's.

Another relevant feature of the dummy-specific cities is that one is able to identify only the effects of actually entered cities. This is a common problem of any discrete-choice model, in which one cannot use as a regressor a dummy variable if for any of the values it takes, there is no variation in the dependent variable (see, for example, Toivanen and Waterson, 2005). This is precisely the case of non-entered cities, all of them with no variation in PRES_{ikt}, that is, in which $\mathrm{PRES}_{ikt} = 0$ for all routes.[29] On the other hand, it is not a good procedure to create one dummy for each actually entered city, as this would be certainly inducing somewhat artificially designed correlation with the dependent variable, due to the obvious fact that only routes from and to actually chosen airports will be entered. The extreme alternative, namely the drop of all city dummies, would probably be inappropriate as it would induce omitted variables bias.

Thus, in order to balance between the gains of controlling for effects which are city-specific and to avoid the aforementioned sort of artificial correlation, I then focused on the *network decisions of any potential newcomer* in the Brazilian domestic market. In fact, given that *there are no secondary airports near major Brazilian cities able to handle midsize jet operations (737s, A320/319, etc.)* (Silva and Espírito Santo, 2003), any major player considering entering the market would not be able to avoid having operations at the airports of some of the most important cities within the country. Indeed, this is a sort of network decision that is expected *ex-ante*, irrespective of the type of operations and specific niche of the potential competitor. This evidence can be regarded as a justification for the inclusion of dummies for the major cities present in the sample, as they constitute the potential mini-hubs for any entering carrier; at the same time, one would not be causing unreasonable correlation with the dependent variable, as the dummies are designed independently of Gol's entry decision.

Therefore, here I used dummies to control for the effects of ten cities, namely, Brasília, Belo Horizonte, Curitiba, Manaus, Fortaleza, Porto Alegre, Recife, Rio de Janeiro, Salvador and Sao Paulo; this was the list of the top-10 cities in terms of total density of traffic from 1998 and 2002 (*source*: Statistical Yearbook of DAC, vol. I).

Descriptive statistics of some variables used in the empirical model are presented in Table 3.[30]

[29] The other extreme would be the case of the sample containing cities with all routes actually entered (that is, $\mathrm{PRES}_{ikt} = 1$ for all routes), and thus generating the same problem of lack of variation in PRES_{ikt} – a case not present in the current data sample.

[30] It is relevant to emphasize that both den_{kt} and sdr_{kt} have zero as minimum. This is because of routes in which air transport operations were either interrupted or there were no direct flights in a given year. This generated the problem of dealing with the logarithm of zero in (6). One way to circumvent this problem is by having the data transformation indicated by Fox (1997): *[to] add a positive constant (called "start") to each data value to make all the values positive*. Hence, "starts" were applied to all observations of both variables in order to permit accomplishing proper estimations.

Table 3
Descriptive statistics

Variable	Designation	Mean			Standard deviation (full sample)
		$PRES_{lkt}=0$	$PRES_{lkt}=1$	Full sample	
$pres_{ikt}$	Route presence – Gol	–	–	0.198	0.398
den_{kt}	Route number of PAX/Year	21,552.220	236,466.475	64,007.379	246,566.718
km_{kt}	Route distance	966.817	1383.838	1049.197	776.293
sdr_{kt}	Route direct seats per PAX	2.827	2.626	2.787	7.362
$cpres_{kt-1}$	City presence – Gol (one lag)	0.309	0.599	0.366	0.482
$tbapres_{kt-1}$	Route presence – Transbrasil (one lag)	0.050	0.113	0.063	0.242

3.3. The issue of endogeneity, instruments and estimator

One relevant issue related to the estimation of (6) is the potential correlation of den_{kt} and sdr_{kt} with the error term ε_{ikt}. In fact, one would expect both variables to be jointly determined with Π^*_{kt} and thus causing simultaneous equations bias to emerge. The correlation would be in the following fashion: if actual profits are higher than predicted profits, that is, a positive ε_{kt}, which stimulates entry, then route density may be higher due to new demand generation permitted by the LCC (a fact reported by Whinston and Collins, 1992), and thus one would have positive correlation between den_{kt} and ε_{kt}.[31] A similar effect is expected to happen with sdr_{kt}: a positive ε_{kt} would cause post-entry reactions by FSC's via increase in capacity (either flight frequency and/or aircraft size) and therefore increase in sdr_{kt} (also reported by Whinston and Collins, 1992). Of course, the opposite may happen in case of a "crowding-out" effect caused by Gol's entry, that is, FSC rivals reducing sdr_{kt} after Gol enters via reduction in the number of flights, in the size of aircrafts or even exiting the market.[32] In both cases, with either positive or negative correlation with the error term, the standard probit estimation would either overestimate or underestimate the true effects on entry as one would not account for post-entry route density and capacity adjustment in the estimation.

As endogeneity is potentially present, one needs to perform a test for exogeneity in the model; the variables under suspicion were den_{kt}, sdr_{kt} and their second-order terms sdr^2_{kt}, $sdr_{kt} * den_{kt}$, $sdr_{kt} * km_k$, den^2_{kt}, $den_{kt} * km_k$. The test employed was the one suggested by Smith and Blundell (1986), which is more suitable for discrete-choice models than, for example, the frequently used Hausman test. It is Chi-squared distributed with m degrees of freedom – m being the number of endogenous variables in the model – and tests the null hypothesis that all explanatory variables are exogenous; a rejection therefore indicates that the standard probit should not be employed. For the present model, the Smith–Blundell statistic was 14.58 (P-value of 0.04), which results in the rejection of the null at 5% of level of significance.[33]

Once exogeneity of den_{kt} and sdr_{kt} (and related terms) is rejected, one needs an instrumental variables estimator for binary dependent variables. Moreover, GMM estimation would be required in case of rejection of the hypothesis of homoskedasticity of ε_{kt}. In order to test for this, a likelihood-ratio test of heteroskedasticity in the discrete-choice framework was performed after a maximum-likelihood heteroskedastic probit estimation. This test requires the specification of an indicator vector of suspected explanatory variables that could affect the unobservables, which, in this case, was set equal to $[sdr_{kt-1}, den_{kt-1}, km_k]$.[34] The null hypothesis of homoskedasticity was not rejected at 10% level of significance – the χ^2 statistic with 3 degrees of freedom was 1.57 (P-value of 0.6671).

As homoskedasticity is not rejected, one possible discrete-choice estimator that control for endogeneity is Amemiya (1978)'s Generalized Least Squares (AGLS); here I employed the AGLS implementation of Newey (1987). In the case of disturbances that are normally distributed, this estimator is consistent, and asymptotically equivalent to the efficient minimum χ^2 estimator (Lee, 1991 and Newey, 1987); also it is shown to be more

[31] The prior expectation here is, as discussed before, that Gol's entry usually induced increase in route density.
[32] For example, Transbrasil exited the market some months after Gol's entry.
[33] The list of instrumental variables used for this test (and for estimations) is described below.
[34] According to Baum et al. (2003), when testing for heteroskedasticity in a simultaneous equation framework, the indicator vector must be exogenous and is typically formed by "either instruments or functions of the instruments".

efficient than other popular two-stage estimators for simultaneous equations with binary response models (for example, the 2SIV estimator of Rivers and Vuong, 1984.[35])

The steps of AGLS estimation are the following: in the first stage, a set of regressions is estimated by OLS to obtain the reduced form parameters and the respective residuals are computed; this is followed by running a probit with the exogenous variables, the predicted endogenous variables and the residuals as regressors; then, in the final stage, a generalized least square estimator is performed in order to obtain efficient estimates of the structural parameters. This estimator requires consistent standard errors correction to account for the first-stage estimation, which is performed here by making use of Newey (1987)'s approach.[36]

The basic procedure for identification here was to employ lagged variables as instruments and test for their validity. The list of instrumental variables was den_{kt-1}, sdr_{kt-1}, stt_{kt-1} (total direct seats available), swe_{kt-1} (total direct seats available during weekends) and asz_{kt-1} (average size of aircraft); it also comprised the respective second-order terms: $(\ln \text{den}_{kt-1})^2$, $(\ln \text{sdr}_{kt-1})^2$, $\ln \text{den}_{kt-1} * \ln \text{km}_k$, $\ln \text{den}_{kt-1} * \ln \text{sdr}_{kt-1}$, $\ln \text{km}_k * \ln \text{sdr}_{kt-1}$. The validity of instruments is supported by the following diagnostics described below.

Firstly, in terms of *relevance* of the instruments, by having a look at the matrix of correlations between endogenous and instrumental variables one can have an idea of the reasonably high correlation among them (Appendix 1, Table 12).

Also in terms of relevance of the instruments, by inspecting the partial R^2 and the F-test of joint significance of the excluded instruments in the first-stage regressions; the minimum R^2 was 0.55 and the minimum F-statistic was 105.20 (P-value of 0.00), which further indicated they are fairly correlated with the endogenous variables (Appendix 1, Table 13).

Since the number of instruments exceeds the number of endogenous regressors I made use of overidentification restrictions tests to check for the *validity* of the instruments proposed (tests of orthogonality, as in Davidson and MacKinnon, 1993; see Baum et al., 2003, for a survey); by regressing a linear probability model in two-stages least squares (LPM/2SLS) one could confirm the validity of instruments. The tests used were the Sargan $N * R^2$ test and the Basmann test, and both failed in rejecting the null hypothesis that the excluded variables are valid instruments (Appendix 1, Table 13).

With the intention of emphasizing the relevance of controlling for endogeneity, I perform comparison between the standard (single stage) probit with the AGLS in the results presentation of Section 4; this is specially useful to have an idea of the magnitude (and sign) of the underlying simultaneous equations bias.

4. Estimation results

Table 4 presents the results obtained from models with and without simultaneous equations bias correction – respectively, AGLS in column (1) and ordinary Probit in column (2):

By analyzing the results presented in column (1), one can see that, in spite of the expected multicolinearity among the terms of the translog profits equation, most variables related to density (den_{kt}), distance (km_k) and rivals' presence (sdr_{kt}) are significant at 5% level. cpres_{kt-1} is very significant and with a positive sign, and thus some evidence is found on the existence of sunk costs at the city level; tbapres_{kt-1} is also significant and with a negative sign, meaning that routes exited by a bankrupt (Transbrasil) were probably not perceived as profitable by the newcomer – and thus one can reject the hypothesis that Transbrasil's exit served as a market opportunity for Gol.

In order to investigate the *full effects* of variables den_{kt}, km_k and sdr_{kt}, one needs a measure that takes into account the coefficients of all related terms, and thus including the quadratic and interactions terms of Eq. (6). Therefore, *arc elasticities* were extracted, calculated by computing the full effect on the estimated probability of an arbitrary percentage change in each variable; the elasticity was considered a better measure than the marginal effects as it is invariant to the unit of measure. Figures were calculated by making use of the formula $(\text{Pr}_1–\text{Pr}_0)/0.10$, where Pr_1 is the predicted probability with the explanatory variables at the sample mean,

[35] Blundell and Smith (1989) and Rivers and Vuong (1988) provide additional discussion on relative efficiency of the AGLS estimator in comparison to others found in the literature.

[36] Stata's routine "ivprob" was used to perform all estimations and standard error corrections in Newey (1987)'s fashion (Harkness, 2001).

Table 4
Estimation results

Dependent variable	PR [ENTRY = 1]	
	(1) AGLS	(2) PROBIT
$\ln \text{den}_{kt}$	0.078_*	0.057
	(0.039)	(0.032)
$(\ln \text{den}_{kt})^2$	0.003	$0.006^{‡}$
	(0.003)	(0.002)
$\ln \text{km}_k$	$0.464^{†}$	$0.562^{†}$
	(0.211)	(0.205)
$(\ln \text{km}_k)^2$	$-0.033^{†}$	$-0.038^{†}$
	(0.015)	(0.015)
$\ln \text{sdr}_{kt}$	-0.183_*	-0.105_*
	(0.072)	(0.054)
$(\ln \text{sdn}_{kt})^2$	0.012_*	$0.006^{†}$
	(0.006)	(0.003)
$\ln \text{den}_{kt} * \ln \text{km}_k$	-0.012_*	−0.007
	(0.006)	(0.005)
$\ln \text{den}_{kt} * \ln \text{sdr}_{kt}$	0.014_*	0.005_*
	(0.006)	(0.003)
$\ln \text{km}_k * \ln \text{sdr}_{kt}$	0.015	0.011
	(0.009)	(0.007)
cpres_{kt-1}	$0.049^{‡}$	$0.061^{‡}$
	(0.023)	(0.021)
tbapres_{kt-1}	$-0.035^{†}$	$-0.042^{†}$
	(0.016)	(0.014)
Control for endogeneity	Yes	No
Second-order terms	Yes	Yes
LR χ^2 statistic	$46{,}200^{‡}$	$460.25^{‡}$
#Predicted = 0/#Actual = 0	673/719	683/719
#Predicted = 1/#Actual = 1	118/177	109/177
Lave–Efron Pseudo-R^2	0.493	0.514
McKelvey–Zavoina pseudo-R^2	0.803	0.750
No. of observations	896	896

Notes: (i) marginal-effects reported; (ii) standard errors in parentheses; (iii) $*$ means significant at 10%, † at 5% and ‡ at 1% level; (iv) city-specific dummies not reported.

except for the one under analysis, which is increased by 10% (represented by 0.10), and Pr_0 is the predicted probability holding all variables at the sample mean.[37] It is important to emphasize that, as here we have a translog specification, which engenders interactions between terms, the elasticity of any variable always depends on the values of the other variables in the model.

Figures in Table 5 are interpreted in the following way: if, for instance, den_{kt} is increased by 10%, the probability of entry (at the sample mean) is increased by 15%. By examining the differences (in percentage) between estimated elasticities between estimators, one can see that a significant bias is generated when endogeneity is not controlled. As expected, there seems to be a positive bias related to density (the difference between estimators is +88%), indicating that this variable is positively correlated with the error term, and, as discussed before, this is likely due to new demand generation caused by LCC entry.[38] Likewise, the positive simultaneity bias caused by not controlling for endogeneity of sdr_{kt} (associated with the +7% difference between elasticities) provides some evidence that LCC entry causes FSC presence to adjust upwards, and therefore providing basis

[37] In terms of the binary explanatory variables, cpres_{kt-1} and tbapres_{kt-1} were set equal to zero and the DC's were considered by extracting the average effect times two (as each route has two endpoint cities).

[38] As mentioned in Section 1, routes entered by Gol had 13.1% increase in traffic density against a 7.0% increase on all 500 top-routes, when comparing figures of 2002 (after entry) with 2000 (previous to entry).

Table 5
Estimated elasticities

Variable	(1) AGLS	(2) PROBIT	(2)–(1) %
denkt	0.151	0.283	88
km_k	0.061	0.194	220
sdr_{kt}	0.166	0.179	7

Notes: (i) figures calculated at the sample mean; (ii) calculated as a 10% increase in each variable at the mean.

for the rejection of the hypothesis of "crowding-out"; this is consistent with Whinston and Collins (1992)'s results of an increase in 25% of incumbents' seats offered in response to low-cost airline entry.[39]

The last coefficient, km_k, has large positive difference between estimators (+220%); although flight distance is not *per se* an endogenous variable, its full effect measured by the elasticity in Table 5 is formed by endogenous variables, namely, the second-order terms $\ln den_{kt} * \ln km_k$ and $\ln km_k * \ln sdr_{kt}$. On account of these interactions, one would expect that, *ceteris paribus*, the true sensitivity of an additional kilometer to be lower in case of higher demand generation and higher presence of competitors – which is caused by the simultaneity bias of, respectively, den_{kt} and sdr_{kt}.[40]

We now turn to the analysis of the signs and magnitudes of the estimated elasticities (the AGLS column). From Table 5 one can see that the elasticities of the original, not log-transformed, variables den_{kt}, km_k, and sdr_{kt} were, respectively 0.151, 0.061 and 0.166, all measured at the sample mean. Apart from the results of den_{kt}, which can be naturally thought of having positive overall effects – that is, the more is a given route's density of traffic the more it is attractive for LCC entry – special attention is required with respect to the analysis of the effects of km_k, and sdr_{kt}.

Firstly, we have an overall positive elasticity of sdr_{kt}, considering everything else held constant at the sample mean. The immediate conclusion implied by this result is that the higher is the presence of the FSC competitors in terms of seats available on direct flights (*per* route passenger) the higher is the propensity to enter of Gol; in other words, the more is the market underserved by direct FSC supply the less is the entry probability. On the one hand, one could interpret this finding as an indication that Gol does not follow the typical LCC practice of avoiding market contact with the legacy carriers but, quite the opposite, prefers behaving like a follower, learning from the others' past entry decisions in order to make its own route choices. The "learning" argument is in line with the results of Toivanen and Waterson (2005): *Structural form estimations show that the positive effect of rival presence on the probability of entry is due to firm learning: rival presence increases the estimate of the size of the market*. Also, this would be clearly suggestive that route presence is quite an indication of underlying profitability, in opposition to Evans and Kessides (1993), which found evidence only for airport presence effects in the US market.

On the other hand, however, one could have the "market niche" argument of the LCC's: by positioning itself close to well-served direct markets, Gol is able to detect market opportunities once not perceived by the FSC's; this is especially true if one observes that, contrary to both SWP and JBP, and as discussed before, Gol provides a wider range of origin-and-destination products with stops and flight connections, and therefore placing in the market as the low fare alternative for less time-sensitive passengers.

Table 6 presents a disaggregation of the elasticity of sdr_{kt} with respect to own values of that variable, with both den_{kt} and km_k held constant; one can observe decreasing but always positive elasticity figures, which means that a point of probability maximization is reached at higher levels of sdr_{kt}. This pattern confirms that Gol has a lower preference for creating new markets or entering underserved routes, contrary to the SWP.

Table 7 presents another disaggregation of the elasticity of sdr_{kt}, with respect to kilometers and density, this time holding sdr_{kt} constant at the mean. Table 7 is quite useful in showing a detailed analysis of Gol's

[39] For an analysis of FSC's price responses to Gol's entry, see Oliveira and Huse (2004).

[40] One has to be cautious with those arguments, however, as here we have only the difference between the estimated effects, and not the real simultaneity bias. The difference between the results of the estimator is indicative of the problems engendered by not controlling for the endogeneity of the regressors, however.

Table 6
Sdr_{kt} disaggregated elasticities (1)

sdr_{kt}	0.70	1.00	1.50	3.00	6.00	10.00
Elasticity	0.82	0.66	0.45	0.14	0.02	0.00

Notes: (i) figures calculated holding km_k and den_{kt} at the sample mean; (ii) calculated as a 10% increase in each variable at the mean; (iii) values of sdr_{kt} are representative of the following percentiles: 0.35, 0.50, 0.65, 0.80, 0.90 and 0.95.

Table 7
Sdr_{ikt} disaggregated elasticities (2)

den_{kt}	km_k						
	350	500	750	1150	1850	2250	2600
1000	2.32	2.31	2.26	2.21	2.20	2.22	2.24
3000	2.55	2.36	2.15	1.98	1.89	1.88	1.90
6000	2.40	2.11	1.83	1.62	1.51	1.50	1.51
15,000	1.85	1.50	1.20	0.99	0.89	0.88	0.90
50,000	0.80	0.53	0.35	0.25	0.21	0.22	0.23
150,000	0.12	0.06	0.03	0.02	0.01	0.02	0.02
300,000	0.01	0.00	0.00	0.00	0.00	0.00	0.00

Notes: (i) figures calculated holding sdr_{kt} at the sample mean; (ii) calculated as a 10% increase in each variable at the mean; (iii) values are representative of the following percentiles: 0.20, 0.35, 0.50, 0.65, 0.80, 0.90 and 0.95.

route-choice preferences regarding opponents' presence. Indeed, it is possible to observe two regimes: one, for the great majority of the routes, of ever positive elasticities – for routes with density below 150,000 pax/year– and one with elasticities that are almost null – associated with very high density routes (for Brazilian standards), above 150,000 pax/year. This likely means that opponent's presence is a good indicator of underlying profitability for low-to-medium sized markets (in terms of density of traffic) but it is irrelevant for high-sized ones. In other words: actual market size is much more observable for the newcomer the higher is traffic density, and for routes in which traffic is rather thin, opponents' presence becomes a better signal for entry.

To sum up on the effects of sdr_{kt}, one has, contrary to the traditional Industrial Organization literature, that rival's market presence does not inhibit entry but, on the contrary, is used as a warning sign for underlying profitability (mainly in markets with lower size). This is consistent with the results of Toivanen and Waterson (2005) which unveiled learning processes regarding entry.[41] There are three explanations for these results: first, as Brazil's very high interest rates are well known for increasing the risk of enterprise, firms usually prefer not taking additional risk of venturing to create new markets; second, the airline market all over the world has been highly volatile and uncertain in the past few years; and third, as regulators were stimulating entry and forcing entry barrier to vanish, it was relatively easy for Gol to enter the same markets of its opponents and without much competitive disadvantage in terms of slots, access to airport facilities, etc.

The other result that needs to be carefully addressed is related to the marginal effects of km_k. A more detailed analysis of this variable is not only essential for proper understanding of the model's most relevant outcomes but also for performing an analysis of Gol's consistency with either SWP or JBP, detailed in Section 1. The positive elasticity of flight haul, presented in Table 5, does not reveal much as it is a rather aggregate figure, measured at the sample mean; once again, one useful alternative is to extract the same measure for a broader set of combinations of density and flight-haul values:

As one can see in Table 8, Gol's propensity to enter a route is marked by diminishing returns of flight haul, and with steadily decreasing effects of density. Again, one can observe two regimes: first, for routes with traffic

[41] In the present case, one could think of Gol learning from the FSC's previous entry decisions, that is, from the number of direct seats available, in order to infer the amount of business traffic on a given route, which is usually associated more with direct flights.

Table 8
Disaggregated elasticities of km_k

den_{kt}	km_k						
	350	500	750	1150	1850	2250	2600
1000	15.59	10.05	6.16	3.62	1.83	1.28	0.92
3000	9.22	5.90	3.54	1.99	0.90	0.56	0.33
6000	6.28	3.95	2.30	1.24	0.49	0.26	0.10
15,000	3.39	2.02	1.09	0.53	0.16	0.04	−0.04
50,000	1.00	0.50	0.22	0.08	0.01	−0.02	−0.04
150,000	0.12	0.04	0.01	0.00	0.00	0.00	−0.01
300,000	0.01	0.00	0.00	0.00	0.00	0.00	0.00

Notes: (i) figures calculated holding sdr_{kt} at the sample mean; (ii) calculated as a 10% increase in each variable at the mean; (iii) values are representative of the following percentiles: 0.20, 0.35, 0.50, 0.65, 0.80, 0.90 and 0.95.

density values up to approximately 50,000 pax/year, where distance has an ever increase effect on entry (although with diminishing returns), probably meaning that Gol is willing to substitute density by kilometers since it is able to force passengers to have stops or connect; this seems to be in line with a modified version of the JBP. And second, for routes with very thick density (higher than 50,000 pax/year), flight haul has no influence on entry; this is the outcome of the same factors affecting the elasticities of sdr_{kt} on the same set of routes, as seen above.

One would claim, however, that Gol changed operational standards from 2002 on, as discussed in Section 2, and probably started to enter a broader range of markets, especially with respect to long-haul routes and flight connections. This might be due to the opportunities emerged after some events of 2001, such as the exit of Transbrasil Airlines, the barriers to expansion at São Paulo city, the DAC's authorization to operate Rio de Janeiro's city-centre airport, and the fiercer incumbents' reactions on short-haul routes (Oliveira and Huse, 2004).

If the above argument is correct, however, the aggregated 2001–2002 regressions of Table 4 would present a rather "average" entry behavior, and disaggregation with respect to time would then be required. In order to perform that, variables $\ln km_k$, $(\ln km_k)^2$, $\ln den_{kt} * \ln km_k$ and $\ln km_k * \ln sdr_{kt}$ were multiplied by a dummy representative of year 2002, in order to test for possible structural change from that year on; thus the following variables were generated: $\ln km_k * d02$, $(\ln km_k)^2 * d02$, $\ln den_{kt} * \ln km_k * d02$ and $\ln km_k * \ln sdr_{kt} * d02$. Table 9 reports the results for the same AGLS estimates but with those variables included:

By making use of the results of Table 9, it is possible to compare the elasticities of km_k across flight distance and route density *disaggregated by year*, in order to inspect how Gol's sensitivity to kilometers changed from 2001 to 2002. Tables 10 and 11 report the results.

Using Table 10 one can assess Gol's entry strategy in its start-up year. In this case a parabolic curve is clearly observed, meaning that the highest probability of entry is located within 1150 and 1850 km for most cases; this could be associated with the SWP; in comparison, the average stage length of Southwest Airlines in 2003 was approximately 1160 km.[42] On the other hand, in Table 11, the estimates for 2002 resulted in an ever-positive flight-haul elasticities for any route density lower than 150,000 pax/year, and this is certainly more in line with JBP. Therefore, we have that a pattern of entry that was in accordance with the SWP, as observed in 2001, seems to be replaced by a propensity to enter a more diversified set of routes, and thus also considering higher flight sectors in 2002 (JBP). In both cases, however, a set of almost null elasticities for thick-density routes was observed.

The aforesaid findings noticeably reject the notion that Gol follows a pure standard of operations like the SWP or the JBP, but, consistently with a recent trend in the LCC segment, preferred to develop a more diversified portfolio of markets. Some evidence is found, however, that, for a great deal of medium-sized markets, Gol behaved more consistently with the SWP, but this was limited to its first year of operations; in contrast,

[42] *Source*: Southwest Airlines Annual Reports (2003).

Table 9
Estimation results disaggregated by year

Dependent variable	PR [ENTRY = 1]
	AGLS
$\ln \text{den}_{kt}$	0.067
	(0.038)
$(\ln \text{den}_{kt})^2$	0.004
	(0.004)
$\ln \text{km}_k$	0.525‡
	(0.222)
$(\ln \text{km}_k)^2$	−0.039‡
	(0.016)
$\ln \text{sdr}_{kt}$	−0.157*
	(0.073)
$(\ln \text{sdr}_{kt})^2$	0.011
	(0.005)
$\ln \text{den}_{kt} * \ln \text{km}_{kt}$	−0.011*
	(0.006)
$\ln \text{km}_k * \ln \text{sdr}_{kt}$	0.012
	(0.009)
$\ln \text{den}_{kt} * \ln \text{sdr}_{kt}$	0.013
	(0.006)
$\ln \text{km}_k * \text{d02}$	−0.050†
	(0.031)
$(\ln \text{km}_k)^2 * \text{d02}$	0.007†
	(0.004)
$\ln \text{den}_{kt} * \ln \text{km}_k * \text{d02}$	0.001
	(0.001)
$\ln \text{km}_k * \ln \text{sdr}_{kt} * \text{d02}$	0.000
	(0.001)
apres_{kt-1}	0.008
	(0.034)
tbapres_{kt-1}	−0.033†
	(0.016)
LR χ^2 statistic	468.36‡
Predicted = 0/Actual = 0	670/719
Predicted = 1/Actual = 1	119/177
Lave–Efron pseudo-R^2	0.504
McKelvey–Zavoina pseudo-R^2	0.810
No. of observations	896

Notes: (i) marginal-effects reported; (ii) standard errors in parentheses; (iii) * means significant at 10%, † at 5% and ‡ at 1% level; (iv) city-specific dummies not reported.

Table 10
Disaggregated elasticities of km_k – 2001

den_{kt}	km_k						
	350	500	750	1150	1850	2250	2600
1000	11.09	6.81	3.83	1.85	0.32	−0.21	−0.58
3000	6.65	4.02	2.17	0.92	−0.09	−0.47	−0.75
6000	4.52	2.67	1.37	0.51	−0.21	−0.51	−0.73
15,000	2.37	1.31	0.61	0.16	−0.24	−0.43	−0.58
50,000	0.61	0.28	0.10	0.00	−0.11	−0.18	−0.25
150,000	0.05	0.02	0.00	0.00	−0.01	−0.02	−0.04
300,000	0.00	0.00	0.00	0.00	0.00	0.00	−0.01

there is evidence that she accomplished a deviation towards a more JBP-like standard of operations, implemented since 2002.

Table 11
Disaggregated elasticities of km_k – 2002

den_{kt}	km_k						
	350	500	750	1150	1850	2250	2600
1000	23.02	14.55	8.83	5.26	2.87	2.17	1.73
3000	13.98	8.84	5.31	3.08	1.59	1.16	0.89
6000	9.76	6.10	3.58	2.01	0.98	0.69	0.50
15,000	5.54	3.32	1.82	0.94	0.41	0.26	0.17
50,000	1.90	0.96	0.42	0.17	0.05	0.03	0.01
150,000	0.29	0.10	0.03	0.01	0.00	0.00	0.00
300,000	0.03	0.01	0.00	0.00	0.00	0.00	0.00

Notes: (i) figures calculated holding sdr_{kt} at the sample mean; (ii) calculated as a 10% increase in each variable at the mean; (iii) values are representative of the following percentiles: 0.20, 0.35, 0.50, 0.65, 0.80, 0.90 and 0.95.

Two caveats must be considered with respect to the above-mentioned results on flight distance: firstly, as discussed before, country idiosyncrasies (for example, unobserved economies of scope) probably influenced Gol in the strategic decision of not to focus only on non-stop short flight markets, but to put into practice a modified version of JBP – that is, also considering long-haul markets but with many stops and connections. Also, it is important to emphasize that, from 2002 on, Gol's pace of expansion ultimately made it the third biggest domestic airline; it is no surprise, therefore, that its entry behavior became more similar to the incumbent majors as Gol began to enter every single dense route across the country, irrespective of other market attributes, such as flight haul or rival's presence.[43] Once the decision of rapid expansion was taken it actually made the airline path-dependent, which can potentially undermine her LCC attributes and cost advantage.

5. Conclusions

This paper aimed at developing an empirical model for the analysis of entry decisions of Gol Airlines, the first low-cost carrier in Latin America. By making use of Amemiya's Generalized Least Squares (AGLS) it was possible to estimate a route-choice model associated with a flexible post-entry equilibrium profits equation, and in which some of the regressors were treated as endogenous.

Results revealed market size and rivals' route presence to be relevant indicators of underlying determinants of profitability; the relevance of rivals' presence is probably in accordance with a learning process (Toivanen and Waterson, 2005). The consistency of Gol's decision making with the pattern of entry classically established by Southwest Airlines – with stronger preference for dense and short-haul routes – was investigated and was not rejected for the start-up year (2001). Unambiguous evidence was found, however, that Gol deviated from this paradigm towards a standard of operations more in accordance with the JetBlue Airways' paradigm (higher average stage length), in 2002, when compared to 2001. This tendency engendered diversification of portfolio of routes, instead of specialization in one single business approach.

The main reason for that deviation is associated with country idiosyncrasies like unobserved economies of scope, which probably influenced Gol in the strategic decision of not to focus only on non-stop short flight sector markets, but to put into practice a modified version of JBP – that is, considering long-haul markets but with many stops and connections.

Acknowledgements

The author would like to thank Michael Waterson, Margaret Slade, Adalberto Febeliano, Jorge Silveira, Robson Ramos, Allemander Pereira Filho, Cristiano Miranda, Cristian Huse, Dante Aldrighi Mendes, the participants of a seminar at IBMEC (São Paulo, 2004), the participants of the annual conference of ESRC

[43] The increase in market contact among LCC's and FSC's is also reported by Boguslaski et al. (2004).

Table 12
Matrix of correlations of variables

Variable	$\ln \text{den}_{kt}$	In sdr_{kt}	$(\ln \text{den}_{kt})^2$	$(\ln \text{sdrj}_{kt})^2$	$\ln \text{den}_{kt} * \ln \text{km}_k$	$\ln \text{den}_{kt} * \ln \text{sdr}_{kt}$	$\ln \text{km}_k * \ln \text{sdr}_{kt}$
$\ln \text{den}_{kt}$	1.000						
$\ln \text{sdr}_{kt}$	0.422	1.000					
$(\ln \text{den}_{kt})^2$	0.374	−0.177	1.000				
$(\ln \text{sdr}_{kt})^2$	0.206	0.943	−0.287	1.000			
$\ln \text{den}_{kt} * \ln \text{km}_k$	0.992	0.420	0.375	0.208	1.000		
$\ln \text{den}_{kt} * \ln \text{sdr}_{kt}$	0.847	0.310	0.630	0.083	0.844	1.000	
$\ln \text{km}_k * \ln \text{sdr}_{kt}$	0.412	0.973	−0.168	0.933	0.427	0.321	1.000
$\ln \text{km}_k$	0.074	−0.003	−0.059	0.061	0.131	0.058	0.194
$(\ln \text{km}_k)^2$	0.067	−0.002	−0.064	0.066	0.124	0.052	0.194
$\ln \text{den}_{kt-1}$	0.795	0.282	0.497	0.104	0.787	0.763	0.270
$(\ln \text{den}_{kt-1})^2$	0.568	−0.018	0.862	−0.154	0.566	0.700	−0.019
In sdr_{kt-1}	0.246	0.734	−0.079	0.695	0.246	0.275	0.729
$(\ln \text{sdr}_{kt-1})^2$	0.083	0.658	−0.189	0.698	0.085	0.102	0.673
$\ln \text{den}_{kt-1} * \ln \text{km}_k$	0.788	0.277	0.497	0.103	0.795	0.762	0.285
$\ln \text{den}_{kt-1} * \ln \text{sdr}_{kt-1}$	0.762	0.347	0.602	0.156	0.762	0.865	0.340
$\ln \text{km}_k * \ln \text{sdr}_{kt-1}$	0.244	0.717	−0.079	0.694	0.259	0.280	0.763
$\ln \text{seats}_{kt-1}$	0.125	−0.425	0.383	−0.371	0.131	0.184	−0.379
$\ln \text{swe}_{kt-1}$	0.667	0.596	0.427	0.460	0.667	0.762	0.604
$\ln \text{asz}_{kt-1}$	0.289	0.017	0.315	0.073	0.309	0.396	0.118

Table 13
Hypothesis tests – AGLS model

Test	Description	H_0	Statistic	*P*-value
Endogeneity	Smith–Blundell test (1986)	All explanatory variables are exogenous	15.49	0.03
Heteroskedasticity	Likelihood-ratio/maximum-likelihood heteroskedastic Probit	Homoskedasticity	1.23	0.75
Relevance of instruments (test of	Partial R^2 of excluded instruments (min)		0.55	
Correlation with included endogenous variable)	*F*-test of excluded instruments -joint significance (min)	Instruments are not relevant	105.20	0.00
Validity of Instruments (overidentification/ orthogonality test of all instruments)	Sargan Test	Instruments are not valid	1.87	0.60
	Bassman test	Instruments are not valid	1.82	0.61

Econometric Study Group (Bristol, 2004), CAPES and FAPESP. Special thanks to Rachel Oliveira. All errors are the author's. Aeronautical Institute of Technology, Brazil.

Appendix 1. Additional statistics

See Tables 12 and 13.

References

Amemiya, T., 1978. The estimation of a simultaneous equation generalized probit model. Econometrica 46, 1193–1205.

Baum, C., Schaffer, M., Stillman, S., 2003. Instrumental variable and GMM: estimation and testing, vol. 545. Department of Economics Working Paper, Boston College.

Ben-Akiva, M., Lerman, S., 1985. Discrete Choice Analysis: Theory and Application to Predict Travel Demand. The MIT Press, Cambridge, MA.

Berry, S., 1992. Estimation of a model of entry in the airline industry. Econometrica 60, 889–917.

Berry, S., Carnall, M., Spiller, P., 2006. Airline hubs: Costs, markups and the implications of customer heterogeneity. In: Lee, D. (Ed.), Advances in Airline Economics, vol. 1. Competition Policy and Antritrust, Elsevier, Cambridge, USA, pp. 183–213.
Blundell, R., Smith, R., 1989. Estimation in a class of simultaneous equation limited dependent variable models. Review of Economic Studies 56, 37–57.
Boguslaski, R., Ito, H., Lee, D., 2004. Entry patterns in the Southwest Airlines route system. Review of Industrial Organization 25 (3), 317–350.
Borenstein, S., 1989. Hubs and high fares: dominance and market power in the U.S. airline industry. Rand Journal of Economics 20, 344–365.
Brander, J., Zhang, A., 1990. Market conduct in the airline industry: an empirical investigation. Rand Journal of Economics 21 (4), 567–583.
Davidson, R., MacKinnon, J., 1993. Estimation and Inference in Econometrics, second ed. Oxford University Press, New York.
Doganis, R., 2001. The Airline Industry in the 21st Century. Routledge, London.
Dresner, M., Lin, J., Windle, R., 1996. The impact of low-cost carriers on airport and route competition. Journal of Transport Economics and Policy 30, 309–328.
Evans, W., Kessides, I., 1993. Localized market power in the U.S. airline industry. Review of Economics and Statistics 75, 66–75.
Fox, J., 1997. Applied Regression Analysis, Linear Models, and Related Methods. Sage Publications, London.
Guimarães, C., 2002. Feita Para Voar. Revista Exame 01–04.
Harkness, J., 2001. IVPROB-IVTOBIT: Stata modules to estimate instrumental variables probit and tobit. Statistical Software Components from Boston College Department of Economics.
Ito, H., Lee, D., 2003. Market density and low cost carrier entries in the US airline industry: implications for future growth. Mimeo.
Joskow, A., Werden, G., Johnson, R., 1994. Entry, exit, and performance in airline markets. International Journal of Industrial Organization 12, 457–471.
Lee, L., 1991. Amemiya's generalized least squares and tests of overidentification in simultaneous equation models with qualitative or limited dependent variables. Discussion Paper Series, Center for Economic Research, Department of Economics, University of Minnesota, Minneapolis, Minnesota, No. 262, p. 12.
Lima, E., 2002. Brazil's daring wings. Air Transport World May Cleveland.
Morrison, S., 2001. Actual, adjacent, and potential competition: estimating the full effect of Southwest Airlines. Journal of Transport Economics and Policy 35, 239–256.
Mullineaux, D., 1978. Economies of scale and organizational efficiency in banking: a profit-function approach. Journal of Finance 33, 259–280.
Newey, W., 1987. Efficient estimation of limited dependent variable models with endogenous explanatory variables. Journal of Econometrics 36, 231–250.
Oliveira, A.V.M., Huse, C., 2004. Localized competitive advantage and price reactions to low cost carrier entry in the Brazilian airline industry. In: Proceedings of the 2004 Air Transport Research Society (ATRS) World Conference, Istambul.
Oum, T., Zhang, A., Zhang, Y., 1997. Inter-firm rivalry and firm-specific price elasticities in deregulated airline markets. In: Oum, T. et al. (Eds.), Transport Economics – Selected Readings. Harwood Academic Publishers, Amsterdam, p. 691.
Richard, O., 2003. Flight frequency and mergers in airline markets. International Journal of Industrial Organization 21 (6), 907–922.
Rivers, D., Vuong, Q., 1984. Limited information estimators and exogeneity tests for simultaneous probit models. Social Science Working Paper, California Institute of Technology, Pasadena, CA, vol. 539.
Rivers, D., Vuong, Q., 1988. Limited information estimators and exogeneity tests for simultaneous probit models. Journal of Econometrics 39, 347–366.
Silva, E., Espírito Santo Jr., 2003. Introducing the low cost-low fare concept in Brazil: Gol airlines. In: Proceedings of the 7th Annual World Conference of the Air Transport Research Society.
Slade, M., 1986. Taxation of non-renewable resources at various stages of production. Canadian Journal of Economics 19, 281–297.
Smith, R., Blundell, R., 1986. An exogeneity test for a simultaneous equation tobit model with an application to labor supply. Econometrica 54, 679–686.
Toivanen, O., Waterson, M., 2005. Market structure and entry: where's the beef? RAND Journal of Economics 36 (3), 680–699.
Tretheway, M., 2004. Distortions of airline revenues: why the network airline business model is broken. Journal of Air Transport Management 10, 3–14.
US Department of Transportation, 2001. Special feature: the low-fare evolution II. Domestic Airline Fares Consumer Report, Q3.
US Department of Transportation, 2002. Special feature: the low-fare evolution. Domestic Airline Fares Consumer Report, Q3.
Whinston, M., Collins, S., 1992. Entry and competitive structure in deregulated airline markets: an event study analysis of people express. Rand Journal of Economics 23, 445–462.
Windle, R., Dresner, M., 1999. Competitive responses to low cost carrier entry. Transportation Research Part E 35, 59–75.
Zalamea, L., 2001. Brazil's Gol Airlines Aims for Low-cost, No-frills Niche. Aviation Week & Space Technology, New York.