Long-distance Inter-city Passenger Transport: Demand and product quality choice

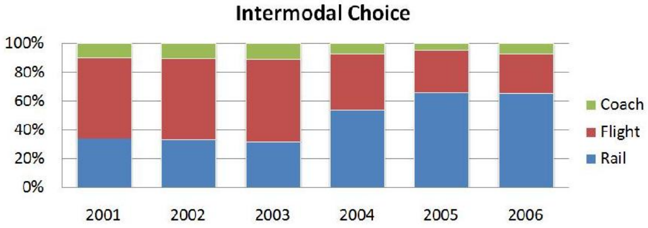
Jisun Baek

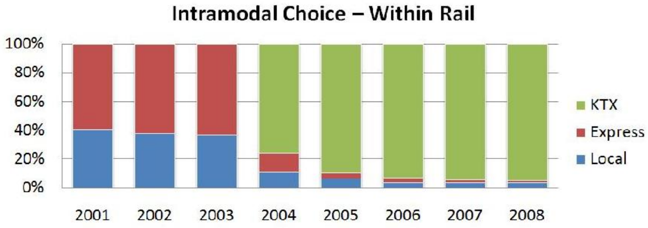
Columbia

October 7 2009

Motivation

Seoul - Busan





Motivation

		Obs	Mean	Std.Dev	Min	Max
Sae-ma-eul (Express)	Before	20	11.85	2.81	6	17
	After	49	15.43	4.80	7	25
Mu-gung-hwa (Local)	Before	48	20.13	5.93	9	28
	After	136	23.65	6.10	9	49

Research Question

- Analysis on the demand system of long distance passenger travel
  - Competition between modes
  - Effect of introducing KTX(high-speed train) in 2004 on consumer welfare
- Quality differentiation in the rail service
  - Frequency choice
  - How many times a train stops from the origin to the destination

Literature Review

- Transport
  - Berry, Carnall, and Spiller (2006) Berry (1994) Berry and Jia (2009) Borenstein and Netz (1999) Brueckner (2004)
  - Wardman (2006) Gagnepain and Ivaldi (2002) Ivaldi and Vibes (2005) Richards and Ben-Akiva (1975)
- Estimation
  - McFadden (1981) McFadden and Train (2000) Anderson and Palma (1992)
  - Berry, Levinsohn, and Pakes (1995) Berry, Levinsohn, and Pakes (2004)
- Quality differentiation
  - Mussa and Rosen (1978) Champsaur and Rochet (1989) Johnson and Myatt (2003)
  - Gabszewicz, Shaked, Sutton, and Thisse (1986)

Industry Background

- Domestic Airlines
  - 2∼4 active firms
  - Only a few destinations are available
  - It takes at most 1 hour to any destination in the sky.
- Intercity Coaches
  - 183 paths in bus connections, varying across years
  - There are 5-10 active firms in long-distance bus, depending on paths
- Railway
  - KTX was introduced on April 2004(Available for a few destinations)
  - 3 train classes are currently available
  - Strong regulations on pricing

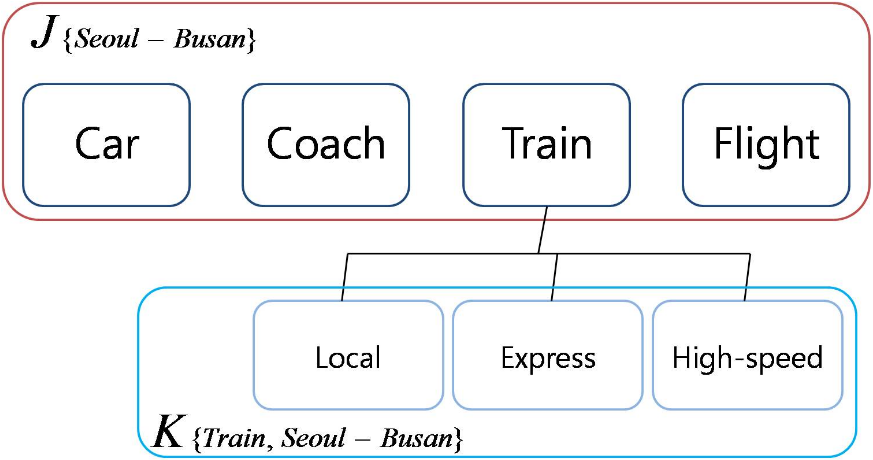
Data Description

- The market level data is observed every month from January 2001 to December 2008. The following list includes the variables observed.
  - Number of passengers for each train class(directional), for domestic airlines(directional), and for coaches(non-directional)
  - Price of each mode
  - Characteristics of each mode including travel time, distance and frequencies, the number of stops to the destination
  - Departure time of day (rail and flight since January 2004)
  - Schedule of rail service (Time tables)
  - Demographical and geographical characteristics[1] of the area that each station belongs to
  - Price index
- Some cost side variables like wage, the unit price of each passenger car, fuel and electricity prices

Consumer Decision Process

- A consumer has to choose a mode, given an origin-destination.
- First, she will choose a mode from her choice set J_r.
- Second, she will choose one of available classes K_{jr} if she takes a mode j.
  - The class is defined by speed.
  - Thus, both flight and coach have only one class of service.
  - Rail has 3 classes of service(KTX, Express, Local) at most, and the available classes vary with markets.

Nested Logit



Choice within Mode

- Consumer i's utility from a class l of mode j

$$u^i_{jl,r} = x_{jl,r}\beta^i + \xi_{jl,r} + \lambda_2 \epsilon^i_{2jl,r} \tag{1}$$

- i will choose $l \in K_{jr}$ if and only if

$$u^i_{jl,r} \geq u^i_{jk,r} \quad \forall k \in K_{jr} \tag{2}$$

.

Choice between Modes

- Let $\delta_{j,r}$ denote the expected attractiveness of a mode $j \in J_r$

$$\delta_{j,r} = \mathrm{E}\left[\max_{l \in K_{jr}} \; u^i_{jl,r}\right] \tag{3}$$

- Consumer i's utility from a mode $j \in J_r$ is

$$U^i_{j,r} = \delta^i_{j,r} + \lambda_1 \epsilon^i_{1j,r} \tag{4}$$

  where $\epsilon^i_{1j,r}$ is the individual heterogeneous taste on each traffic mode and λ_1 is a nested logit parameter which reflects the heterogeneity of different modes.

- i will choose $j \in J_r$ if and only if

$$U^i_{jr} \geq U^i_{kr} \quad \forall k \in J_r \tag{5}$$

Market Shares : within a mode

- A vector of parameters, β^i follows the distribution $\Phi^i(\cdot)$
- x_{jl} is the observed characteristics of l
- ξ_{jl} is the unobserved characteristics of l
- Assume $\epsilon^i_{2jl,r} \sim$ type I extreme value
- The probability that i chooses a mode l from K_{jr}, $s^i_{l|j,r}$ is

$$s^i_{l|j,r} = \frac{\exp\left(\frac{x'_{jl,r}\beta^i + \xi_{jl,r}}{\lambda_2}\right)}{\sum_{q \in K_{jr}} \exp\left(\frac{x'_{jq,r}\beta^i + \xi_{jq,r}}{\lambda_2}\right)} \tag{6}$$

Market Shares

- Assuming $\epsilon^i_{1j,r} \sim$ type I extreme value, Prob(i chooses a mode j from J_r), $s^i_{j,r}$ is

$$s^i_{j,r} = \frac{\exp\left(\delta^i_{j,r}/\lambda_1\right)}{1 + \sum_{q \in J_r} \exp\left(\delta^i_{q,r}/\lambda_1\right)}$$

- The probability that i chooses (j, l) combination given r, $s^i_{jl,r}$ is

$$s^i_{jl,r} = s^i_{l|j,r} \cdot s^i_{j,r} \tag{7}$$

Proportion of Traveler type

- Assume that there are 2 different types(L,B) of consumers distinguished by $\beta^i \sim \Phi^i$.(Berry et al., 2006). Therefore, the distribution of β^i is the mixture of Φ^L and Φ^B.
- Let κ_r denote the proportion of consumer type L in a market r, which varies with markets
- κ_r can explained by observed and unobserved characteristics of market r

$$\kappa_r = g(z_r'\gamma + \nu_r)$$

  where $g(\cdot)$ is a proper CDF.
- The total market share of a combination (j, l) for a market r is

$$s_{jl,r} = \kappa_r \int s^L_{jl,r}(x, \xi, \beta^L) d\Phi^L + (1 - \kappa_r) \int s^B_{jl,r}(x, \xi, \beta^B) d\Phi^B$$

Firms Behavior

Given the demand,

- Step1 : Firms choose the quality of products
  - Airlines and Coach company choose the frequency(Berry & Jia, 2009)
  - Rail company chooses the frequency and the number of stops(or at which stations a train stops)
- Step2 : Firms choose the price of each product

Cost Function

- Total Cost to produce a combination of mode j and a class l in a market r

$$C_{jl,r} = c_{jl}(w_{jl,r}, f_{jl,r}, h_{jl,r}, q_{jl,r}) + \omega^1_{jl,r} q_{jl,r} + \omega^2_{jl,r} f_{jl,r} + F_{jl,r} \quad (8)$$

- w_{jr} : a vector of exogenous cost shifters
- $f_{jl,r}$: frequency
- $h_{jl,r}$: Quality of service except frequency
- $q_{jl,r}$: Number of ticket sold
- $F_{jl,r}$: Fixed cost for a combination (j, l) in a market r
- $\omega^1_{jl,r}$ $\omega^2_{jl,r}$: Idiosyncratic component
- Function form of $c_{jl}(\cdot)$ may differ across j and l

Cost Function

- Let $j = 1$ denote Rail, $j = 2$ denote Airline, and $j = 3$ denote Coach
- $C_{j,r}$ denote the total cost spent by a firm providing mode j in a market r

$$
\begin{aligned}
C_{1,r} &= \sum_{l \in K_{1r}} c_{1l}(w_{1l,r}, f_{1l,r}, h_{1l,r}, q_{1l,r}) + \omega^1_{jl,r} q_{jl,r} + \omega^2_{1l,r} f_{1l,r} + F_{1l,r} \\
C_{2,r} &= c_{21}(w_{21,r}, f_{21,r}, h_{21,r}, q_{21,r}) + \omega^1_{jl,r} q_{jl,r} + \omega^2_{21,r} f_{21,r} + F_{21,r} \\
C_{3,r} &= c_{31}(w_{31,r}, f_{31,r}, h_{31,r}, q_{31,r}) + \omega^1_{jl,r} q_{jl,r} + \omega^2_{31,r} f_{31,r} + F_{31,r}
\end{aligned} \tag{9}
$$

Frequency Choice

- Airline and Coach company's problem ($j = 2, 3$)

$$\begin{aligned} \text{Max}_{f_{j1,r}} \quad & p_{j1,r} q_{j1,r} - c_{j1}(w_{j1,r}, f_{j1,r}, h_{j1,r}, q_{j1,r}) \\ & -\omega^1_{jl,r} q_{jl,r} - \omega^2_{j1,r} f_{j1,r} - F_{j1,r} \end{aligned} \tag{10}$$

- Rail Company's problem ($j = 1$)
  - f^t_{1l} :Frequency of a train l running on track t
  - $f_{1l,r} = f_{1l,r'} \equiv f^t_{1l}$ if r and r' are on t
  - For each t and l,

$$\begin{aligned} \text{Max}_{f^t_{1l}} \quad \sum_{r \in t} \quad & p_{1l} q_{1l,r} - c_{1l}(w_{1l,r}, f^t_{1l}, h_{1l,r}, q_{1l,r}) \\ & -\omega^1_{1l,r} q_{1l,r} - \omega^2_{1l,r} f^t_{1l} - F_{1l,r} \end{aligned} \tag{11}$$

Stop Choice

- When the quality of product is lowered,
  - Single-product firm : the revenue of firm decreases.

$$\Delta R = p \cdot \Delta q < 0 \tag{12}$$

  - Multi-product firm : Change in the revenue of firm is ambiguous.

$$\Delta R = \underbrace{p_1 \cdot \Delta q_1}_{<0} + \underbrace{\sum_{j>1} p_j \cdot \Delta q_j}_{\geq 0} \tag{13}$$

Stop Choice : Example

$$t_1 \leftrightarrow t_2 \leftrightarrow t_3 \leftrightarrow t_4 \tag{14}$$

- When a train 1 skips t_2, there are 6 markets, $t_1 \leftrightarrow t_3$ $t_1 \leftrightarrow t_4$ and $t_3 \leftrightarrow t_4$
- The number of stops between t_1 and t_4 is 1, and the one between t_1 and t_3 is 0.
- If the train additionally stops at t_2 to increase the number of stops between t_1 and t_4, then the number of markets is now 12, and the number of stops between t_1 and t_3 is 1.

Stop Choice

- M_l^t : Set of markets where l is sold before the quality change
- $M_l'^t$: Set of markets where l is sold after the quality change
- When the service quality of train class 1 is lowered(i.e. a train stops at more stations) in the rail industry,
  - The revenue from $l > 1$ in $M_l^t \backslash (M_1'^t \backslash M_1^t)$ increases because the quality of 1 is lowered
  - The revenue from $l > 1$ in $M_l^t \cap (M_1'^t \backslash M_1^t)$ decreases because of the competition with 1
  - The revenue from 1 in M_1^t decreases because the quality of 1 is lowered
  - The revenue from 1 in $M_1'^t \backslash M_l^t$ increases because 1 is newly sold in these markets

Stop Choice

- Equivalently, the change in the revenue from a track t is

$$\begin{aligned}\Delta R \;=\;& \sum_{l\neq 1}\sum_{r\in M_l^t\backslash(M_1'^t\backslash M_1^t)} p_{lr}\Delta q_{l,r} && (\geq 0)\\ &+\sum_{l\neq 1}\sum_{r\in M_l^t\cap(M_1'^t\backslash M_1^t)} p_{lr}\Delta q_{l,r} && (\leq 0)\\ &+\sum_{r\in M_1^t} p_{1,r}\Delta q_{1,r} && (\leq 0)\\ &+\sum_{r\in M_1'^t\backslash M_1^t} p_{1,r}\Delta q_{1,r} && (\geq 0)\end{aligned} \tag{15}$$

- Rail company compares the revenues to make a train stops at one station

Stop Choice

- Consider a track t, which has $N(t)$ stations on it.

$$t_1 - t_2 - t_3 - \cdots - t_{N(t)-1} - t_{N(t)} \tag{16}$$

- A market is a directional pair of two stations.
- Arrange stations $t_2 \cdots t_{N(t)-1}$ in order of its size. Let $t^{[1]}$ be the station involved the largest demand, and $t^{[N(t)-2]}$ be the one involved the smallest demand.
- The probability that the rail company chooses to make a train l stop at $t^{[1]}$

$$\psi(l, t^{[1]}|\theta) \equiv \text{Prob}[R^t(l, t^{[1]}|\theta) > R^t(l, 0|\theta)|\text{other train classes}] \tag{17}$$

  - $R^t(l, t^{[1]})$: Revenue when it stops at $t^{[1]}$
  - $R^t(l, 0)$: Revenue when it directly connect from t_1 to $t_{N(t)}$

Stop Choice : Example

There are stations on line 1 track.

116th - 110th - 103rd - $\cdots$ $\cdots$ -23rd - 18th - 14th

Station	42nd	34th	96th	14th	72nd	59th	50th	$\cdots$
Express	Stop	Stop	Stop	Stop	Stop	Not	Not	Not
Local	Stop	Stop	Stop	Stop	Stop	Stop	Stop	$\cdots$

Stop Choice

- Let $y(l, t_n)$ be an indicator such that

$$y(l, t^{[n]}) = 1 \quad \text{if} \quad R^t(l, t^{[n]}|\theta) \geq R^t(l, t^{[n-1]}|\theta) \tag{18}$$

- $\sum_{n=2}^{N(t)-1} y(l, t^{[n]})$: Number of stops two end-point on a track t
- The rail company determines

$$\begin{aligned} y(l, t) &\equiv (y(l, t^{[1]}), y(l, t^{[2]}), \cdots, y(l, t^{[N(t)-2]})) \\ \text{where} & \ \{(y_1, y_2, \cdots, y_{N(t)-2}) | y_k \in \{0, 1\} \ \& \\ & \ y_1 \geq y_2 \geq \cdots \geq y_{N(t)-2}\} \end{aligned} \tag{19}$$

Pricing

- Profit of each firm for a market r, $\pi_{j,r}$

$$
\begin{aligned}
\pi_{1,r} &= \sum_{l \in K_{1,r}} p_{1l,r} q_{1l,r} - c_{1l}(w_{1l,r}, f_{1l,r}, h_{1l,r}, q_{1l,r}) - \omega^{1}_{jl,r} q_{jl,r} - \omega^{2}_{1l,r} f \\
\pi_{2,r} &= p_{2,r} q_{2,r} - c_{2}(w_{2,r}, f_{2,r}, h_{2l,r}, q_{2,r}) - \omega^{1}_{jl,r} q_{jl,r} - \omega^{2}_{2,r} f_{2,r} - F_{2,r} \\
\pi_{3,r} &= p_{3,r} q_{3,r} - c_{3}(w_{3,r}, f_{3,r}, h_{3l,r}, q_{3,r}) - \omega^{1}_{jl,r} q_{jl,r} - \omega^{2}_{3,r} f_{3,r} - F_{3,r}
\end{aligned}
\tag{20}
$$

- Airline and Coach company's problem ($j = 2, 3$)

$$
\mathrm{Max}_{p_{j1,r}} \; p_{j1,r} q_{j1,r} - c_{j1}(w_{j1,r}, f_{j1,r}, h_{j1,r}, q_{j1,r}) - \omega^{1}_{jl,r} q_{jl,r} - \omega^{2}_{j1,r} f_{j1,r} - F_{j1,r} \tag{21}
$$

Pricing

- Rail company's problem ($j = 1$)
  - Because of the regulation, the price of rail service must be a function of base rate(ρ_l) and the distance($d_{l,r}$)
  - $p_{1l}(\rho_l, d_{l,r})$: Price of train class l in a market r
  - Rail company can choose 3 numbers, ρ_1, ρ_2, ρ_3

$$\mathrm{Max}_{\rho_l} \quad \sum_r \quad \begin{array}{l} p_{1l}(\rho_l, d_{l,r})q_{1l,r} - c_{1l}(w_{1l,r}, f_{1l,r}, h_{1l,r}, q_{1l,r}) \\ -\omega^1_{1l,r}q_{1l,r} - \omega^2_{1l,r}f_{1l,r} - F_{1l,r} \end{array} \tag{22}$$

Moment Condition : Demand

- Given a market r, the expected market share of a mode j and a class l is

$$E[s_{jl,r}(\xi,\beta,\gamma)] = \int s_{jl,r}(\xi,\beta,\gamma) f(\nu_r) d\nu_r$$

where $f(\cdot)$ is a pdf of ν_r.

- With NS of ν_r from $f(\nu_r)$ (Berry et al., 2004),

$$\bar{s}_{jl,r}(\xi,\beta,\gamma) = \frac{1}{NS}\sum_{ns=1}^{NS} s_{jl,r}(\xi,\beta,\gamma)$$

Moment Condition : Demand

- The parameters can be estimated through the system

$$s^O_{jl,r} = \bar{s}_{jl,r}(\xi, \beta, \gamma)$$

where $s^O_{jl,r}$ is the observed share of (j, l).

- Assume that $\xi(\cdot)$ is calculated by the inversion of the above equation.
- Construct a moment condition based on

$$E(\xi(x_r, \beta, \gamma)|iv_r) = 0 \qquad (23)$$

where iv_r is instrumental variables.

Moment Condition : Pricing

- Moment Conditions from Pricing FOC
  - Airlines and Coaches

$$E(\omega^1_{j,r}|iv_{j,r}) = 0 \tag{24}$$

  - Rail company

$$E(\sum_r \omega^1_{1l,r}|iv_j) = 0 \quad l = 1, 2, 3 \tag{25}$$

Moment Condition : Frequency Choice

- Moment Conditions from Frequency Choice FOC
  - Airlines and Coaches

$$E(\omega_{j,r}^2 | iv_{j,r}) = 0 \tag{26}$$

  - Rail company

$$E(\sum_{r \in t} \omega_{1l,r}^2 | iv_j) = 0 \tag{27}$$

Moment Condition : Stop Choice

- The conditional probability mass function of $y(l, t^{[n]})$ given $y(l, t^{[n-1]}) = 1$ is

$$\begin{aligned} & \Upsilon(y(l, t^{[n]}) | y(l, t^{[k]}) \quad \forall k < n) \\ \equiv \; & \psi(l, t^{[n]})^{y(l, t^{[n]})} (1 - \psi(l, t^{[n]}))^{1 - y(l, t^{[n]})} \quad \text{if} \quad y(l, t^{[n-1]}) = 1 \end{aligned} \tag{28}$$

- The conditional probability mass function of $y(l, t^{[n]})$ given $y(l, t^{[n-1]}) = 0$ is degenerate because only one possible outcome exists in the sample space of $y(l, t)$.

$$\Upsilon(y(l, t^{[n]}) | y(l, t^{[k]}) \quad \forall k < n) = 1 \quad \text{if} \quad y(l, t^{[n-1]}) = 0 \tag{29}$$

Moment Condition : Stop Choice

- The joint probability mass function of $y(l,t)$

$$\begin{aligned} & \Upsilon(y(l,t^{[1]},\cdots,y(l,t^{[n]})) \\ = \; & \Upsilon(y(l,t^{[n]})|y(l,t^{[k]}) \quad \forall k < n) \times \Upsilon(y(l,t^{[1]}),\cdots,y(l,^{[n-1]})) \\ = \; & \left(\psi(l,t^{[n]})^{y(l,t^{[n]})}(1-\psi(l,t^{[n]}))^{1-y(l,t^{[n]})}\right)^{y(l,t^{[n-1]})} \\ & \times \Upsilon(y(l,t^{[1]}),\cdots,y(l,^{[n-1]})) \end{aligned} \tag{30}$$

where $\Upsilon(y(l,t^{[1]}),\cdots,y(l,^{[n-1]}))$ is the joint probability.

- By induction, the joint probability mass function of $y(l,t)$ is, therefore,

$$\begin{aligned} & \Upsilon(y(l,t^{[1]}),y(l,t^{[2]}),\cdots,y(l,t^{[N(t)-2]})) \\ = \; & \Upsilon(y(l,t^{[1]}))\cdot\Upsilon(y(l,t^{[2]})|y(l,t^{[1]})) \\ & \times\cdots\times\Upsilon(y(l,t^{[N(t)-2]})|y(l,t^{[k]}) \quad \forall k < N(t)-2) \end{aligned} \tag{31}$$

Moment Condition : Stop Choice

- Equivalently,

$$\Pi_{n=1}^{N(t)-2} \left[\psi(l, t^{[n]})^{y(l,t^{[n]})} \left(1 - \psi(l, t^{[n]})\right)^{1-y(l,t^{[n]})} \right]^{y(l,t^{[n-1]})} \quad (32)$$

where $y(l, t^{[0]}) = 1$ for all l and all t.

- the log-likelihood function of a multinomial distribution from the station choice outcome for a track t and a train l, $\ln L(l, t)$

$$\ln L(l,t) = \sum_{n=1}^{N(t)-2} y(l, t^{[n-1]}) \times \left[y(l, t^{[n]}) \ln \psi(l, t^{[n]}) + (1 - y(l, t^{[n]})) \ln(1 - \psi(l, t^{[n]})) \right] \quad (33)$$

Moment Condition : Stop Choice

- The parameter θ maximizes the log-likelihood

$$\mathrm{Max}_{\theta} \sum_{t=1}^{T} \ln L_t(\theta) \tag{34}$$

- The first order condition of the above problem becomes a part of the moment conditions for GMM.

Further Study

- Prove that the above process brings the optimality of firms' choice
- Identification ($\because$ Endogenous choice sets)
- The schedule by hours and departure time preference

Anderson, S. P., & Palma, A. D. (1992). Multiproduct firms: A nested logit approach. *The journal of indstrial economics*, *40*(3), 261-276.

Berry, S. (1994). Estimating discrete-choice models of product differentiation. *Rand journal of economics*, *25*(2), 242-262.

Berry, S., Carnall, M., & Spiller, P. (2006). Airline hubbing, costs and demand. In D. Lee (Ed.), (Vol. vol1. Competition Policy and Anti-Trust, p. 183-214). Elsevier Press.

Berry, S., & Jia, P. (2009). Tracing the woes: An empirical analysis of the airline industry. *Working paper*.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, *60*(4), 889-917.

Berry, S., Levinsohn, J., & Pakes, A. (2004). Estimating differentiated products demand systems from a combination of micro and macro data: The market for new vehicles. *Journal of political economy*, *112*(1), 68-105.

Borenstein, S., & Netz, J. (1999). Why do all the flights leave at 8am?: Competition and departure-time differentiation in airline market. *International journal of industrial organization*, *17*, 611-640.

Brueckner, J. K. (2004). Network structure and airline scheduling. *Journal of industrial economics*, *52*, 291-312.

Champsaur, P., & Rochet, J.-C. (1989). Multiproduct duopolists. *Econometrica*, *57*(3), 533-557.

Gabszewicz, J. J., Shaked, A., Sutton, J., & Thisse, J.-F. (1986). Segmenting the market: The monopolist's optimal product mix. *Journal of economic theory*, *39*, 273-289.

Gagnepain, P., & Ivaldi, M. (2002). Incentive regulatory policies: the case of public transit systems in France. *Rand journal of economics*, *33*(4).

Ivaldi, M., & Vibes, C. (2005). Intermodal and intramodal competition in the long-haul passenger transport market. *IDEI Report*(9).

Johnson, J. P., & Myatt, D. P. (2003). Multiproduct quality competition: Fighting brands and product line pruning. *The American economic review*, *93*(3), 748-774.

McFadden, D. (1981). Structural analysis of discrete data with econometric analysis. In C. F. Manski & D. McFadden (Eds.), (p. 198-272). The MIT Press.

McFadden, D., & Train, K. (2000). Mixed mnl models for discrete response. *Journal of Applied Econometrics*, *15*(5), 447-470.

Mussa, M., & Rosen, S. (1978). Monopoly and product quality. *Journal of economic theory*, *18*, 301-307.

Richards, M. G., & Ben-Akiva, M. E. (1975). *A disaggregate travel demand model*. Saxon house and Lexington books.

Wardman, M. (2006). Demand for rail travel and the effects of external factors. *Transportation research part E*, *42*, 129-148.