Airline Entry and Exit with Learning*

Mingao Xie

I Introduction

Market entry is always risky, the uncertainty can come from demand, cost, competitors and even policy. In the end, the risk lies in the lack of information, firm can never know anything perfectly. Uncertainty can lead to unprofitability and ”failure”, and firm wants to fail fast if ”failure” is inevitable and exit as soon as possible. Demand uncertainty is one main source of uncertainty. Information acquisition is one way to reduce uncertainty and most of the learning is costly. Before entering a big new market, firm can test products in smaller market, or hire forecast team to predict the market demand; even with that the uncertainty can never be fully resolved. Hitsch (2006) reports in the U.S. ready-to-eat breakfast cereal industry, close to 70% of all new products are scrapped within two years after the introduction. In U.S airline industry, the failure rate is also high, nearly 30% of all non-stop routes are terminated within six months operation. Free learning is also possible, firm can gather more information from other player's realized market result. For example, if an airline carrier wants to fly a route between airport A and airport B and does not know the real demand for that route, but noticed the other carrier who flied that route didn't end up very well, may choose not to fly that route because the history has proved that route unprofitable. But because of cost heterogeneity, one route that is unprofitable for one carrier can be profitable for another carrier who has lower cost. Relating it to an interesting phenomenon that in the U.S airline industry, low cost carrier is sometimes more interested in flying the abandoned (by legacy carriers) routes and failure rates are actually lower in the abandoned market, it looks like abandoned market is not ”bad” at all. This raises the question of what can a low cost carrier learn from legacy carrier's history? The most likely explanation is market demand. Many routes were terminated because the

*This prject is still in progress.

realized demand is much lower than the predicted one. However, the learning from other's history comes with the curse, the fact that the market being abandoned itself is a signal of unprofitability, so the low cost airline has to face a fundamental trade off between reduced uncertainty in market demand and unprofitability. This research aims to find the mechanism behind this.

A Related Literature

Learning demand is nothing new. Many have been done in theory and application I list the closest paper (to my research) on learning demand. [**Hitch(2006)**] studies the decision problem of a ready-to-eat cereal firm who is uncertain about the demand and develops a learning model where the firm learns the unknown demand parameter from the realized product shares. His model shows that the value of reducing uncertainty can be large, and that explains why firm is still actively launching new product even though the point estimate of profits is negative. [**jeon2022learning**] studies the role of demand uncertainty in explaining cyclical investment fluctuations in the container shipping industry, in her model, demand is AR(1) and firm learns the AR(1) parameters from firm's own realized market share. All those papers are about learning demand, whether static demand or dynamic demand, but not so much has been done in learning from your competitor's realized market share and how to infer that to firm's own expected future profits. And that could be the potential contribution of this project.

Airline industry has been well studied, but mostly on competition, merger, demand estimation, real time pricing strategy and not so much have been done in market selection and learning from rivals. [**Berry(2010)**] compare the airline industry in 1999 and 20006 and find passengers displayed a stronger preference for nonstop flights and are becoming more price sensitive. If the change in preference is consistent, this could explain the success of low cost carriers because they charge a lower ticket price and mainly fly non-stop routes. [**Berry(1996)**] shows hubbing airline has ability to raise prices to those price-inelastic business travelers but this high prices do not provide "monopoly umbrella" to other non-hub airlines. Another more related paper on airline entry is [**Berry(1992)**] where he assumes the more profitable firm enters the airport first and airport presence can be an indication of underlying profitability. [**Rupp(2023)**] provide a long-term perspective on airline performance and find steady decrease in prices and increase in market share of low-cost carriers. Overall, because of the competition and rise of LCC, consumer welfare increased dramatically.

B U.S Airline Industry and Route Selection

This section provides overview of the U.S. airline industry. After the Airline Deregulation Act of 1978, airlines were given more freedom to set their own routes, fares, and schedules, leading to increased competition and lower prices for consumers. This resulted in a significant expansion of the airline industry, with new airlines entering the market and existing airlines expanding their operations. JetBlue entered the market in 1998 as low-cost carrier; Southwest also entered as low-cost carrier in 1971 and became the most successful airline in the U.S. However, other smaller carriers struggled to compete with larger, more established airlines which led to consolidation within the industry with many smaller airlines merged into big airlines or went bankrupt. Delta acquired Northwest Airline in 2008, creating one of the world's largest airlines. Continental Airlines merged to United Airlines in 2010 which enhanced United Airline's network and made it the largest airline by traffic at that time. In 2013, the merger between American Airlines and US Airways made it the world's largest airline by fleet size and destination served. Some other airlines including Eastern Airlines and Trans World Airlines (TWA) went bankrupt because of increased competition and high operating costs. Overall, deregulation has led to a more dynamic and competitive airline industry, with lower fares and more choices for consumers. However, the fierce competition and less opportunity make it harder for small airline to strive and for big airline to be sustainable and profitable.

Airlines are trying their best to avoid competition. One airline industry analyst says "The only way to be competitive in the airline industry is to not compete with other airline." One common way is to collude with other players, collusion is usually hard to observe and prove. Another way is to deter entry, for example, New York remains a highly constrained market with plenty of airlines sitting on the outside looking for more slots and slot squatting is an anti-competitive thing going on in those slot constrained airports. The other way to avoid competition is to enter the market that your competitors probably won't be interested in. For example, legacy carriers use hub-to-spoke system and most of the routes are either connected to their bubs directly or indirectly through a connecting airport, and they probably won't be interested in flying a route that is too far away from their web. This could be an opportunity for smaller airlines who use point-to-point system and are more flexible in choosing routes and most of the low cost carriers are using the system. Beside this, another strategy is to fly the routes that were abandoned by the competitors. Those abandoned markets are usually the market with small demand or low willingness to pay to support the big airlines. However, the "bad" market could still work for low cost carriers if they have cost advantage. CEO of Breeze Airline (an ultra-low-cost-carrier) Neeleman said "We can get trip costs that are 20% to 25% below where the other guys are. We can go into markets that may not make sense for them but make all the sense in the world for us." Table 1 lists some examples of LCC entering the abandoned routes soon after monopolist abanoned that route.

II Data

Two main data sources are Airline Origin and Destination Survey (DB1B) and T100, both are published by US Department of Transportation (DOT). The DB1B data is a 10% sample of airline tickets from reporting carriers. The data provides detailed information on origin, destination airports, ITIN fares, operating carriers, distance on a quarterly base. The main focus of this research is the entry and learning of LLC (low cost carrier) and most of the LCC routes are non-stop flight, so I select the routes where only non-stop service is available. T-100 Domestic Segment contains domestic non-stop segment monthly data, including carrier, origin, destination, aircraft type and service class for transported passengers, available capacity. I use the T-100 data to filter the targeted routes and market.

I also use Census data to compute market size because market size is defined as the average population of origin and destination cities. U.S. Census provides demographic information including population on the county level, it is conducted every ten years and linear interpolation is needed to compute the annual population.

A Sample Selection

Most papers define a market as a directional city to city pair, but because a big city tends to have multiple airports and LCC sometimes strategically choose to occupy a secondary airport, I define a market as a non-directional airport to airport pair. For example, the flight from IAD (Dulles International Airport) to JFK (John F. Kennedy International Airport) and the flight from JFK to IAD are considered as the same (route). Same as Berry & Jia (2010), market size is computed as the geometric mean population of the origin city and the destination city.

I use the data from 1999 to 2018. Same as Rupp, Kuiken & Williams (2023), I classify each airline into one of the three categories: Global Network Carriers (GNC[1]), Low Cost Carriers (LCC[2]), and Other. There are about 200 carriers in the Other category but their overall volume takes less than 20% of the total volume, so in this study I only focus on GNC and LCC[3] and their interaction.

There are some airports in a small town with market size smaller than 1000 people, and market like this will not be the main interest in this project, so in the demand estimation part I select the route with market size greater than 850,000 people in 2006[4].

B Data Summary

Figure 1 shows the evolution of number of domestic non-stop routes in the US starting from 2001, the number of non-stop routes increased drastically. Defining a route as monopoly at year t if only one carrier flies that route at year t, Figure 2 shows the change of number of routes by different groups, overall the number of routes with monopoly, duopoly increased during the past 20 years but number of routes with monopoly increased the most. Figure 3 shows the proportion change, the proportion of routes with monopoly increased from 35% in 1999 to 41% in 2018 while the proportion of routes with more than 2 players dropped by 5%.

For the actual volume, Figure 4 shows the total number of passengers transported increased except for the years after the financial crisis or recession in 2000 and 2008, more passengers imply a growing demand for air travel. Although GNC is always the biggest contributor to the volume, its contribution is diminishing and the volume share of LCC more than doubled, a growing proportion of passengers chose LCC and are taking passengers away from GNC. Figure 5 shows the general trend of volume share. Load factor is the ratio of actual sold seats to the number of available seats and a higher load factor indicates better efficiency in utilizing the available capacity of the airline's fleet. From Figure 6, GNC has the highest load factor and load factor for all three groups increased. 7 shows the number of routes served by those GNC, LCC and Other. GNC flied around 4000 routes in 1999 and gradually decreased the number of routes served. Meanwhile LCC increased the number of markets served and start from 2012 those two groups served about the same amount of markets,

[1]GNCs include American Airlines, American West Airlines, Delta Airlines, Alaska Airlines, Continental Airlines, Hawaiian Airlines, Northwest Airlines, United Airlines, and US Airways

[2]LCCs include AirTran Airways, ATA Airlines, JetBlue Airlines, Midwest Airlines, National Airlines, Southwest Airlines, Virgin America, Frontier Airlines, Allegiant Air, Breeze Airways, Spirit Airlines, Sun Country Airlines, and Avelo Airlines.

[3]LCC in this project includes Ultra-Low Cost Carriers (ULCC): Frontier, Allegiant, Breeze, Spirit, Sun Country, and Avelo.

[4]same as Berry & Jia (2010)

around 3000 routes.

From 1999 to 2018, there had been more than 22400 unique routes and 5808 of them are short-lived (existed in US history for less than a year). 16073 of routes have never been served by LCC and 13445 of routes have never been flied by GNC, so in theory airline carrier has many potential routes to fly. Figure 8 shows the number of markets entered by LCC and GNC, at the beginning of 2000, GNC entered more markets than LCC but started from 2006 LCC entered more markets than GNC. This can be rationalized by the fact that LCC was the relatively new player at the beginning of 21 centry and was expanding the network while GNC has been in market since 1980 and has established the network well. From Figure 9, despite the fact that LCCs were entering a large amount of markets each year, they were not interested in exploring routes that had never been served by any other players before, the reason of that is not clear, it could be LCC is more risk averse and can not afford too many failure, or their demand forecast ability is much worse than GNC and they are relying on GNC to provide signals. On the contrary, Figure 10 shows GNC was actively exploring new routes at the beginning of 2000 but then very quickly slowed down the exploration pace. This makes sense because the number of potential routes is not infinity, in the end most, if not all, of the profitable routes will be entered. Figure 11 give s direct comparison of completely new markets explored by GNC and LCC, given the fact that GNC and LCC do not differ too much in volume and LCC some years entered more markets than GNC, this indicates GNC and LCC have different entry pattern or market selection. Figure 12 compares the entry pattern of GNC and LCC each year and the difference is obvious, 40% to 40% of the GNC entered routes are completely new routes, meanwhile the proportion is only about 20% for LCC.

Once a profitable route is discovered, it will soon be entered by other airlines. For routes explored by LCC, 8.68% were entered by other airlines within one year; and 15.47% were entered by followers in three years. For GNC explored routes the proportion is higher, 15.65% were entered by other airlines within one year and 28.53% for three years. If we define a route is a failure if that route has only been served for less than six months (not necessarily have to be consecutive months), GNC has a failure rate of 33.50% and LCC has a failure rate of 39.30%. However, if limit to routes that were abandoned by GNC, LCC's failure rate decreased to 34.70%. The decrease in failure rate is counterintuitive, because we usually expect an abandoned market to be less profitable and more likely to fail. The learning story could be a potential answer to this mystery, even though market is abandoned, entrant actually now knows this market better.

Table 2 and Table 3 summarize the overall descriptive statistics for LCC and GNC non-stop flights from 2001 to 2010. GNC's average ITIN fare is \$351.7, about twice as much as LCC's. But note that the average distance and market size for those two group does not differ that too much, although LCC has a slightly smaller average market size and shorter distance, which is consistent with finding in other papers that LCCs focus on shorter routes and smaller markets. Figure 13 tracks the average ITIN fare, distance, market share and tour for LCC, GNC and other from 2001 to 2010, overall the gap between ITIN fare between GNC and LCC is persistent and stable. And the mean route distance of group Other is only half of the mean route distance of GNC and LCC, it somehow indicates that GNC and LCC are competing on the same track while Other is providing different product.

III Demand Estimation

A Utility function

Discrete choice model for a specific route:

$$u_{ijt} = \alpha p_{jt} + x_{jt}\beta + \xi_{jt} + \epsilon_{jt} \tag{1}$$

where x_{jt} is a vector of product characteristics which include distance, distance squared, tourist[5], airline carrier fixed effect. p_{jt} is ticket price of carrier j at year t, ξ_{jt} is unobserved characteristics of product j to econometrician and ϵ_{jt} is logit error and is iid.

The logit demand for route j at year t:

$$\ln(s_{jt}) - \ln(s_{0t}) = \alpha p_{jt} + x_{jt}\beta + \xi_{ijt} \tag{2}$$

We can observe market share from the data, and use that to back up the left hand side of equation (2), and then use linear regression to get α and price elasticity.

[5]tourist is a dummy variable and equals one if either origin or destination airport locates in Las Vegas or Florida.

B Instruments

Price can be correlated with unobserved ξ_{jt} so α can be endogenous. A classic instrument is the number of products in a market, but I define each route as a market, so by definition the number of product in a market is 1. To capture the competitiveness of the market, the instrument includes the number of rivals that also fly that route.

Another set of IVs are variables that affect costs but do not affect demand. I use the quarterly jet fuel prices per gallon as another instrument. The number of destinations to which a carrier flies from the origin airports and the destination airports reflect the size of operation and can affect the cost of the flight, so another set of instruments includes the number of routes that airline also serves from the origin airport, the number of routes that airline also serves from the destination airport.

C Results

Estimates are on the aggregate level. Table 4 shows the coefficients in 2001 and in 2006. The second and forth columns are estimates from OLS, third and fifth columns are estimates form 2SLS. Simple comparison of ITIN fare coefficients between 2001 and 2006 shows a more negative price coefficient. Among all the airlines in the sample, legacy carriers including United Airline, Delta, American Airline, Continental Airlines (merged into UA in 2012) give a higher utility while low cost carriers including JetBlue, Spirit Airline, Sun Country Airlines have a large negative fixed effect.

Figure 14 shows the price elasticity from 2001 to 2010. My estimates of price elasticity is larger (in absolute value) than the estimates from other papers. Berry (2010) records estimates of price elasticity -1.55 in 1999 and -1.67 in 2006 while my estimate is -4.97 in 2006. Rupp, Kuiken and Williams (2023) track the top 250 airports in the US throughout 1998 to 2022 and find the elasticity ranges from -19.07 to -0.09 with median be -0.79. White (2023) did a similar research and built a tax calculator as IV and finds elasticity is -1.71 for the median market. I can come up with several potential reasons why my estimates are different from theirs, first we define market differently; some papers define market as directional flight while I define it as non-directional flight; Berry's earlier paper define market as city-to-city pair while I define market as airport-to-airport pair; Rupp's paper defines market as one-way origin-and-destination pair, regardless if there is layover. Our sample selection is also different, Rupp's paper selects the top 250 airports while I select the routes with market size greater than 850,000 in 2006. Because of all those differences, the passenger composition is different, in Berry's paper, the proportion of tourists[6] passengers is 59% and business travelers passengers is 41% while in my data the proportion of tourists passengers[7] is close to 80%, if we believe leisure passengers are more price sensitive than business passengers, it is not very surprising that my price elasticity is more negative than Berry's estimates. The final potential reason could be my IV is not good enough (correlation between endogenous variable and IV is not big enough or IV itself is not exogenous). To improve my IV, I will check the tax rate in different airports and see if adding that will improve the performance of IV.

IV Learning Model

A Setup

The model will be a discrete infinite time horizon model with two players. It is a sequential move game and at the beginning of each period, mover has to make entry/exit decision. Re-entry is allowed and there is no entry cost. After mover's entry/exit decision, firms in the market (if any) set prices and realize profit. There is a new market with an unknown parameter λ_t in consumer's utility function and λ_t is time variant. Utility of consumer i choosing firm j at period t:

$$u_{ijt} = \lambda_t + \alpha P_{tj} + x_j\beta + \epsilon_{ijt} \tag{3}$$

where p_{tj} is the price set by firm j, x_j is the firm specific characteristics, ϵ_{itj} is iid utility shock and is Type I Extreme Value distributed. Assume λ_t is AR(1):

$$\lambda_t = \gamma_0 + \gamma_1\lambda_{t-1} + \eta_t \tag{4}$$

where η_t is iid shock, $\eta_t \sim N(0, \sigma_\eta^2)$, and $\gamma_1 \in (0, 1)$. At period 0:

$$\lambda_0 = \gamma_0 + \eta_0 \tag{5}$$

At the beginning of period t, firms observe a same but noisy signal of λ_{t-1}: $\tau_t = \lambda_{t-1} + \omega_t$. ω_t is iid and $\omega_t \sim N(0, \sigma_\omega^2)$. Imagine both airlines have their forecasting team, did some market survey and

[6] Tourists is labeled as type 1 and business travelers is labeled as type 2 in his paper.

[7] In my code, tourists is coded as holding the coach class tickets, business travelers are those holding the first class and business class ticket.

τ_t is the signal they received regarding the last period's demand λ_{t-1}. The signal will help them forecast the future demand. Assume firms know $\{\gamma_1, \sigma_\eta, \sigma_\omega\}$ and the only unknown parameter for firms is γ_0. For econometricians, we do not know any parameters but have to estimate them. Firm's belief is about γ_0. They re-estimate γ_0 based on the received new signal and the realized market share (if any) in the previous period. If either firm was in the market then they observe λ, the true market demand. And if neither was in the market, they observe the signal τ_t. Note that if either firm was in the market, they still receive signal, but the signal provides no additional information about γ_0, and they ignore signal. Assume fixed cost and marginal cost are public information. Also assume both firms have the same belief and belief is public information. Will relax this assumption to information asymmetric later.

Notation:

- History $h_t = \{h_{\tau t}, h_{Et}\}$
  - $h_{\tau t} = \{\tau_1, \tau_2, \tau_3 ..., \tau_t\}$, $h_{\tau t}$ is a sequence of signals that firms used to forecast future λ. Note that if at any period s, there was at least one firm in the market, τ_{s+1} will be replaced by λ_s because firms can figure out λ_s from the realized market share in period s.
  - $h_{Et} = \{(E_{11}, E_{12}), (E_{21}, E_{22}), ..., (E_{t-1,1}, E_{t-1,2})\}$ is the history of firms presence in the market. E_{tj} can take value 0 or 1. For example, $(E_{21}, E_{22}) = (1, 0)$ means during period 2, firm 1 was in the market and firm 2 was not in the market.
- Mover j's action set: $a_{tj} = \{exit,\ entry\}$. These two words are abused here, for the mover who is already an incumbent, "entry" means stay in the market; and for the mover who is potential entrant, "exit" simply means keep staying away. Action determines if firm wants to produce in that period.

Sufficient statistics: history h_t can be long and add unnecessary computation burden. Because the role of history in this model is to provide belief for firms, a sufficient statistics for history will be belief and I do not have to store a huge set of history as state variable. And state variable for firm j will be $s_{tj} = \{b_t, \lambda_t, E_{t-1,j}, E_{t-1,-j}\}$, the first term is belief of λ_t, second term is an indicator of if firm j was in market in lat period, third term indicates if firm -j was in market in last period.

Timeline:

1. At period 1, the firm with higher fixed cost is called to move first. Non-mover firm waits.

2. Both firms receives the noisy signal $\tau_1 = \lambda_0 + \omega_1$ and form expectation of λ_1. Mover decides if he should enter, if enters, sets prices that maximizes the period profits. At the end of period 1, profit realization (if any) is observed by both firms.

3. At period 2, both firms again receive a noisy signal $\tau_2 = \lambda_1 + \omega_2$. Non-mover in period 1 becomes mover in this period, both firms update belief based on signal and previous period's profit realization. Then the mover makes entry decision. Post entry/exit decision, firms in the market set prices to maximize period profits.

4. Game continues and never ends.

B Belief updating

Firm's belief is over λ_t, or more precisely, is over the term γ_0. The belief in each period is described by a normal probability density $p_t(\lambda_0) = N(\lambda_0|\mu_t, \sigma_t^2)$, so belief is fully described by two state variables μ_t and σ_t^2 where μ_t is the mean of normal distribution and σ_t^2 is the variance of the distribution.

My current plan is to use Bayesian updating and the belief update part is not fully solved yet, but a simpler version can be easily solved. For example, if $\gamma_1 = 0$, which means $\lambda_t = \gamma_0 + \eta_t$ and prior belief is $(\mu_t, \sigma^2 t)$. If the firm was in market at t, λ_t is revealed and the posterior belief is

$$\mu_{t+1} = \mu_t + \frac{\sigma_t^2}{\sigma_t^2 + \sigma_\eta^2}(\lambda_t - \mu_t)$$

and

$$\sigma_{t+1}^2 = \sigma_t^2(1 - \frac{\sigma_t^2}{\sigma_t^2 + \sigma_\eta^2})$$

If the firm was not in market at t, the firm receives the noisy signal $\lambda_t + \omega_t$ and the posterior belief is

$$\mu_{t+1} = \mu_t + \frac{\sigma_t^2}{\sigma_t^2 + \sigma_\eta^2 + \sigma_\omega^2}(\lambda_t + \omega_t - \mu_t)$$

and

$$\sigma_{t+1}^2 = \sigma_t^2(1 - \frac{\sigma_t^2}{\sigma_t^2 + \sigma_\eta^2 + \sigma_\omega^2})$$

It is easy to notice that if the firm observes λ_t instead of $\lambda_t + \omega_t$, the firm can update his belief faster, and the learning speed improvement depends on the variance of the noise in the signal. The general belief updating part is remained to be developed.

C Value function

When it is firm j's turn to move, firm j first determines if he should be in the market or walk away. If he decides to be in the market, he sets prices that maximizes his expected net present value. Mover j's value function:

$$V_{tj}(s_{tj}) = \max\{E[V_{tj}(s_{tj})|\text{enter}],\ E[V_{tj}(s_{tj})|\text{exit}]\} \tag{6}$$

and optimal decision a^*_{tj}:

$$a^*_{tj} = \underset{a_{tj}=\text{enter,exit}}{\text{argmax}}\ E[V_{tj}(s_{tj})|a_{tj}] \tag{7}$$

$E[V_{tj}(s_{tj})|\text{enter}]$ is the value of being in the market at t:

$$E[V_{tj}(s_{tj})|\text{enter}] = E[\pi_{tj}(s_{tj})] + \beta E[W_{t+1,j}(s_{t+1,j})|s_{tj}, a_{tj} = \text{entry}] \tag{8}$$

and state transition:

$$\begin{aligned}
\lambda_{t+1} &= \gamma_0 + \gamma_1 \lambda_{t-1} + \eta_t \\
E_{t,j} &= 1 \\
E_{t,-j} &= E_{t-1,-j}
\end{aligned}$$

Note that the problem is stationary because I assume firms use current period's belief b_t in forecasting future demand, and continuation value depends on b_t but not b_{t+1}, b_{t+2} and so on.

$E[V_{tj}(s_{tj})|\text{exit}]$ is the value of walking away at t:

$$E[V_{tj}(s_{tj})|\text{exit}] = 0 + \beta E[W_{t+1,j}(s_{t+1,j})|s_{tj}, a_{tj} = \text{exit}] \tag{9}$$

and state transition:

$$\lambda_{t+1} = \gamma_0 + \gamma_1 \lambda_{t-1} + \eta_t$$
$$E_{t,j} = 0$$
$$E_{t,-j} = E_{t-1,-j}$$

Note that it is a sequential move game, mover in this period will become a non-mover in next period and non-mover's value W_{tj}:

$$W_{tj}(s_{tj}) = E[\pi_{tj}(s_{tj})] + \beta E[V_{t+1,j}(s_{t+1,j})|s_{tj}, a^*_{t,-j}] \tag{10}$$

with state transition:

$$\lambda_{t+1} = \gamma_0 + \gamma_1 \lambda_{t-1} + \eta_t$$
$$E_{t,j} = E_{t-1,-j}$$
$$E_{t,-j} = a^*_{t,-j}$$

where $a^*_{t,-j}$ is rival firm -j's optimal choice because rival is mover in this period and

$$a^*_{t,-j} = \underset{a_{t,-j}=\text{enter,exit}}{\text{argmax}} E[V_{t,-j}(s_{t,-j})|a_{t,-j}] \tag{11}$$

D Profit function

Profit if λ_t was known:

$$\tilde{\pi}_{tj}(s_t; \lambda_t) = (P_{tj} - MC_j)S_{tj} - FC_j \tag{12}$$

where S_{tj} is the market share of firm j at time t:

$$S_{tj} = \frac{exp(\delta_{tj})}{1 + \sum_j exp(\delta_{tj})} \tag{13}$$

and δ_{tj} is the mean utility:

$$\delta_{tj} = \lambda_t + \alpha P_{tj} + x_j \beta \tag{14}$$

Take FOC of equation (12) w.r.t P_{tj}:

$$S_{tj} + (P_{tj} - MC_j)\frac{\partial S_{tj}}{\partial P_{tj}} = 0 \tag{15}$$

And the NE price $P_t^* = \{P_{t1}^*, P_{t2}^*\}$ are the prices such that equation (15) is satisfied for all firms which are in the market at time t simultaneously. If there is no competitor in market at t, P_t^* is a singleton instead of a vector. Profit given optimal pricing is $\tilde{\pi}_{tj}(s_t, e_{tj}; \lambda_t, P_t^*)$.

Finally, the $E[\pi_{tj}(s_t, e_{tj})]$ term in equation (6) is defined as

$$E[\pi_{tj}(s_t, e_{tj})] = E[\tilde{\pi}_{tj}(s_t; \lambda_t, P_t^*)] \tag{16}$$

The expectation on the RHS is over λ_t because that is only source of uncertainty in period profit function.

E Solution Strategy

Value function iteration should work.

V Simulation of two simpler models

A A simple infinite periods monopoly model

This model is an infinite, discrete time model. There is a single firm who has to make entry/exit decision at the beginning of each period. There is no entry cost and re-entry is allowed. There are two possible states of the world, H or L with transition probability $P_H = \text{Prob}(S_{t+1} = H|S_t = H)$ and $P_L = \text{Prob}(S_{t+1} = L|S_t = L)$. Also assume $P_H = P_L = 0.8$. Assume $\pi(S_t = H) = \pi(H) > 0 > \pi(S_t = L) = \pi(L)$ where $\pi(S_t)$ is the profit function when firm is in the market. b_t is firm's belief of being in good state and $b_t = \text{Prob}(S_t = H)$, and the state variable for the firm will be b_t. At the beginning of period t, there is probability ϕ that firm can receive a signal which reveals the value of previous period's state: S_{t-1}. If firm is in market at period t, S_t will be revealed. Meaning there are two ways for firm to learn the true value of S_t: he can either enter the market in period t to get S_t for sure, or stay out and hope he is lucky enough to receive the signal.

Value function:

$$V(b_t) = \max\{V(b_t, \text{entry}), V(b_t, \text{exit})\} \tag{17}$$

and

$$a_t^* = \underset{a_t = \text{entry, exit}}{\text{argmax}} \; V(b_t, a_t) \tag{18}$$

And

$$\begin{aligned} V(b_t, \text{entry}) &= E[\pi(b_t)] + \beta E[V(b_{t+1})|\text{entry}] \\ &= b_t[\pi(H) + \beta V(b_{t+1} = 0.8)] + (1 - b_t)[\pi(L) + \beta V(b_{t+1} = 0.2)] \end{aligned} \tag{19}$$

The second line of equation (3) follows from the fact that once entered the market, S_t is revealed and b_{t+1} will be either 0.8 or 0.2, depending on S_t is H or L.

And

$$\begin{aligned} V(b_t, \text{exit}) &= 0 + \beta E[V(b_{t+1})|\text{exit}] \\ &= 0 + \beta\big[\phi[b_t V(b_{t+1} = 0.8) + (1 - b_t)V(b_{t+1} = 0.2)] + \\ &+ (1 - \phi)V(b_{t+1} = 0.8b_t + 0.2(1 - b_t))\big] \end{aligned} \tag{20}$$

Equation (4) shows that the value of exit incorporates the possibility firm may receive the signal and the true value of S_t is revealed even though he was not in market at t; and also the possibility that he won't receive the signal and has to form the belief b_{t+1} based on b_t. Note that belief $b_t \in [0.2, 0.8]$.

Requirement:

1. $V(b = 0.8) = \max\{V(b = 0.8, \text{entry}), V(b = 0.8, \text{exit})\} = V(b = 0.8, \text{entry})$. Otherwise, if firm with the most optimistic belief does not want to enter, market will never be entered.

2. $V(b = 0.2) = \max\{V(b = 0.2, \text{entry}), V(b = 0.2, \text{exit})\} = V(b = 0.2, \text{exit})$. Otherwise, if firm with the least optimistic belief wants to enter, market will never be empty.

This model can be solved using value function iteration. Figure 15 shows the value of the firm, the value increases in firm's belief. Firm's optimal decision rule is simple, when belief is smaller than the cutoff (0.382 in my setting), value of staying out is greater than the value of entry and firm does not enter, when firm becomes more confident and has a belief higher than the cutoff, firm enters. Figure 16 shows the simulation results of 100 periods. The main takeaway is the signal can reduce firm's probability of making Type I error (enter the bad market), but could increase firm's probability of making Type II error (did not enter the good market). Another interesting thing, when firm knows he has high probability of receiving the signal, he loses some information acquisition motive and only enters the market when he has a high belief.

B A simple three periods duopoly model

The model setting is quite different from the monopoly model. There are two firms, firm 1 and firm 2. There are 2 possible states, H and L. In each period, there is some probability that period is High state, but firm does not know that real probability. Instead, firms know two candidates but do not know which one is true. Let two lotteries represent the two candidates: lottery 1: $(p_1, 1 - p_1)$ and lottery 2: $(p_2, 1 - p_2)$. Lottery 1 means there is probability p_1 the period be High state and probability $1 - p_1$ that period be Low state. Similar interpretation for lottery 2. Both firms have same belief b_t which represents their belief of lottery 1 being the real one. Profit for firm 1 when High state is (π_{MH1}, π_{DH1}), and when Low state is (π_{ML1}, π_{DL1}) (D means duopoly and M means monopoly). At the beginning of each period, each firm has 50% chance being called to make entry/exit decision. If anyone is the market, both firm will observe the market realization and update their belief b based on that using Bayesian updating. This is a 3 periods game, at the beginning of period 4, both firms are asked to leave the market forever. Finally , the state variable $s_t = \{b_t, e_t\}$ where b_t is belief of lottery 1 is the true lottery, $e_t = (e_{-jt}, e_{jt})$ represents the rival -j's existence and firm j's existence in market before anyone is called to play. So e_t can take four combinations $(0, 0), (1, 0), (0, 1), (1, 1)$.

The following part is the value functions for firm 1, the value function for firm 2 is similar. Value function when none of the firms is in market at the beginning of $t = 3$, meaning $e_3 = (0, 0)$:

$$V_3(b_3, 0, 0) = \frac{1}{2} \times 0 + \frac{1}{2} \max\{0, b_3(p_1\pi_{MH} + (1 - p_1)\pi_{ML}) + (1 - b_3)(p_2\pi_{MH} + (1 - p_2)\pi_{ML})\} \quad (21)$$

Value function when both of them are in market, meaning $e_3 = (1, 1)$:

$$\begin{aligned} V_3(b_3, 1, 1) = \frac{1}{2} &\times (\text{Prob(Firm 2 enters)}E[\pi_D] + (1 - \text{Prob(Firm 2 enters)})E[\pi_M]) + \\ &\frac{1}{2}\max\{0, b_3(p_1\pi_{DH} + (1 - p_1)\pi_{DL}) + (1 - b_3)(p_2\pi_{DH} + (1 - p_2)\pi_{DL}\} \end{aligned} \tag{22}$$

where is first $\frac{1}{2}$ part represents the expected utility when rival is called to play, and the second $\frac{1}{2}$ part is the expected when this firm is called to play. Prob(Firm 2 enters) is the probability of rival deciding to enter conditional on rival is called to play and Prob(Firm 2 enters)=1 if for firm 2, $b_3(p_1\pi_{DH} + (1 - p_1)\pi_{DL}) + (1 - b_3)(p_2\pi_{DH} + (1 - p_2)\pi_{DL} > 0$.

Value function when firm 1 is in market but firm 2 is not, meaning $e_3 = (0, 1)$:

$$\begin{aligned} V_3(b_3, 0, 1) = \frac{1}{2} &\times (\text{Prob(Firm 2 enters)}E[\pi_D] + (1 - \text{Prob(Firm 2 enters)})E[\pi_M]) + \\ &\frac{1}{2}\max\{0, b_3(p_1\pi_{MH} + (1 - p_1)\pi_{ML}) + (1 - b_3)(p_2\pi_{MH} + (1 - p_2)\pi_{ML}\} \end{aligned} \tag{23}$$

Value function when firm 1 is not in market but rival is in market, meaning $e_3 = (1, 0)$:

$$V_3(b_3, 1, 0) = \frac{1}{2} \times 0 + \frac{1}{2}\max\{0, b_3(p_1\pi_{DH} + (1 - p_1)\pi_{DL}) + (1 - b_3)(p_2\pi_{DH} + (1 - p_2)\pi_{DL})\} \tag{24}$$

Using backward induction, can back up firm's entry/exit strategy at the last period, and from that back up firm's strategy at period 2 and so on.

Value function at period 2 for firm 1 is generalized to this:

$$\begin{aligned} V_2(b_2, e_2) = \frac{1}{2} &\times E[\pi(b_2, e_2) + \beta V_3(b_2, e_3)|\text{rival is the mover}] + \\ &\frac{1}{2} \times E[\pi(b_2, e_3) + \beta V_3(b_2, e_3)|\text{agent is the mover}] \end{aligned} \tag{25}$$

And at period 1, no firm is in the market yet, so $e_1 = (0, 0)$, value function for firm 1:

$$\begin{aligned} V_1(b_1, e_1 = (0, 0)) = \frac{1}{2} &\times E[0 + \beta V_2(b_1, e_2)|\text{rival is the mover}] + \\ &\frac{1}{2} \times E[\pi(b_1, e_1) + \beta V_2(b_1, e_2)|\text{agent is the mover}] \end{aligned} \tag{26}$$

Belief updating is Bayesian, once a high state is observed, it make firms believe lottery 1 is more likely to be true(recall lottery 1 gives a higher probability for H), and b will increase. Let b' be the posterior belief and b be the prior belief, after observed H:

$$b' = \frac{bp_1}{bp_1 + (1-b)p_2} \tag{27}$$

This model can be easily solved by backward induction. The model parameters are set as $p_1 = 0.8, p_2 = 0.4$, firm 1's profit profile $\{\pi_{MH} = 3, \pi_{ML} = -3, \pi_{DH} = 1, \pi_{DL} = -4\}$ and firm 2's profit profile $\{\pi_{MH} = 2, \pi_{ML} = 1, \pi_{DH} = -1, \pi_{DL} = -4\}$. The parameter values are such that firm 2 has advantage in low state because it can survive as a monopolist in low state while firm 1 can not. At period 1, result shows if called to play, firm 1 will only enter if his belief is greater than 0.238 while firm 2 will always enter. Figure 17 and 18 shows the entry decision in period 2 and period 3, as time approaches the end of the game, firm 1 behaves more cautiously and never enter as long as firm 1 is in market. Figure 19, 20 and 21 show the value of two firms at three periods. There is sudden jump in firm's value which is caused by the sudden change in rival's entry strategy. When it is closer to the end of the game (period 3 in this model), firm 2 prefers to stay in market before being called to play, while in period 2 firm 2 prefers not to have market presence before being called to play, this is due to the fact firm 2 can never be profitable in duopoly while firm 1 can still be profitable in duopoly if market is good enough, and to avoid the possible duopoly in period 3, firm 2 will choose not to enter market in period 2 if called to play when belief is high.

VI Appendix

Entrant	Monopolist who left	Route	Abandoned Year-M	Entry Year-M
Midwest Airlines	Northwest Airlines	Sacramento - Green Bay	2007 – 02	2007 – 12
Southwest	Northwest Airlines	Raleigh/Durham - Chicago	1999 – 03	1999 – 06
Allegiant Air	Delta	Des Moines - Hartford	2002 – 11	2003 – 09
Southwest	Delta	Des Moines - Memphis	2015 – 07	2016 – 02
JetBlue	United Airlines	San Jose - Fort Lauderdale	2016 – 11	2017 – 04
Breeze	Spirit	Orlando - Charleston	2023 – 05	2023 – 05
Avelo	Frontier	Wilmington New Castle Airport	2022 – 06	2023 – 02
Avelo	American Airline	Tweed New Haven Airport	2021 – 10	2021 – 11
Southwest	Continental Airlines	Cleveland - Birmingham	1999 – 05	1999 – 08

Table 1: Examples of LCC entering the abandoned markets

Statistic	N	Mean	St. Dev.	Min	Max
ITIN fare	2,506,358	351.7	288.5	26	1,999
ITIN yield	2,506,358	0.4	0.3	0.01	4.1
Distance	2,506,358	1,238.2	718.3	55	4,983
Market size	2,506,358	3,905,982.0	3,529,995.0	808,481.0	12,872,322.0
Tour	2,506,358	0.2	0.4	0	1

Table 2: Descriptive statistics for GNC, averaged over 2001-2010

Statistic	N	Mean	St. Dev.	Min	Max
ITIN fare	703,468	178.7	106.6	26	1,995
ITIN yield	703,468	0.2	0.1	0.01	3.5
Distance	703,468	1,118.7	649.6	17	4,983
Market size	703,468	3,438,503.0	3,536,290.0	809,653.5	12,872,322.0
Tour	703,468	0.2	0.4	0	1

Table 3: Descriptive statistics for LCC, averaged over 2001-2010

	Dependent variable:			
	Mean utility			
	OLS 2001	*2SLS* 2001	*OLS* 2006	*2SLS* 2006
	(1)	(2)	(3)	(4)
ITIN Fare	-0.0002^{***}	-0.0080^{***}	-0.0002^{***}	-0.0232^{***}
	(0.00001)	(0.0003)	(0.00001)	(0.0012)
Distance (1,000 miles)	-1.4951^{***}	1.2024^{***}	-0.6866^{***}	3.1257^{***}
	(0.0116)	(0.1096)	(0.0093)	(0.1970)
Distance2 (1,000,000 miles)	0.3574^{***}	-0.0404^{**}	0.1357^{***}	-0.2756^{***}
	(0.0039)	(0.0177)	(0.0031)	(0.0253)
Tourist place (FL/LAS)	0.0169^{***}	-0.5985^{***}	0.4190^{***}	-0.9539^{***}
	(0.0064)	(0.0277)	(0.0047)	(0.0726)
PSA Airlines			-0.4170^{***}	-0.9443^{***}
			(0.0405)	(0.1899)
American Airline	2.1492^{***}	2.0038^{***}	2.1336^{***}	1.8464^{***}
	(0.0142)	(0.0297)	(0.0172)	(0.0813)
Alaska Airlines	3.4996^{***}	1.5962^{***}	3.2393^{***}	-0.3378^{*}
	(0.0241)	(0.0904)	(0.0207)	(0.2044)
JetBlue	2.5697^{***}	0.1735	2.2153^{***}	-2.6927^{***}
	(0.0235)	(0.1066)	(0.0202)	(0.2647)
Continental Airlines	2.4446^{***}	2.4014^{***}	2.6637^{***}	2.5300^{***}
	(0.0160)	(0.0330)	(0.0180)	(0.0841)
ANA and JP Express	-2.3018^{***}	-2.0534^{***}		
	(0.2397)	(0.4934)		
Delta	2.1102^{***}	1.3806^{***}	2.1485^{***}	-0.0934
	(0.0148)	(0.0421)	(0.0179)	(0.1403)
ExpressJet	-0.6435^{***}	-1.2142^{***}	0.8393^{***}	-1.0753^{***}
	(0.0314)	(0.0685)	(0.0312)	(0.1740)
AirTran Airways	3.4359^{***}	1.7589^{***}	2.6711^{***}	-0.8833^{***}
	(0.0233)	(0.0820)	(0.0209)	(0.2039)
Gojet Airlines			0.7518^{***}	0.8642^{***}
			(0.0320)	(0.1490)
Frontier Airlines	2.2148^{***}	0.8361^{***}	2.2451^{***}	-0.9336^{***}
	(0.0260)	(0.0765)	(0.0244)	(0.1962)

Continues on the next page

Table 4: Base Case Parameters Estimates for 2001 and 2006

	Dependent variable:			
	Mean utility			
	OLS 2001	*2SLS* 2001	*OLS* 2006	*2SLS* 2006
	(1)	(2)	(3)	(4)
Hawaiian Airlines	4.4741***	1.8767***	3.5819***	−1.7775***
	(0.0385)	(0.1300)	(0.0319)	(0.3083)
American West Airlines	2.6576***	1.4093***	2.3806***	0.0495
	(0.0172)	(0.0609)	(0.0186)	(0.1458)
Jade Cargo International	0.5988***	−0.0823		
	(0.0750)	(0.1566)		
Envoy Air	0.5296***	0.2222***	0.7112***	−0.0832
	(0.0158)	(0.0347)	(0.0177)	(0.0914)
Vanguard Airlines	1.6389***	−0.7230***		
	(0.0274)	(0.1094)		
Spirit Airlines	2.0777***	−0.1409	1.7919***	−1.9655***
	(0.0276)	(0.1048)	(0.0215)	(0.2142)
Northwest Airlines	2.4227***	1.8176***	2.0252***	1.0201***
	(0.0170)	(0.0425)	(0.0186)	(0.1004)
National Airlines	1.7406***	0.2807***		
	(0.0323)	(0.0881)		
American Eagle	−0.6512***	−0.6859***	−0.2555***	−0.8588***
	(0.0479)	(0.0985)	(0.0200)	(0.0977)
SkyWest			0.7660***	0.1310
			(0.0196)	(0.0964)
Horizon Air	1.5327***	0.0296	1.4746***	−1.6753***
	(0.0713)	(0.1585)	(0.0323)	(0.2185)
AirBridgeCargo Airline	0.7850***	0.5780***	1.2568***	1.3177***
	(0.0228)	(0.0476)	(0.0210)	(0.0978)
Shuttle America			0.5521***	−0.3048***
			(0.0224)	(0.1128)
Sun Country Airlines	2.5036***	0.0776	2.4677***	−2.5280***
	(0.0342)	(0.1193)	(0.0352)	(0.3004)
Trans World Airlines	2.0294***	1.2906***		
	(0.0196)	(0.0499)		

Continues on the next page

	Dependent variable: mean utility			
	OLS	*2SLS*	*OLS*	*2SLS*
	2001	2001	2006	2006
	(1)	(2)	(3)	(4)
ATA Airlines	2.6751***	0.4839***	2.2757***	−2.3379***
	(0.0205)	(0.0966)	(0.0252)	(0.2605)
USA 3000 Airlines			2.4217***	−0.4161*
			(0.0340)	(0.2132)
United Airline	2.5059***	2.2578***	2.4466***	2.8854***
	(0.0144)	(0.0311)	(0.0175)	(0.0845)
US Airways	2.1889***	2.0169***	2.4349***	1.3467***
	(0.0157)	(0.0330)	(0.0182)	(0.1009)
Southwest Airlines	3.1752***	1.0605***	2.8555***	−1.8207***
	(0.0156)	(0.0898)	(0.0178)	(0.2500)
Mesaba Airlines	−1.9588***	−1.7778*	−0.3147***	−1.1312***
	(0.4480)	(0.9219)	(0.0609)	(0.2861)
Mesa Airlines			0.6404***	−0.4236***
			(0.0207)	(0.1100)
Midwest Airlines	1.1209***	0.3200***	1.1878***	0.9155***
	(0.0414)	(0.0910)	(0.0272)	(0.1269)
Air Wisconsin Airlines	0.6043***	0.3771***	1.0489***	0.9726***
	(0.0279)	(0.0580)	(0.0227)	(0.1058)
Republic Airlines			1.0924***	−0.3320**
			(0.0282)	(0.1496)
ExpressJet Airlines			1.2790***	1.2073***
			(0.0216)	(0.1004)
Quarter 1	0.2676***	0.4599***	−0.0330***	−0.1618***
	(0.0061)	(0.0147)	(0.0047)	(0.0230)
Quarter 2	0.3078***	0.4644***	0.0596***	0.4148***
	(0.0061)	(0.0140)	(0.0046)	(0.0279)
Quarter 3	0.1816***	0.1224***	0.0763***	0.1868***
	(0.0062)	(0.0130)	(0.0046)	(0.0220)
Constant	−8.5021***	−7.3174***	−8.2390***	−3.3503***
	(0.0142)	(0.0553)	(0.0173)	(0.2593)
Observations	276,970	276,970	437,400	437,400
R^2	0.3209	−1.8762	0.3268	−13.5360
Adjusted R^2	0.3208	−1.8765	0.3268	−13.5373
Residual Std. Error	1.10 (df = 276933)	2.26 (df = 276933)	1.07 (df = 437360)	4.96 (df = 437360)
F Statistic	3,634.4750*** (df = 36; 276933)		5,444.8130*** (df = 39; 437360)	
Note:				*p<0.1; **p<0.05; ***p<0.01

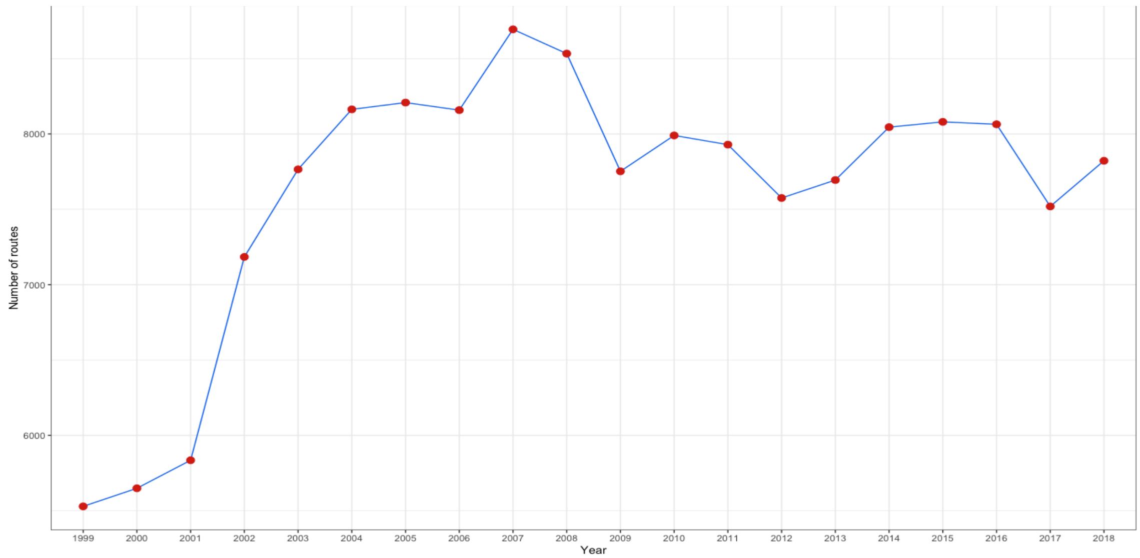

Figure 1: Total number of non-stop routes in the U.S., 1999 - 2018

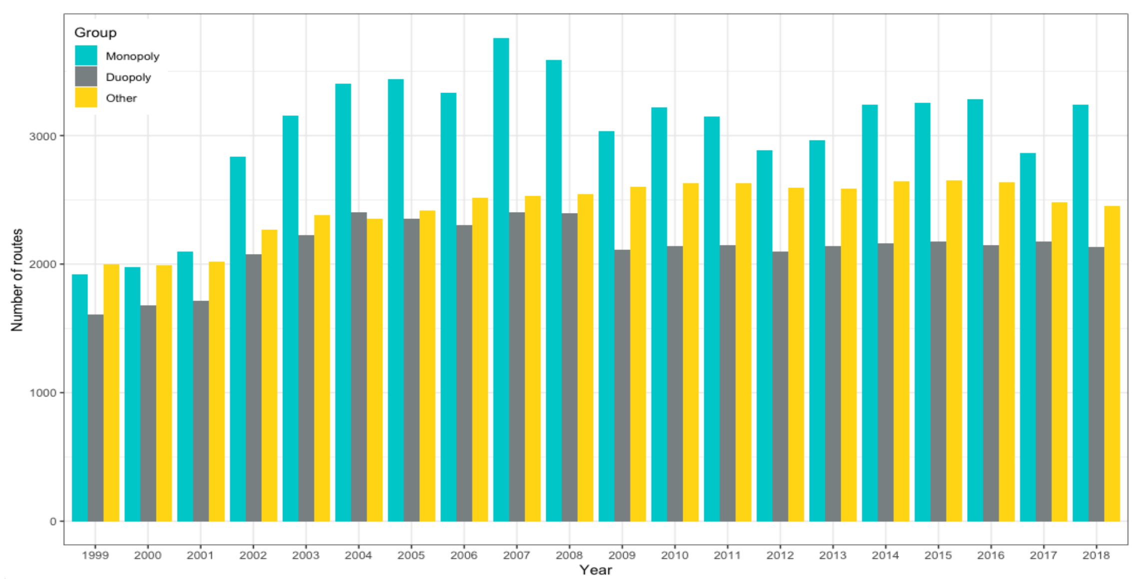

Figure 2: Number of non-stop routes in the U.S. by monopoly, duopoly and other

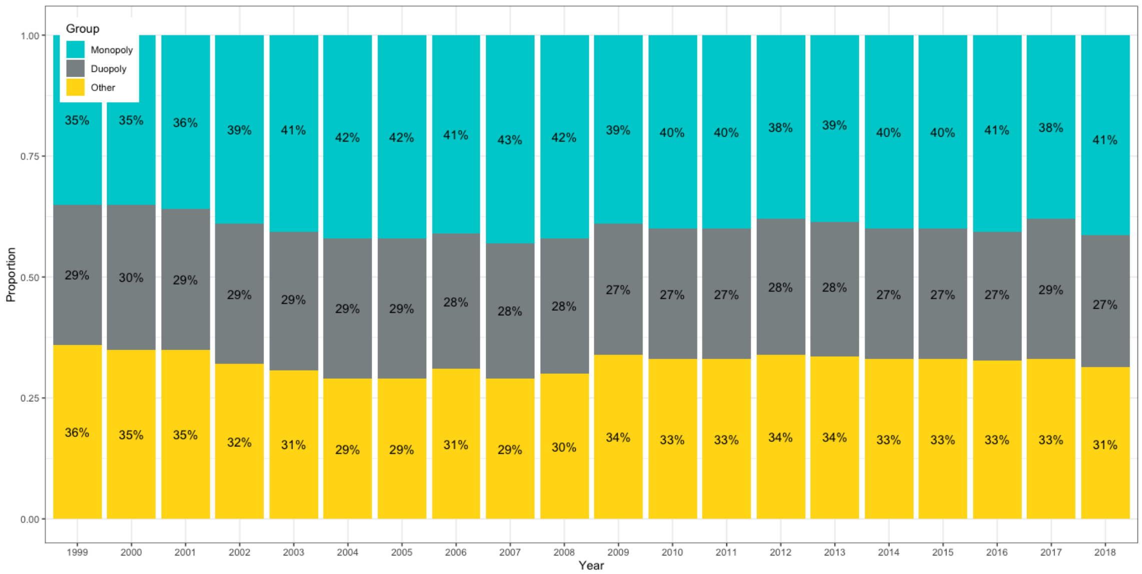

Figure 3: Composition of non-stop routes by monopoly, duopoly and other

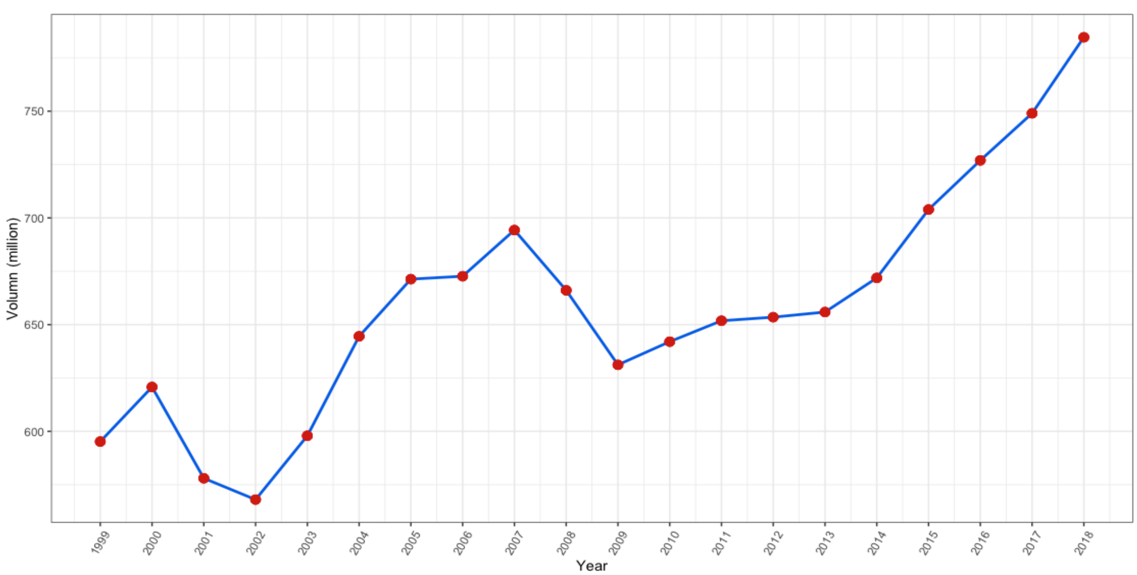

Figure 4: Total volume for non-stop flights: 1999 to 2018

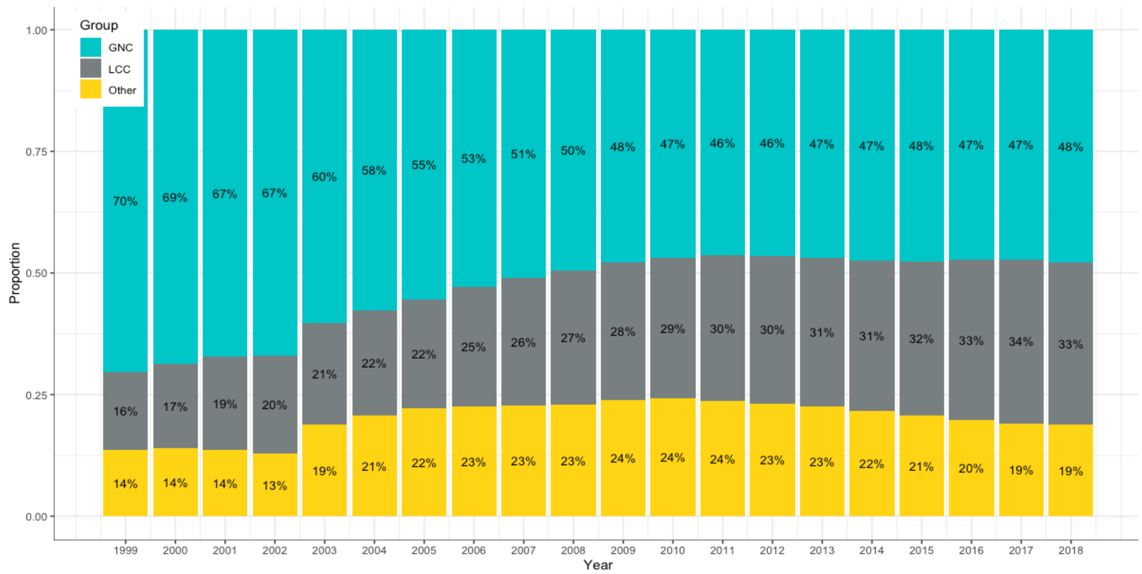

Figure 5: Total volume share of GNC, LCC and Other: 1999 - 2018

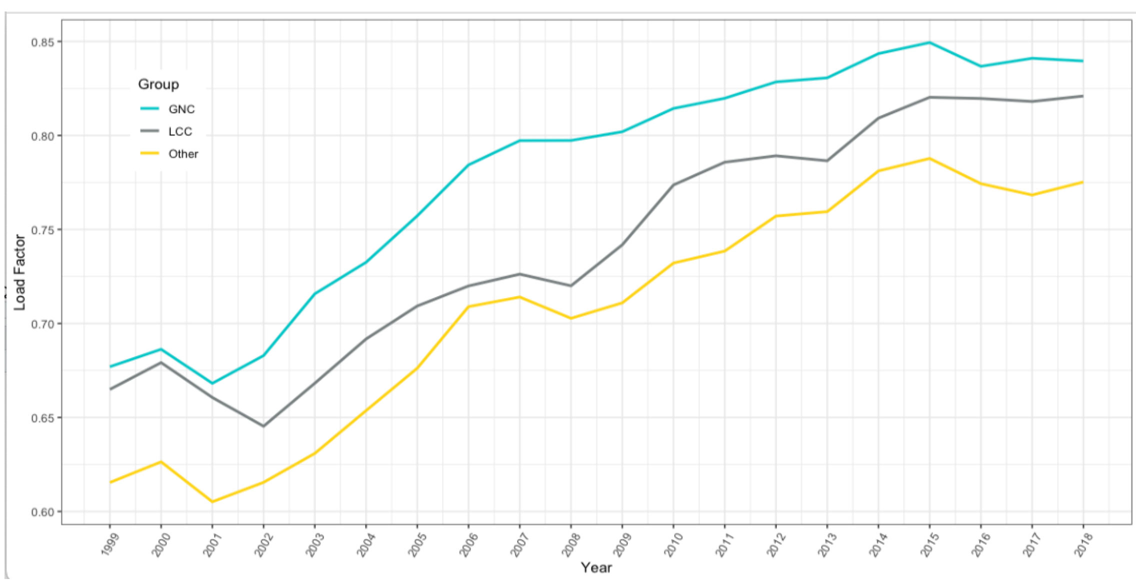

Figure 6: Load factor of GNC, LCC and Other: 1999 - 2018

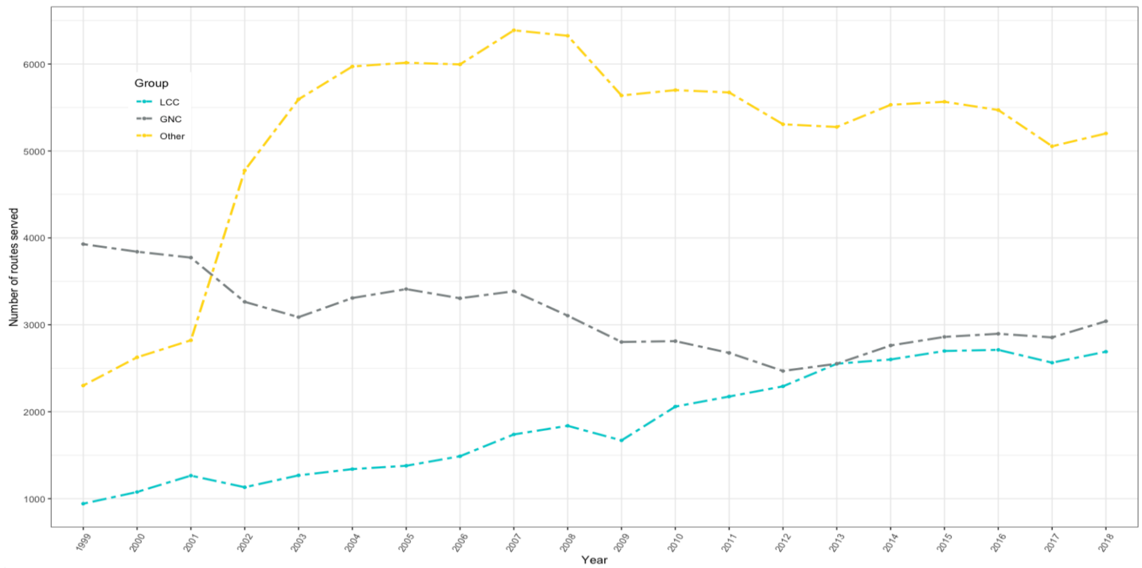

Figure 7: Number of markets served by LCC, GNC and Other: 1999 - 2018

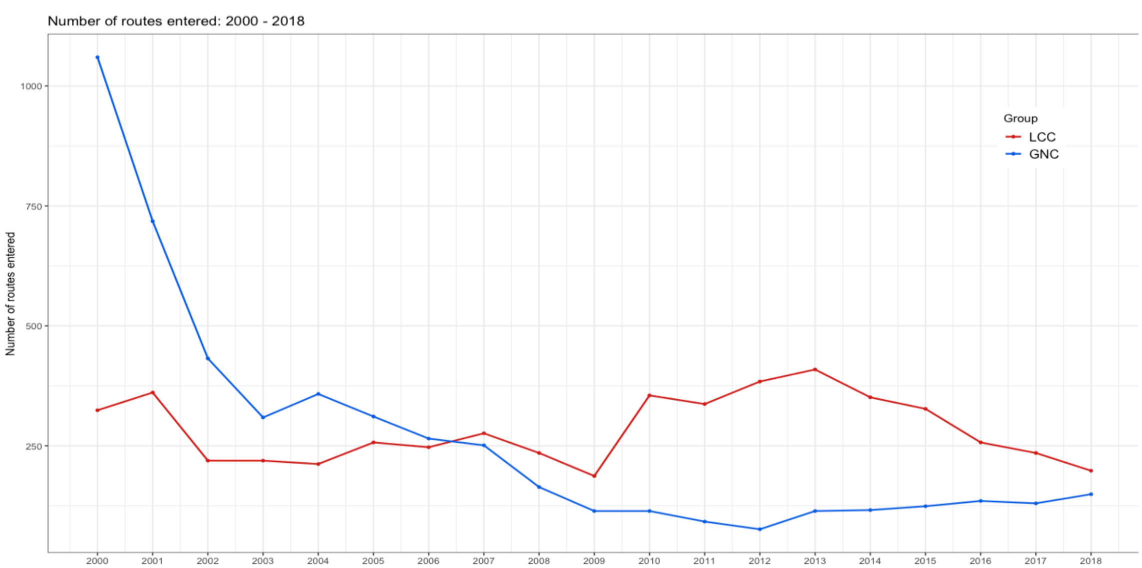

Figure 8: Number of markets entered by LCC and GNC: 2000 - 2018

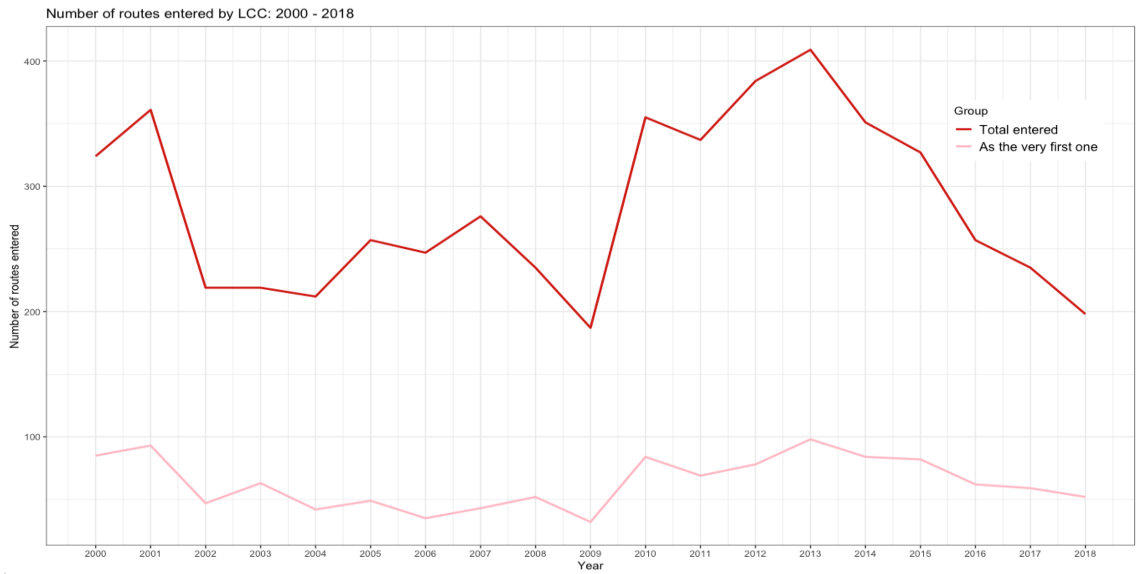

Figure 9: Number of markets entered by LCC: 2000 - 2018

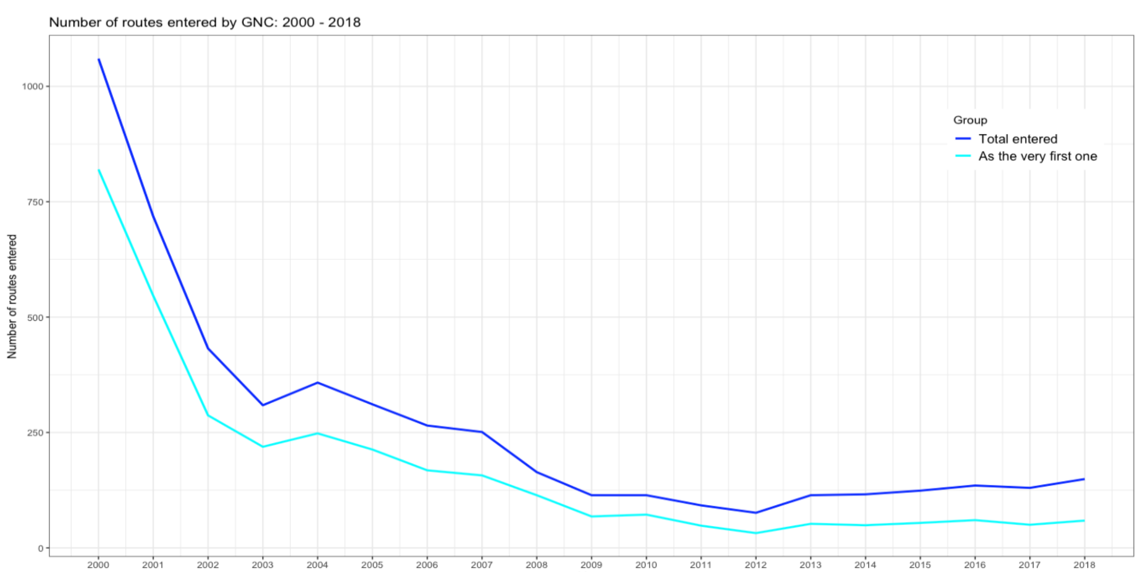

Figure 10: Number of markets entered by GNC: 2000 - 2018

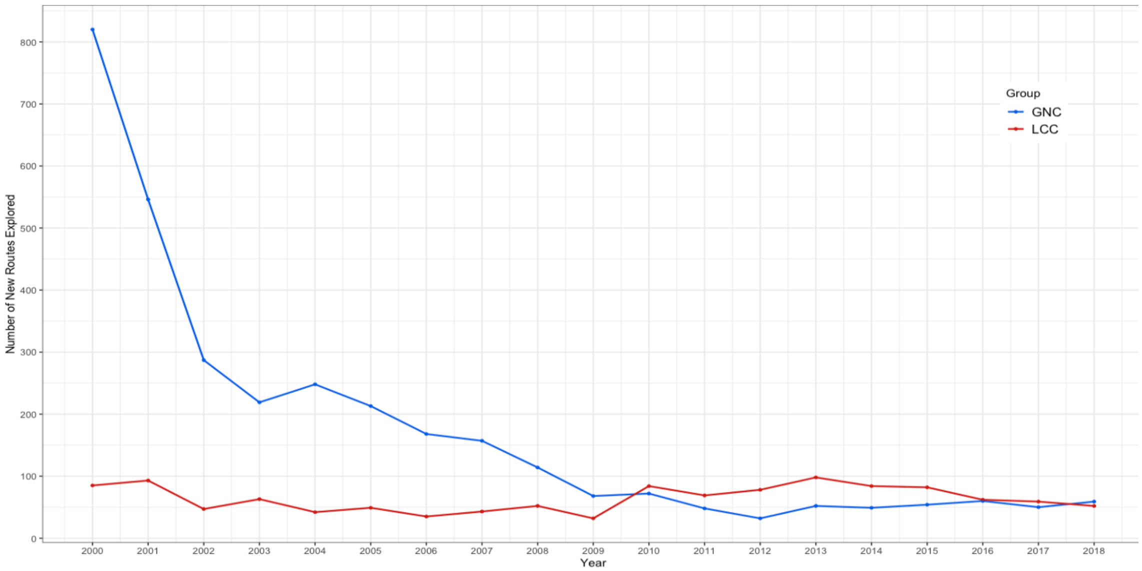

Figure 11: Number of completely new markets entered by GNC and LCC: 2000 - 2018

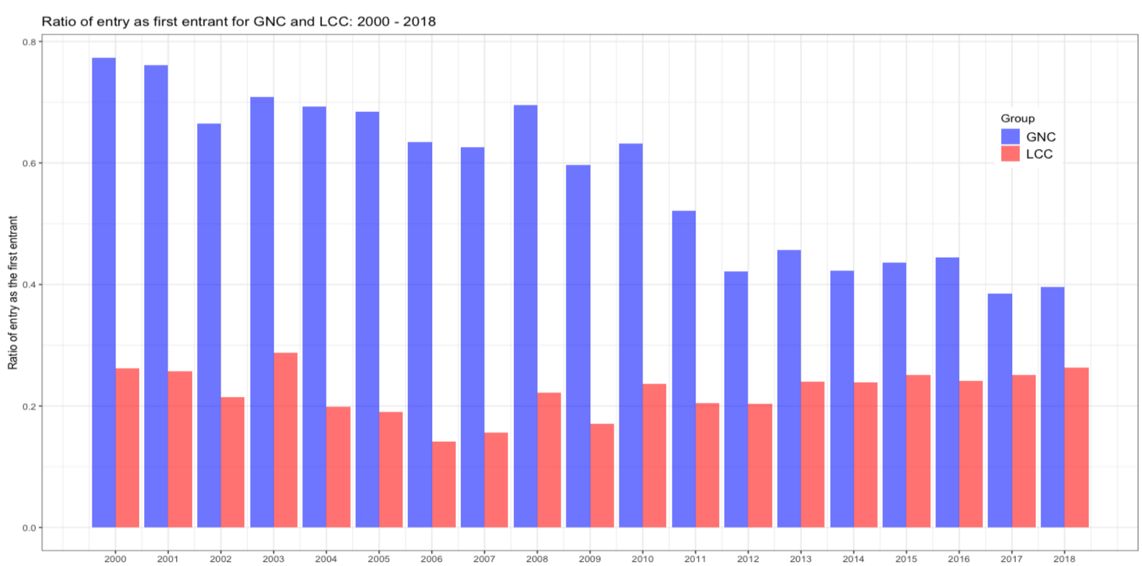

Figure 12: Proportion of number of new markets entered by LCC and GNC: 2000 - 2018

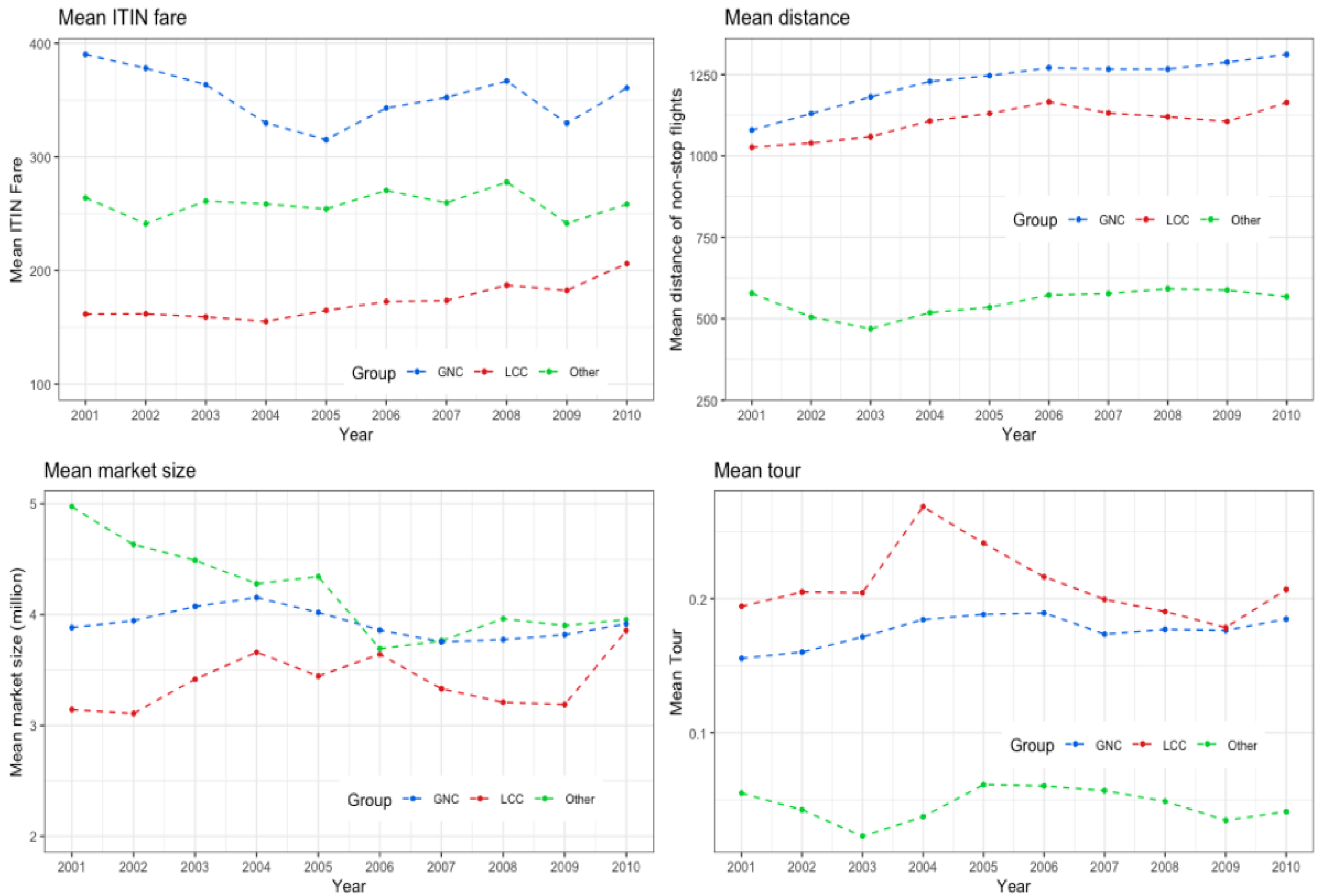

Figure 13: Summary statistic for GNC, LCC and Other

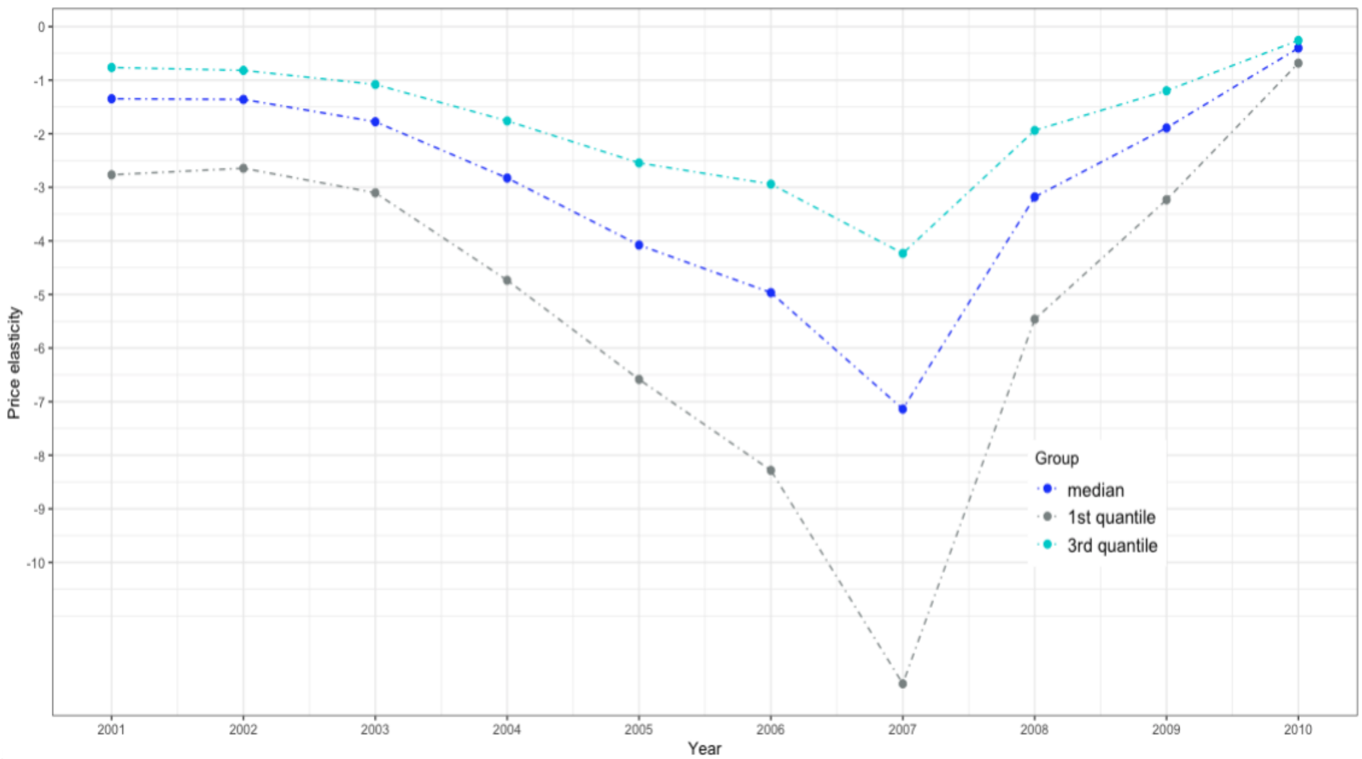

Figure 14: Price elasticity:2001 - 2010

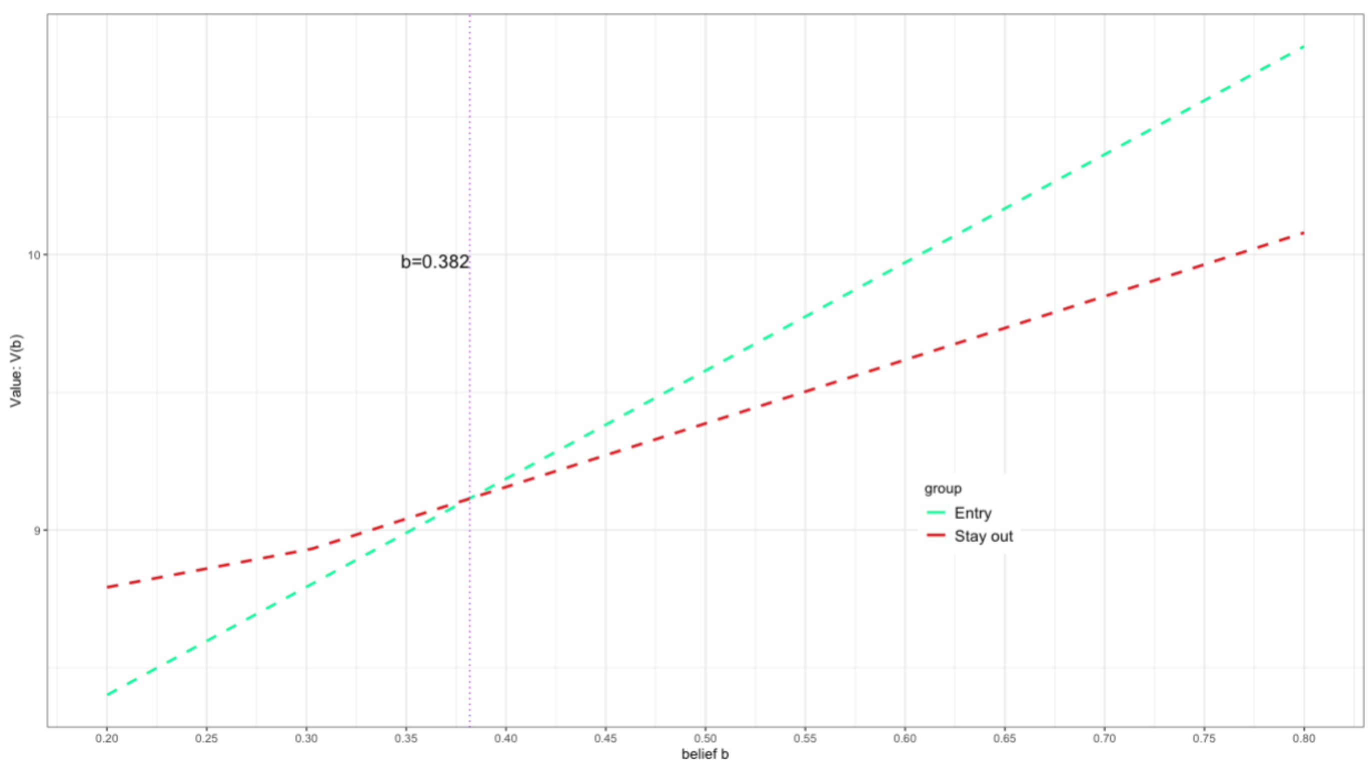

Figure 15: Monopoly model: value

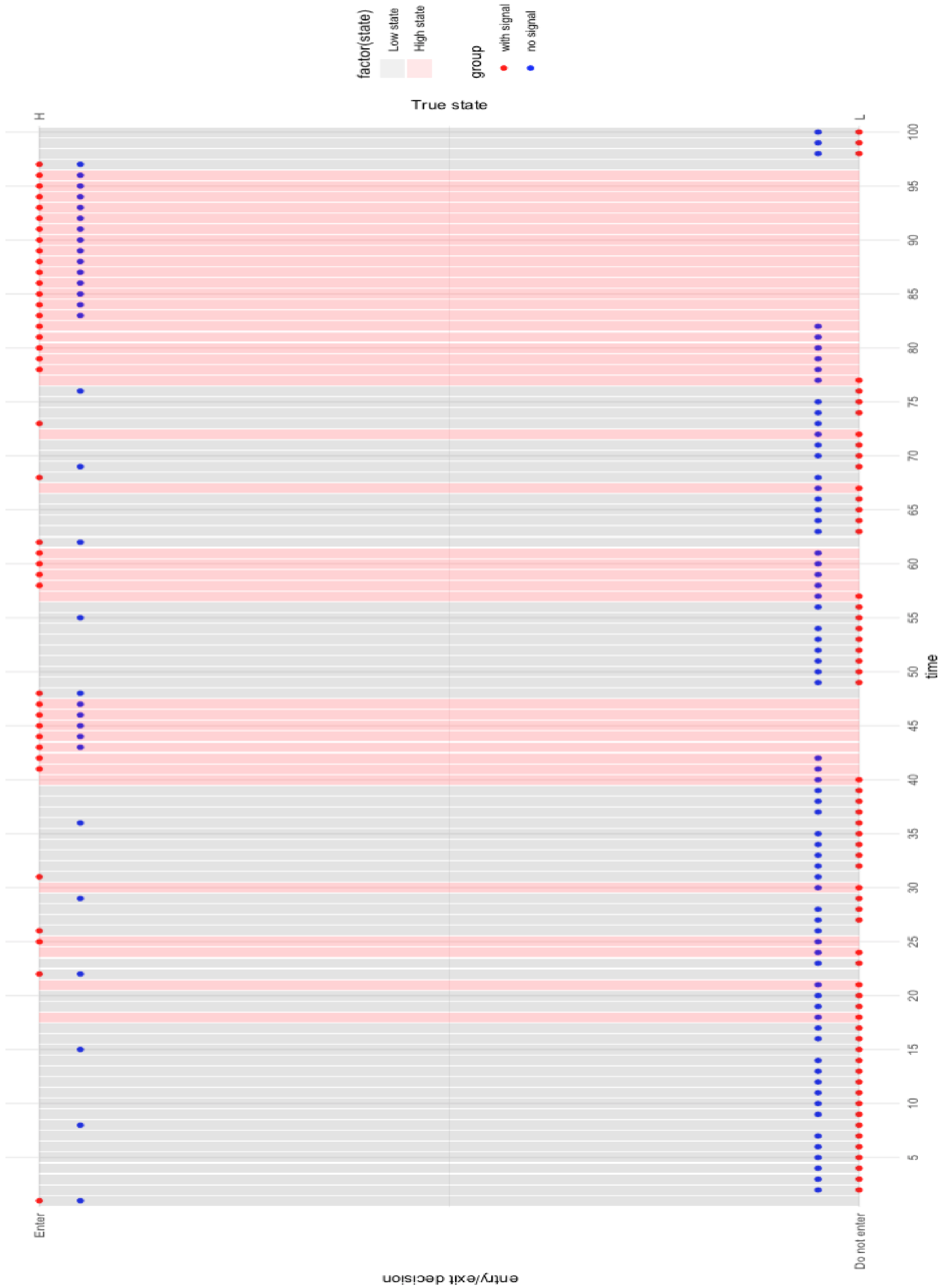

Figure 16: Monopoly model: simulation for 100 periods

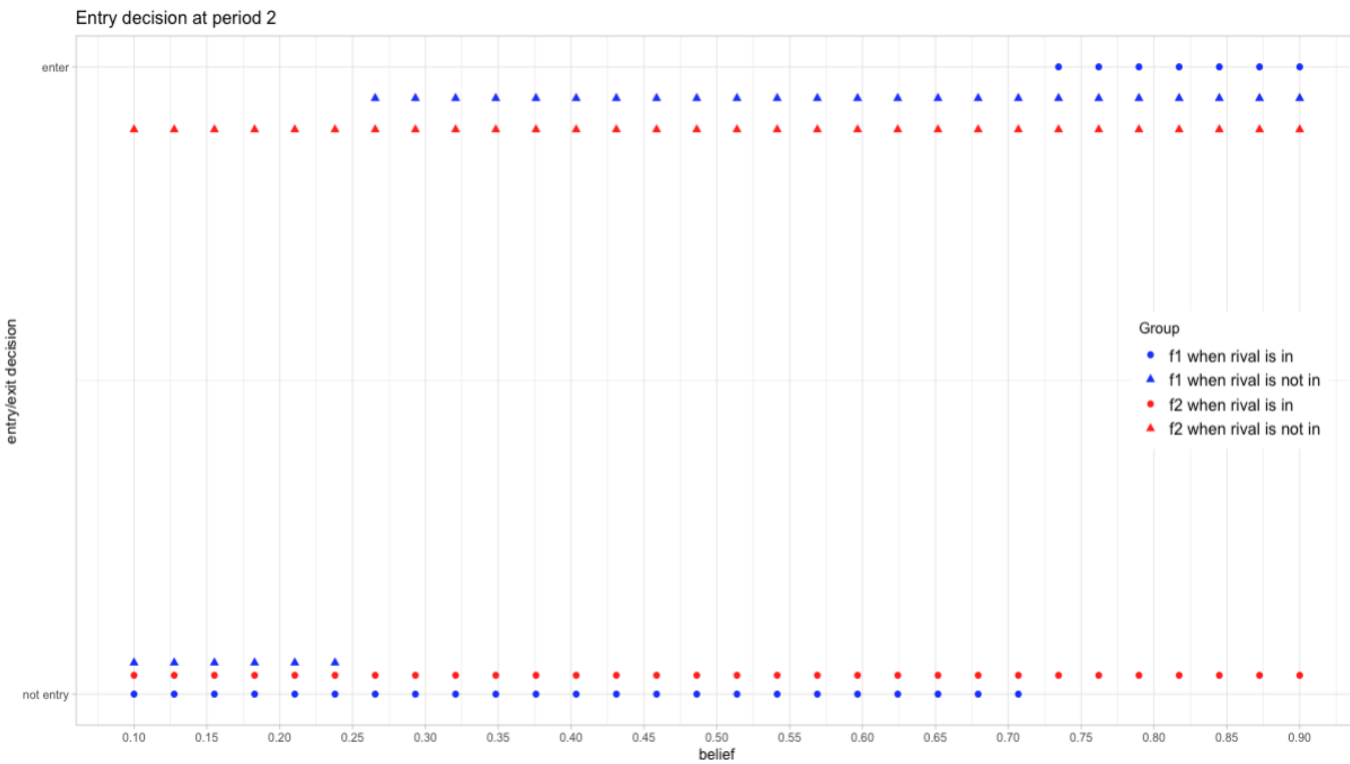

Figure 17: Duopoly model: entry decision at period 2

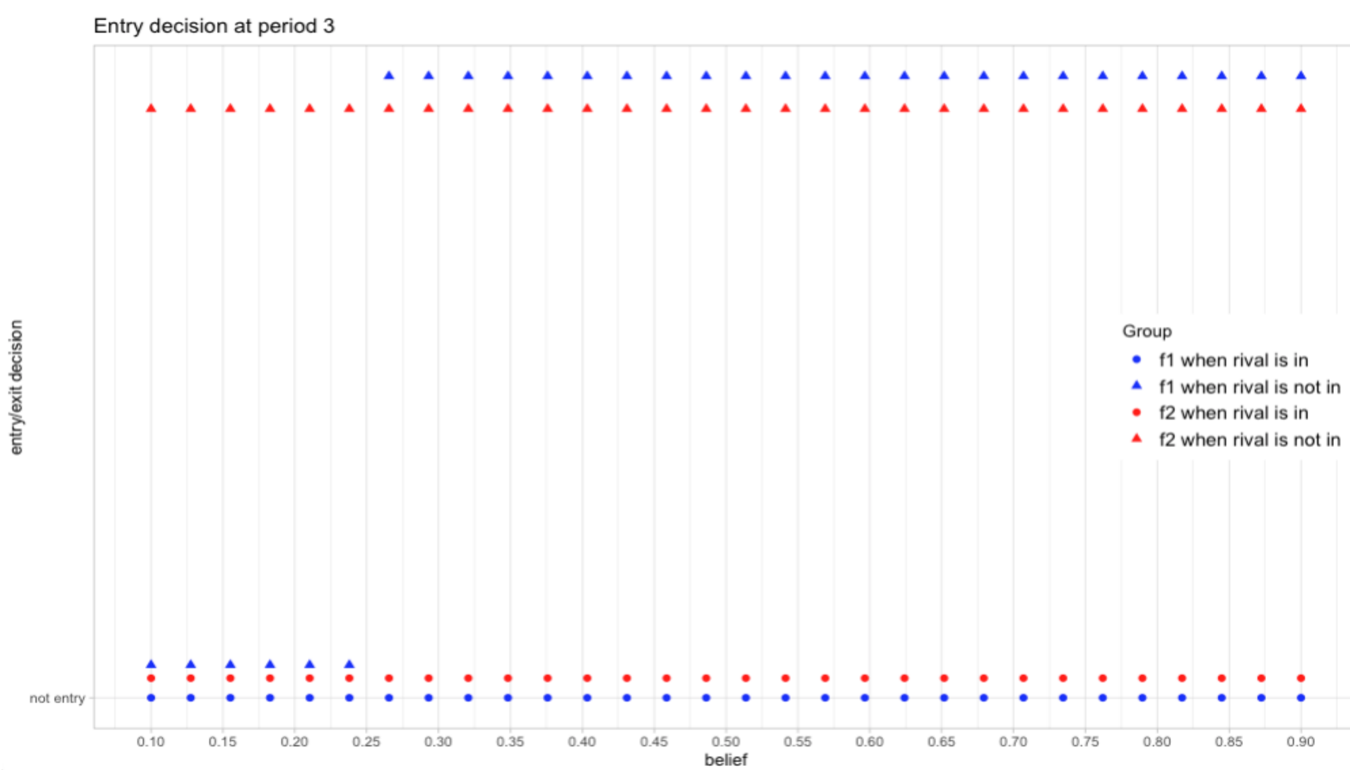

Figure 18: Duopoly model: entry decision at period 3

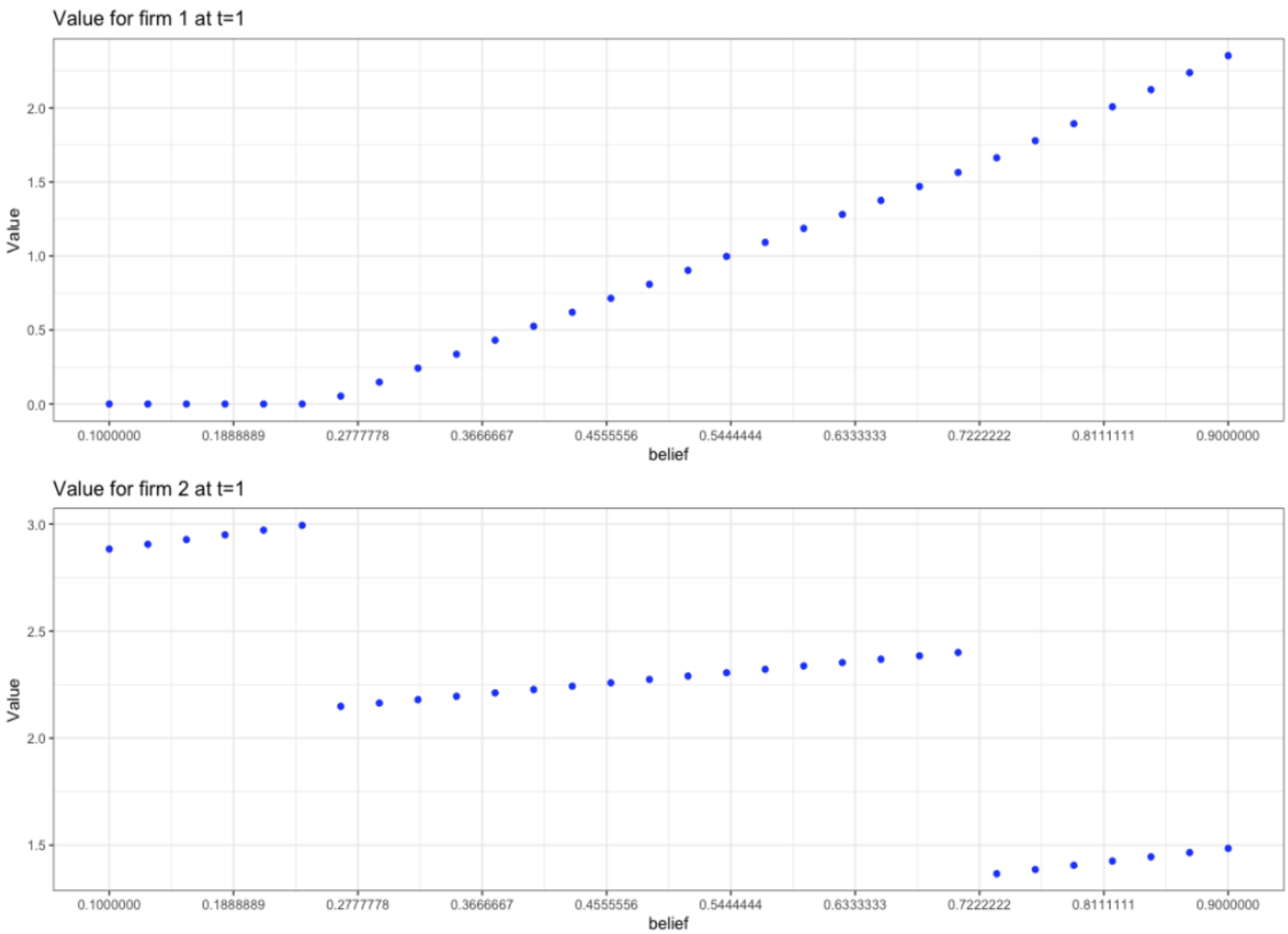

Figure 19: Duopoly model: value of firms at the beginning of period 1

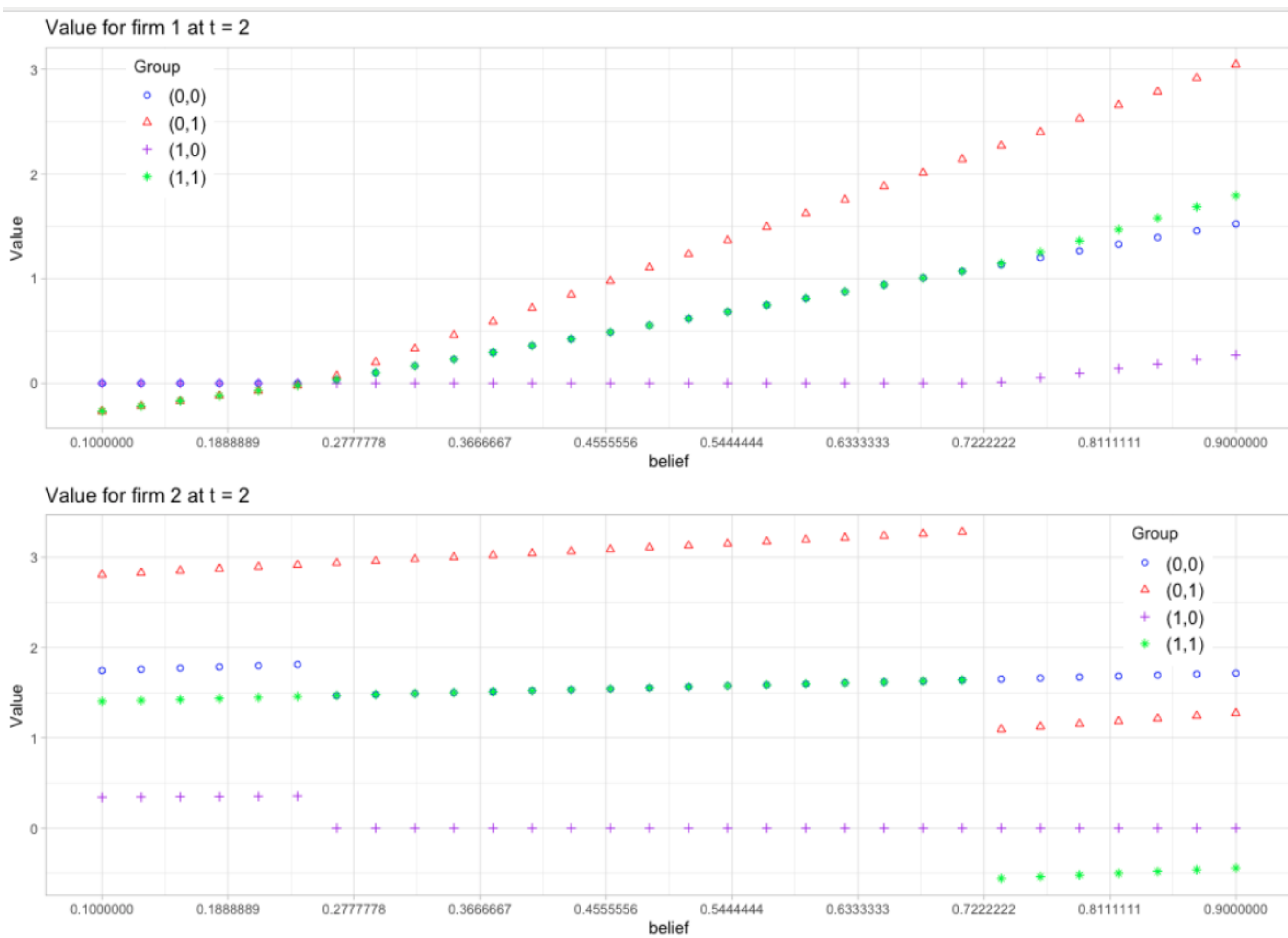

Figure 20: Duopoly model: value of firms at the beginning of period 2

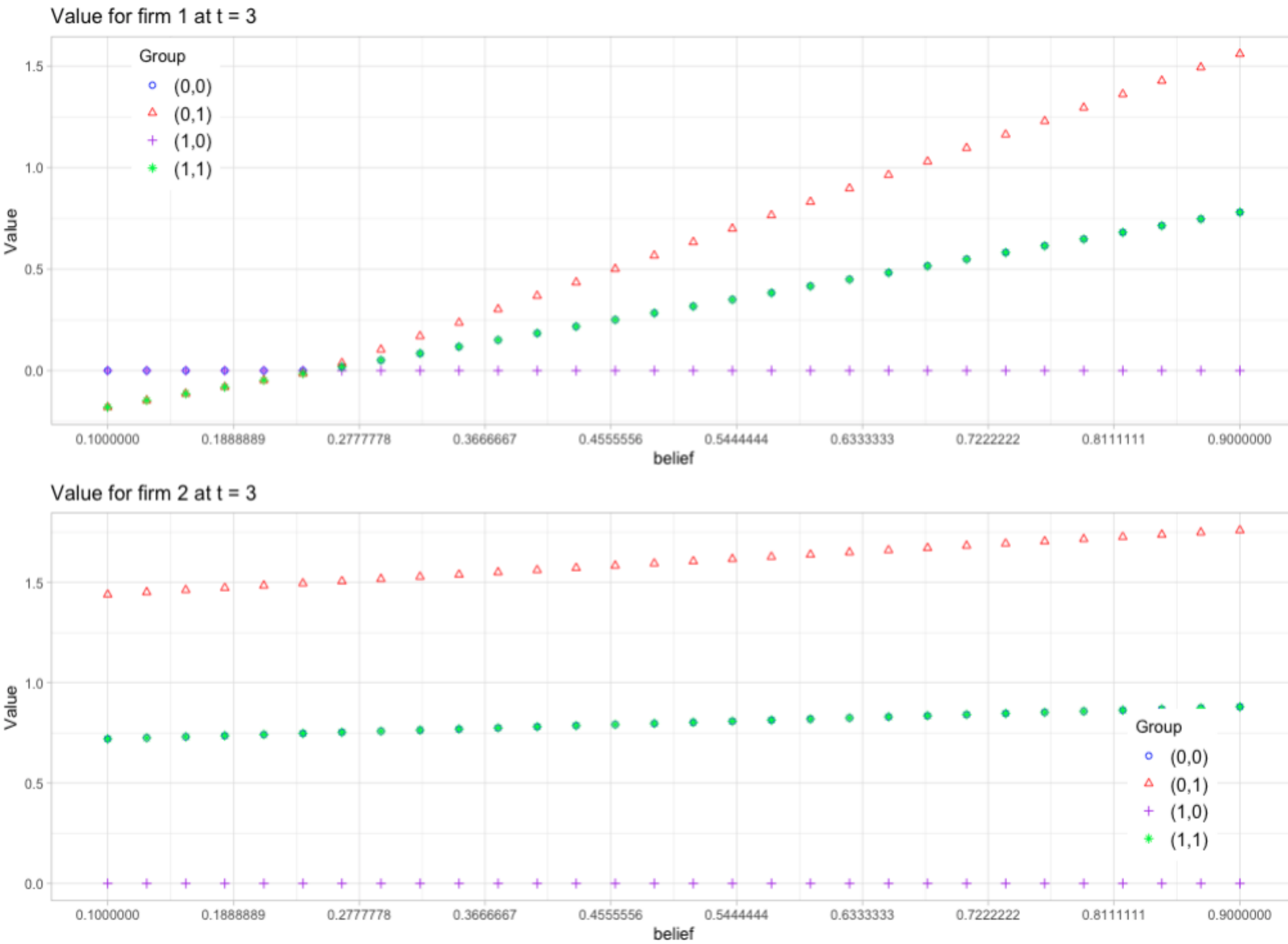

Figure 21: Duopoly model: value of firms at the beginning of period 3