Online Appendix

Primal-Dual Value Function Approximation for Stochastic Dynamic Intermodal Transportation with Eco-labels

Arne Heinold, Frank Meisel
School of Economics and Business, Christian-Albrechts-Universität, Kiel, Germany, {arne.heinold;meisel}@bwl.uni-kiel.de

Marlin W. Ulmer
Department of Business Administration, Otto-von-Guericke-Universität, Magdeburg, Germany, marlin.ulmer@ovgu.de

---

Contents

Appendix A.1: Rail corridor data and results

Table A1 provides detailed information about the rail terminals and Table A2 provides descriptive statistics for the five rail corridors that are considered in the experimental study. Table A3 shows experimental results for each corridor.

Table A1: Rail terminals used in the experiments.

id	country	city	longitude	latitude	NUTS2 code
1	Austria	Enns	14.464040	48.221026	AT31
2	Austria	Vienna	16.419189	48.169662	AT13
3	Austria	Wels	14.061931	48.181697	AT31
4	Belgium	Brussels	4.379132	50.878214	BE10
5	Belgium	Liège	5.613670	50.644345	BE33
6	Bulgaria	Sofia	23.309374	42.717075	BG41
7	Bulgaria	Vidin	22.907837	44.005436	BG31
8	Croatia	Zagreb	16.010578	45.757174	HR04
9	Denmark	Copenhagen	12.565903	55.672019	DK01
10	Estonia	Tallinn	24.738283	59.440381	EE00
11	France	Avignon	4.838317	43.936140	FRL0
12	France	Bordeaux	-0.550549	44.789377	FRI1
13	France	Calais	1.848871	50.940848	FRE1
14	France	Lille	3.001735	50.634354	FRE1
15	France	Lyon	4.847635	45.727409	FRK2
16	France	Paris (Noisy-le-Sec)	2.458354	48.896651	FR10
17	France	Paris (Vaires-Torcy)	2.637652	48.875173	FR10
18	France	Paris (Valenton)	2.460646	48.758967	FR10
19	Germany	Berlin	13.369713	52.525051	DE30
20	Germany	Braunschweig	10.539963	52.252608	DE91
21	Germany	Bremen	8.813076	53.082983	DE50
22	Germany	Cologne	7.013148	50.957723	DEA2
23	Germany	Frankfurt am Main	8.662662	50.106530	DE71
24	Germany	Halle / Leipzig	12.318222	51.622882	DEE0
25	Germany	Hamburg	10.007076	53.552812	DE60
26	Germany	Hamm	7.809032	51.678358	DEA5
27	Germany	Hanover	9.741539	52.376319	DE92
28	Germany	Koblenz	7.590497	50.350254	DEB1
29	Germany	Lübeck	10.670178	53.867052	DEF0
30	Germany	Magdeburg	11.627278	52.130682	DEE0
31	Germany	Mainz	8.259620	50.000914	DEB3
32	Germany	Munich	11.560487	48.140116	DE21
33	Germany	Nuremberg	11.082447	49.446139	DE25
34	Germany	Regensburg	12.100189	49.012202	DE23
35	Germany	Rostock	12.124909	54.138555	DE80
36	Germany	Schwerte	7.562163	51.442592	DEA5
37	Greece	Athens	23.667535	37.961970	EL30
38	Greece	Thessaloniki	22.891929	40.665426	EL52
39	Hungary	Budapest	19.087919	47.468587	HU10
40	Hungary	Komarom	18.115012	47.749390	SK02
41	Italy	Bologna	11.344115	44.506286	ITH5
42	Italy	Cervignano A.G.	13.343164	45.823843	ITH4
43	Italy	Florence	11.251361	43.787270	ITI1
44	Italy	Milan	9.117494	45.511339	ITC4
45	Italy	Napoli	14.288077	40.853414	ITF3
46	Italy	Novara	8.624298	45.452662	ITC1
47	Italy	Padova	11.880506	45.418542	ITH3
48	Italy	Roma	12.517239	41.892000	ITI4
49	Italy	Rosarno	15.897764	38.423280	ITF6
50	Italy	Turin	7.679198	45.062136	ITC1
51	Italy	Venice	12.220498	45.484107	ITH3
52	Italy	Verona	10.949910	45.427006	ITH3
53	Latvia	Riga	24.156870	56.943266	LV00
54	Lithuania	Kaunas	24.061114	54.913948	LT00
55	Norway	Oslo	10.761967	59.908491	SE31
56	Poland	Poznan	16.911403	52.401943	PL41
57	Poland	Warsaw	21.051751	52.252203	PL12
58	Portugal	Grandola	-8.554498	38.181846	PT18
59	Portugal	Poceirao	-8.743508	38.634455	PT17
60	Portugal	Sines	-8.843050	37.939595	PT18
61	Romania	Arad	21.325276	46.189729	RO42
62	Romania	Bucharest	26.041328	44.468362	RO32
63	Romania	Calafat	22.925101	43.993994	RO41
64	Romania	Cernavoda	28.071656	44.296271	RO22
65	Romania	Constanta	28.583681	44.165262	RO22
66	Romania	Craiova	23.817774	44.328616	RO41

— Continued on next page —

— Continued Table A1 —

67	Romania	Drobeta Turnu Severin	22.643940	44.622050	RO41
68	Romania	Timisoara	21.207632	45.750793	RO42
69	Slovakia	Bratislava	17.147929	48.171346	SK01
70	Slovenia	Ljubljana	14.614992	46.059614	SI02
71	Spain	Algeciras	-5.449048	36.126428	ES61
72	Spain	Barcelona	2.192621	41.385915	ES51
73	Spain	Cartagena	-0.974878	37.605154	ES62
74	Spain	Cordoba	-4.789366	37.888424	ES61
75	Spain	Madrid	-3.674798	40.386403	ES30
76	Spain	Murcia	-1.027458	38.007272	ES62
77	Spain	Tarragona	1.253371	41.111766	ES51
78	Spain	Valencia	-0.370571	39.441337	ES52
79	Spain	Valladolid	-4.726878	41.642059	ES41
80	Sweden	Gothenburg	11.998932	57.719147	SE23
81	Sweden	Malmo	13.000177	55.609093	SE22
82	Sweden	Stockholm	18.057823	59.330709	SE11
83	Sweden	Trelleborg	13.170864	55.369546	SE22
84	United Kingdom	Folkestone	1.159358	51.097539	UKJ4
85	United Kingdom	Glasgow	-4.257533	55.859122	UKM8
86	United Kingdom	London	-0.123176	51.531637	UKI2

Table A2 Descriptive statistics for the considered rail corridors.

source	target	avg. num. orders	terminals	dist. between stops (km)
Constanta	Cartagena	237	65, 64, 62, 66, 67, 68, 61, 39, 8, 70, 42, 51, 47, 52, 44, 46, 50, 15, 11, 72, 77, 78, 76, 73 (count: 24)	48.1, 185.5, 210.5, 113.8, 210.4, 57.4, 25.5, 251.5, 44.4, 35.9, 100.8, 28.0, 85.3, 152.2, 42.0, 99.7, 88.7, 232.8, 304.7, 92.4, 257.3, 99.1, 54.5 (avg: 122.6)
Glasgow	Athens	261	85, 86, 84, 13, 14, 4, 5, 22, 28, 31, 23, 33, 34, 3, 1, 2, 69, 40, 39, 61, 68, 67, 66, 63, 7, 6, 38, 37 (count: 28)	13.7, 89.3, 56.8, 105.3, 124.3, 104.0, 16.8, 92.8, 100.0, 37.3, 226.8, 100.4, 118.9, 38.5, 183.7, 66.2, 46.2, 94.1, 222.3, 57.4, 210.4, 113.8, 111.8, 16.4, 270.6, 209.6, 210.0 (avg: 112.5)
Oslo	Sines	224	55, 80, 81, 9, 29, 25, 21, 27, 26, 36, 22, 5, 4, 14, 16, 17, 18, 12, 79, 59, 58, 60 (count: 22)	93.1, 146.6, 21.4, 63.8, 62.5, 116.7, 122.2, 176.1, 34.2, 81.3, 83.2, 104.0, 124.3, 231.7, 14.3, 24.7, 110.5, 198.2, 46.2, 57.1, 77.4 (avg: 94.7)
Stockholm	Rosarno	180	82, 81, 83, 35, 19, 24, 34, 32, 52, 47, 41, 43, 48, 45, 49 (count: 15)	46.1, 33.3, 159.6, 227.5, 134.3, 200.7, 138.0, 89.4, 85.3, 123.9, 92.2, 260.2, 236.0, 416.5 (avg: 160.2)
Tallinn	Algeciras	230	10, 53, 54, 57, 56, 19, 30, 20, 27, 26, 36, 22, 5, 4, 14, 16, 17, 18, 12, 79, 75, 74, 71 (count: 23)	56.5, 44.8, 122.2, 305.4, 173.1, 145.6, 82.8, 66.6, 176.1, 34.2, 81.3, 83.2, 104.0, 124.3, 231.7, 14.3, 24.7, 110.5, 198.2, 189.4, 172.6, 48.8 (avg: 117.7)

Table A3 Average objective's lower-bound gap for each corridor.

		objective costs			objective eco-labels		
source	target	single stage	*VFA CO*	*VFA EL*	single stage	*VFA CO*	*VFA EL*
Constanta	Cartagena	13.21	3.87	6.79	79.43	23.37	10.46
Glasgow	Athens	10.55	3.36	5.81	65.29	24.71	11.56
Oslo	Sines	17.06	4.48	8.09	71.50	17.83	9.73
Stockholm	Rosarno	8.85	4.08	6.51	55.49	16.95	9.73
Tallinn	Algeciras	12.89	3.84	6.37	75.97	20.97	10.34
average		12.51	3.93	6.71	69.54	20.77	10.36

Appendix A.2: Illustration of an order's possible routings

Figure A1 illustrates an order's possible routings in which the dashed lines represent truck transportation, used in the pre- and post-carriage to/from the terminals and for the direct (road-only) routing whereas the solid lines represent the transportation via train. In our illustrative example, three terminals qualify for loading (I, II, III) and two terminals for unloading (X, XI), resulting in six possible intermodal paths for this order.

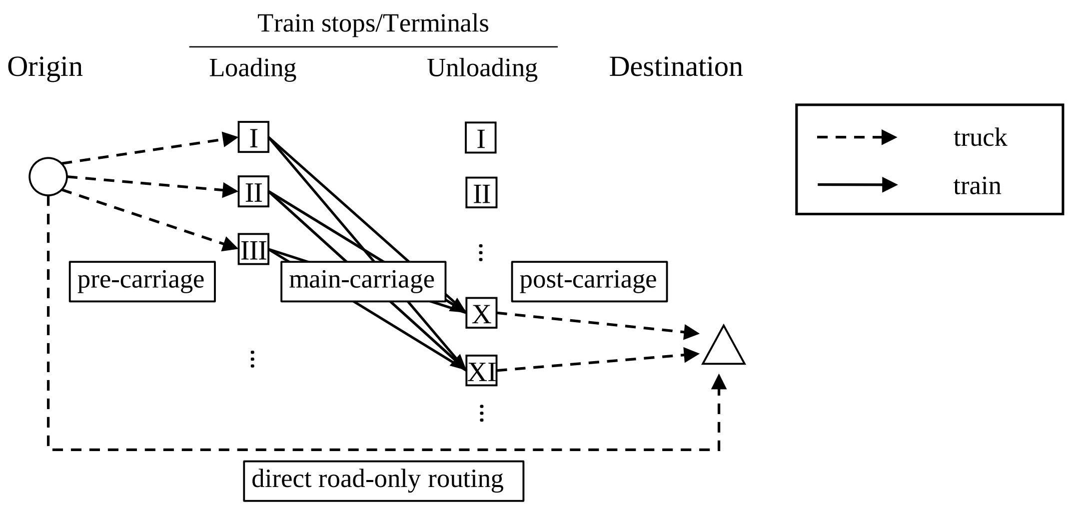

Figure A1 Illustration of an order's possible routings.

Appendix A.3: Notation of the sequential decision process

An overview of the notation that is used in the sequential decision process is provided in Table A4.

Table A4 Overview of the notation used in the sequential decision process.

name	description
General notation	
$\bar{e}_o$	upper emission limit induced by eco-label λ_o of an order $o \in \mathcal{O}_t$
q_o	payload quantity of an order $o \in \mathcal{O}_t$
t_p^1 / t_p^2	loading/unloading terminal of a path $p \in \mathcal{P}_t$
λ_o	eco-label of an order $o \in \mathcal{O}_t$
$\mathcal{T}$	set of decision points / train stops / terminals
$\mathcal{O}_t$	set of orders in decision point $t \in \mathcal{T}$
$\mathcal{P}_t$	set of paths in decision point $t \in \mathcal{T}$
$\mathcal{P}_t^o$	set of paths of an order $o \in \mathcal{O}_t$ in decision point $t \in \mathcal{T}$
$\mathcal{T}_p$	set of terminals visited by a path $p \in \mathcal{P}_t$
State	
$\boldsymbol{f_t}$	binary vector describing which orders violate their eco-label in a state S_t
$f_t(o)$	=1, if order $o \in \mathcal{O}$ violates its eco-label in a state S_t, =0 else
$\boldsymbol{r_t}$	binary vector describing which paths are used in a state S_t
$r_t(p)$	=1, if path $p \in \mathcal{P}$ is used in a state S_t, =0 else
S_t	state in decision point $t \in \mathcal{T}$
Decision	
$e_t^x(p)$	overall emissions (truck and train) of path $p \in \mathcal{P}$ under the decision in a state S_t, =0 else
$l_t^x(t')$	load of the train at terminal $t' \in \mathcal{T}$ under the decision in a state S_t, =0 else
$L^{\max}$	capacity of the train
$\boldsymbol{x_t}$	binary vector describing which paths are used under the decision in a state S_t
$x_t(p)$	=1, if path $p \in \mathcal{P}_t$ is used under the decision in a state S_t, =0 else
$\mathcal{X}(S_t)$	set of feasible decisions in a state S_t
Reward	
c_p	overall costs (truck and train) of a path $p \in \mathcal{P}_t$
$e^{train}(l_t^x(t'))$	train emissions under a load $l_t^x(t')$
e_p^{truck}	truck emissions of a path $p \in \mathcal{P}_t$
$\boldsymbol{f_t^x}$	binary vector describing which orders violate their eco-label under a decision $\boldsymbol{x_t}$
$f_t^x(o)$	=1, if order $o \in \mathcal{O}$ is routed in accordance to its eco-label under a decision $\boldsymbol{x_t}$, =0 else
$R(\cdot)$	reward function
$R^C(\cdot)$	reward function for objective costs
$R^E(\cdot)$	reward function for objective eco-labels
α	parameter to weight $R^C(\cdot)$ and $R^E(\cdot)$ in the objective function
Post-decision state	
$\mathcal{P}_t^x$	set of paths in a post-decision state S_t^x
$\mathcal{P}^{o,x}$	set of available paths of an order $o \in \mathcal{O}$ in a post-decision state S_t^x
S_t^x	post-decision state in decision point $t \in \mathcal{T}$
Stochastic information	
$\boldsymbol{\hat{\omega}_{t+1}}$	stochastic information in decision point $t+1 \in \mathcal{T}$
$\tilde{\mathcal{O}}_{t+1}$	set of new orders in decision point $t+1 \in \mathcal{T}$
$\tilde{\mathcal{P}}_{t+1}$	set of new paths in decision point $t+1 \in \mathcal{T}$
Ω	set of stochastic information
Transition	
$S^M(\cdot)$	system model function to transform to the next state
Optimal policy	
$V_t(S_t^x)$	value of post-decision state S_t^x
π	policy describing a decision for each decision point
Π	set of all policies

Appendix A.4: Generic algorithm for VFA with basis functions

Algorithm 1 outlines a generic framework for VFA with basis functions. Step 1 is to initialize the feature weights θ_a^0, which is done by using 0 for all features in our experiments. This results in an initially myopic policy. The problem is then solved in N iterations under Step 2, where each iteration simulates a single trip of the train along its route for a sampled set of orders appearing at each terminal, i.e., in each iteration we compute one trajectory of the problem. Thereby, N is a sufficiently large iteration counter at which the approximations converged sufficiently. It is parameterized by experiment. At each iteration $n = 1, 2, \ldots, N$, we randomly draw an information $\boldsymbol{\hat{\omega}_1^n}$ that yields the currently known orders for the initial state S_1^n (in which the train is at the first terminal), see Step 2a. From here, we solve the problem in a forward programming manner once for each decision point $t = 1, 2, \ldots, T-1$. In particular, we find a decision $\boldsymbol{x_t^n}$ in Step 2b by solving the MIP with the latest values of the features' weights θ_a^{n-1} (the weights obtained after the previous iteration). The MIP is formally described at the end of this section. The decision and the value of the corresponding post decision state are stored in Step 2c as they are used later in the updating process. Then, in Step 2d, we draw an outcome of the random information $\boldsymbol{\hat{\omega}_{t+1}^n}$ to obtain new orders and transform to the next state S_{t+1}^n using the transition function $S^M(\cdot)$. Finally, after making the last decision in the last relevant decision point $T-1$ (the train's penultimate stop), we update the weights of the features in Step 2e using the stored values from Step 2c. The value of $S_t^{n,x}$ can be calculated as the difference in accumulated reward in t and T.

Step 1: Choose initial values for weights θ_a^0 of all features $a \in \mathcal{A}$,
Step 2: **for** *n = 1,2,. . .,N* **do**
  Step 2a: Draw information $\boldsymbol{\hat{\omega}_1^n} \in \Omega$ for initial state S_1^n.
  for *t = 1,2,. . .,T-1* **do**
    Step 2b: Find primal decision $\boldsymbol{x_t^n}$ by solving the MIP model using weights θ_a^{n-1}.
    Step 2c: Store decision $\boldsymbol{x_t^n}$, post-decision state $S_t^{n,x}$, and accumulated reward.
    Step 2d: Draw information $\boldsymbol{\hat{\omega}_{t+1}^n} \in \Omega$ and transform to next state: $S_{t+1}^n = S^M(\cdot)$.
  end
  Step 2e: Update weights θ_a^n for all features $a \in \mathcal{A}$ using the values from step 2c.
end

Algorithm 1: Generic algorithm for VFA with basis functions.

Appendix A.5: Analysing overall emissions

We report in this section results from minimizing overall emissions of a solution and their interplay with the objectives of minimizing costs and eco-label violations. We do this by considering an objective 'emissions' in which overall emissions are minimized. For this, we define $R_t^{CO2}(S_t, \boldsymbol{r_t}, \boldsymbol{x_t}) = \sum_{p \in \mathcal{P}_t} e_t^x \cdot (x_t(p) - r_t(p))$ as the reward that is used in the sequential decision process. The pareto frontiers for this objective and objectives costs and eco-labels is shown in Figures A2a and A2b. The pareto frontiers are calculated like the pareto frontier in paper. For reasons of readability, we only show results for the dual solution as well as *VFA EL* and *VFA CO* for objectives eco-labels and costs, respectively. We make two important observations. First, our VFAs also learn under the objective of minimizing overall emissions. In particular, *VFA CO* and *VFA EL* both result in a lower-bound gap of about 4% under the pure emission objective, indicating that these routings are very close to the dual solution assuming an ex-post decision making under perfect information. Second, the shapes of the pareto frontiers constitute the trade-off between the different objectives. In particular, achieving low overall emissions comes along with higher overall costs as well as solutions in which less orders are routed in accordance to their requested eco-labels. The figures also reveal that the trade-off between emissions and eco-labels is much stronger than it is between emissions and costs, e.g., the lower-bound gap of eco-labels is higher compared to the lower-bound gap of costs under a pure emission objective. There are two reasons for this. First, objectives emissions and costs generally tend to point towards similar routings (e.g., Kopfer et al. 2014) and, second, we treat the fulfillment of eco-labels in a binary manner (yes or no) and not as a relative measure. The latter reason results in a decision making in which the carrier does not consider the magnitude at which an order fails to meet its requested eco-label. For a discussion of this issue we refer to Heinold and Meisel (2020) for an intermodal transportation problem in which the emissions exceeding an order's emission limit are minimized too.

(a) Pareto frontiers for varying weights of objectives costs and emissions.

(b) Pareto frontiers for varying weights of objectives eco-labels and emissions.

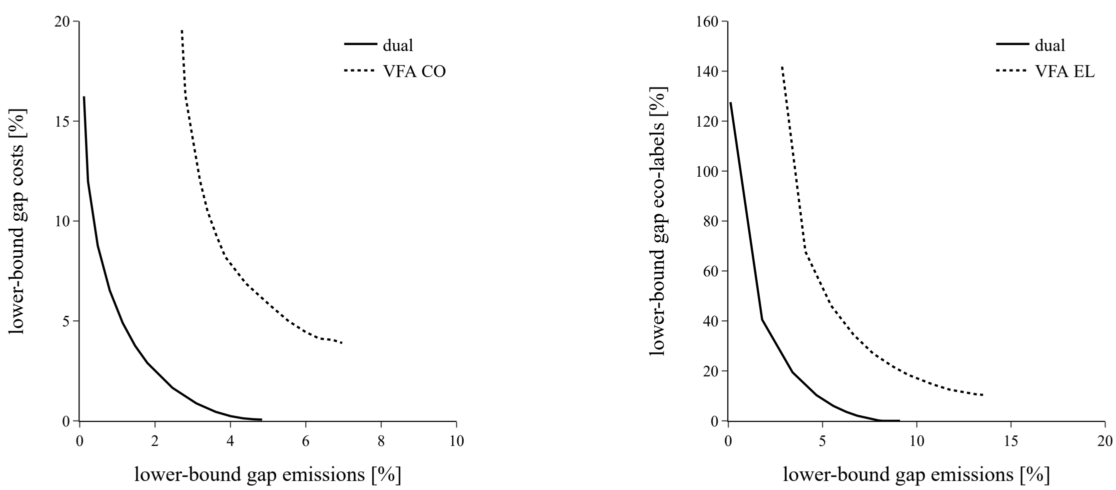

Figure A2 **Pareto frontiers for objectives emissions vs. costs and objectives emissions vs. eco-labels.**

Appendix A.6: The intermodal share under different solution approaches

The intermodal share under different solution approaches is shown in Figure A3.

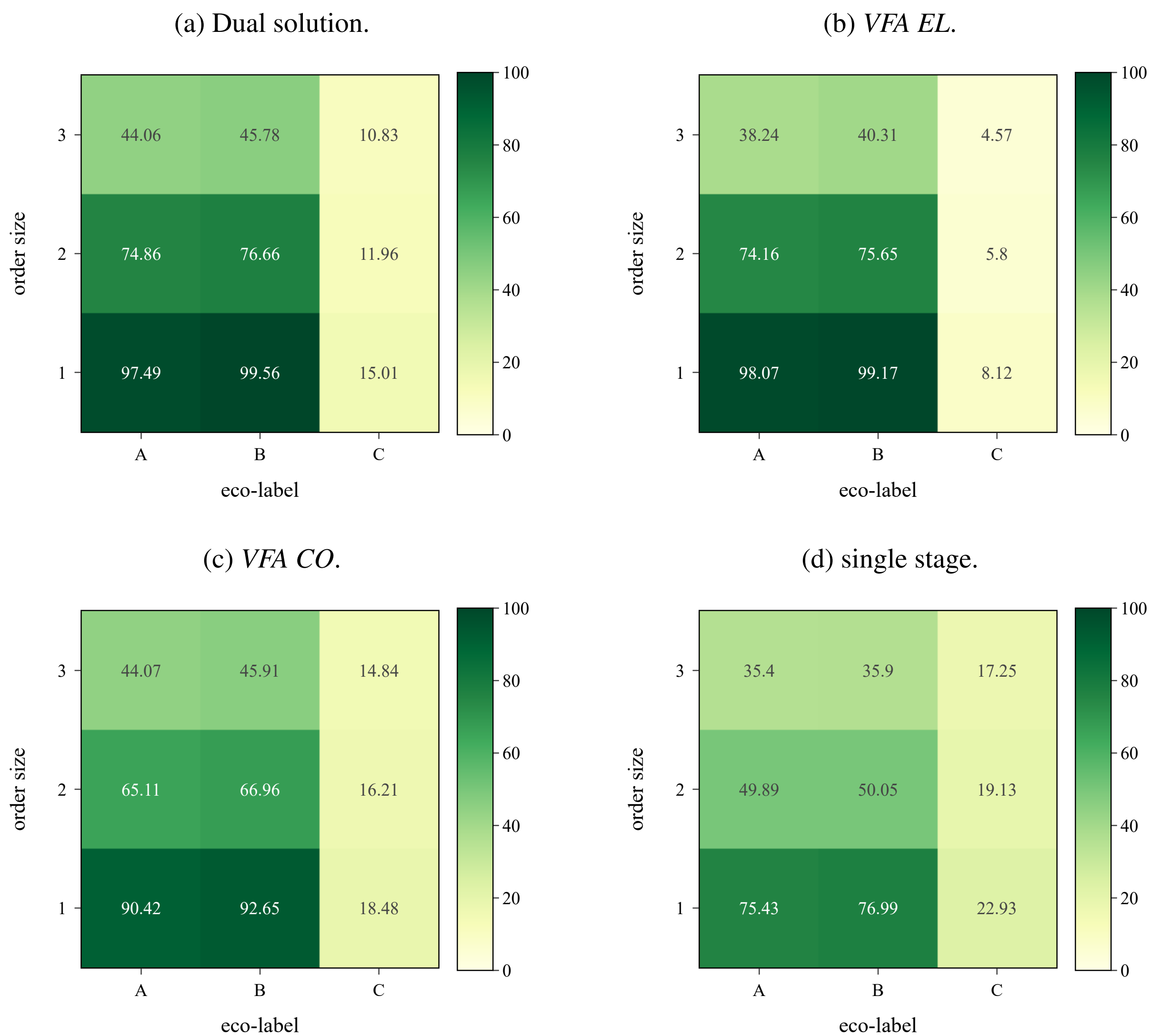

Figure A3 The intermodal share under objective eco-labels for combinations of an order's eco-label and size.

Appendix A.7: Analyzing the fit of the feature sets

We report in Table A5 the coefficient of determination (R^2) resulting from a regression between the optimal value (calculated ex-post like the dual trajectory) and the value from the VFA approach in the final decision point T. Thus, it is a general indicator for the fit of a feature set by measuring whether it is suited to predict the value in the final decision point, or not. The table shows values for both performance measures under both objective for the VFA approaches as well as the benchmark single-stage heuristic. For objective eco-labels, the values for all feature sets and for both measures are high (R^2=1 means a perfect fit) and differences between VFA approaches are small. In contrast, for objective costs, differences between the sets are once again small, but R^2 is clearly higher for the (corresponding) measure costs than for measure eco-labels. This means that the objective costs does not work well for complying with eco-labels, whereas the objective eco-labels also leads to relatively good values for the measure costs. This observation is also in line with the finding from the pareto frontier in the experimental study in which the gaps for eco-labels are higher under the pure cost objective compared to the gaps for costs under the pure eco-label objective. This is also shown in the results for the single-stage heuristic in which R^2 is higher for the measure costs than it is for the measure eco-labels under both objectives. Here, we note that R^2 values of the single-stage heuristic are generally high for both measures and both objectives, although the results in our experimental section have shown that the approach does not perform so well when looking at the lower-bound gap. This confirms that the R^2 as calculated in this section works as an indicator for the fit of a policy but has a limited use when it comes to predicting its actual performance (e.g., Mes and Rivera 2017).

Table A5 Coefficient of determination (R^2) of the solution approaches under both objectives.

	single stage	*VFA EL*	*VFA CO*	*VFA All*
objective eco-labels				
eco-labels	0.7818	0.9567	0.9289	0.9252
costs	0.9806	0.9250	0.9109	0.9064
objective costs				
eco-labels	0.7380	0.7495	0.7636	0.7546
costs	0.9860	0.9835	0.9967	0.9901

To further test the fit of our feature sets, we look at the performance of the VFAs for test problem instances that somewhat differ from those that were used in the learning. In particular, we test the performance of the basis functions that were learned under the default setting (i.e., order size q_o is 1, 2, or 3 FTL with equal probability) on a setting in which all orders have an identical equal size of either 1, 2, or 3 FTL. We show results for objectives costs and eco-labels in Figures A4 and A5, respectively. We present the box plots for solution approaches single stage, *VFA CO* (for objective costs) and *VFA EL* (for objective eco-labels), where '*' denotes that the VFA is learned under the default setting, whereas learning and testing distribution is the same otherwise. It can be seen that the performance of the VFAs is better than the benchmark heuristic in all of the considered experiments. Furthermore, the VFAs learned under the default distribution perform similar to the VFAs learned under the equal-size distributions. In particular, *VFA CO** and *VFA EL** perform as good as *VFA CO* and *VFA EL* for $q = 2$ and $q = 3$ and for $q = 1$ and $q = 2$ under objectives costs and eco-labels, respectively. Overall, we conclude that our feature sets also work well in settings with other distributions.

(a) All orders with $q = 1$. (b) All orders with $q = 2$. (c) All orders with $q = 3$.

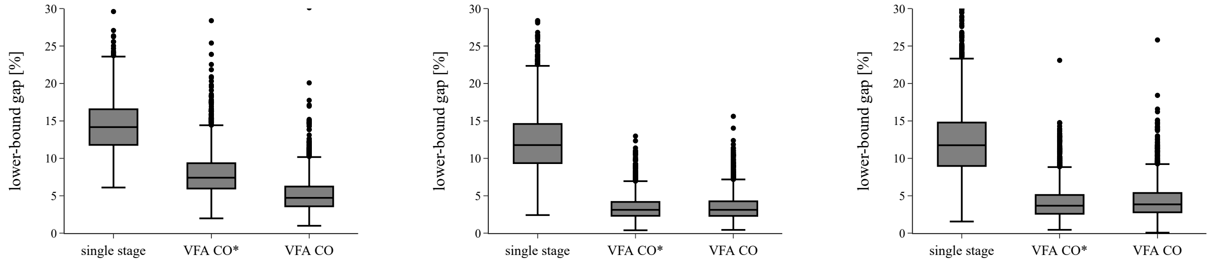

Figure A4 **Results for objective costs and approaches single stage, *VFA CO*, and *VFA CO** under different settings of equal-size orders.**

(a) All orders with $q = 1$. (b) All orders with $q = 2$. (c) All orders with $q = 3$.

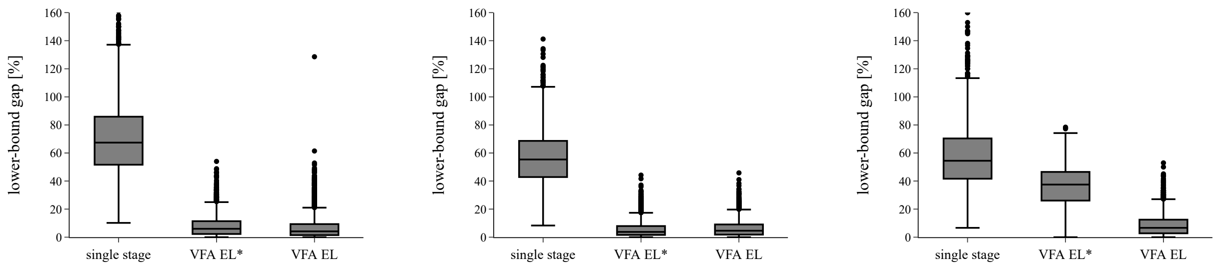

Figure A5 **Results for objective eco-labels and approaches single stage, *VFA EL*, and *VFA EL** under different settings of equal-size orders.**

References

Heinold, A. and Meisel, F. (2020). Emission limits and emission allocation schemes in intermodal freight transportation, *Transportation Research Part E: Logistics and Transportation Review* **141**: 101963.

Kopfer, H. W., Schönberger, J. and Kopfer, H. (2014). Reducing greenhouse gas emissions of a heterogeneous vehicle fleet, *Flexible Services and Manufacturing Journal* **26**(1): 221–248.

Mes, M. R. K. and Rivera, A. P. (2017). Approximate dynamic programming by practical examples, *in* R. J. Boucherie and N. M. van Dijk (eds), *Markov Decision Processes in Practice*, Springer International Publishing, Cham, pp. 63–101.