Reinforcement Learning in Carpooling Markets with Dutch Auctions

Deniz Kayar

Rhine-Waal University Kleve

About me

- Deniz Kayar
- Computer Scientist, TU Dortmund
- Rhine-Waal University, Kleve, NRW
- Research Assistant, Faculty Society & Economics, Prof. Thomas Pitz
- Laboratory for experimental economics

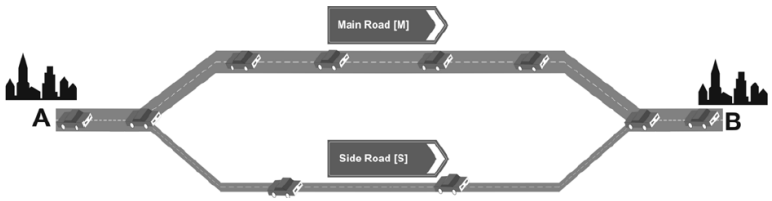
Overview

- Funding initiative „FH Struktur“
- Research project „Incentive and Information Facilities within Multimodal Mobility Concepts“
- Prof. Dr. Thomas Pitz, Prof. Dr. Hasan Alkas, Prof. Dr. Jörn Sickmann, Wolf Gardian
- Multi Agent Simulation
- Markets
- Auctions
- Reinforcement Learning

Motivation

- „Commuters route choice-behaviour“ (2006)
- Rainhard Selten, Thorsten Chmura, Thomas Pitz, Sebastian Kube, Michael Schreckenberg
- Minority Game, Main road & Side road
- Multi agent Simulation with Reinforcement Learning:
  - Nash Equlibrium is not the best strategy for dynamic systems like these
  - Reinforcement Learning gave better results
  - Good Prediction for route choice and change-over

Selten, R., Chmura, T., Pitz, T., Kube, S. & Schreckenberg, M. (2006). Commuters route choice-behaviour. *Games and Economic Behaviour, vol. 58, issue 2, 394-406.* doi:10.1016/j.geb.2006.03.012
URL http://dx.doi.org/10.1016/j.geb.2006.03.012 2, 4



Dutch Auctions in Carpooling

- Commodity „car seat“ temporarily available similar to dutch flower market
- Dutch Auction competes against Fixed Price
- Determine with a Reinforcement Learning approach which mechanism gives higher payoffs.

Simulation Steps

- Agents: **Carowner (CO)** and **Not Carowner (NCO).**
- NCOs want to travel from one starting point to an endpoint. They can either travel by **train** or take a **lift (car seat)** offered by COs.
- Multiple Rounds: In each round each agent gets a new uniformly distributed **reserveration price** (for example between 1 and 100 €).
- Reservation price → CO: minimum selling price, NCOs: maximum buying price
- Selling mechanisms: **Dutch Auction** or **Fixed Price.**
- **Train price** : For example 100 €, upper bound.
- Each agent picks **randomly** one of the two selling mechanisms in each round (Dutch or Fixed).
- Auctions (COs) **progress stepwise** and are matched with buyers (NCOs)
- **Reinforcement Learning**: After each round the probability for picking one of the mechanisms changes according to the payoff of the previous rounds.

Fixed Price

- Two lists: **COs** and **NCOs**
- Car seats are sold **immediately** withou waiting for an auction mechanism to take place.
- The selling price of a car seat is identical with the reservation price **r_x** of the corresponding **car owner x**
- The first Not Car Owner **y** in the NCO list with a reservation price **r_y >= r_x,** gets the car seat and the the trade is complete
- It there is no r_y with r_y >= r_x, the offer will be closed and no agent was found to travel in x‘s car.

Dutch Auction

- Again two lists: **COs** and **NCOs**
- Start price of a dutch auction is slightly higher than the maximum possible reservation price of 100 € (e.g. 110 €)
- In a dutch auction a price is decreased step wise and the size of these steps (out of **[minStepSize,maxStepSize**],here [1,5]) is **randomly determined** at the beginning of a simulation. This number remains constant during the simulation.
- If the current price of an auction falls below the reservation price of the seller the auction is finished **unsuccessfully**.
- If the current price reaches a point, where it is equal or lower than the reservation price of a NCO, then this NCO gets the seat and the auction is finished **successfully**
- In each time step the list of auctions and NCOs is **permuted randomly**.

Matching

- Carowners, that sell their car seats via Dutch Auction can only be matched with buyers, that also chose Dutch Auction as a buying mechanism.
- 4 Lists, 2 for Dutch Auctions and 2 for Fixed Prices

Reinforcement Learning

Initialization:
For each agent i let $[q^1_{i,1}, ..., q^1_{i,n}]$ be the initial propensities, where n is the number of possible strategies.

Period 1:
Each agent i chooses strategy x with probability $\frac{q^1_{i,x}}{\sum_{y=1}^{n} q^1_{i,y}}$.

Periods t+1:
For each agent i, let a^t_i be the payoff of player i in period t and x the chosen strategy in period t. Then the propensities for period $t+1$ are calculated as follows:

CASE I ($a^t_i \geq 0$):

$$q^{t+1}_{i,y} := \begin{cases} q^t_{i,y} + a^t_i, & \text{if } y = x \\ q^t_{i,y}, & \text{if } y \neq x \end{cases}$$

CASE II ($a^t_i < 0$):

$$q^{t+1}_{i,y} := \begin{cases} q^t_{i,y}, & \text{if } y = x \\ q^t_{i,y} + |a^t_i|, & \text{if } y \neq x \end{cases}$$

Each player i chooses strategy x with probability $\frac{q^{t+1}_{i,x}}{\sum_{y=1}^{n} q^{t+1}_{i,y}}$.

Simulation runs

- One simulation
- Multiple simulations
- Multiple simulations with varying CO/NCO ratios

- Single Run:
  - 25 COs, 100 NCOs, 1000 Periods
- Multirun:
  - Identical with single run, repeated a 100 times
- Multirun with varying CO/NCO runs:
  - 100 simulation runs with 500 periods each
  - Variation of COs: 10, 20, 50, 100, 200
  - NCO count stays constant: 100.

One Simulation Run

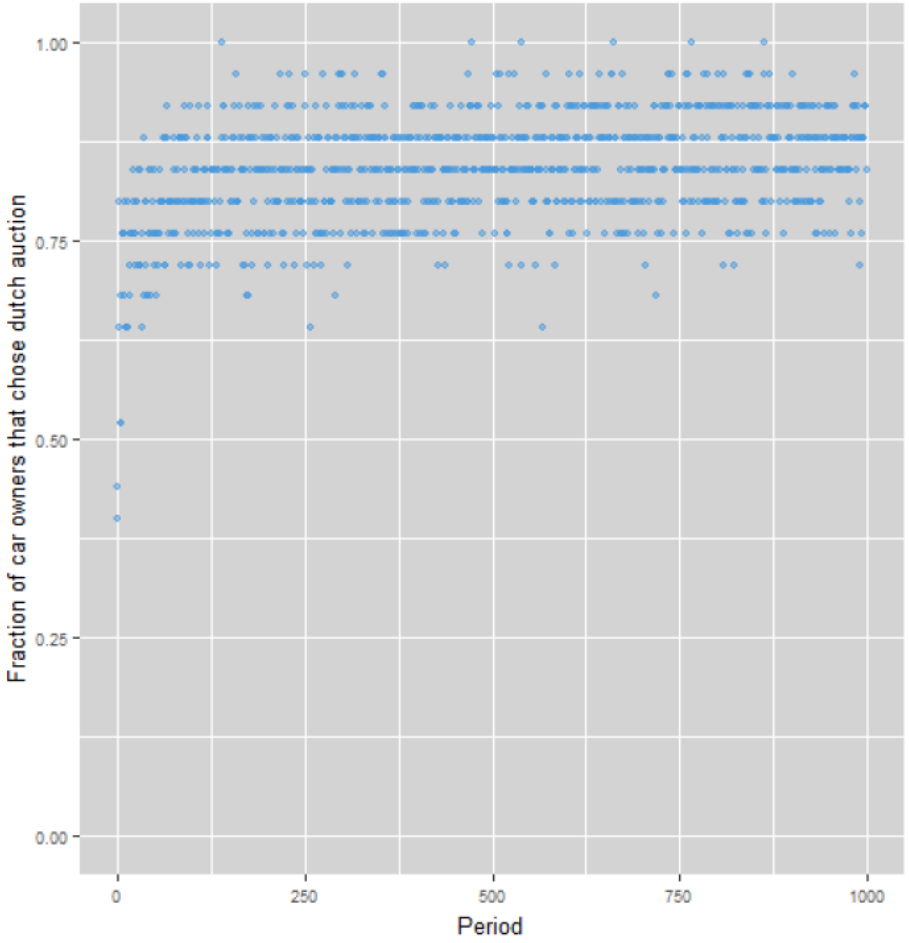


Figure 1: Fraction COs who chose Dutch Auction (1 simulation run)

One Simulation Run

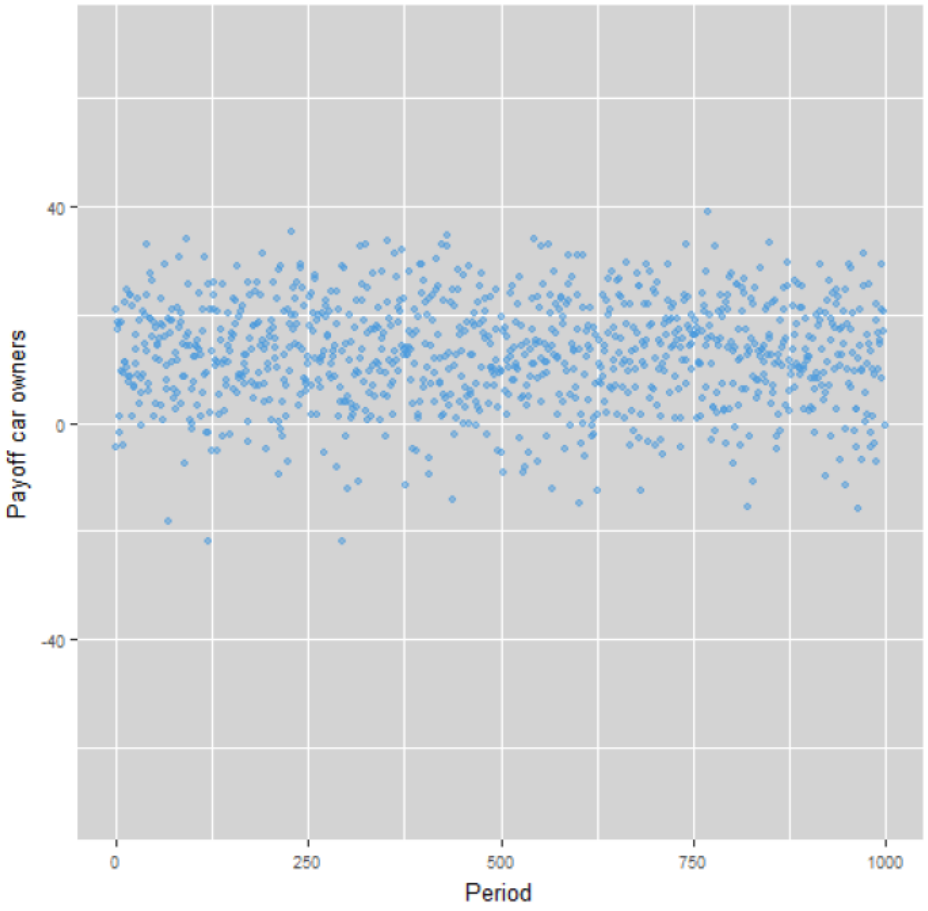


Figure 2: Payoff for COs (1 simulation run)

One Simulation Run

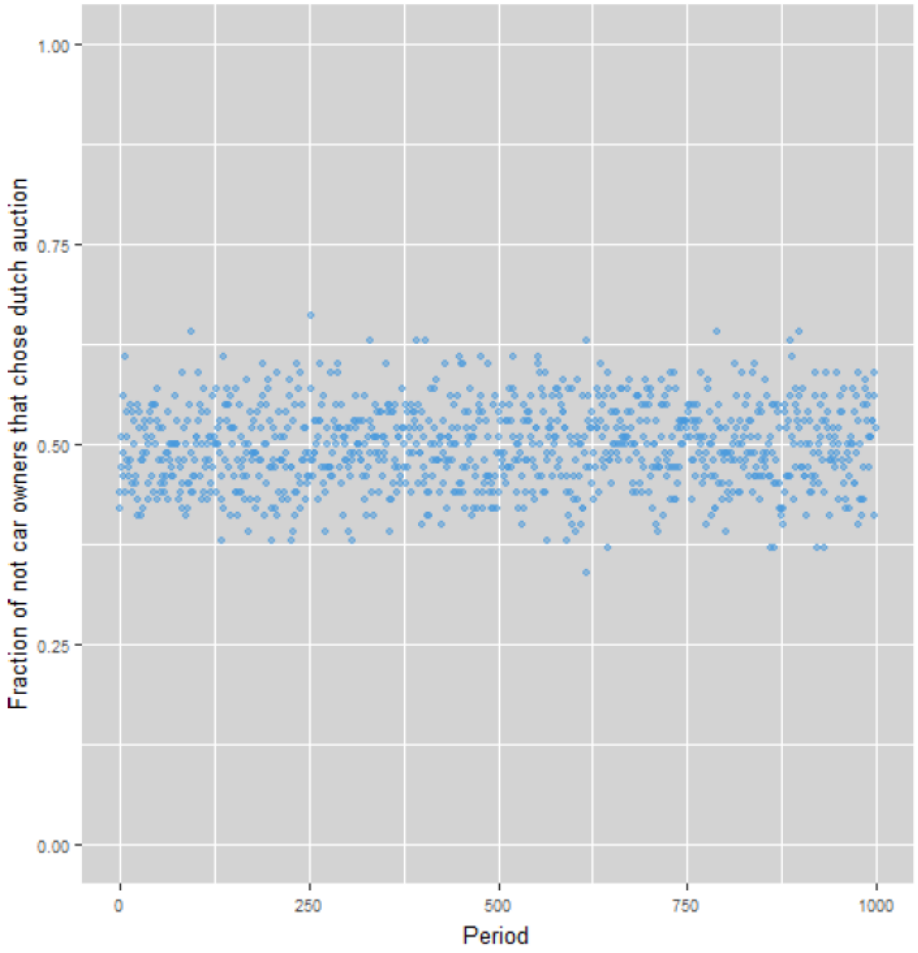


Figure 3: Fraction of NCOs who chose Dutch Auctions (1 simulation run)

One Simulation Run

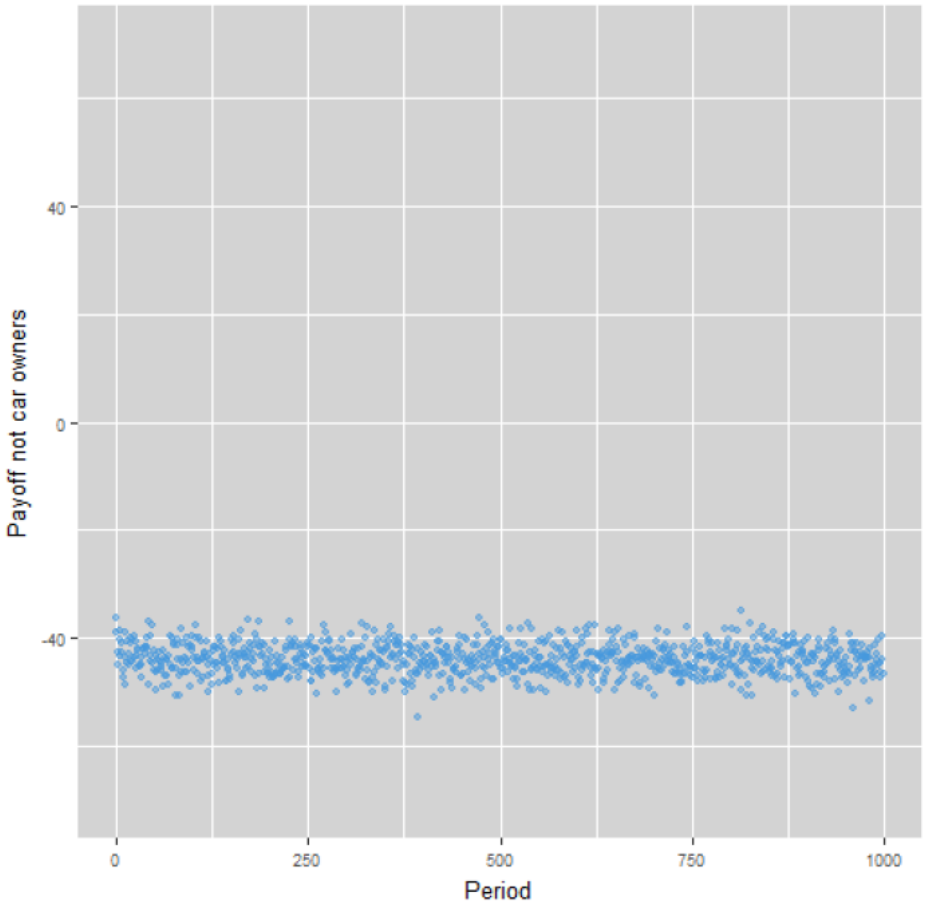


Figure 4: Payoff for NCOs (1 simulation run)

Multiple Simulation runs

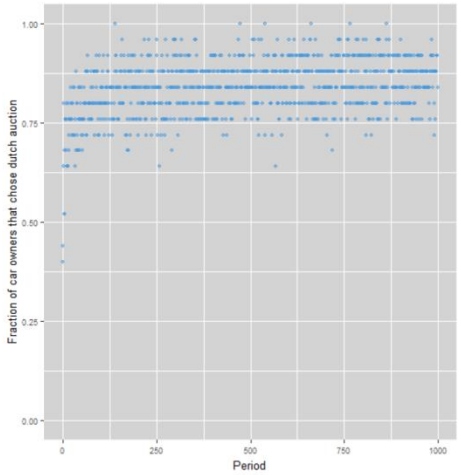


Figure 1: Fraction COs who chose Dutch Auction (1 simulation run)

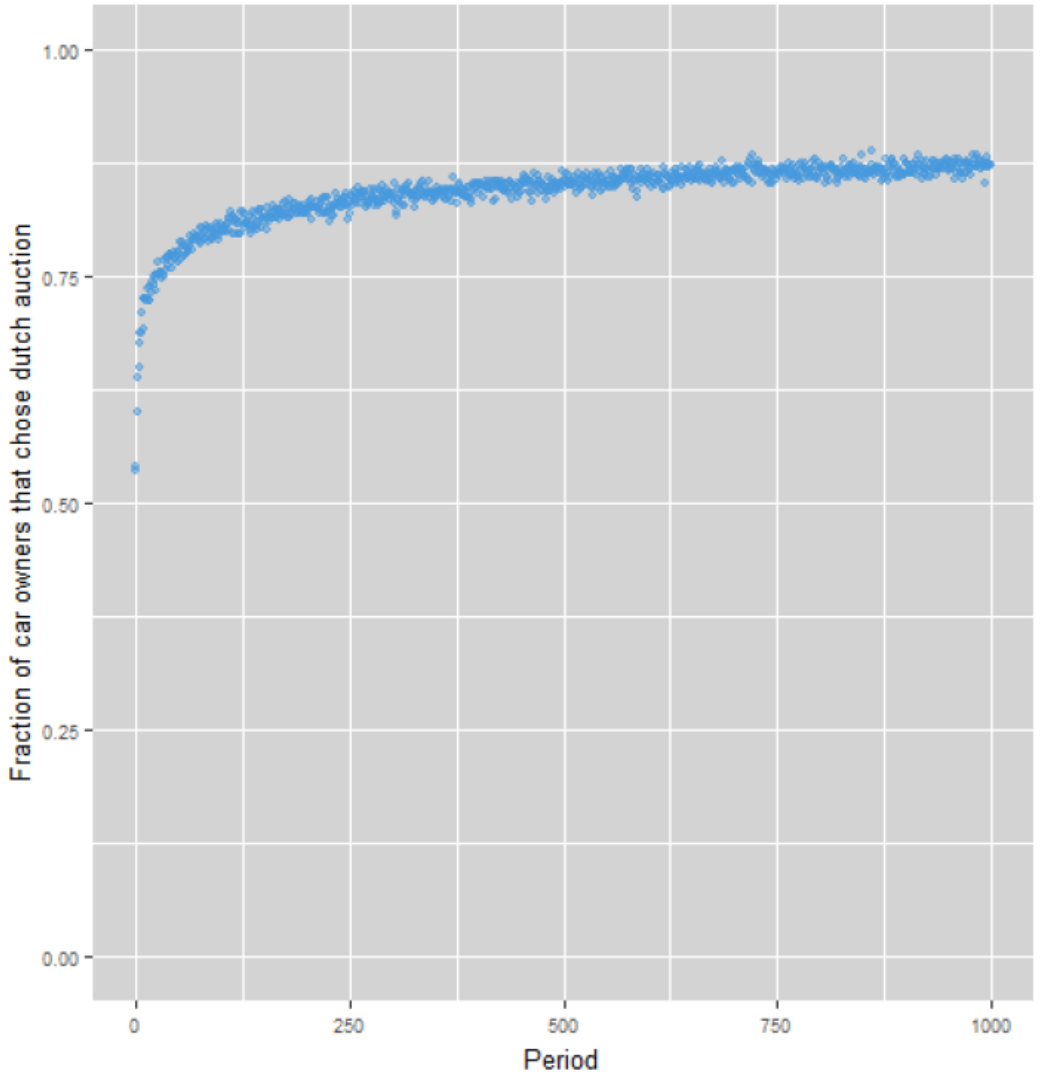


Figure 5: Mean fraction of COs who chose Dutch Auction (100 simulation runs)

Multiple Simulation runs

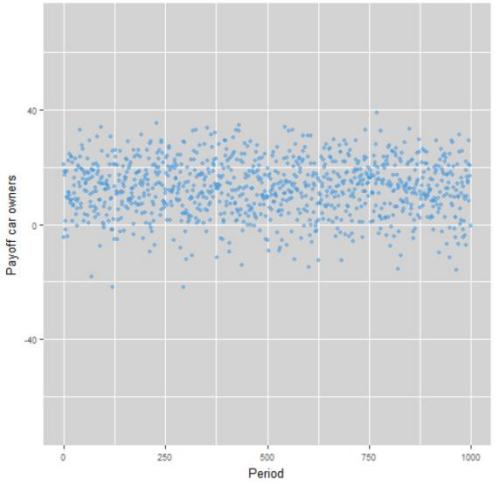


Figure 2: Payoff for COs (1 simulation run)

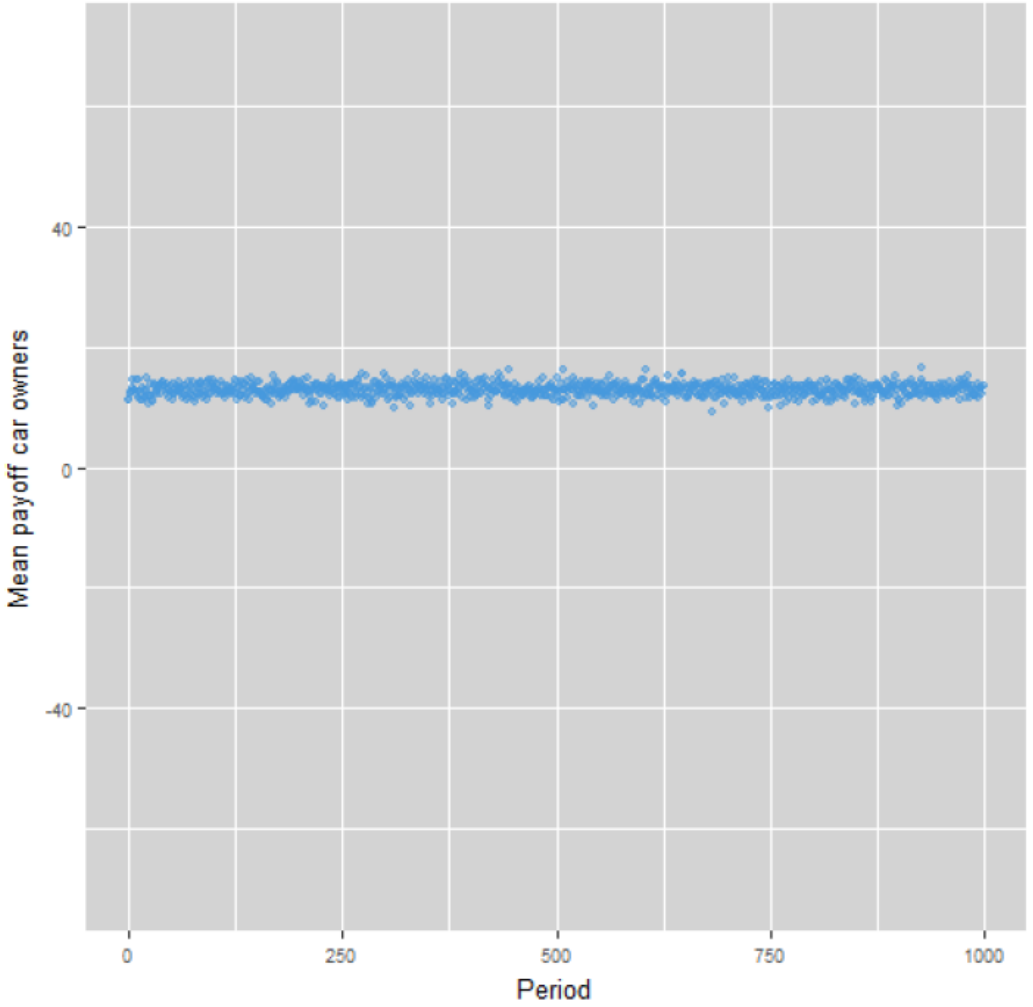


Figure 6: Mean payoff for COs (100 simulation runs)

Multiple Simulation runs

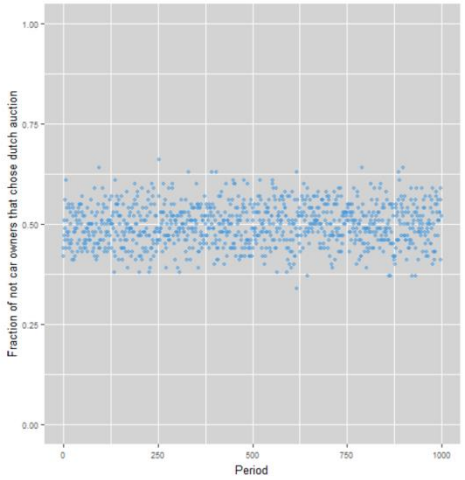


Figure 3: Fraction of NCOs who chose Dutch Auctions (1 simulation run)

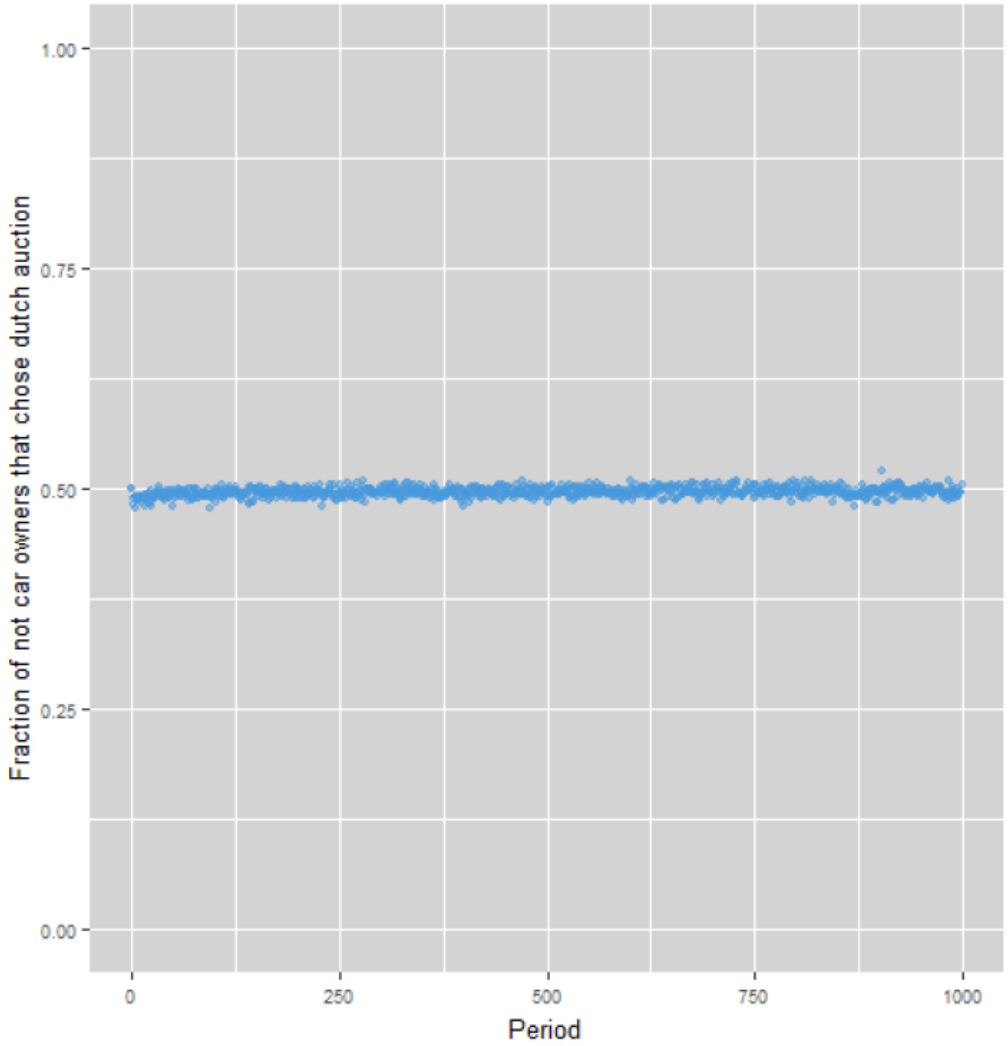


Figure 7: Mean fraction of NCOs who chose Dutch Auctions (100 simulation runs)

Multiple Simulation runs

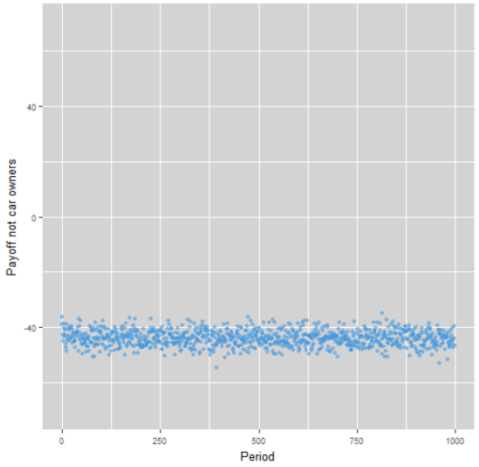


Figure 4: Payoff for NCOs (1 simulation run)

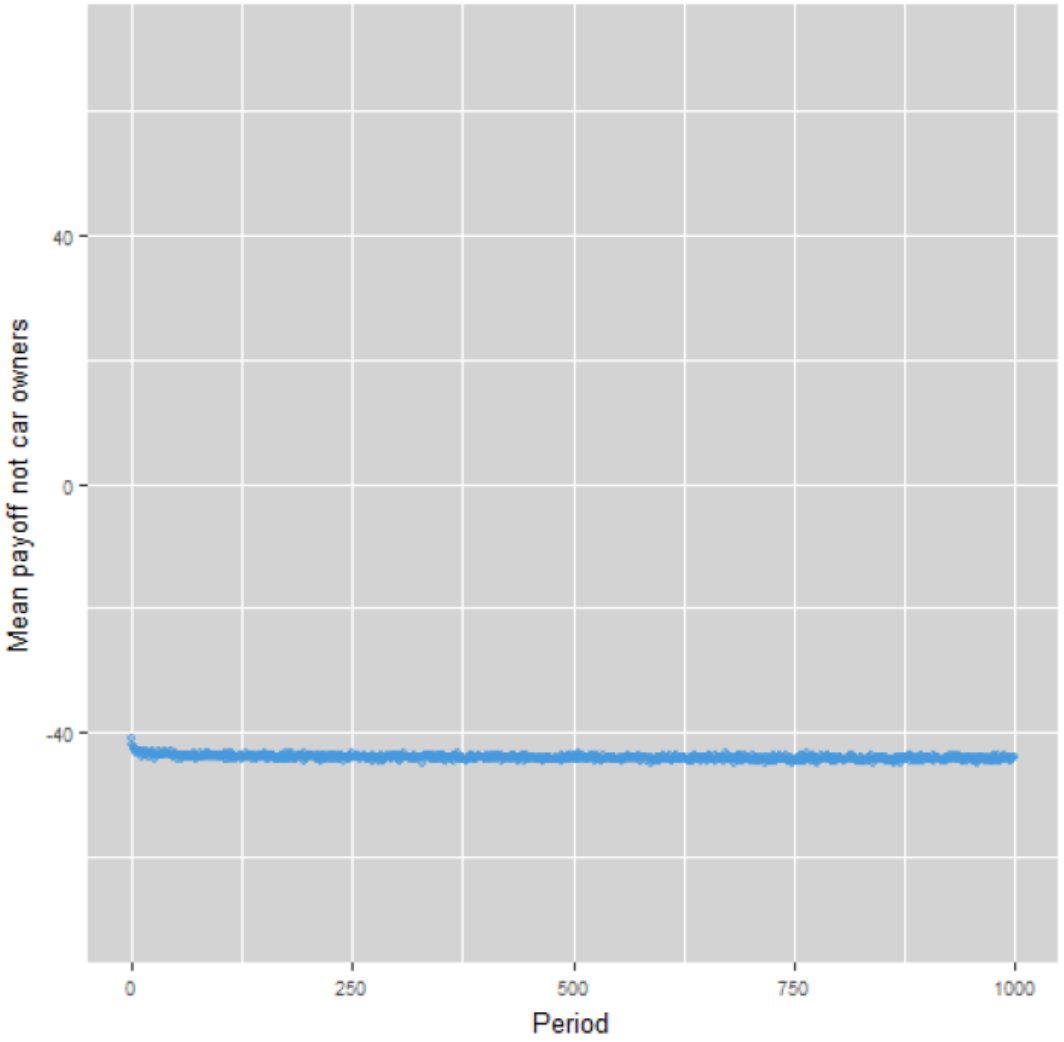


Figure 8: Mean payoff for NCOs (100 simulation runs)

Varying CO/NCO ratio

- Multirun with varying CO/NCO runs:
  - 100 simulation runs with 500 periods each
  - Variation of COs: 10, 20, 50, 100, 200
  - NCO count stays constant: 100.

Figure 7: Mean fraction of NCOs who chose Dutch Auctions (100 simulation runs)

COs	NCOs	FoundBuyers	TookTrain	NCO Dutch	CO Dutch	$\overline{\pi}_i(\sigma)$	$\overline{\pi}_j(\sigma)$
10	100	9.08	90.92	0.49	0.97	30.49 (16.26)	-48.09 (2.96)
20	100	17.14	82.86	0.49	0.89	18.14 (10.02)	-46.19 (2.89)
50	100	39.36	60.64	0.49	0.71	0.74 (6.09)	-35.48 (3.33)
100	100	67.57	32.43	0.47	0.59	-13.93 (3.68)	-18.47 (3.67)
200	100	89.59	10.41	0.45	0.55	-30.08(2.48)	-0.65 (4.58)

Table 2: Results for final periods with varying CO counts (COs: Number of COs; NCOs: Number of NCOs; Found-Buyers: Mean number of COs, that found a buyer; TookTrain: Mean number of NCOs, that took the train; NCO Dutch: Mean fraction of NCOs that Dutch Auction; CO Dutch: Mean fraction of COs that chose Dutch Auction; $\overline{\pi}_i$: Mean payoff for COs; $\overline{\pi}_j$: Mean payoff for NCOs)

Varying CO/NCO ratio

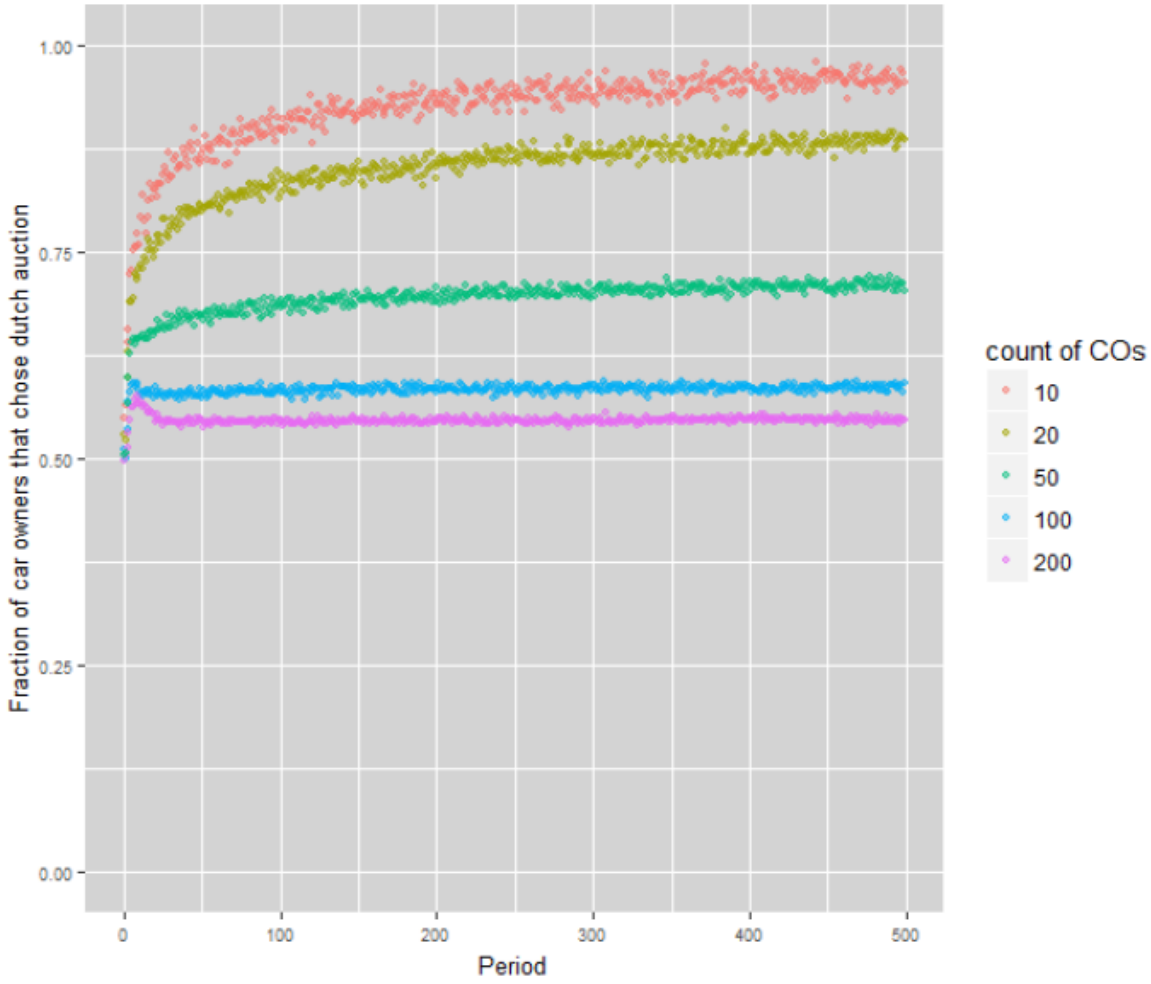


Figure 9: Fraction of COs who chose Dutch Auction with varying CO/NCO ratios

Varying CO/NCO ratio

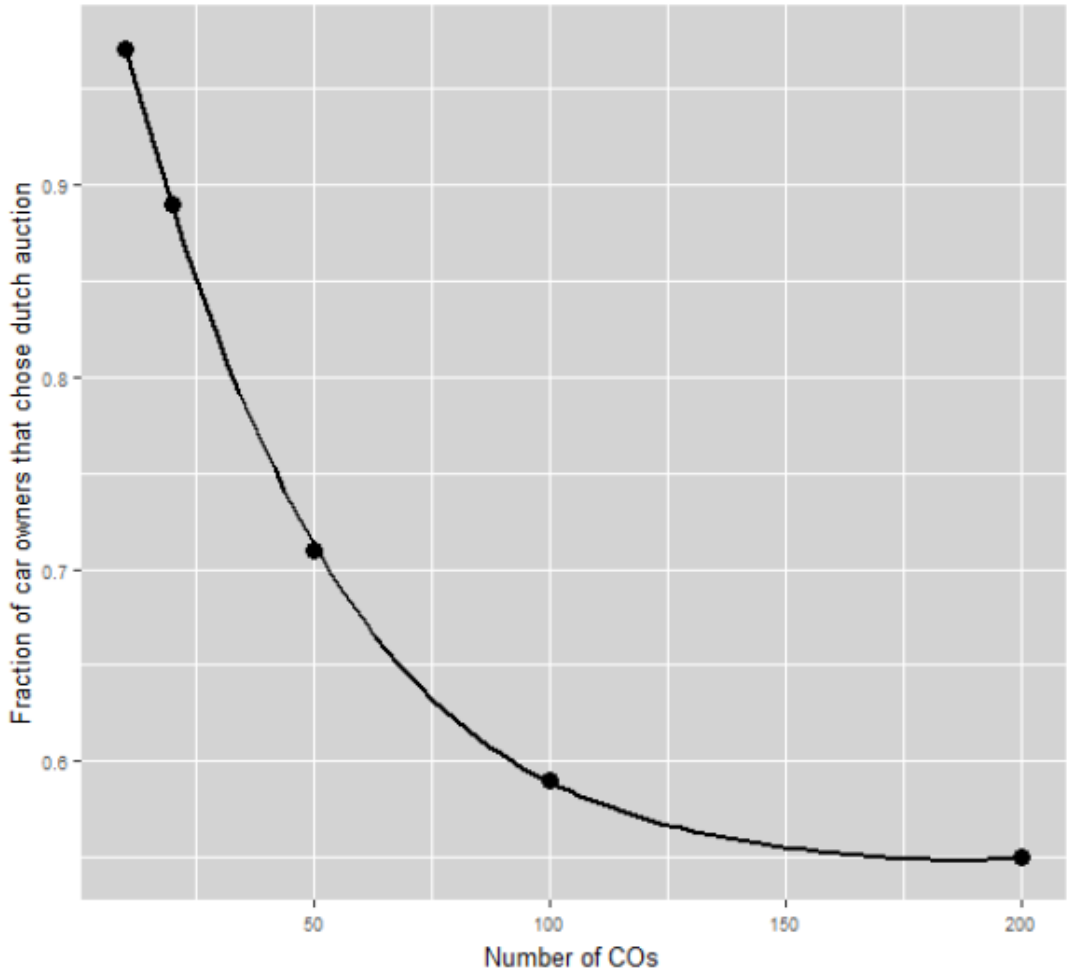


Figure 10: Mean fraction of COs who chose Dutch Auction with varying CO/NCO ratios

Varying CO/NCO ratio

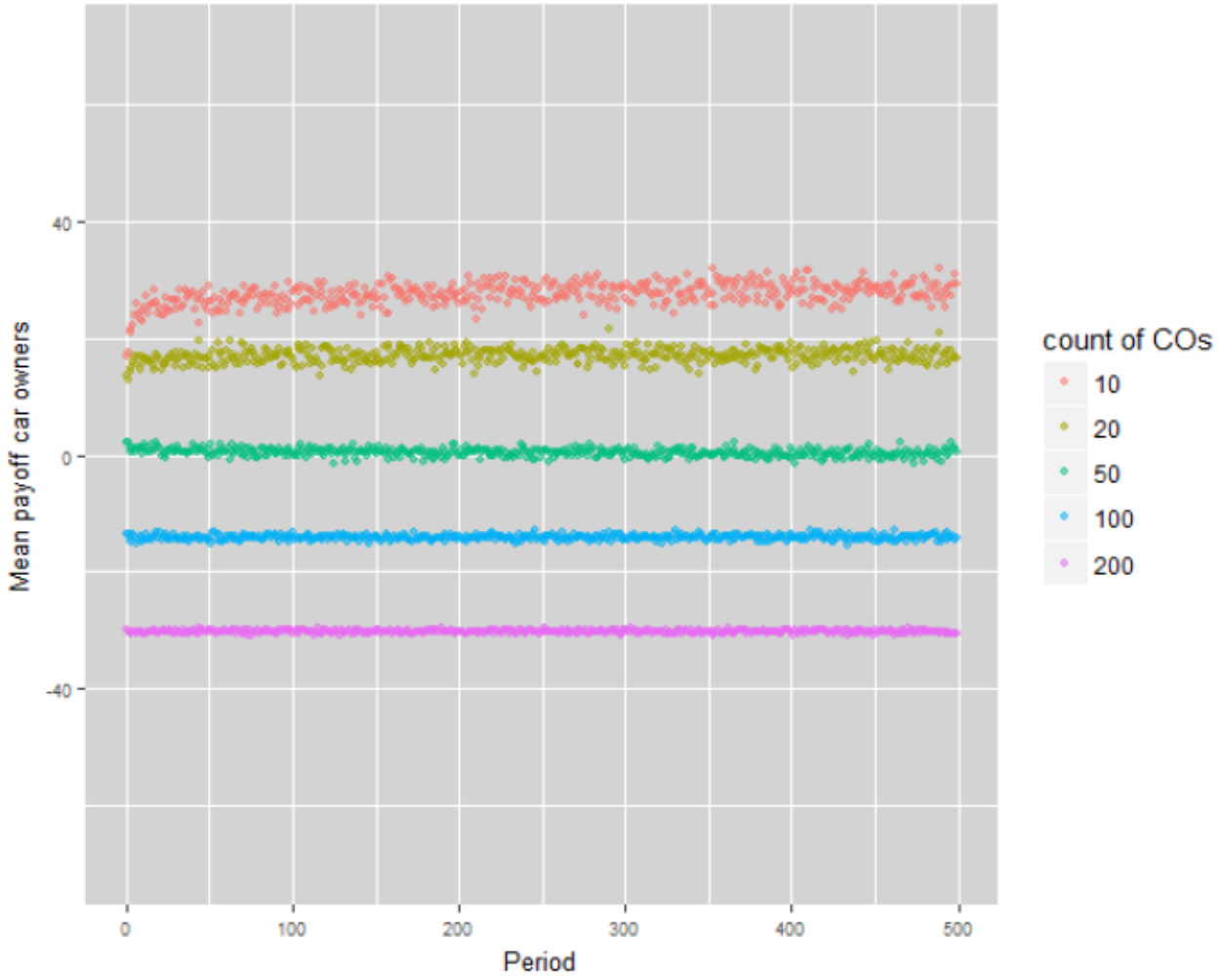


Figure 11: Mean payoff for COs with varying CO/NCO ratios

Varying CO/NCO ratio

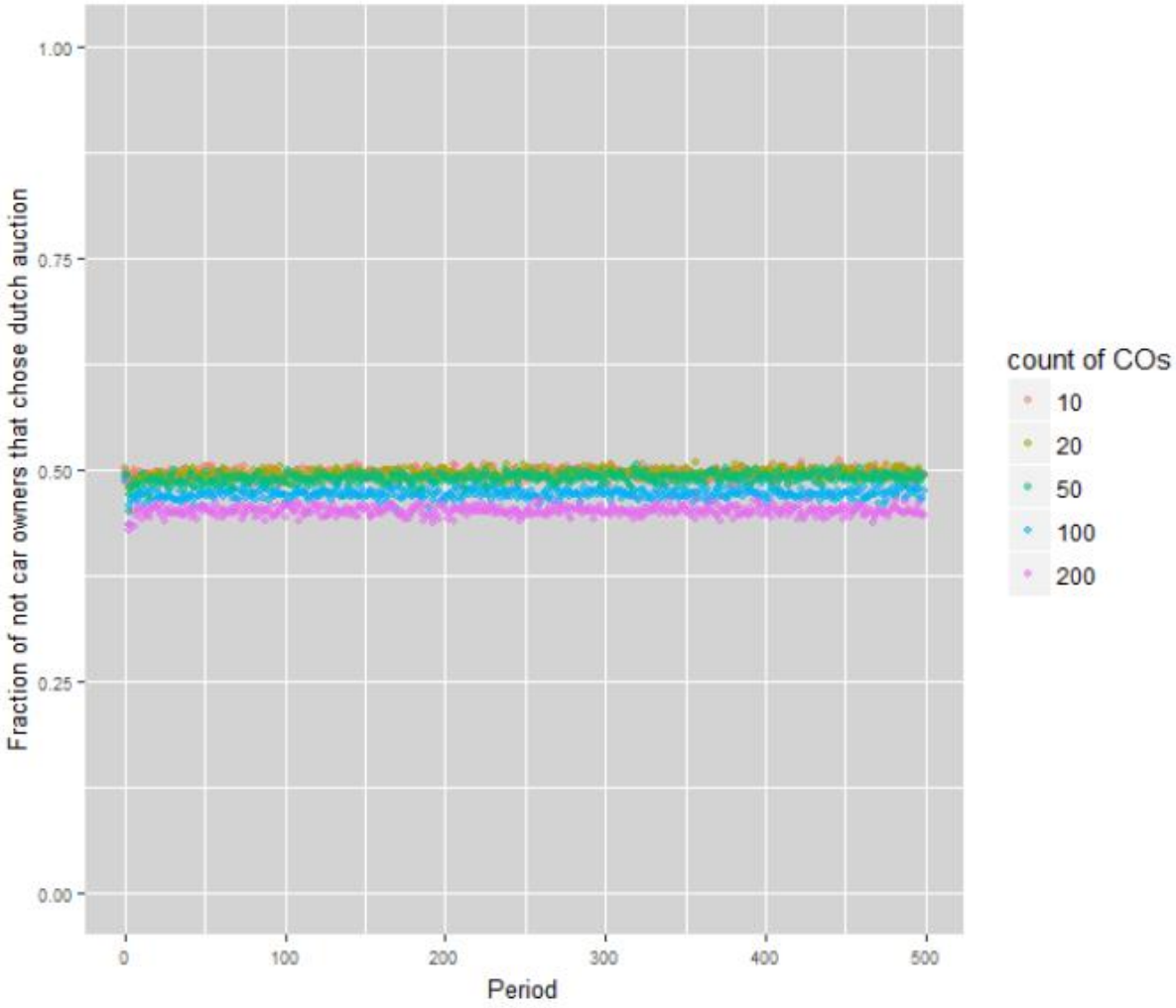


Figure 12: Fraction of NCOs who chose Dutch Auctions with varying CO/NCO ratios

Varying CO/NCO ratio

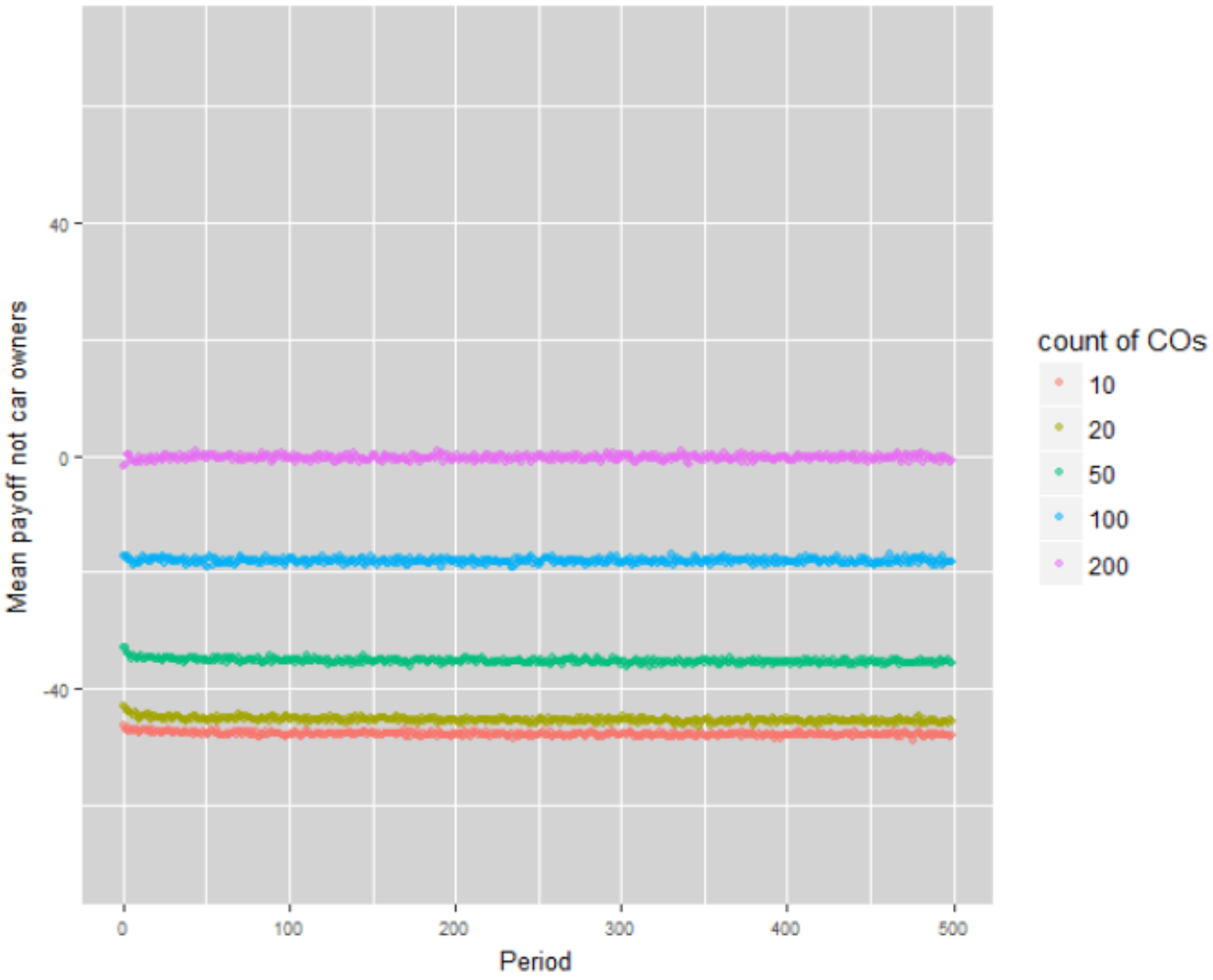


Figure 13: Mean payoff for NCOs with varying CO/NCO ratios

Conclusion

- Under the chosen parameter initializations we found that in our multi agent market simulation **Dutch Auction is clearly preferred by sellers**. Buyers on the other hand have a slight preference for the Fixed Price mechanism.
- The degrees of preferences are **strongly connected to the seller/buyer ratio** of the market. The less sellers are available, the more the sellers will prefer the Dutch Auction. The preference for fixed price for buyers is only observable if there are many sellers available relative to the number of buyers.
- Under the given parameter initializations the results imply, that the use of a Dutch Auction mechanism in the analyzed carpooling markets would **increase the number of sellers** and in conclusion the number of offers. The higher number of offers could lead to **more options** for the buyers. This might have a **growing effect** for the car pooling market and its impact as an alternative for transportation.