Not for Publication in Journal

Online appendix for

Hedging Strategies: Electricity Investment Decisions under Policy Uncertainty

Jennifer Morris[*], Vivek Srikrishnan,[**] Mort Webster[**] and John Reilly[***]

[*] Corresponding Author. Joint Program on the Science and Policy of Global Change, Massachusetts Institute of Technology, Cambridge, MA, USA. Email: holak@mit.edu.

[**] John and Willie Leone Family Department of Energy and Mineral Engineering, Pennsylvania State University, State College, PA, USA.

[***] Joint Program on the Science and Policy of Global Change, Massachusetts Institute of Technology, Cambridge, MA, USA.

Appendix A: Exact Dynamic Programming Results

Numerical implementation of the exact DP requires discretizing the continuous decision space. The non-carbon generation's share of new electricity production is discretized into steps of 5% over the range of 0% to 50% in Stage 1, and over 0% to 80% in Stage 2.[1] The Stage 1 electricity emissions reduction decision is implemented as the maximum Stage 1 emissions growth rate allowed in each CGE period, from 0.8 to 1.2 in steps of 0.05. Applying a given growth rate to the 2010 base year emissions creates an emissions cap that is imposed in 2015, and applying the growth rate to the 2015 emissions cap determines a cap that is imposed in 2020. Those "self-imposed" caps are incorporated as emissions quantity constraints within the CGE model.

For the exact DP solution, we use a discrete approximation of the continuous probability distribution over future emissions caps. Specifically, a discrete three-point probability distribution over three emissions cap scenarios is assumed. The emissions cap scenarios are: (1) no emissions cap, (2) an emissions cap of 20% below the cumulative no policy emissions ("-20% Cap"), and (3) an emissions cap of 40% below the cumulative no policy emissions ("-40% Cap"). Each of these scenarios is assigned an associated probability, which collectively sum to one.

The procedure for solving the DP-CGE model consists of two steps. First, the CGE model simulates each possible path (each combination of decisions and uncertain outcomes) over both stages, calculating the total consumption for each stage. The decisions and uncertain outcomes are exogenously imposed on the CGE model, which then endogenously chooses all other output quantities, including the shares of natural gas and coal generation. Second, backward induction is performed using the consumption (social welfare) values for each decision stage and the probabilities of the respective uncertain outcomes.

The focus of the analysis is on the optimal near-term decisions about new electricity investment shares and emissions reductions under different scenarios of uncertainty in the future emissions limit. Table A1 presents the optimal Stage 1 decisions under several alternative scenarios.

As a basis of comparison, we first consider the results under perfect information about the future emissions cap; i.e., knowing which cap will exist in Stage 2, what should we do in Stage 1? The three deterministic scenarios in Table A1 are: (1) no emissions cap (D1), (2) a cap requiring a 20% reduction of cumulative emissions (D2), and (3) a cap requiring a 40% reduction of cumulative emissions (D3). The nomenclature of the probability distributions in Table A1 follows the format: .XX No Cap = probability of no emissions cap, .XX -20% Cap = probability of cap with 20% reduction, and .XX -40% Cap = probability of cap with 40% reduction.

[1] These ranges were chosen after extensive testing of the model under a wide range of scenarios found that Stage 1 non-carbon share never rose above 50% and Stage 2 never rose above 80%.

Table A1: Optimal Stage 1 Decisions for Deterministic and Four Illustrative Uncertain Scenarios

			Stage 1 Decisions			
			Share of New Investment			
	Policy Scenario	Probability Distribution	Non-carbon	Coal	Gas	Emissions Reductions
Deterministic	D1	1.00 No Cap 0.00 -20% Cap 0.00 -40% Cap	0%	37%	63%	0%
	D2	0.00 No Cap 1.00 -20% Cap 0.00 -40% Cap	5%	30%	65%	6%
	D3	0.00 No Cap 0.00 -20% Cap 1.00 -40% Cap	35%	14%	51%	24%
Uncertain	U1	0.33 No Cap 0.33 -20% Cap 0.33 -40% Cap	20%	17%	63%	18%
	U2	0.10 No Cap 0.40 -20% Cap 0.50 -40% Cap	25%	19%	56%	18%
	U3	0.50 No Cap 0.40 -20% Cap 0.10 -40% Cap	5%	20%	75%	13%
	U4	0.40 No Cap 0.10 -20% Cap 0.50 -40% Cap	25%	19%	56%	18%

If it is known that there will be no cumulative limit on emissions, there is no reason to invest in higher-cost non-carbon technologies or to reduce emissions in Stage 1. If there will be a -20% cap with certainty, it is optimal to invest in some non-carbon electricity production (5%) and to reduce emissions (by 6% relative to reference emissions) in Stage 1 in anticipation of the Stage 2 policy. Given the parameterization of the fixed factor resource constraint, a 5% non-carbon investment share is the minimum amount required for this constraint to be non-binding on the rate of technology expansion in the following period, thereby providing flexibility to expand non-carbon penetration to the desired amount in Stage 2 (assuming no other constraint on the growth of the non-carbon share[2]). If there will be a -40% cap in Stage 2 with certainty, it is optimal in Stage 1 to invest more aggressively in non-carbon electricity production (35%) and to more aggressively reduce emissions (by 24%). The results of the two deterministic scenarios with emissions caps are simply the solutions to the optimal timing problem under perfect information; the solution is the welfare-maximizing allocation of emissions reductions and substitution to non-carbon electricity production in both stages to meet each cumulative target.

It is important to note that the resulting emissions from the electric sector are the net effect of both the new electricity investment decisions (i.e., the relative share of new electricity production from non-carbon sources) and the decisions about how to operate existing (vintaged) electricity capital. At the margin, emissions can be reduced by shifting another unit of capital to non-carbon from conventional, or by substituting gas for coal as an input to conventional electricity production, or by changing the operation of vintage capacity (substitution between vintages or reducing output).

2 For the exact DP, we assume that the maximum rate of non-carbon growth between stages allows the share of non-carbon in new investment to increase by no more than 50 percentage points (0.5) from Stage 1 to Stage 2 (for example, if 20% of investment is in the non-carbon technology in Stage 1, then the max non-carbon share in Stage 2 is 70%). This growth limit is explored with the ADP approach in Section 4.1.2.

For the deterministic cases, the resulting optimal Stage 2 decision is a constant. For both cap scenarios, the majority of the required emissions reductions take place in Stage 2 (87% for the -20% cap and 76% for the -40% cap). Greater emissions reductions take place in Stage 1 under the -40% cap, relative to the other scenarios, to help distribute the cost of the cap over time, as has been shown by other studies focusing on the intertemporally optimal emissions path (e.g., Manne & Richels 1995; Wigley *et al.*, 1996; Bosetti *et al.*, 2009). These studies show that pathways involving modest reductions in early years followed by sharper reductions later on are most cost-effective, and that the more stringent the cumulative emission reduction goal, the more near-term reductions are optimal.

The results for four illustrative uncertainty scenarios are also presented in Table A1, each defined by a probability distribution over the three emissions caps. The decision problem is to choose the Stage 1 non-carbon share of new electricity production and Stage 1 emissions reductions conditional on the given probability distribution of the cumulative emissions limit. In the absence of an objective source for this policy uncertainty, the probabilities assigned to future emission caps represent beliefs about how likely each cap is. The results of interest are the optimal Stage 1 decisions. The optimal Stage 2 decisions depend on which cap is realized, and therefore there are three different possible sets of optimal decisions for Stage 2, each conditional on one of the emissions caps being realized. The explicit consideration of uncertainty in the future emissions cap results in a near-term hedging strategy in terms of the amount of non-carbon electricity production (and therefore the amount of capital stock allocated to non-carbon electricity) and the level of emissions reductions.

In the first illustrative uncertainty scenario (U1), in which each cap is assumed to have equal (1/3) probability, the optimal Stage 1 decision is that 20% of the new electricity production should be non-carbon. This production share is distinct from the optimal Stage 1 share in any of the three deterministic scenarios. Nor is this relative share the average or a linear combination of the results from the three deterministic scenarios. The resulting optimal strategy can only be solved for using the explicit stochastic formulation.

The Stage 1 optimal non-carbon production share depends on the assumed probability distribution over future emissions caps. In the second uncertain scenario (U2), in which the "no cap" scenario is less likely, the optimal Stage 1 share for non-carbon electricity is 25%. A 25% share of non-carbon generation in electricity is also optimal in the fourth uncertain scenario (U4), in which "no cap" and the -40% cap are both more likely than the -20% cap. In that case, the higher probability of the tighter cap justifies more near-term investment in the non-carbon technology. In contrast, when "no cap" and the -20% cap are the most likely outcomes (U3), only a 5% production share of non-carbon electricity is optimal in Stage 1.

The Stage 1 decision about emissions reductions similarly depends on the probability distribution of future policies. Under uncertainty, the optimal strategy is to reduce emissions more than in the deterministic -20% cap scenario (the "middle" case), but not as much as in the deterministic -40% cap scenario (Table A1). In scenario U3, with a lower probability on the -40% cap, the Stage 1 cumulative emissions are reduced by 13% relative to reference. In the other uncertainty scenarios shown here, an 18% reduction in Stage 1 is optimal.

The illustrative scenarios described above demonstrate that the optimal Stage 1 decisions depend critically on the beliefs about the relative likelihood of future emissions caps. In general, the optimal first stage strategy under uncertainty is to allocate a greater share of capital and labor to non-carbon electricity and to reduce emissions more than is optimal under the deterministic expected value scenario. In other words, the solution to the "average case" would under-invest in non-carbon generation and would emit too much carbon.

Systematically exploring the space of possible 3-point probability distributions over future emissions caps, we find that it is rarely optimal to produce no electricity from non-carbon sources in Stage 1. Only when there is zero probability of the -40% cap and a low probability (20% or less) of the -20% cap is it optimal to exclusively build fossil-based electricity capacity in the near-term. In all other scenarios, it is optimal to produce at least 5% of the new electricity from the non-carbon technology, and to shift the inputs for conventional production from coal to natural gas. As explained above, the 5% non-carbon production share is sufficient to make the fixed factor constraint non-binding in Stage 2. Even with a relatively low probability of the -40% cap, this investment is a cost-effective hedge against the uncertainty in the future cap. If there is at least a 20% probability of a -40% cap, then it is optimal to produce 20% of new electricity from the non-carbon technology. Similarly, unless the probability of no cap in Stage 2 is zero, Stage 1 emissions should be reduced by 6% or more relative to reference. If there is at least a 20% probability of the -40% cap, then it is optimal to reduce Stage 1 emissions by 18% or more.

Appendix B: Approximate Dynamic Programming Algorithm

There are two broad classes of ADP algorithms: (1) iterative approaches, which are based on value iteration or policy iteration (examples include TD-learning and Q-learning), and (2) linear programming based approaches (Schweitzer and Seidman, 1985; de Farias and Van Roy, 2003). Here, we present an ADP solution method that combines the strengths of both classes of algorithms, starting with the linear programming approach. A well-known approach for solving an exact DP problem is to use linear programming (Manne, 1960; De Ghellink, 1960). Solving the DP described in Section 3 is equivalent to solving the following linear program:

$$\min c'J$$
$$\text{Subject to: } E_\xi[g(x,u,\xi)+\gamma J(f(x,u,\xi))] \le J(x) \quad \forall x \in X, u \in U(x) \tag{B1}$$

where x is the state vector, u is the control vector, ξ is the stochastic process, γ is the discount factor, f is the system equation of motion, and c is vector of state-relevance weights (positive). The main idea is that if J is a feasible solution, then $J \le J^*$.

An approximate version of this can be constructed by approximating the cost-to-go function J with a set of basis functions Φ and coefficients r in a smaller dimensional space than the original full state space (Schweitzer and Seidman, 1985; de Farias and Van Roy, 2003).

$$J(x) = \sum_i \Phi_i(x) r_i = \Phi r \tag{B2}$$

The vector of coefficients r can be found by solving:

$$\min c'\Phi r$$
$$\text{Subject to: } E_\omega\left[g(x,u,\omega)+\alpha\Phi\left(f(x,u,\omega)\right)r\right] \le \Phi(x)r \quad \forall x \in X, u \in U(x) \tag{B3}$$

In this approximation, there are only $R \cdot I$ variables (number of basis functions and number of stages). However, the number of constraints remains the same as in (B1) (one constraint for each state-control pair).

A further modification was presented in de Farias and Van Roy (2004), in which the constraints to be included in the LP are sampled to solve:

$$\min c'\Phi r$$
$$\text{Subject to: } E_\omega\left[g(x,u,\omega)+\alpha\Phi\left(f(x,u,\omega)\right)r\right] \le \Phi(x)r \ \ \forall (x,u)_k \sim P(X,U(x)) \tag{B4}$$

The main drawback of the formulation presented in equation (B4) is that the convergence results proven in de Farias and Van Roy (2004) are based on an idealized choice of the probability distribution used to sample the constraints; specifically, this choice assumes knowledge of an optimal policy.

Here, we demonstrate a more efficient and generalizable solution algorithm, building on the approach of de Farias and Van Roy. Specifically, we borrow the main advantage of the value-iteration class of approaches (e.g., Powell, 2011), which iteratively improves the approximation of the cost-to-go function by using the previous approximation to inform the next set of state-action pairs to sample. To implement this iterative approach, we make two additional modifications. First, rather than the cost-to-go function, which is a function of the state x, we use Q-factors from Q-Learning algorithms (Bertsekas and Tsitsiklis, 1996), which are the cost of each state-action pair.

The second modification is to use a Metropolis Hasting algorithm (Rubinstein and Kroese, 2008) that accepts or rejects a state-control pair sample based on how promising the sample is according to the Q-factor estimate. To do that, we sample uniformly across the state-action space to determine a candidate state-action pair (x_i', u_i'), we compute the value of its associated Q-factor $(\Phi_i(x_i', u_i')r_i)$, and compute the ratio of this Q-factor relative to the range of possible Q-factor values in the latest iteration $(\underline{\Phi_i(x_i,u_i)r_i}, \overline{\Phi_i(x_i,u_i)r_i})$:

$$\overline{\phi} = \frac{\Phi_i(x_i', u_i')r_i - \underline{\Phi_i(x_i, u_i)r_i}}{\overline{\Phi_i(x_i, u_i)r_i} - \underline{\Phi_i(x_i, u_i)r_i}} \tag{B5}$$

We then draw a sample $\overline{\xi}$ from a probability distribution Ξ_t. We accept the state-action pair if $\overline{\phi} \geq \overline{\xi}$. To ensure convergence to the optimal solution, the variance of the probability distribution Ξ_t must decrease with t, the probability of accepting any sample must be strictly positive, and the probability of accepting any sample with associated $\overline{\phi} \geq 1$ must be 1. To ensure that these conditions hold, we determine the probability of accepting a proposed sample based on a sample drawn from $\Xi_t \sim U(-\frac{1}{t}, 1)$. In other words, we accept a sample with probability $\alpha = \min\left(\max(\overline{\phi}, 0) + \frac{1}{t}, 1\right)$. The performance of the algorithm improves when the state-control pairs associated with binding constraints in the previous iteration are retained in future iterations. We call this algorithm a *Q*-factor Adaptive Relaxed Linear Problem (QARLP) algorithm.

Table B1 shows the performance of the QARLP algorithm as applied to our problem, and compares the results with those from an exact DP approach at different resolutions. For this table, the ADP approach uses a continuous uniform distribution for future climate policy discretized by the unit interval while the DP uses a discrete three-point distribution in which each of the three policies highlighted in Appendix A are equally likely. Non-carbon growth limits of both 0.5 and 0.3 were tested. The results illustrate the curse of dimensionality associated with DP. Indeed, the runtime for the DP with 0.01 grid resolution (matching the precision of the QARLP results) and the 0.3 non-carbon growth limit was over seven times that of the QARLP method, and with a 0.5 non-carbon growth limit the DP at this resolution failed to converge within a 24 hour walltime limit. While the absolute value error from the 0.05 grid resolutions appears small (and the runtime is significantly less than that for QARLP), we would expect both the runtime and the error to increase as the complexity of the problem increases (e.g., through an increase in the number of stages or the number of decision variables and exogenous shocks). A further illustration of the performance of the algorithm is given in Figure B1, which shows the convergence of the maximum Q-Value (i.e., the optimal consumption) as a function of iterations. Note that the algorithm converges within 50 iterations; the remaining oscillations are due to the random sampling nature of the algorithm.

Table B1. QARLP vs. DP

Grid Size	Growth Limit	DP Results		QARLP Results	
		Value ($B)	Walltime (min)	Value ($B)	Walltime (min)
0.10		91758	0.75		
0.05	0.3	91781	2.16	91787	149.6
0.01		91805	678.71		
0.10		91772	0.79		
0.05	0.5	91788	2.43	91800	256.6
0.01		-	> 1440		

Note: Results in Section 3.2 utilized a 0.05 grid size and a non-carbon growth limit of 0.5.

Figure B1. Convergence Results: The maximum Q-Value (the optimal consumption) as a function of iteration.