Online Appendix

A Bilevel Approach to Multi-Period Natural Gas Pricing and Investment in Gas-Consuming Infrastructure

1. Different Bilevel Reformulations

In addition to the approach taken in the paper to linearize the bilinear term in the upper-level player's objective function, we considered several modeling alternatives. Although a review of the literature led us to believe that each would prove inferior to exact linearization, we nevertheless conducted a comparative analysis to determine the relative performance of each. The computational results presented below provide insights for those working on similar problems where an exact linear reformulation might not be possible.

The first approach is to simply leaving the bilinear term as is in the objective function and use a solver that is able to handle non-convex problems. For all of our runs, we used Gurobi 9.1.2, which we can use in this case as well since it is able to solve problems to optimality with both bilinear terms in the objective function and the constraints (Gurobi Optimization, 2020). For another optimization model dealing with generation expansion planning using this solver's capability of handling bilinear terms directly, see Guo et al. (2022). Another approach we tested was based on McCormick envelopes to approximate the bilinear term (McCormick, 1976). This approach replaces a bilinear term with one variable representing the approximated value, and four linear constraints: two convex under-estimators and two convex over-estimators such that a bilinear term xy can be replaced by w when the following set of inequalities are enforced.

$$w = xy$$

$$w \ge \underline{x}y + x\underline{y} - \underline{x}\underline{y}$$

$$w \ge \overline{x}y + x\overline{y} - \overline{x}\overline{y}$$

$$w \le \overline{x}y + x\underline{y} - \overline{x}\underline{y}$$

$$w \le \underline{x}y + x\overline{y} - \underline{x}\overline{y}$$

$$(1a)$$

In (1a), $(\underline{x},\overline{x})$ and $(\underline{y},\overline{y})$ represent upper and lower bounds on x and y, respectively. For examples of McCormick envelopes being used in bilevel models in the energy domain, see Zhao et al. (2019); Zeng et al. (2020).

Another approach we tested was to replace the bilinear terms with piece-wise linear approximations. In this method, we replaced xy with its exact reformulation $m^2 - n^2$, where $m = \frac{x+y}{2}$ and $n = \frac{x-y}{2}$. We then replaced the quadratic terms, i.e., m^2 and n^2 , with their piece-wise linear approximations. For further discussion of this method and an application in a bilevel setting, see Jayadev et al. (i2022).

Each of the approaches was run on smaller instances of our problem for two different planning horizons: 10 years and 15 years. This enabled us to observe how each approach scaled with the increasing number of time periods. All models were solved with an optimality gap of $1e^{-4}$ with the same parameter values used in the REFR scenario on a personal computer with an Intel[®] CoreTM i-7 processor, clocking at 1.30 GHz with 32 GB of RAM.

Table 1 presents the computational results from the runs we made with the KKT approach using four different treatments of the bilinear term: exact linearization (LNRZ), leaving it as is (BLNR), using McCormick envelopes (MCRM), and using a piece-wise linear approximation (PCWS). We observed both in 10-year and 15-year runs, solutions from LNRZ, BLNR, and PCWS produce the same

revenue for the producer. MCRM, however, produces significantly lower revenue for the producer in both runs. When we observe the run times of these cases, we see that LNRZ significantly outperforms the other three approaches for both planning horizons. After LNRZ, BLNR appears to be the best option as it reaches the optimal solution in the second shortest time converging to the same upper-level revenue value.

Table 1.	CPU times	(s) for a	a specified	optimality g	gap and	revenue of the	producer at the solution
----------	-----------	-----------	-------------	--------------	---------	----------------	--------------------------

Gap	LNRZ		BLNR		MCRM		PCWS	
Сар	10-year	15-year	10-year	15-year	10-year	15-year	10-year	15-year
10.00%		1		165		150		-
5.00%				200		195		-
1.00%				225		225		-
0.00%	0.26	2.18	5.69	226.11	3.96	226.48	21.81	-
Revenue	215.481	214.767	215.481	214.767	189.309	186.637	215.481	190.681^{1}

In a second set of tests, we tried to leverage the solver's ability to handle bilinear terms directly. Our first insight was that the KKT-based lower-level reformulation coupled with the BLNR approach does not yield scalable solutions for the full planning horizon (20 years). However, we also tested another approach in which Gurobi could be used to handle the bilinear terms. In that reformulation, a duality-based approach of the lower-level problem is used instead of the KKT-based approach. This involves replacing the lower-level LP with its primal constraints, dual constraints, and the strong duality condition instead of the KKT conditions. It is well known that these three sets of constraints are sufficient for optimality and so can be used in place of the lower-level LP. See Garcia-Herreros et al. (2016) for an example of a bilevel model that is converted to a single-level model using this approach.

We now present the duality-based reformulation of our bilevel problem.

$$\begin{aligned} \max \sum_{r,t} \delta_t \big(FUEL_t^{NG} - PROD_t \big) n_{rt}^{NG} + \delta_t \big(FUEL_t^{NC} - PROD_t \big) n_{rt}^{NC} \\ \text{s.t.} & HRR \times FUEL_t^{NG} = FUEL_t^{NC} & \forall t \in \mathcal{T} \\ & FUEL_t^{NG} \geq \big(1 - MAXDEC \big) FUEL_{t-1}^{NG} & \forall t \in \mathcal{T}, t \neq 2020 \\ & FUEL_t^{NG} \leq \big(1 + MAXINC \big) FUEL_{t-1}^{NG} & \forall t \in \mathcal{T}, t \neq 2020 \\ & FUEL_{2020}^{NG} \leq \overline{FUEL}_{2020}^{NG} & \forall t \in \mathcal{T}, t \neq 2020 \\ & \sum_{f} n_{rt}^f + v_{rt}^{DCHR} - v_{rt}^{CHR} = DEM_{rt} & \forall t \in \mathcal{T}, r \in \mathcal{R}_S \\ & n_{rt} - \sum_{t'} ENC_{rtt'}^f m_{t'}^f \leq INIT_{rt}^f & \forall f \in \mathcal{F}, r \in \mathcal{R}, t \in \mathcal{T} \\ & - \sum_{f,t'} CREN_{rtt'}^f m_{t'}^f - \sum_{t'} CREN_{rtt'}^{BT} m_{t'}^{BT} \leq CRIN_{rt}^{BT} + \sum_{f} CRIN_{rt}^f - REQC_{rt} \\ & \forall r \in \mathcal{R}, t \in \mathcal{T} \\ & \forall r \in \mathcal{R}, t \in \mathcal{T} \end{aligned}$$

¹PCWS timed out at 1000 seconds with 99.24% optimality gap. It found a non-optimal feasible solution by then; however, it failed to close the optimality gap before the time limit.

$$\begin{split} \sum_{r} v_{rt}^{DCHR} - \sum_{r} EFF_{t} v_{rt}^{CHR} &\leq 0 \\ \sum_{r} v_{rt}^{DCHR} - \sum_{r} EFF_{t} v_{rt}^{CHR} &\leq 0 \\ \nabla v_{rt}^{CHR} - \sum_{f} v_{rt}^{DCHR} - \sum_{r} EFF_{t} v_{rt}^{CHR} &\leq 0 \\ \nabla v_{rt}^{CHR} - \sum_{f} v_{rt}^{DCHR} - \sum_{f} v_{rt}^{DCHR} &\leq 0 \\ \nabla v_{rt}^{CHR} - \sum_{f} v_{rt}^{DCHR} - \sum_{f} v_{rt}^{DCHR} &\leq 0 \\ \nabla v_{rt}^{CHR} - \sum_{f} v_{rt}^{DCHR} - \sum_{f} v_{rt}^{DCHR} - v_{rt}^{DCHR} -$$

This method, however, requires using the strong duality condition in the constraint set of the model, which in turn brings bilinear terms into the constraints. To test whether this approach is better than the linearization approach (now called KKT-LNRZ) we used in the paper, we ran the duality-based BLNR approach (now called Dual-BLNR for more clarity) for the first 7 scenarios previously discussed and the 5 scenarios discussed in Section 2 below for the full planning horizon. The runtimes for both KKT-LNRZ and Dual-BLNR are given in Table 2 using an optimality gap of 0.5% and 1000 seconds of CPU time as the stopping criteria.

The results demonstrate that whether one method outperforms the other depends on the parameter values. What we see is that KKT-LNRZ shows less variation in runtimes while the Dual-BLNR approach shows a greater variation, especially considering the extremely slow convergence of the

²Timed out at 1000 seconds at 2.15% optimality gap.

Table 2. CPU times (s) of each scenario with the linearization and duality-based approaches

Scenario	KKT-LNRZ	Dual-BLNR
REFR	9.13	249.14
ITAX	68.52	41.95
LTAX	54.70	43.49
HTAX	29.09	85.92
DCRD	447.17	64.45
LCRD	520.02	453.65
HCRD	142.97	131.09
ITDC	38.08	114.95
HTHC	25.34	1000^{2}
HTLC	14.52	214.71
LTHC	29.40	62.83
LTLC	136.16	134.69

HTHC scenario. The runtime differences when KKT-LNRZ outperforms Dual-BLNR tend to be greater than the runtime differences when the performance is the other way around. These results also demonstrate that although instance sizes are the same, runtimes can be significantly affected by changes in the value of in a few parameters.

2. Additional Scenarios

This appendix supplements the analysis presented in the paper with additional scenarios focusing on (1) the effects of multiple policies and (2) restrictions on future nuclear investments. In the five scenarios that follow we investigate different combinations of the carbon tax and carbon capture credit levels.

- Low Carbon Tax and Low Carbon Capture Credit Scenario (LTLC): Combination of the LTAX and LCRD scenarios.
- High Carbon Tax and Low Carbon Capture Credit Scenario (HTLC): Combination of the HTAX and LCRD scenarios.
- Low Carbon Tax and High Carbon Capture Credit Scenario (LTHC): Combination of the LTAX and HCRD scenarios.
- High Carbon Tax and High Carbon Capture Credit Scenario (HTHC): Combination of the HTAX and HCRD scenarios.
- Increasing Carbon Tax and Decreasing Carbon Capture Credit Scenario (ITDC): Combination of the ITAX and DCRD scenarios.

Several of the scenarios presented in the paper led to investment in nuclear power over the planning horizon. Except for the NSPT scenario, those investments were usually small-scale compared to the investments in other technologies. It is important to acknowledge, though, that the future of nuclear power generation is not solely governed by cost dynamics, but has a social and political component that our model does not capture. To observe how our model results change when new nuclear investments are not allowed, we consider the following three scenarios.

- Reference Scenario without Nuclear (NN_REFR): All parameters are the same as the REFR scenario, except nuclear investments are set to zero.
- High Carbon Tax Scenario without Nuclear (NN_HTAX): All parameters are the same as the HTAX scenario, except nuclear investments are set to zero.
- No Strategic Pricing with High Carbon Tax Scenario without Nuclear (NN_NSPT): All parameters are the same as the NSPT scenario, except nuclear investments are set to zero, and natural gas prices are set to NN_REFR levels.

To allow the model to compensate for decreased capacity due to the absence of the nuclear power option, we increase the 15 GW per time period investment cap in each of the other fuels to 17.5 GW under no nuclear investment scenarios.

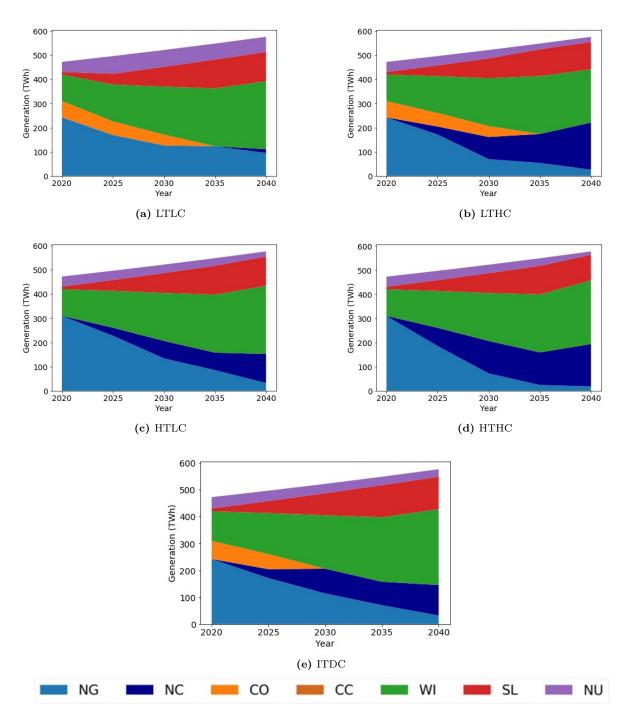
2.1. Multiple policy scenarios

Figure 1 illustrates the combined effects of the electricity generation mixes when carbon tax scenarios and carbon capture credit scenarios are coupled. Comparing the LTLC scenario (Figure 1a) with its component scenarios, i.e., LTAX and LCRD discussed in the paper, we observe that the combined behavior in the presence of both a low tax and a low credit is akin to the LTAX scenario, yet it yields slightly different results. In particular, the coupled policies bring natural gas with CCS in the generation mix in 2040. Although the percentage of this technology is very small in the overall mix, it is still significant because no natural gas with CCS was available in the LCRD scenario. Therefore, whereas a low credit was not sufficient to induce any electricity generation from sources with CCS, an additional low tax can help exceed this threshold.

The LTHC scenario (Figure 1b) realizes a very significant increase in natural gas with CCS, with this generation type amounting to 33.87% of the total electricity generation in 2040 as opposed to 18.78% when there is the same level of carbon capture credit but no carbon tax (HCRD). This difference comes at the expense of natural gas without CCS, which represents only 4.45% of power generation in 2040 under LTHC. The HTLC scenario results (Figure 1c) demonstrate that the effect of carbon tax supporting an otherwise insufficient carbon capture credit observed in the LTLC scenario can be quite powerful when the carbon tax level is high. Consequently, the share of natural gas with CCS in 2040 is 20.85% in the HTLC scenario, whereas it is non-existent in both of its component scenarios. This comes partially at the expense of solar, whose share of 26.58% in HTAX decreases to 20.98% in HTLC when a low credit is introduced. This is caused by the fact that the combined effects in HTLC is net positive for natural gas with CCS. Here, a \$80/MtCO₂ carbon tax is levied on the 10% CO₂ it emits, whereas a \$30/MtCO₂ carbon capture credit is given for the 90% it captures.

The HTHC scenario (Figure 1d) provides the most aggressive policies to replace carbon-intensive technologies with either their CCS versions or with renewables. By 2040, power generation from natural gas without CCS occurring drops sharply to only 3.17%. Finally, the ITDC scenario (Figure 1e) demonstrates a gradual replacement of coal with natural gas with CCS. It shares the common characteristic of favoring renewables and CCS over natural gas without CCS over the planning horizon with other scenarios that impose a substantial carbon tax.

So far, we mostly compared different multiple policy scenarios against their component scenarios; however, the capacity expansion results presented in Figure 2 enable us to better compare these scenarios among themselves. It is noteworthy that in the majority of these scenarios, except for LTHC and HTHC, the investment in wind is at its maximum, and the investment in solar is high in



 ${\bf Fig.~1.}$ Generation by technology for multiple policy scenarios

all cases. Despite these similarities, they show great variation in how much is invested in natural gasbased sources with and without CCS. This demonstrates how the pricing decisions of the strategically acting natural gas producer affect investments at the lower level.

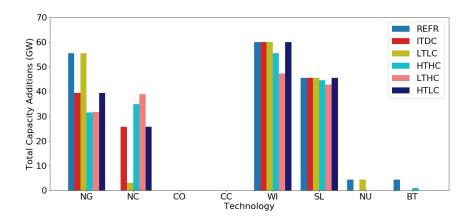


Fig. 2. Total capacity investments in different technologies for multiple policy scenarios.

The highest level of total investment in natural gas-based infrastructure occurs in LTHC scenario with 31.69 GW and 38.89 GW invested in no-CCS and CCS versions, respectively. The absence of a high carbon tax in this scenario leads to less renewable investments than in other scenarios, which is compensated by high natural gas investments. When the tax is increased in this scenario, we move to the HTHC scenario which realizes the lowest level of regular natural gas-based investments, and a lower level of natural gas with CCS investments than those of LTHC. This relationship of the lower tax version of the two scenarios with identical carbon capture credits having higher investments in both types of natural gas-based power plants, however, does not carry over when we fix the carbon capture credit at a low level and compare LTLC with HTLC. In this case, although the increased tax again leads to a lower level of regular natural gas investments, it leads to a higher level of capacity expansion in natural gas with CCS. This difference can be explained by the fact that unlike their high credit counterparts, both low credit scenarios induce the same level of renewable investments. Thus the low tax version of the same credit scenario does not have the deficit in capacity that is brought by lower renewable investments, which enables it to invest less in natural gas with CCS.

Optimal prices under the multiple policy scenarios are depicted in Figure 3. Here we observe

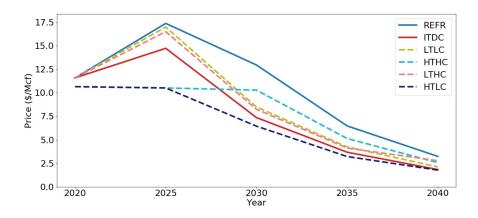


Fig. 3. Natural gas price charged by the producer in multiple policy scenarios.

the prices in all multiple policy scenarios are lower than the reference prices. Except for the HTHC scenario, prices also follow a general order by scenario, where LTLC and LTHC prices stay at similar levels with LTLC prices, usually only slightly higher (this changes in 2040), then followed by ITDC prices, which are followed by HTLC prices as the lowest. To analyze this order, it is important to outline the different effects our tested policies can have on prices.

In the paper, we establish that the natural gas producer takes advantage of the fact that the carbon tax affects coal more than it affects natural gas, and when this tax is high enough, it initially charges a lower price and maintains the low price to displace coal entirely from the generation mix. We again see this effect in both scenarios with high carbon tax, HTLC and HTHC. However, in the HTHC scenario, higher carbon capture credits enable the producer to maintain these initially lower prices longer without having to forego a significant portion of natural gas consumed by the electric utility. Since this effect is absent in low tax scenarios, i.e., LTLC and LTHC, their prices generally follow a higher trajectory compared to HTLC. ITDC, having a tax level higher than those of the low tax scenarios, but not high enough to entirely displace coal, follows a path in between.

Although we are able to explain the main behavior of the prices in multiple policy scenarios, it is important to elaborate on the price effects of different carbon capture values. Specifically observing the price behavior of LTLC and LTHC, we do not see a discernible difference between prices over the planning horizon. However, the outcomes of these two scenarios are vastly different. In fact, the LTHC scenario generates 30.05% more electricity from natural gas-based sources than the LTLC scenario. Therefore, although higher credit gives the natural gas player more room to increase prices, it chooses not to do so until the last time period to be able to induce more natural gas consumption at the lower level, which is the strategy that maximizes its revenue.

Figure 4 summarizes the upper and lower-level objective function values in each of the multiple policy scenarios and how they compare against the REFR scenario. We observe all scenarios incurring a higher total discounted cost for the lower-level electric utility than the REFR benchmark. Even in the LTHC scenario, which provides a high benefit for carbon capture and a low cost for tax, the lower-level cost is higher than REFR. This is due to the scale of carbon tax and carbon capture credit. Regular natural gas and coal infrastructure is already in the model in high capacities, and carbon emitting resources get used in significant amounts to satisfy demand, whereas infrastructure with CCS needs to be built from scratch, and its use does not reach a high enough level to offset the cost incurred by the carbon tax. The highest cost for the lower-level problem is realized in the HTLC scenario at \$482B, which is followed closely by HTHC at \$481B. When we look at producer revenues for these two cases, we see HTHC's total revenue is at \$247B, whereas HTLC's is at \$224B.

With respect to the upper-level revenues, we observe the highest values for HTHC. This is an interesting result given that we also have a scenario that provides the same carbon capture credit with less tax. It is important to recall the discussion in the paper describing how a higher carbon tax could be more beneficial for the natural gas producer than a low tax if a high tax is enough to make the natural gas player displace coal with tactical prices, whereas low tax is not. We observe the same outcome here, this time with an additional boost from high carbon capture credit, which brings the revenue even above that of REFR. The next two highest total revenues are realized in the LTHC scenario with \$229B, and in the HTLC scenario with \$224B, suggesting that high credit has a more powerful effect on revenues than high tax. The only scenario that leads to lower revenue than the REFR benchmark is LTLC, mostly because of the absence of a significant CCS investment in the lower level. In that case, revenue is mostly driven by the tax alone.

Figure 5 shows how the multiple policy scenarios affect CO₂ emissions. When we compare the

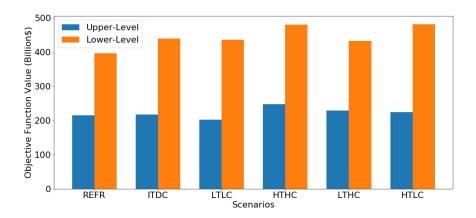


Fig. 4. Objective function values of upper and lower-level players in multiple policy scenarios.

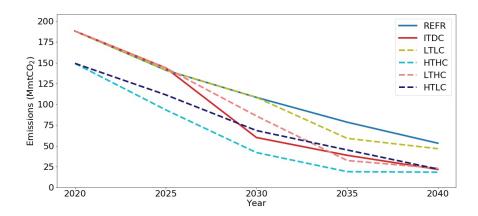


Fig. 5. CO₂ emissions due to the power sector in multiple policy scenarios.

point reached by each scenario in 2040 in terms of emissions, we observe that the scenarios are ordered by the level of emission-cutting values. Containing two high emissions-cutting policies, HTHC leads to the lowest emissions. Containing one high and one low emissions-cutting policy, HTLC and LTHC both end at a similar level above HTHC. With two low emissions-cutting scenarios, LTLC converges above HTLC and LTHC, and below the REFR benchmark. Meanwhile emissions under ITDC start relatively high in 2020, and end around HTHC and one-high, one-low scenarios in 2040. Since renewable development happens at a more steady pace in all scenarios, kinks in the emission graphs mostly represent either the phase-out of coal, or a sudden deployment of high natural gas with CCS.

2.2. No nuclear investment scenarios

Figure 6 compares the evolution of the generation mix of the three previously analyzed scenarios to their no-nuclear counterparts and Figure 7 illustrates how capacity investments change when new nuclear investments are disabled. Analyzing these two sets of graphs in tandem, we observe that in all cases, coal increases its share in the generation mix, even inducing additional investments in the NN_REFR scenario. As expected, we see a higher renewable investment resulting in higher shares of renewables in the generation mix when nuclear capacity is not expanded. Total shares of rewnewables in the generation mix in 2040 reach 80.01%, 87.33%, and 87.26% for NN_REFR, NN_HTAX, and NN_NSPT scenarios, respectively. Disabling nuclear investments has a negative effect on the share of

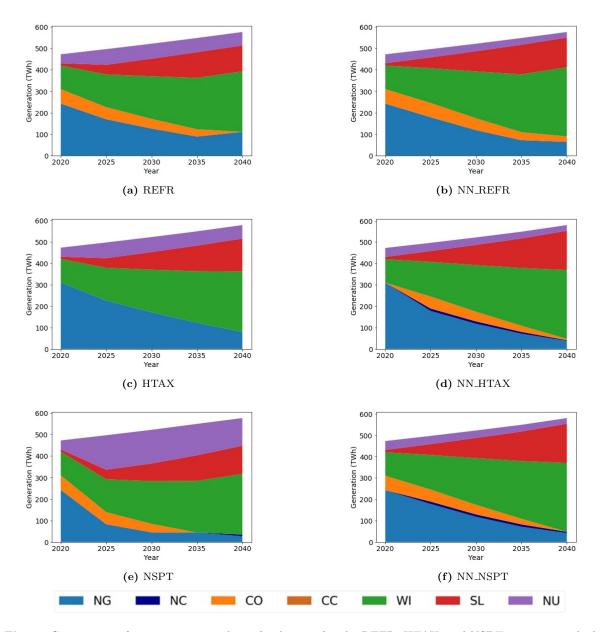


Fig. 6. Comparison of generation mixes by technology under the REFR, HTAX, and NSPT scenarios with their no-nuclear counterparts.

natural gas-based technologies in the REFR and HTAX scenarios, whereas we observe significantly higher natural gas use in the NN_NSPT scenario. Considering the fact that the NSPT scenario is the most nuclear-intensive scenario we investigated, this result makes sense as the large baseload provided by nuclear in that scenario cannot be replaced by renewables alone.

Figure 8 plots natural gas prices over the planning horizon for REFR, HTAX, and NSPT and their no-nuclear counterparts. From the figure, we see that disabling nuclear causes higher prices for both the NN_REFR and NN_HTAX scenarios. This means that the natural gas producer is able to take advantage of the utility's reduced ability to respond to natural gas prices. Disabling nuclear leaves fossil sources as the only non-intermittent generation options that can be invested in. The natural gas producer takes advantage of this fact, which is further beneficial because of the high costs associated with generating electricity from coal, especially under a carbon tax. By being able to charge much higher prices for natural gas, the producer offsets the revenue loss by decreased natural gas use in the NN_REFR and NN_HTAX scenarios. Thus, no-nuclear versions of each scenario lead to higher total

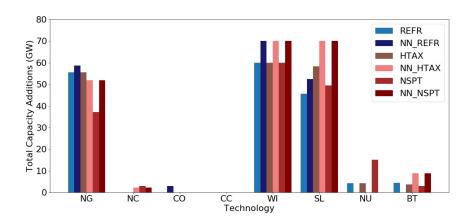


Fig. 7. Comparison of total capacity investments in different technologies under the REFR, HTAX, and NSPT scenarios with their no-nuclear counterparts.

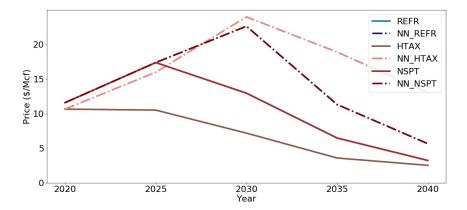


Fig. 8. Comparison of natural gas price charged by the producer under the REFR, HTAX, and NSPT scenarios with their no-nuclear counterparts.

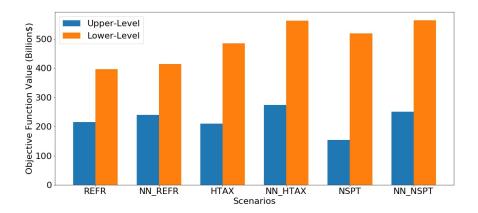


Fig. 9. Comparison of objective function values of upper and lower-level players under the REFR, HTAX, and NSPT scenarios with their no-nuclear counterparts.

revenues for the producer, and higher total costs for the utility, as evidenced in Figure 9.

Lastly, Figure 10 depicts how CO₂ emissions are affected by disabling investments in nuclear. Although the share of renewables increases under no-nuclear scenarios, total CO₂ emissions in all three scenarios generally increase over the planning horizon except for NN_HTAX in 2040 and NN_REFR in 2035. This is mainly caused by an increase in coal usage when nuclear investments are restricted.

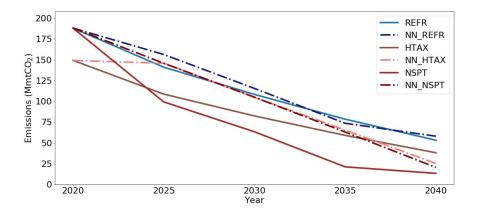


Fig. 10. Comparison of CO₂ emissions due to the power sector under the REFR, HTAX, and NSPT scenarios with their no-nuclear counterparts.

References

Garcia-Herreros, P., Zhang, L., Misra, P., Arslan, E., Mehta, S., Grossmann, I.E., 2016. Mixed-integer bilevel optimization for capacity planning with rational markets. Computers & Chemical Engineering 86, 33–47.

Guo, C., Bodur, M., Papageorgiou, D.J., 2022. Generation expansion planning with revenue adequacy constraints. Computers & Operations Research 142, 105736.

Gurobi Optimization, 2020. Non-Convex Quadratic Optimization. Available at: https://www.gurobi.com/resource/non-convex-quadratic-optimization/, Last accessed: 3/19/2022.

Jayadev, G., Leibowicz, B.D., Bard, J., Çalcı, B., i2022. Strategic interactions between liquefied natural gas and domestic gas markets: A bilevel model. Computers & Operations Research 144, 105807.

- McCormick, G.P., 1976. Computability of global solutions to factorable nonconvex programs: Part I—Convex underestimating problems. Mathematical Programming 10, 147–175.
- Zeng, B., Dong, H., Xu, F., Zeng, M., 2020. Bilevel programming approach for optimal planning design of EV charging station. IEEE Transactions on Industry Applications 56(3), 2314–2323.
- Zhao, C., Wan, C., Song, Y., 2019. An adaptive bilevel programming model for nonparametric prediction intervals of wind power generation. IEEE Transactions on Power Systems 35(1), 424–439.