Supporting Information:

Analysis of Carbon Capture Strategies for Refineries Considering Extraordinary Future Economic and Policy Uncertainties

Fang Li[a], Liang Jing[b], Sean McCoy[a], Sara Hastings-Simon[c] and Joule Bergerson*[1a]

[a] Department of Chemical and Petroleum Engineering, University of Calgary, 2500 University Drive NW, Calgary, Alberta, Canada T2N 1N4

[b] Energy Traceability Technology, Technology Strategy & Planning Department, Saudi Aramco, Dhahran, Saudi Arabia 34465

[c] Department of Earth, Energy, and Environment, University of Calgary, 2500 University Drive NW, Calgary, Alberta, Canada T2N 1N4

*corresponding author: jbergers@ucalgary.ca

Table of Contents

List of Figures

List of Tables

Supplementary Method 1 – PRELIM and Carbon Capture Module

PRELIM Introduction

The Petroleum Refinery Lifecycle Inventory Model (PRELIM) is a Microsoft Excel based inventory model to evaluate the gate-to-gate environmental impact of refineries' operations. PRELIM version 1.6 (1) is employed in this study to evaluate the energy consumption and greenhouse gas (GHG) emissions of the U.S. refining sector. PRELIM v1.6 is currently capable of simulating three types of refineries: hydroskimming, medium conversion, and deep conversion refineries. For the medium conversion and deep conversion refineries, three combinations of process units are available for each. In contrast, only one combination is available for the hydroskimming refinery, as detailed in Table S1

PRELIM considers different types of crude that a refinery process, which influences the throughputs of various process units, thereby affecting the energy consumption and GHG emissions of the entire refinery. PRELIM necessitates a specific format for the crude assay, which must include detailed information such as crude qualities (e.g., API (American Petroleum Institute) gravity, sulfur content by weight percentage), and the distribution of crude components based on distillation temperature (i.e., true boiling point (TBP) distribution). Currently, PRELIM v1.6 stores over 600 crude assays from oilfields across the world based on public or industrial data.

The production of a variety of refinery products (17 products) could be simulated by PRELIM v1.6. However, only a subset of these products is simulated in this work due to the lack of data, including blended gasoline, jet fuel or aviation turbine fuel (Jet-A/AVT), ultra-low sulfur diesel (ULSD), fuel oil, coke, liquid heavy ends (LHE), liquified petroleum gas (LPG) and lubricants (Lube).

By selecting the refinery configuration, inputting crude assays, and specifying target product slates (i.e., the composition of different refinery products), PRELIM adjusts the material flows between process units. It aligns actual product slates with target slates using an embedded linear programming optimization tool. This approach enables the derivation of an energy consumption profile and GHG emissions that more accurately reflect real refinery operations.

Carbon Capture Module

In this study, a carbon capture module based on PRELIM framework is further developed to evaluate the GHG emissions reduction potential and the economic performance for capturing CO_2 from various refinery's process units. A list of individual and combination of process units that are considered for capture is listed in Table S17.

The carbon capture module applies a point-to-point post-combustion capture strategy, which includes individual absorbers that are assigned to those individual furnaces/boilers, a common regenerator and a common compressor. The captured CO_2 is highly purified (i.e., 99.5%) and

compressed (i.e., 2220 psia), making it ready for transportation. However, CO_2 transportation and storage phase are not included in the carbon capture module, as shown in Figure S1.

CO_2 concentrations in different exhausted gas streams are calculated based on fuel composition, which are either extracted from literature (2–9) (i.e., natural gas, internal liquid fuel, coke, tail gas from hydrogen purification unit – pressure swing adsorption (PSA) - of steam methane reformer (SMR)) or fed with industrial data (i.e., refinery fuel gas) and stored in the carbon capture module. The specific electricity and steam consumption for capturing CO_2 from different gas streams with different CO_2 concentrations are presented in Table S14.

The carbon capture module considers a multi-stage centrifugal CO_2 compression with inter-stage cooling for dehydration. For each stage of compression, the polytropic head for CO_2 compression is calculated based on the equation Eq. S1 extracted from the literature (10).

$$W^{real\ gas} = H_R^{real\ gas} = \frac{H_p^{real\ gas}}{\eta_p} \approx \left[\frac{8.314}{MolWt}\right]\left[\frac{n}{n-1}\right]\frac{T_s Z_{avg}}{\eta_p}\left[\left(\frac{P_d}{P_s}\right)^{\frac{n-1}{n}} - 1\right] \quad \text{Eq. S1}$$

Where, $W^{real\ gas}$ is the polytropic work for compressing the CO_2 gas, $H_R^{real\ gas}$ is the actual head required for CO_2 compression, $H_p^{real\ gas}$ is the polytropic head required for CO_2 compression, η_p is the polytropic efficiency, $MolWt$ is the molecular weight of CO_2, *n* is the polytropic exponent, T_s is the temperature at the compressor inlet, Z_{avg} is the average of the gas compressibility at the suction and discharge conditions, P_d and P_s are the pressure of CO_2 at discharge and suction conditions.

The polytropic exponent could be calculated by Eq. S2 as follows:

$$n = \frac{ln\left(\frac{P_d}{P_s}\right)}{ln\left(\frac{Z_s T_s P_d}{Z_d T_d P_s}\right)} \quad \text{Eq. S2}$$

Where, Z_s and Z_d are the gas compressibilities at the suction and discharge conditions, T_s and T_d are the gas temperatures at the suction and discharge conditions.

Table S18 summarizes all the parameters and the calculated polytropic work for CO_2 compression in each stage.

General Computing Process for the Carbon Capture Module

Figure S10 depicts how the carbon capture module calculating process is integrated into PRELIM. A detailed description of the computing process is presented below.

Step 1: PRELIM will first calculate the GHG emissions and energy consumption baseline for individual process units to adapt the designated product slate.

Step 2: Deciding if carbon capture is deployed for individual units. If not, the emission/energy baseline will be used to calculate the final results for the entire refinery process & products.

Step 3: If carbon capture is deployed for selected units, the carbon capture module then calculates a new emissions profile for each unit based on the CO_2 content in the flue gas flow. Moreover, the energy consumption for the carbon capture & compression unit will be calculated based on the total amount of CO_2 captured in the process. The energy consumption for capturing and compressing CO_2 is allocated to each unit on a mass basis.

Step 4: Once the total emissions and energy consumption is determined for the entire refinery process, new emissions factors are generated for different energy sources (such as natural gas heat, refinery fuel gas heat, steam, etc.) and units (including fluid catalytic cracker (FCC), residue fluid catalytic cracker (RFCC), SMR, etc.). The newly generated values replace the default emissions factors and energy consumption in PRELIM for energy sources and individual units, and the emissions intensity for refinery processes and products is thus determined.

Supplementary Method 2 - Refinery Product Slates, Crude Proxies, Configurational Crude Input and Cost Analysis under Different Future Transportation Fuel Demand Scenarios

Refinery Product Slates

The National Renewable Energy Laboratory's Electrification Future Study (NREL's EFS)(11) has modelled the future transportation fuel demands for the U.S. based on different EV adoption rate scenarios (see detailed description for each scenario in Table S10 and the estimated amount of transportation fuel in Figure S11). We selected 2019 as our baseline year and two other specific years – 2035 and 2050 – as critical timestamps.

In order to estimate the future production of refinery products in the U.S, it is essential to consider the anticipated fuel demand for the transportation sector. A computation procedure similar to Kim (12) is established to formulate the expected refinery product slates, which is detailed stepwise below.

Step 1. Determine the future demand for selected refinery products based on EFS. As mentioned in the previous section (Supplementary method 1), a subset of refinery products is considered in this study, including blended gasoline, Jet-A/AVT, ULSD, fuel oil, coke, LHE, LPG and Lube. EFS provided future demands data for multiple refinery products, including diesel, gasoline, jet fuel, LPG, lubricants and residual fuel oil from the transportation sector. However, these products are not necessarily all consumed by the transportation sector. For example, 77% diesel produced by the refining sector is consumed by transportation sector, the rest is either consumed by industrial sector (14%) or residential, commercial, power generation purposes (9%)(13). Assuming the consumption proportions for each transportation fuel (i.e., gasoline, jet fuel, diesel) remain constant, we can readily estimate the total future supply of these fuels in the U.S. market. Additionally, fuel oil (i.e., heating oil) and diesel, both categorized as distillate fuel oils, are closely related. The EIA has documented the historical ratio of total fuel oil supplied to total diesel supplied, allowing for an estimation of total supplied fuel oil once the diesel figures are known. Likewise, LHE, also known

as residual fuel oil, predominantly used by marine vessels and not suitable for direct combustion in vehicles or jet engines, has its consumption in the transportation sector documented historically by the EIA. This data facilitates the estimation of the total LHE supplied in the U.S. market. Although coke is not directly modeled by EFS, it is considered a critical part of the product slate as it represents the heavier fractions of crude that cannot be converted into transportation fuels. The total supplied coke can be calculated from its average proportion of total refinery products over time, a figure also documented by the EIA. The historical contribution of fuel to supply listed above is summarized and presented in Table S2.

Step 2. Determine the amount of refinery products that are produced by local refinery. It is apparent that the U.S. domestic fuel market does not follow a localization pattern – understood here is the practice of producing goods locally and then consuming them within the same geographic area. For example, there are 1.22 Gbbl/y of gasoline supplied to PADD 1 in 2019 while there is only 95 mmbbl/y of gasoline that are locally produced in the same area. To determine the product slate of a refinery for future transportation fuel demand scenarios, it is essential to ascertain the quantity of locally produced refinery products. The amount of a specific fuel produced locally should equate to the total supply of that fuel, adjusted by the net movement (i.e., exports minus imports) and any change in stock levels. To incorporate historical trends in local production, we assume that the arithmetic average of the ratio of locally produced fuel volume to total fuel supply over a given time interval will remain constant in the future. Therefore, the volume of locally produced specific fuel can be calculated based on this assumption, as shown in Eq. S3 and Eq. S4. Table S3 summarizes the proportion of locally produced refinery products to total fuel supplies.

$$\varphi_{local\ produced} = \frac{1}{m}\sum_{1}^{m}\left[\frac{\left(Q_{total\ supplied}^{\ hist} - Q_{import}^{\ hist} + Q_{export}^{\ hist} - Q_{stock\ change}^{\ hist}\right)}{Q_{total\ supplied}^{\ hist}}\right] \quad \text{Eq. S3}$$

$$Q_{local\ produced}^{\ future} = Q_{total\ supplied}^{\ future} \times \varphi_{local\ produced} \quad \text{Eq. S4}$$

Where, $\varphi_{local\ produced}$ is the arithmetic mean value of the historical contribution of locally produced fuel to total fuel supplies, $Q_{total\ supplied}^{\ hist}$, $Q_{import}^{\ hist}$, $Q_{export}^{\ hist}$ and $Q_{stock\ change}^{\ hist}$ are EIA reported fuel volume according to the applications on an annual basis, *n* is the total number of years being considered, $Q_{local\ produced}^{\ future}$ denotes the future refinery produced transportation fuels by volume, $Q_{total\ supplied}^{\ future}$ denotes the total supplied transportation fuels by volume.

Table S3 summarizes the calculated $\varphi_{local\ produced}$ based on EIA's data.

Step 3. Repeat step 1 and 2 for other regions.

Blending Crude Proxies.

PRELIM requires user input on what type of crude that refineries will be processing in order to determine the energy consumption and GHG emissions. And there are obvious regional differences

in terms of crude qualities across U.S. For example, the API gravities of crudes process by different PADDs ranges from 28.6 to 35.5 while the average API for entire U.S. is 32.9. Cooney et al.(14) developed a linear programming optimization method to assemble a crude proxy that matches the projected API gravity and sulfur content by blending different crude assays stored in PRELIM inventory. The same method is applied to formulate crude proxies that will be used in this study. The procedure is described below.

Step 1. A business-as-usual (BAU) scenario for future crude qualities in each PADDs is extrapolated using the exponential triple smoothing (ETS) forecasting function embedded in Microsoft Excel, based on historical crude quality data reported by EIA(15). Two distinct scenarios are contemplated for future projections: heavy versus light crude in relation to API gravity and sweet versus sour crude concerning sulfur content, to assess potential variances in GHG emissions resulting from the processing of different types of crude. The upper and lower bounds for the ETS according to a chosen 95% projection interval are employed for these bifurcated scenarios. The light crude scenario employs the upper bound of the fitted API gravity, while the sweet crude scenario incorporates the lower bound of the fitted sulfur content. The opposite is true for the heavy and sour crude scenarios. The projected API gravity & sulfur content for these two scenarios will be served as targeted API and targeted sulfur in following steps.

Step 2. The U.S. refineries blend their own crude slates using both domestic and imported crudes. The volume fraction as well as the qualities of imported crudes from different countries are documented by EIA(16). Country level imported crude proxies could be assembled using the objective function and the constraints developed by Cooney et al. (14), as shown in Eq. S5.

$$Minimize{:}f_0(\theta_i) = \left|\frac{calculatedAPI}{targetAPI}\right| \times 3.29 + \left|\frac{calculatedSul}{targetSul} - 1\right| \quad \text{Eq. S5}$$

Where θ_i is the volume fraction of crude assay *i* from a specific country. $targetAPI$ and $targetSul$ are projected API gravity and sulfur content from ETS, respectively. $calculatedAPI$ and $calculatedSul$ are volume weighted aggregated API gravity and sulfur content determined by individual crude assays *i* and its volume fraction θ_i, as shown in Eq. S6.

$$\begin{cases} calculatedAPI = \dfrac{141.5}{\sum_i (SpGr_{crude\ i} \times \theta_i)} - 131.5 \\ calculatedSul = \sum_i (Sul_{crude\ i} \times \theta_i) \end{cases} \quad \text{Eq. S6}$$

Two constraints are applied for the optimization, as shown in Eq. S7.

$$\begin{cases} \theta_i \le 1 \\ \sum_i \theta_i = 1 \end{cases} \quad \text{Eq. S7}$$

Where, $SpGr_{crude\ i}$ is the specific gravity of crude i. $Sul_{crude\ i}$ is the sulfur content (wt%) of crude i.

Step 3. Assuming that crude qualities and volume fraction remain constant for imported crudes from 2019 onwards, the target domestic crude qualities could then be determined by subtracting volume-weighted API & sulfur content of imported crudes from projected API and sulfur. Once the target API and sulfur content is determined, the volume fraction of each domestic crude assays could be determined to blend a domestic crude proxy.

Step 4. Once country level crude proxies (including both imported and domestic crudes) are determined, we can apply the optimization again to determine the volume fractions for each country to match the target API and sulfur content and hence assemble the crude proxy that is assumed to be processed by the refining sector of different PADDs.

Step 5. Repeat step 1 to step 4 for other regions (PADDs). The blended crude proxies are presented in Dataset S4.

Configurational Crude Input

The configurational crude input is the total actual volume of crude that will be processed by the atmospheric tower (AT) for all refineries that shares the same configurations. Although PRELIM does not require the refinery's throughput to determine the emissions intensity per barrel of crude ($kgCO_2eq/bbl$) or per megajoule of product ($kgCO_2eq/MJ$), it is necessary to input the configurational crude input to aggregate emissions intensities across various refinery configurations. This aggregation is essential for calculating the total GHG emissions within a region, which typically hosts refineries of differing configurations. Furthermore, the volume proportions of crude processed by different refinery configurations also change over time, a result of the combined impact of evolving crude quality and product demand. Kim(12) developed an optimization model which used the standard refinery complex index (SRCI) to project the future trend of configurational refinery crude input for a specific region. This study applies a similar approach to derive the configurational crude input for future scenarios. Detailed stepwise procedure is outlined below.

Step 1. Determine the standardized complexity factor. The complexity factor is first developed by Nelson(17–19) which is used to assess the complexity of refinery's process unit by comparing the cost per capacity for building new individual process units and the cost per capacity for building new AT. Complexity factor is determined by Eq. S8 and is published by Oil and Gas Journal (OGJ)(20) on a yearly basis.

$$CF_{Unit} = \frac{Cost_{Unit}/Capacity_{Unit}}{Cost_{AT}/Capacity_{AT}} \quad \text{Eq. S8}$$

Where CF_{Unit} is the complexity factor for individual process unit. $Cost_{Unit}$ and $Cost_{AT}$ are the cost of building a new process unit and a new AT, respectively. $Capacity_{Unit}$ and $Capacity_{AT}$ are the capacity of the new-built process unit and AT, respectively.

Historically, CF has changed due to the change in cost per capacity for refinery process units. In our study, we applied a constant CF (i.e., Standardized CF, SCF), which was taken from Kaiser(21) and is shown in Table S19.

Step 2. Determine the individual and configurational Standardized Refinery Complexity Index (SRCI). Once the SCF is determined, the SRCI for individual refineries in the U.S. could be calculated using Eq. S9.

$$SRCI_{Refinery} = \sum \left(\frac{Capacity_{Unit}}{Capacity_{ADU}} \times SCF_{Unit} \right) \quad \text{Eq. S9}$$

OGJ published unit capacities for every refinery in the U.S. from 1993 to 2019, and thus the historical SRCI for individual refineries could be calculated. The configurational SRCI for configuration *i* is the volume weighted average SRCI for a total of *j* individual refineries with configuration *i* (see Eq. S10) over *k* years, where *k* is the total number of years from 1993 to 2019.

$$SRCI_{Config,\,i} = \sum_{k} \sum_{j} \left(SRCI_{Refinery,\,i}^{\,j,k} \times \frac{Q_{Refinery,\,i}^{\,j,k}}{\sum_{k} \sum_{j} Q_{Refinery,\,i}^{\,j,k}} \right) \quad \text{Eq. S10}$$

Where, $Q_{Refinery,\,i}^{\,j,k}$ is the crude volume input (bbl/d) for an individual refinery *j* with configuration *i* at year k. The configurational SRCI is presented in Table S5.

Step 3. Develop the multi-linear regression model to determine the target future regional SRCI. Since the configurational SRCI is derived from historical individual refinery SRCI, which is varying depending on the changing crude qualities and product demand over time in the U.S(12). Consequently, this allows for a hierarchical system that ranks refineries with different configurations based on their ability to process crudes of varying qualities and produce products that meet the demands of the U.S. market. The rationale of the SRCIs set foundation for developing a multi-linear regression model that correlates the regional SRCI with the yield of different products and crude quality. In order to derive the regional SRCI, we first need to derive the total crude input (i.e., crude input) for each refinery configuration for each year by simply aggregating crude input for each individual refinery of configuration *i* together for each year. Then, the regional SRCI for year *k* can be defined as the sum product of the configurational SRCI for configuration *i* and the volume proportion of the crude input for each configuration *i* (see Eq. S11).

$$SRCI_{Region}^{\,k} = \sum_{i} \left(SRCI_{Config,\,i} \times \frac{Q_{Config,\,i}^{\,k}}{\sum_{i} Q_{Config,\,i}^{\,k}} \right) \quad \text{Eq. S11}$$

Where, $Q_{Config,\,i}^{\,k}$ is the total crude input for all refineries of configuration *i* at year k.

Since EIA documents crude qualities(15) (i.e., API gravity) as well as the product yields (22) (e.g., gasoline %, diesel %, jet %, etc.) for each year, we can develop the linear regression model to

correlate the regional SRCI for different region *n* (i.e., in this case, PADD *n*), $SRCI_{Region,n}$, and API and product yields (See Eq. S12).

$$SRCI_{Region,n} = c_n + a_{1,n} \times API_n + a_{2,n} \times Gas_n + a_{3,n} \times Diesel_n + a_{4,n} \times Jet_n \quad \text{Eq. S12}$$

Where n denotes the PADD number, c_n is the constant intercept, a_1, a_2, a_3, a_4 are the regression coefficients, $SRCI_{Region}$ is the regional Standardized Refinery Complexity Index, API is the regional crude input API degree, Gas, Diesel, Jet are the regional gasoline yields, diesel yields and jet yields respectively. The constant and regressions coefficients for different PADD are list in Table S6.

The developed multi-linear regression model hence allows the projection of the regional SRCI using future crude APIs and products' yields in different U.S. PADDs.

Step 4. Determine future configurational crude input using optimization. In order to determine the crude input of different refinery configurations in future, an optimization model adopted from Kim(12) is utilized. The regression model projects the target regional SRCI that is expected to process crude with certain quality as well as to produce the product slate that adapts to the transition of the fuel market. On the other hand, the regional SRCI can be defined as the sum product of the configurational SRCI for configuration *i* and the volume proportion of the crude input for each configuration *i*, as we mentioned in Eq. S11. Let us define the multi-linear regression model (i.e., Eq. S12) derived target future regional SRCI in PADD *n* for future year *k* as $SRCI_{Region,n}^{Target,k}$. Let $x_{i,n}^{k}$ be the future configurational crude input for configuration *i* in PADD *n* for future year *k*. Hence the regional SRCI in PADD *n* for future year *k* could also be write as shown in Eq. S13.

$$SRCI_{Region,n}^{k} = \sum_i \left(SRCI_{Config,i} \times \frac{x_{i,n}^{k}}{\sum_i x_{i,n}^{k}} \right) \quad \text{Eq. S13}$$

Assuming that the regional SRCI, $SRCI_{Region,n}^{k}$, will tend to match the target regional SRCI, $SRCI_{Region,n}^{Target,k}$, indicating that refineries with different configurations will adapt their operations – understood here that they will adjust their shares in the regional total crude processing capacity – so that the jurisdiction is able to process the crude with a changing quality while to produce the desired product slates that meet the regional demands. Hence, the objective function can be written as:

$$Minimize: f_0\left(x_{i,n}^{k}\right) = \left| SRCI_{Region,n}^{Target,k} - SRCI_{Region,n}^{k} \right| \quad \text{Eq. S14}$$

Subject to

$$f_1\left(x_{i,n}^{k}\right) = \left| \sum_i x_{i,n}^{k} - X_n^k \right| \leq \left(0.05 \times X_n^k\right), k \in \{2035, 2050\} \quad \text{Eq. S15}$$

$$f_2\left(x_{i,n}^k\right) = \left|\sum_i \left(y_{i,n}^k \times \frac{x_{i,n}^k}{\sum_i x_{i,n}^k}\right) - Y_n^k\right| \le \left(0.05 \times Y_n^k\right), k \in \{2035, 2050\} \quad \text{Eq. S16}$$

$$f_3\left(x_{i,n}^k\right) = x_{i,n}^k \le \left[\left(1 + \Delta p_{i,n}^{ub}\right) \times X_{i,n}^{2019}\right], k = 2035 \quad \text{Eq. S17}$$

$$f_4\left(x_{i,n}^k\right) = x_{i,n}^k \le \left[\left(1 + 2 \times \Delta p_{i,n}^{ub}\right) \times X_{i,n}^{2019}\right], k = 2050 \quad \text{Eq. S18}$$

$$f_5\left(x_{i,n}^k\right) = x_{i,n}^k \ge \left[\left(1 + \Delta p_{i,n}^{lb}\right) \times X_{i,n}^{2019}\right], k = 2035 \quad \text{Eq. S19}$$

$$f_6\left(x_{i,n}^k\right) = x_{i,n}^k \ge \left[\left(1 + 2 \times \Delta p_{i,n}^{lb}\right) \times X_{i,n}^{2019}\right], k = 2050 \quad \text{Eq. S20}$$

Constraints must be considered for the optimization model to acquire reasonable results and are explained as follows.

The first constraint is that the region's total refinery crude input should match the total refinery production, as shown in Eq. S15. EIA documented data (23, 24) suggested a linear correlation exists between the refinery's input and output, as shown in Figure S12.

Let us define the volume of the total refinery output in PADD *n* as On, thus the linear correlation could be written as:

$$X_n = b_n \times O_n \quad \text{Eq. S21}$$

Where, *n* denotes the PADD number, X is the total regional refinery input by volume, O is the regional total refinery output, b is the linear regression coefficient. The value of coefficient b_n is presented in Table S20.

Such relationship creates a certain range (i.e., ±5%) that allows the configurational refinery crude input can vary in order to achieve the optimization goal, as shown in Eq. S15.

$$f_1\left(x_{i,n}^k\right) = \left|\sum_i x_{i,n}^k - X_n^k\right| \le \left(0.05 \times X_n^k\right), k \in \{2035, 2050\} \quad \text{Eq. S15}$$

Where *i* denotes the refinery configuration, *n* denotes the PADD number, *k* denotes the future year (i.e., 2035 or 2050), x is the future configurational crude input, X is the total regional refinery crude input by volume.

The second constraint is to match the regional volume-weighted product yields, such as gasoline, diesel, and jet fuel, with the product yields derived from previous section (Supplementary method 2) (see Eq. S4). PRELIM can generate product slates for each refinery configuration that matches the target regional slates as closely as possible by adjusting the material flows between process units. These PRELIM-derived product slates are then aggregated based on the volume share of each refinery configuration throughput at a regional level and match with the target slates (i.e., ±5%).

$$f_2\left(x_{i,n}^k\right) = \left|\sum_i \left(y_{i,n}^k \times \frac{x_{i,n}^k}{\sum_i x_{i,n}^k}\right) - Y_n^k\right| \le \left(0.05 \times Y_n^k\right), k \in \{2035, 2050\} \quad \text{Eq. S16}$$

Where *i* denotes the refinery configuration, *n* denotes the PADD number, *k* denotes the future year (i.e., 2035 or 2050), Y is the target regional product's yields derived from previous section

(Supplementary method 2), y is the product's yields derived from PRELIM for different configuration in specific region, x is the future configurational crude input.

The third to sixth constraints from Eq. S17 and Eq. S20 specify the upper and lower bounds for the future configurational throughput. These bounds are determined based on the historical data obtained from EIA and OGJ sources(20, 25).

To set the upper bounds for 2035, the mean configurational refinery crude input increases over the past 15 years, compared to 2019, is assumed to be the maximum percentage increase. The upper bounds for each configuration are then determined as the 15-year average percentage increases based on the 2019 refinery crude input. Similarly, the upper bounds for 2050 are set as twice the 15-year average percentage increase based on the 2019 refinery crude input.

The lower bounds for the years 2035 and 2050 are also determined using the same approach. Specifically, they are set as the 15-year averaged percentage decreases twice the 15-year averaged percentage decreases based on 2019 throughput, respectively. The historical average numbers are used to reflect how much a refiner is willing to change their throughput to adapt to market changes such as shifts in crude qualities and product demand.

For any single year q from 2005 to 2019 with refinery crude input $X_{i,n}^{q}$, if $\left(X_{i,n}^{q} - X_{i,n}^{2019}\right) > 0$, then,

$$\Delta p_{i,n}^{ub} = \frac{1}{m}\sum_{1}^{m}\left[\frac{\left(X_{i,n}^{q} - X_{i,n}^{2019}\right)}{X_{i,n}^{2019}}\right] \quad \text{Eq. S22}$$

$$f_3\left(x_{i,n}^{k}\right) = x_{i,n}^{k} \le \left[\left(1 + \Delta p_{i,n}^{ub}\right) \times X_{i,n}^{2019}\right], k = 2035 \quad \text{Eq. S17}$$

$$f_4\left(x_{i,n}^{k}\right) = x_{i,n}^{k} \le \left[\left(1 + 2 \times \Delta p_{i,n}^{ub}\right) \times X_{i,n}^{2019}\right], k = 2050 \quad \text{Eq. S18}$$

Where *i* denotes the PADD number, *j* denotes the refinery configuration, m is the total number of years with $\left(X_{i,n}^{q} - X_{i,n}^{2019}\right) > 0$ from 2005 to 2019, *k* denotes the future year (i.e., 2035 or 2050), x is the future configurational crude input, $\Delta p_{i,n}^{ub}$ is the mean configurational refinery crude input increase over the past 15 years compared to the year 2019, $X_{i,n}^{2019}$ is the configurational refinery crude input in each PADD for the year 2019.

For any single year q from 2005 to 2019 with refinery crude input $X_{i,n}^{q}$, if $\left(X_{i,n}^{q} - X_{i,n}^{2019}\right) < 0$, then,

$$\Delta p_{i,n}^{lb} = \frac{1}{m}\sum_{1}^{m}\left[\frac{\left(X_{i,n}^{q} - X_{i,n}^{2019}\right)}{X_{i,n}^{2019}}\right] \quad \text{Eq. S23}$$

$$f_5\left(x_{i,n}^{k}\right) = x_{i,n}^{k} \ge \left[\left(1 + \Delta p_{i,n}^{lb}\right) \times X_{i,n}^{2019}\right], k = 2035 \quad \text{Eq. S19}$$

$$f_6\left(x_{i,n}^{k}\right) = x_{i,n}^{k} \ge \left[\left(1 + 2 \times \Delta p_{i,n}^{lb}\right) \times X_{i,n}^{2019}\right], k = 2050 \quad \text{Eq. S20}$$

Where *i* denotes the PADD number, *j* denotes the refinery configuration, m is the total number of years with $\left(X_{i,n}^{q} - X_{i,n}^{2019}\right) < 0$ from 2005 to 2019, *k* denotes the future year (i.e., 2035 or 2050), x is the future configurational crude input, $\Delta p_{i,n}^{lb}$ is the mean configurational refinery crude input decrease over the past 15 years compared to the year 2019, $X_{i,n}^{2019}$ is the configurational refinery

crude input in each PADD for the year 2019. The mean configurational refinery crude input change for different configurations is summarized in Table S21.

Step 4. Repeat step 1 to step 3 for other regions (PADDs).

It is important to note that we simulated the crude input solely for the years 2035 and 2050. Although this method has the advantage of estimating annual operational changes in the refining sector based purely on publicly available data concerning crude qualities and product yields, real refineries often adjust their crude input based on refinery margins on a weekly, if not daily, basis. This discrepancy underscores that our optimization model is better suited for estimating changes in crude input over longer intervals. In this study, the chosen interval is 15 years, during which the correlation between configurational throughput changes and variations in crude qualities and product yields becomes significantly more pronounced.

There is additional limitation of the SRCI framework. The SRCI framework applied in this study is calibrated using historical refinery data and is intended to represent regional refinery complexity within a transportation-fuels–oriented operating regime. Historically, U.S. refineries have been predominantly focused on transportation fuel production, with gasoline, diesel, and jet fuel together accounting for approximately 85% of total refinery output according to EIA data. Consequently, the SRCI regression relates crude quality (API gravity) and transportation fuel yields (gasoline, diesel, and jet) to regional refinery complexity, consistent with the scope and assumptions of the modeled scenarios.

However, historical data may not fully capture potential future structural shifts in refinery product slates. For example, a scenario in which a greater share of the naphtha fraction is redirected from gasoline blending toward petrochemical production would represent a departure from past operating trends. The current SRCI specification does not explicitly include petrochemical-specific process units such as steam cracking, nor does it incorporate petrochemical feedstock yields as explanatory variables. Although petrochemical feedstocks accounted for approximately 2.0–5.7 vol% of total U.S. refinery output between 2011 and 2019, and national production declined from approximately 146 million barrels per year in 2005 to 109 million barrels per year in 2019, these historical shares do not preclude future reconfiguration under changing demand conditions (24). Data limitations also exist at the regional level; for example, no petrochemical feedstock production is reported for PADD 4 during the 2011–2019 period.

Capturing a material shift toward refinery–petrochemical integration would therefore require re-specification and recalibration of the SRCI framework, rather than extrapolation of historical relationships beyond their validated domain.

CO_2 Avoidance Cost Analysis.

The CO_2 avoidance cost equation, adopted from Gale et al. (2), necessitates the estimation of capital expenditures (CapEx) and operational expenditures (OpEx) for various process units at different scales under assorted carbon capture scenarios. This approach requires establishing a

cost baseline from a reference process. In this study, the reference process for carbon capture refers to the Base Case 04-03 developed by Gale et al. (2). The background information for the reference case can be found in Table S22.

It is designed to capture CO_2 from multiple process heaters in a large complex refinery, and the flue gas stream is generated by AT and VDU (i.e., Vacuum Distillation Unit). Table S23 summarizes the equipment list, size metrics and the equipment cost for the reference process.

The cost of each piece of equipment could be estimated using the factorial method, as shown in Eq. S24.

$$C_e = a + bD^r \quad \text{Eq. S24}$$

Where C_e is the equipment cost, a and b are coefficients, D is the size metric, and r is the exponent. Eq. S24 could be used for estimating equipment cost that are presented in Table S23, and selected a, b, D and r that are used in this study could be found in Table S24.

Noted Eq. S24 is only valid when the size metric D falls between D_H and D_L. For circumstances where D exceeds D_H, extra equipment will be considered. Table S25 compared the estimated cost for the reference configuration, which is estimated based on a factorial approach, and the base case 01-01 by Gale et al. (2), which is capturing CO_2 from an on-site power plant. All costs are calibrated to the year 2019.

The function of equipment cost estimation for the absorption section is shown in Eq. S25.

$$\begin{aligned} Cost\ of\ absorption_{new} \\ = Cost\ of\ absorber_{new} + Cost\ of\ other\ items_{new} = \\ + \left(\frac{Flue\ gas\ mass\ rate_{new}}{Flue\ gas\ mass\ rate_{ref}}\right)^{exponent} \times Cost\ of\ other \end{aligned} \quad \text{Eq. S25}$$

Linear correlations have been developed between the diameter and the height of the absorber and the flue gas mass rate based on (2) and are shown in Eq. S26 and Eq. S27.

$$Absorber\ height = Flue\ gas\ mass\ rate\ \times 0.0213 + 33.63 \quad \text{Eq. S26}$$

$$Absorber\ diameter = Flue\ gas\ mass\ rate\ \times 0.011 + 3.0757 \quad \text{Eq. S27}$$

In order to avoid unrealistic estimation of the size of the absorber, the maximum height for the absorber is set to 50 m. Hence, the absorber's max capacity and the absorber's max diameter could be calculated as 768.54 t/h and 11.53 m. If the total flue gas mass rate exceeds the max capacity, then extra trains of the absorber will be considered to serve the needs for processing the additional flue gas stream.

The function of equipment cost estimation for the regeneration section is shown in **Eq. S28**. Linear correlations have been developed between the height of the stripper and the CO_2 flow rate to compression based on literature data and are shown in **Eq. S29**.

$$Cost\ of\ regeneration_{new} = \left(\frac{CO_2\ Flowrate\ to\ compression_{new}}{CO_2\ Flowrate\ to\ compression_{ref}}\right)^{exponent} \times \left(St \right. \quad \text{Eq. S28}$$

$$Stripper\ height = \ CO_2\ Flowrate\ to\ compression\ \times 0.0628 + 16.684 \quad \text{Eq. S29}$$

In order to avoid unrealistic estimation of the size of the stripper, the maximum height for the stripper is set to 50 m. Hence the max capacity of the stripper could be calculated as 530.51 t/h. If the total CO_2 flow rate to the stripper exceeds the max capacity, then extra trains of stripper will be considered to serve the needs for processing the extra stream of CO_2 stream.

Modeling the dimensions of the carbon capture process unit requires a rigorous analysis of mass transfer and thermodynamics, including equilibrium CO_2 concentrations in the gas phase, estimation of mass transfer coefficients, effective mass transfer area, and gas superficial velocity. However, as discussed in this paper, CO_2 sources are distributed throughout the refinery, with operating conditions that vary based on crude input and desired product slates. Given the extensive number of cases simulated in this study (a total of 19,610 data points), a simplified approach (i.e., linear regression) was adopted to estimate the dimensions of major carbon capture process units. Additionally, the linear regression is based solely on data from Gale et al. study, where the reported flue gas flow rates range from 61.3 to 1,150 t/h. The dataset indicates a strong correlation between flue gas flow rate and absorber height but does not show such relationship between the ratio of CO_2 concentrations in the inlet and outlet gas streams and absorber height. Similarly, a strong correlation is observed between flue gas flow rate and absorber diameter, as well as CO_2 flow rate and stripper height. As previously mentioned, a maximum height limit was imposed on absorbers and strippers to ensure that the dimensions remain within the range of the dataset used to develop linear regression equations. However, in certain scenarios, the flue gas flow rate from specific process units falls below the lower bound of Gale's dataset (61.3 t/h). In such cases, extrapolating the data may lead to an overestimation of carbon capture costs. Furthermore, Julio et al. (26) suggest that the height of commercial absorbers should typically range from 20 to 36 m. This implies that many of the cases in our study with very low flue gas flow rates would likely not be

considered viable for carbon capture in reality. Even if they were, the costs would be prohibitively high due to the limited availability of equipment with the required dimensions. Danaci et al. (27) also shows that when flue gas flow rate is extremely low (3.6 t/h), the CO_2 capture cost could reach over 1,000 \$/t CO_2 captured. These considerations have been captured in the linear regression equations by including the intercept.

Moreover, a sensitivity analysis was conducted to assess the impact of variations in unit dimensions on CO_2 avoidance costs, considering cases where dimensions are 50% and 150% of those estimated using linear regression (see Figure S23). The mean CO_2 avoidance costs under these three scenarios (i.e., 50%, 100%, and 150% of the estimated dimensions) are 290.3, 293.4, and 301.3 \$/t, respectively. These results confirm that even significant changes in absorber and stripper dimensions have a minimal effect on CO_2 avoidance costs.

Eq. S30, Eq. S31, Eq. S32 and Eq. S33 are functions of equipment cost estimation for CO_2 compression, interconnection, cooling towers and wastewater treatment, respectively.

$$Cost\ of\ compression_{new} = Cost\ of\ compression_{new}^{capacity\ dependent} + Cost\ of\ c \times \left(Cost\ of\ compression_{ref}^{capacity\ dependent}\right) + 600{,}00 \quad \text{Eq. S30}$$

$$Cost\ of\ interconnection_{new} = \left(\frac{CO_2\ Flowrate\ to\ compression_{new}}{CO_2\ Flowrate\ to\ compression_{ref}}\right)^{exponent} \times \left(Cost\ of\ interconnec\right) \quad \text{Eq. S31}$$

$$Cost\ of\ cooling\ tower_{new} = \left(\frac{Number\ of\ cells_{new}}{Number\ of\ cells_{ref}}\right)^{exponent} \times \left(Cost\ of\ cooling\ tower_{ref}\right) \quad \text{Eq. S32}$$

$$Cost\ of\ wastewater\ treatment_{new} = \left(\frac{Wastewater\ inlet_{new}}{Wastewater\ inlet_{ref}}\right)^{exponent} \times \left(Cost\ of\ wastewater\ treat\right) \quad \text{Eq. S33}$$

Noted that the capacity for a single cell of the cooling tower is set to 2500 m^3/h. The capacity parameter and the exponent mentioned in Eq. S25, Eq. S28, and Eq. S30 to Eq. S33 are presented in Table S26.

The other minor equipment cost is set to be 10% of the total cost for the major equipment mentioned above. Location factors are transposed from the Chemical Engineering Design book(28), containing 4 U.S regions: Gulf Coast, Mid-West, West Coast and East Coast. PADD 1 and 2 refers to East Coast, PADD 3 refers to Gulf Coast, PADD 4 refers to Mid-West, and PADD 5 refers to West Coast. Table S15 summarizes the location factors that are used in this study.

Other key performance indicators (KPIs) that are necessary for calculating CapEx and OpEx are included in Table S7 and Table S8, respectively.

The Chemical Engineering Plant Cost Index (CEPCI) used to estimate future cost is extrapolated from literature and is presented in Table S27. Hence, the CapEx can be estimated using Eq. S34.

$$CapEx_{new} = C_{TPC}^{ref} \times (1 + R_{other}) \times LF \times \frac{CEPCI_{new}}{CEPCI_{ref}} = \left(C_{e}^{ref} \times (1 \quad \text{Eq. S34}$$

$$LF \times \frac{CEPCI_{new}}{CEPCI_{ref}}$$

Where, $CapEx_{new}$ refers to the capital expenditure at the target year (i.e., 2035 or 2050), C_{TPC}^{ref} refers to the total plant cost at the baseline year (i.e., 2019), C_{e}^{ref} refers to the equipment purchase cost at the baseline year, R_{other} refers to the ratio of other costs to the total plant cost, $R_{construction}$ refers to the ratio of construction cost to the total equipment purchase cost, R_{EPC} refers to the ratio of EPC service cost to the total equipment purchase cost, $CONTG$ refers to the project contingency, LF refers to the location factor, $CEPCI$ refers to the Chemical Engineering Plant Cost Index.

The CO_2 avoidance cost metric as defined in this study (and consistent with most others) excludes carbon pricing effects and potential revenues from CO_2 sales or utilization. CO_2 avoidance costs are intended to be used as a means of comparing the cost of reducing emissions using a given mitigation option with other like options or with an externally imposed carbon price (29). While they may include the effect of policies such as investment or production tax credits (e.g., Section 45Q tax credits), we do not include such policies in our avoidance cost calculations.

Furthermore, we exclude impacts of CO_2 sales or utilization on investment decisions, as the current and foreseeable market size for CO_2 as a commodity remains limited and highly uncertain. In North America, the largest end use of externally sourced CO_2 is enhanced oil recovery (EOR), with total demand on the order of ~60 $MtCO_2$/yr, of which only approximately 16% is supplied by industrial sources(30). Prices for commodity CO_2 are typically determined through long-term bilateral contracts and vary widely depending on purity requirements, transaction volumes, and local supply–demand conditions. For example, CO_2 used in ammonia and other industrial applications is often priced in the range of 3–15 $/$tCO_2$, whereas niche applications requiring high purity and small show volumes can command substantially higher prices(31). Given the relatively small scale of these markets and the high uncertainty in achievable CO_2 prices, incorporating CO_2 sales revenues into avoidance cost curves would affect only a small fraction of total captured CO_2 and would be unlikely to materially alter the relative ranking of capture options, while potentially introducing additional uncertainty.

From a Canadian policy perspective, the federal government has implemented an industrial carbon pricing framework - the Output-Based Pricing System (OBPS) - which applies to large industrial facilities in most provinces. Under the OBPS, facilities face a carbon price on emissions that exceed facility-specific emissions-intensity benchmarks. The federal government establishes minimum national carbon pricing benchmarks, which were set at CAD 20/tCO_2 in 2019 and increased by CAD 10/tCO_2 annually to CAD 50/tCO_2 in 2022. The benchmark was subsequently updated to

increase by CAD 15/tCO_2 annually starting in 2023, reaching CAD 170/tCO_2 by 2030 (32). Assuming an average annual exchange rate of 0.72 USD per CAD, the corresponding carbon prices are approximately 14.4, 36.0, 57.6, and 122 \$/tCO_2 in 2019, 2022, 2024, and 2030, respectively. Figure 2 indicates that carbon price levels of approximately 122 \$/tCO_2 are required to enable substantial emissions reductions from the U.S. refining sector—approximately 16.6 $MtCO_2eq/y$ (17.6%)—through deployment of CCS on FCC units at medium and deep conversion refineries, as well as SMR units at deep conversion refineries in PADD 3. Interpreted in the Canadian policy context, this price level is comparable to the federal benchmark scheduled for 2030. However, the scheduled benchmark does not necessarily reflect the effective compliance price faced by regulated facilities. In practice, effective carbon prices in some provinces have been substantially below the headline benchmark. For example, recent credit market conditions in Alberta have resulted in effective compliance prices below 20 CAD/tCO_2 (33). While recent federal–provincial discussions have indicated an intention to strengthen effective pricing toward CAD 130/tCO_2, no firm implementation date has been established, and future effective price trajectories remain uncertain.

Supplementary Method 3 - Real Option Analysis

Probability of Individual Refinery's Strategic Options.

In option pricing theory, the Black-Scholes model assumes that the stock price follows a Brownian motion with a drift (i.e., geometric Brownian motion) (34). The stock price that follows the geometric Brownian motion will satisfy the following stochastic differential equation:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \qquad \text{Eq. S35}$$

Where S_t here represents the stock price, μ is drift, σ is volatility, and W_t is a Wiener process (Brownian motion). Binomial lattice model is a discrete time model that eventually converges to the Black-Scholes equation, where the random walk of the stock price on the binomial lattice approximates the GBM process (35). There exists a close analogy between financial option and real option analysis, according to He (36).

Three individual refineries with different configurations (i.e., hydroskimming, medium conversion: FCC & GOC, deep conversion: FCC & GOC) are considered in the RO analysis. The crude input for these refineries in 2019 is assumed to be the average (i.e., arithmetic mean) for refineries with the same configurations across the U.S., thereby representing typical U.S. refineries for that year. Future (i.e., 2020 to 2050) crude input for these refineries are calculated based on the ratio of the total projected crude input for each refinery configuration in the U.S., as forecasted by the optimization technique described in the previous section (Supplementary method 2) under varying fuel demand scenarios, to the total crude input in 2019, as shown in Eq. S36. Note for each

individual refinery we assume three fuel demand scenarios (i.e., low, medium and high), which correspond to BL to LD35/50, BL to MD35/50 and BL to HD35/50 scenarios.

$$Q_{Crude,j}^{t} = Q_{Crude,j}^{2019} \times \frac{x_j^t}{x_j^{2019}} \quad \text{Eq. S36}$$

Where $Q_{Crude,j}^{t}$ and $Q_{Crude,j}^{2019}$ are the crude input of the selected individual refinery with configuration j ($j = 0, 3, 6$) in year t ($t \in (2020, 2050)$) and year of 2019, respectively. x_j^t and x_j^{2019} are the total crude input of refineries in U.S. with configuration *j* in year t and year of 2019, respectively.

It is noteworthy that the crude input for certain refinery archetype, such as hydroskimming (configuration 0), is projected to increase in the future under specific fuel demand scenarios. However, it is widely recognized that U.S. refiners are unlikely to expand their operations in the face of declining fuel demand and more stringent carbon policies (37). Consequently, a constraint has been implemented ensuring that the selected individual refinery maintains its crude input constant whenever the ratio $\frac{x_j^t}{x_j^{2019}} \geq 1$, indicating that no new investments will be made to expand the refinery's operations. The total production and the yield of transportation fuels are optimized using the embedded product matching tool in PRELIM to align with the EIA-documented product slates in 2019 and the projected refinery product slates from 2020 onwards for each refinery. Detailed information on each individual refinery is provided in Table S9. Sensitivity analysis on the likelihood of different options for a medium typical U.S. refinery under a medium transportation fuel demand is also considered in Section 3.2.

Consider a state space A, which contains a vector a_t denoting the refinery's current status including volume of crude input Q_{Crude}, volume of refinery product m, $Q_{Product\_m}$, mass of GHG emissions M_{GHG}, natural gas consumptions for refinery process units and carbon capture unit E_{NG} and E_{NG_CC}, electricity consumptions for refinery process units and carbon capture unit E_{Elec} and E_{Elec_CC}, fixed operating cost U_{Fix} and capital investment in carbon capture I_{CC} at time t. State space A can thus be expressed as equation Eq. S37. The action b_t that refinery could take at time t, are from a set of actions B as stated in equation Eq. S38. Hence the refinery's cash flow at state a_t when action b_t is taken could be estimated as equation Eq. S39

$$A = \left\{\left[Q_{Product_m}^{t}, Q_{Crude}^{t}, M_{GHG}^{t}, E_{NG}^{t}, E_{NG\_CC}^{t}, E_{Elec}^{t}, E_{Elec\_CC}^{t}, U_{Fix}^{t}, I_{CC}^{t}\right]\right\} \quad \text{Eq. S37}$$

$$B = \left\{ \begin{bmatrix} 1.Continue\ operation; \\ 2.Deploying\ carbon\ capture; \\ 3.\ Shutdown\ the\ refinery \end{bmatrix} \right\} \quad \text{Eq. S38}$$

$$CFL(a_t)_{b_t} = \begin{cases} \left(\sum \pi^{t}_{Product_m} Q^{t}_{Product_m}\right) - \pi^{t}_{Crude} Q^{t}_{Crude} - \pi^{t}_{Carbon} M^{t}_{GHG} - \pi^{t}_{NG} E^{t}_{NG} - \pi^{t}_{Elec} E^{t}_{Elec} - U^{t}_{Fix} \\ \left(\sum \pi^{t}_{Product_m} Q^{t}_{Product_m}\right) - \pi^{t}_{Crude} Q^{t}_{Crude} - \pi^{t}_{Carbon} M^{t}_{GHG} - \pi^{t}_{NG}\left(E^{t}_{NG} + E^{t}_{NG_{CC}}\right) - \pi^{t}_{Elec}\left(E^{t}_{Elec} + E^{t}_{Elec_{CC}}\right) - U^{t}_{Fix} - I^{t}_{CC} \\ 0 \end{cases} \quad \text{Eq. S39}$$

$$CFL(a_t)_{b_t} = \begin{cases} R^{t}_{Products} - C^{t}_{Crude} - C^{t}_{Carbon} - C^{t}_{NG} - C^{t}_{Elec} - C^{t}_{Fix},\ b_t = 1 \\ R^{t}_{Products} - C^{t}_{Crude} - C^{t}_{Carbon_{CC}} - C^{t}_{NG_{CC}} - C^{t}_{Elec_{CC}} - C^{t}_{Fix} - I^{t}_{CC},\ b_t = 2 \\ 0,\ b_t = 3 \end{cases} \quad \text{Eq. S40}$$

Where, $CFL(a_t)_{b_t}$ is the cash flow of the refinery at time t with state a_t when action b_t is taken. π^{t}_{Crude}, π^{t}_{Carbon}, π^{t}_{NG}, π^{t}_{Elec}, $\pi^{t}_{Product_m}$ are prices for crude oil, carbon emissions, natural gas, electricity and product m, respectively.

Let $S(a_t)_{b_t}$ be the present value of the refinery at time t with state a_t when action b_t is taken. Let $\delta(a_t)_{b_t}$ be the instantaneous dividend distribution rate at time t. Thus $S(a_t)_{b_t}$ and $\delta(a_t)_{b_t}$ could be expressed as Eq. S41 and Eq. S42, respectively.

$$S(a_t)_{b_t} = \sum_{i=t}^{T} \frac{CFL(a_t)_{b_t}}{(1+\mu)^i},\ t \in (0,T), b_t \in (1,3) \quad \text{Eq. S41}$$

$$\delta(a_t)_{b_t} = \frac{CFL(a_t)_{b_t}}{S(a_t)_{b_t}},\ t \in (0,T), b_t \in (1,3) \quad \text{Eq. S42}$$

Where μ is the risk-adjusted discount rate (i.e., 12%), T is the maturity time (i.e., expiry time) of options.

When $b_t = 1$, $S(a_t)_1$ is equivalently the present value of the project without option (i.e., continue operation) and hence could be served as the underlying asset value in RO (38). The prices of crude oil, natural gas, refinery products and electricity have strong correlations and are both affected by factors beyond market conditions (e.g., geopolitics) and thus beyond the capacity of this study. Hence, we utilized the prices estimated by the Annual Energy Outlook (AEO) (39) on reference case for cash flow calculation. The carbon price, on the other hand, is assumed to follow a GBM (40), such that the random walk of the underlying asset value will also follow GBM and traditional option pricing method is applicable (41). The drift and volatility for the GBM of carbon price are both

constant and are taken from McKeller (summarized in Table S28), while the starting carbon prices are taken from the European Union's (EU) Emissions Trading Scheme (ETS) (42).

10,000 trajectories of carbon prices are generated (Figure S12) and will be used for a Monte Carlo simulation for the binomial lattice. Detailed statistics information could also be found in Table S30. Assuming that the cash flow will be paid by the end of each time step, which in this case is 1-year, and if the underlying asset value in node *j* at time t is $S(a_t,j)$, the underlying asset value of its child nodes could be expressed by Eq. S43 to Eq. S46.

$$S(a_{t+1},j)^{+} = \left(S(a_t,j) - S(a_t,j)\delta(a_t)\right) \times u \quad \text{Eq. S43}$$

$$S(a_{t+1},j+1)^{-} = \left(S(a_t,j) - S(a_t,j)\delta(a_t)\right) \times d \quad \text{Eq. S44}$$

$$u = e^{\sigma\sqrt{\Delta t}} \quad \text{Eq. S45}$$

$$d = e^{-\sigma\sqrt{\Delta t}} \quad \text{Eq. S46}$$

Where u and d are up and downward factors respectively. σ is the volatility of the underlying asset value, which could be estimated using the method developed by Cobb et al. (43), Δt is the time step (i.e., 1 year), *j* is the node number at time t ($j \in (1, t)$) and the plus/minus superscripts indicate whether the random walk of the underlying asset value goes up or down.

Hence, the random walk of the underlying asset value on the binomial lattice in node *j* at time t is governed by Eq. S47.

$$S(a_t,j) = S(a_0)_1 \times u^{t-j} \times d^{j} \times \prod_{k=1}^{t-1}\left[1 - \delta(a_k)_1\right] \quad \text{Eq. S47}$$

Let b_T^* be the optimal decisions that the refinery will take at expiry date T and $V(i,T)_{b_T^*}$ be the valuation function of the optimal decision. Hence, the optimal decision and the expected present value of the project at time t could be computed using backward induction by Eq. S48 to Eq. S50.

$$\begin{cases} b_T^* = argsup\left\{V(i,T)_{b_T}\right\} \\ V(i,T)_{b_T^*} = sup\left\{V(i,T)_{b_T}\right\} \end{cases} \quad \text{Eq. S48}$$

$$\begin{cases} V(i,t)_{b_t^*} = max\begin{pmatrix} \left(pV(i,t+1)_{b_{t+1}^*} + (1-p)V(i+1,t+1)_{b_{t+1}^*}\right) \times e^{-rf\cdot\Delta t} + CFL(a_t)_1, \\ S(a_t,j) + \Delta S(a_t), \\ 0 \end{pmatrix} \\ V(i,T)_{b_T^*} = max\begin{pmatrix} S(a_T,j) + S(a_T)_1, \\ S(a_T,j) + S(a_T)_2, \\ 0 \end{pmatrix} \end{cases} \quad \text{Eq. S49}$$

$$p = \frac{e^{rf \cdot (\Delta t)} - d}{u - d}$$ Eq. S50

Where, $V(i,t)_{b^*_t}$ is the valuation function of the optimal decision at time t, p is the risk-neutral probability, rf is the risk-free interest rate, $\Delta S(a_t)$ is the present value difference of the refinery after carbon capture is deployed.

This study uses a different real option model than those by Reinelt et al. (44) and McKellar et al. (45) use different methods for real options analysis—namely dynamic programming vs contingent claims(36). The dynamic programming approach also uses expected asset values but explores all future pathways leading to the current time node. Therefore, the probability used those studies is the "real-world probability," representing the likelihood of each pathway occurring among all possible future pathways. This is why the discount rate is incorporated—because the probability does not capture risk. The rationale for having a different discount rate and risk-free interest rate in this study is that the underlying asset value in 2019 (i.e., node 0) is calculated based on the present value of the future discount cash flow of the refinery, which is the best unbiased estimator of the market value of the project if it were a traded asset (38). The 12% discount rate is a reflection of the company's perspective on the project's economics based on the perfect foresight of future price projections.

However this discount rate cannot be used to calculate the expected present value calculation during the backward induction, because the refinery holds the right but not the obligation to invest - the 12% discount rate would change if the refinery deferred investment (46). He (36) also noted that, mathematically, the two approaches (dynamic programming versus contingent claim) to real options should yield the same results unless a constant discount rate is assumed. In that case, the constant rate must equal the risk-free interest rate to ensure alignment of results, implying a market price of risk equal to zero. This reasoning also aligns with the original binomial option pricing theory by Cox et al (35). The Cox–Ross–Rubinstein (CRR) model is a simplified option pricing framework that does not require investors' subjective probabilities regarding whether the asset value will move up or down. It also does not depend on investors' attitudes toward risk, since the model assumes a constant risk-free interest rate to prevent arbitrage opportunities. However, if the interest rate changes with the up/down states of the portfolio, the model's assumptions no longer hold.

It should also be noted that execution of the option to invest (i.e., deploy carbon capture) and the option to abandon (i.e., shut down the refinery) is non-reversible. Since the binomial lattice is considered path-dependent, certain constraints must be utilized so that once the decision is made, the investment will be locked in subsequent periods (41).

For a specified carbon price trajectory, a single binomial lattice is generated Figure S13. A Monte Carlo simulation is then employed, sampling various carbon price trajectories that adhere to the GBM process to create a set of binomial lattices (i.e., 10,000 runs). This setup allows for the

563 calculation of the likelihood of different options at a specific time t by dividing the number of nodes
564 executing option b_t by the total number of nodes at that time.

565 **Additional Carbon Price Trajectory Scenarios**

566 In addition to the GBM-based carbon price scenario presented in the manuscript, we construct two
567 alternative carbon pricing scenarios in which carbon price changes are governed by a Bernoulli
568 process. In these scenarios, at each decision point the carbon price may increase by a discrete
569 amount with a probability drawn from a specified range. Once a price increase occurs, the carbon
570 price is held constant for a fixed period (i.e., two years), after which another Bernoulli trial
571 determines whether an additional increase occurs.

572 The magnitude of each price increase is drawn from a uniform distribution between 10–30 $/tCO$_2$
573 (equivalent to 5–15 $/t·yr), and the probability of an increase is drawn from a uniform distribution
574 between 10% and 50%. A soft cap on the carbon price of 250 $/tCO$_2$ is imposed such that, once
575 the simulated carbon price reaches or exceeds this level, the probability of further increases is
576 reduced to 0.1–0.5%, effectively limiting additional growth. A threshold year (2035) is also specified
577 to control whether price increases are permitted to occur before or after this point in time. If price
578 increases are allowed to occur before the threshold year, the scenario is referred to as an
579 anticipated carbon price (A-CP) scenario; if price increases are deferred until after the threshold
580 year, it is referred to as a deferred carbon price (D-CP) scenario. In the A-CP scenario, the
581 combination of the soft cap and reduced post-cap probabilities produces a plateau in carbon prices
582 after the threshold year, whereas in the D-CP scenario, carbon prices remain relatively low prior to
583 the threshold year and increase thereafter.

584 These carbon price trajectories are not intended to represent contemporary carbon pricing
585 regulations. Rather, they are designed to achieve three modeling objectives. First, they generate
586 long-term average carbon price levels comparable to those produced by the GBM-based trajectory
587 by 2050, enabling consistent comparison across scenarios. Second, the Bernoulli-based structure
588 with fixed holding periods reflects the stepwise nature of carbon price adjustments commonly
589 observed in regulatory settings. Third, the plateau behavior—particularly in the A-CP scenario—is
590 used to approximate potential mean-reverting behavior in carbon prices under political and
591 regulatory constraints. Given the early stage of carbon pricing implementation across jurisdictions,
592 the long-term equilibrium level of carbon prices remains uncertain, as does whether prices will
593 continue to rise indefinitely or stabilize at some future level. As discussed in the manuscript, current
594 carbon prices are generally insufficient to incentivize large-scale deployment of industrial mitigation
595 technologies such as CCS. Accordingly, we assume that carbon prices eventually converge toward
596 levels higher than current values. For consistency across scenarios, this convergence level is set
597 at 250 $/tCO$_2$, which is close to the mean carbon price realized in the real-options analysis
598 presented in the manuscript (hereafter referred to as R-CP).

Sensitivity Analysis on Probability of Individual Refinery's Strategic Options.

A sensitivity analysis was conducted to evaluate the probability of different strategic options for a medium conversion refinery operating under a medium fuel demand scenario. This analysis varied five key parameters: the risk-neutral interest rate, crack spread, CCS utility consumption, CCS capital costs and carbon price. Each parameter was assessed to determine its influence on the timing and likelihood of implementing strategic decisions at the refinery, with three scenarios—High, Reference, and Low - defined for each to systematically explore their impact on operational choices. In RO analysis, future cash flows are discounted using the risk-free interest rate, typically the 10-year U.S. Treasury bond rate, which assumes investors are indifferent to risk. This allows them to concentrate solely on expected returns without adjusting to risk preferences. In this study, the reference case for the risk-free interest rate is the 10-year U.S. Treasury bond rate as of 2019, which was 3%. The high and low scenarios are set at 7% and 1%, respectively, representing the highest and lowest points of the bond rate over the past 20 years. These variations in the risk-neutral interest rate could indicate how shifts in monetary policy might influence refinery decisions. Crack spread, which refers to the difference between the price of refined petroleum products and the cost of crude oil, reflects the prevailing market conditions of the oil and gas industry. The Annual Energy Outlook (AEO) models these prices under various scenarios, including high, reference, and low oil price scenarios. Consequently, the high, reference, and low crack spread cases for the sensitivity analysis in this study are aligned with these corresponding price scenarios from the AEO. The CCS utility cost and CCS capital cost can reflect the technological advances in carbon capture technology. The higher and lower bond of the CCS utility and capital cost is set to be 150% and 50% of reference cases. Similarly, carbon price, which represents the tax paid by refiners on carbon emissions, also set to be vary between 150% and 50% of the reference case. The selected parameters for sensitivity analysis are summarized in Table S16.

Expected Total GHG Emissions & Refinery Production.

It is important to note that RO is applied only to individual refineries in this study. This approach is used because once a refinery decides to execute a strategic option, such as shutting down, it inevitably alters the market dynamics—specifically, the supply-demand balance and subsequently the price of transportation fuels in the U.S. market. Consequently, these changes will also influence the likelihood of different strategic options being viable for other refineries. Nevertheless, we can still gain insight into the potential changes in total GHG emissions from the three selected refinery configurations in the U.S. by using expected GHG emissions as an indicator. The expected GHG emissions at time t is defined by Eq. S51.

$$E(GHG_t) = \sum_i \sum_b \left(GHG_{i,t}^{\ b} \times p_{i,t}^{\ b}\right) \quad \text{Eq. S51}$$

Where, *i* denotes the selected refinery configurations, t denotes time between 2019 and 2050, b denotes three considered strategic options, $E(GHG_t)$ is the expected total GHG emissions from the three selected refinery configurations at time t, $GHG_{i,t}^{\ b}$ is the GHG emissions induced by strategic option b for configuration *i* at time t, $p_{i,t}^{\ b}$ is the likelihood of the strategic option b for configuration *i* at time t.

Similarly, we can also define the expected production for different refinery products based on the likelihood of different options, as shown in Eq. S52.

$$E(Q_t) = \sum_i \sum_b \left(\sum_y Q_{y,i,t}^{\ b} \times p_{i,t}^{\ b}\right) \quad \text{Eq. S52}$$

Where, *i* denotes the selected refinery configurations, t denotes time between 2019 and 2050, b denotes three considered strategic options, y denotes different refinery products, $E(Q_t)$ is the expected total production from the three selected refinery configurations at time t, $Q_{y,i,t}^{\ b}$ is the production of product y produced by refinery configuration *i* applying strategic option b at time t, $p_{i,t}^{\ b}$ is the likelihood of the strategic option b for configuration *i* at time t.

Supplemental Result 1 – Product Slates, Crude Qualities and Crude Input for U.S. Refineries

Estimated U.S. Regional Volume of Production for Different Refinery Products under Different Transportation Fuel Demand Scenarios.

In the medium petroleum fuel demand (MD) scenario, the total production (including those not intends for transportation usage) of U.S. refining sector is estimated to decrease by 20% in 2035 (from 7.20 Gbbl/y to 5.76 Gbbl/y) and further decrease by 11% in 2050 (from 5.76 to 5.13 Gbbl/y). Comparing the low petroleum fuel demand (LD) scenario to the MD scenario, it is found that the total refinery production is equivalent to 92% of the MD in 2035 (5.32 to 5.76 Gbbl/y) and only 72% in 2050 (3.69 to 5.13 Gbbl/y). Conversely, in the high fuel demand (HD) scenarios, the total refinery production is estimated to be 118% (6.80 to 5.76 Gbbl/y) and 134% (6.87 to 5.13 Gbbl/y) of the MD35 and MD50 scenario, respectively. The decline in gasoline and diesel demands is identified as the primary driver for the decrease in refinery production. According to NREL's EFS (11), gasoline and diesel demand are expected to decrease, while jet fuel demand may remain stable or increase towards 2050. Figure S3 shows that, across all scenarios, gasoline production is projected to decrease by 16% (from 3.69 to 3.09 Gbbl/y) to 47% (from 3.69 to 1.96 Gbbl/y) in 2035 and 21% (from 3.69 to 2.90 Gbbl/y) to 80% (from 369 to 0.747 Gbbl/y) in 2050. Furthermore, the study finds

that PADD 1 refineries produced the most gasoline in the US, making it the region with the highest gasoline production reduction in the future. For instance, the total gasoline production in PADD 1 is expected to decline from 1.20 to 0.405 and 0.255 Gbbl/y HD35 and LD35 scenarios, respectively. The reduced gasoline production in PADD 1 accounts for 55% (LD35) to 132% (HD35) of the total gasoline production reduction.

Estimated U.S. Crude Input Qualities.

Historical data shows that trends of crude APIs are different between PADDs. For example, crudes that are being processed on the East Coast (PADD 1), Gulf Coast and West Coast (PADD 5) are getting lighter over the last few decades. A possible explanation for this trend could be the increasing production of domestic lighter crude (i.e., API > 40) production, which is mainly produced in Texas. For example, over 65% of the volume of crudes that are processed by PADD 3 refineries are imported in 2011, and this percentage has decreased to 30% by 2018. On the contrary, refineries in Mid-West (PADD 2) and Rocky Mountains (PADD 4) are increasingly intaking Canadian crudes that are usually considered heavy, due to geographical advantages. The sulfur content, unlike API, seems to be universally increasing over time across PADDs. Since predicting such dynamic change of crude production/movement is beyond the capacity of this study, we developed two scenarios for both APIs and sulfur contents to capture potential future trend of crude qualities. Detailed results are presented in Figure S2 and Table S4.

Estimated U.S. Configurational Crude input.

The increased adoption of EVs in the U.S. is expected to result in a decline in refineries' throughput as they adjust to changing fuel demands. Under a MD35 scenario, fuel demands are projected to decline from 5.54 to 3.74 Gbbl/y in 2035 and further to 3.04 Gbbl/y in 2050. Consequently, the projected U.S. refineries' throughput will decrease from 16.8 to 13.6 MMbbl/d in 2035 and further decline to 12.2 MMbbl/d in 2050 with heavy/sour crude input. A more severe reduction in refineries' throughput could occur if a greater number of EVs are sold in 2050 (i.e., LD50 scenario), with the total refinery crude input dropping to 8.79 MMbbl/d in 2050, equivalent to only 52.3% of the 2019 level. However, the total refinery crude input could potentially increase from HD35 to HD50 (16.0 vs 16.2 MMbbl/d), as production in PADD 3 and 5 is projected to increase from 2035 to 2050 (see Figure S4). Such an increase is largely driven by differing rates of change in product demand and production across regions between the HD35 and HD50 scenarios (Table S2). For example, in PADD 1, gasoline demand decreases by 78.8 MMbbl/y, yet production drops by only 27.4 MMbbl/y. Similarly, in PADD 3, ULSD production increases by 52.9 MMbbl/y, even though demand rises by just 16.2 MMbbl/y. These regional imbalances between demand and production for specific products contribute to the observed discrepancy: while total fuel demand declines by 33 MMbbl/y, overall refinery production increases by 67 MMbbl/y.

Deep conversion refineries will continue dominating U.S. refining sector regardless of changes in crude quality and product slates. For instance, the volume proportion of hydroskimming refineries will increase by 7.34% (3.02% to 10.4%) if it is for a LD35 scenario with a light and sweet crude input, whilst the volume share of deep conversion refineries could drop by 13.6% (70.6% to 57.0%). The volume shares of each refinery configuration are likely to remain relatively stable, indicating that the configurational portfolio is unlikely to change drastically if refiners continue to follow their previous investment patterns. The quality of crude will slightly impact future refinery crude input. In general, deep conversion refineries are more favorable for heavy/sour crude input compared to light/sweet input. Under a MD35 scenario, the volume percent of crude processed by deep conversion refineries are 7.1% higher for light and sweet crude input compared to heavy and sour crude input (62.3% vs. 55.2%).

The data presented above indicates a significant decrease in the utilization rate of refineries in the United States in the future. Under a MD50 scenario, the refinery crude input in 2050 is projected to be 72.7% of 2019, resulting in a utilization rate slightly above 70% for the entire U.S. refining sector. Refiners are typically required to maintain a utilization rate above 65-70% for safety reasons, and therefore, the shutdown of certain process units or even entire refineries is likely to occur.

As stated by Jiang (47) and Chen (48), the survival of individual refineries will depend on age, size, complexity index, geographic location, interest rate, crude and product prices, and others. The documented data from OGJ (20) shows that 41 shutdowns (including temporary shutdowns) have occurred since 1999, with 27 of them happening for refineries with a capacity between 10,000 to 99,999 MMbbl/d (see Table S31). Refinery configuration also impacts the refiner's decision to shut down their facilities, with 26 out of 41 shutdowns happening for configuration 0 only. Additionally, 27 of those shutdowns occurred in PADD 3 (a total of 12 refineries) and 5 (a total of 15 refineries), indicating that most shutdowns may happen in these regions as their markets are more competitive than other regions.

Hence, it is reasonable to assume that U.S. refiners may prefer to first shut down some of their hydroskimming refineries with lower capacity, as the financial losses would be smaller. Medium/Deep conversion refineries are too significant to fall, and thus refiners are more likely to downgrade/upgrade their process units rather than shut down the entire refinery to adapt to future market demands.

Supplemental Result 2 – GHG Emissions from U.S. Refineries

U.S. Refining Sector GHG Emissions/Energy Consumption Baseline in 2019.

The U.S. refining sector emitted over 212 Mt CO2eq/y in 2019, according to preliminary estimates by PRELIM (see Figure S14). The majority of these emissions, approximately 56.1% (i.e., 119 Mt CO_2eq/y), result from fuel combustions, while approximately 13.6% (28.9 Mt CO_2eq/y) result from FCC coke burn-off. The sector's total on-site energy consumption was estimated to be 2828 PJ/y,

732 less than the EPA's previous estimate of 3400 PJ/y for 2006. PRELIM's calculation of GHG
733 emissions accounts for the assumption that all hydrogen and steam consumed by refineries are
734 produced on-site through steam methane reformer and steam boiler facilities, and electricity is
735 exclusively purchased from natural gas power plants. In addition, emissions resulting from natural
736 gas upstream activities, including extraction and transportation, as well as electricity generation,
737 are included in the estimates. Although there is a slight variation between PRELIM's estimation and
738 the EPA's reported emissions, PRELIM's estimated emissions are relatively consistent with the
739 EPA's reported emissions for 2019 (212 Mt CO_2eq/y versus 165 Mt CO_2eq/y). Therefore, to
740 maintain internal consistency in this study, we have chosen to use PRELIM's estimates as our
741 baseline. The configuration of the U.S. refining sector varies across PADD, with medium conversion
742 refineries (configuration 1) being the dominant configuration in PADD 1, and deep conversion
743 refineries (configuration 4 for PADD 2 and PADD 4, configuration 6 for PADD 3 and PADD 5) being
744 the predominant configuration in the other four PADDs. However, deep conversion refineries are
745 responsible for the majority of greenhouse gas (GHG) emissions in the United States, contributing
746 to approximately 73.1% (155 Mt CO_2eq/y) of emissions. Among deep conversion refineries,
747 configuration 6 refineries account for almost two-thirds of these emissions (101 Mt CO_2eq/y), with
748 configuration 4 refineries contributing the remaining one-third (50.3 Mt CO_2eq/y). In contrast,
749 medium and hydroskimming refineries only account for 26% (55.6 Mt CO_2eq/y) and 1.98% (4.21
750 Mt CO_2eq/y) of GHG emissions, respectively. Configuration 1 and configuration 3 are responsible
751 for the highest GHG emissions among medium conversion refineries, contributing 25.1 and 22.1
752 Mt CO_2eq/y, respectively. The predominance of FCC units indicates that the majority of U.S.
753 refineries are oriented toward gasoline production, which may pose challenges for refiners
754 attempting to reduce gasoline output. Although the share of each type of refinery configuration is
755 similar, regional GHG emissions remain relevant to total refinery production. PADD 3, which emits
756 the most among the five regions, accounts for 52.8% (112 Mt CO_2eq/y) of total U.S. refining sector
757 emissions, with deep conversion refineries contributing 57% (63.4 to 112 Mt CO_2eq/y) of regional
758 emissions. PADD 4, on the other hand, has the lowest emissions of all the regions, emitting only
759 7.91 Mt CO_2eq/y.

760 **U.S. Refining Sector GHG Emissions under Different Transportation Fuel Demand**
761 **Scenarios.**

762 The estimated GHG emissions from the U.S. refining sector under varying transportation fuel
763 demand scenarios and crude input qualities are presented in Figure S15 and Table S13. The
764 analysis indicates that GHG emissions for light/sweet crudes are comparable to those for
765 heavy/sour crudes. As fuel demand decreases from 5.54 to 4.71 Gbbl/y (i.e., from the BL to LD35
766 scenario), GHG emissions decline from 212 to 208 Mt CO2eq/y for heavy/sour crude and from 212
767 to 198 Mt CO2eq/y for light/sweet crude. However, under both crude quality conditions, sectoral
768 GHG emissions exhibit a slight increase (from 208 to 211 Mt CO2eq/y and from 198 to 204 Mt

CO2eq/y for heavy/sour and light/sweet crude, respectively) as fuel demand further declines from 4.71 Gbbl/y to 4.58 Gbbl/y. This increase is primarily attributed to the projected rise in crude input volumes from the LD35 to LD50 scenario (i.e., from 15,767 to 16,218 mmbbl/y for light/sweet crude).

Additionally, sectoral GHG emissions for light/sweet crude consistently remain lower than those for heavy/sour crude. For instance, the U.S. refining sector is estimated to emit 204 Mt CO2eq/y when processing light/sweet crude, which is approximately 7 Mt CO2eq/y lower than the emissions observed when processing heavy/sour crude. This disparity arises because processing light/sweet crude requires less energy, as it involves less heat to convert long-chain hydrocarbons into short-chain, high-value products and less hydrogen to remove sulfur content (See Figure S16).

Supplementary Result 3 - CO_2 Avoidance Costs for Different Carbon Capture Strategies.

The CO_2 avoidance cost varies significantly depending on different process units, refinery configurations, locations, years, and fuel demand scenarios. However, it appeared that the lowest CO_2 avoidance cost will always be achieved by FCC units in deep conversion refineries from PADD3, regardless of year and fuel demand scenarios. For example, the lowest CO_2 avoidance costs when facing a heavy/sour crude input are 86.4, 86.9, 85.2, 84.3, 84.9 and 84.7 $/tCO2, for LD35, LD50, MD35, MD50, HD35 and HD50 scenarios respectively. In fact, FCCs are estimated to be the cheapest process units to deploy carbon capture. For example, most cases (118 out of 172, 68.6%) of minimum CO_2 avoidance costs are achieved by FCC under a heavy/sour crude scenario, ranging from 84.2 to 440 $/t$CO_2$. This is attributed to two reasons. First of all, the large amount of GHG emissions produced by the units sparse the cost in most cases. Secondly, the CO_2 concentration in the exhaust gas from coke combustion is high, which reduced the size & energy consumption for the carbon capture unit.

Following FCC, SMR also presents significant potential for deploying carbon capture at a lower cost. 26 out of 172 cases (i.e., 15.1%) has the lowest CO2 avoidance cost achieved by capturing from SMR. For instance, in HD35 scenario, the CO2 avoidance cost for SMR could be as low as 121 $/tCO2, achieved in a medium conversion refinery in PADD 3. Among these 26 cases, the avoidance cost by capturing from SMR ranges from 120 – 239 $/t CO2. A major factor contributing to the low cost of carbon capture on SMR is the high CO2 concentration in its exhaust streams, a characteristic similar to that of the FCC. Still, our estimates on the CO2 avoidance cost for capturing CO2 from SMR are also higher than other literature (2, 49–57). In our modeling framework, CO_2 is assumed to be captured downstream of the reformer furnace, where tail gas from the pressure swing adsorption (PSA) unit is blended with natural gas and combusted to supply process heat. Typically, 3 gas streams are considered for carbon capture in SMR: i) the syngas stream to the hydrogen purification unit, characterized by the highest CO_2 partial pressure (~2.8 MPa); ii) the tail gas from the purification unit, with the highest CO_2 concentration (potentially up to ~50 mol%); and

iii) the flue gas from the reformer furnace, which has the highest CO_2 mass flow (3, 58). Capturing CO_2 from streams i) and ii) offers lower capture costs due to favorable pressure and CO2 concentrations but limits the overall capture rate to approximately 50–60%. In contrast, capture from stream iii) enables higher overall capture rates - up to 90% -but incurs higher costs. Specifically, the reported CO_2 avoidance cost ranges from €47–66/t (i.e., $57 – 81/t in 2019) for options i) and ii), while capture from the flue gas stream results in a cost of €70/t (i.e., $87/t in 2019).

Under the MD50 scenario (medium transportation fuel demand projected for 2050), the total emissions reduction potential from carbon capture across U.S. refineries is estimated at 94.0 $MtCO_2/y$. SMR contributes approximately 10.8 $MtCO_2/y$, or 11.6% of the total—broadly consistent with values reported in the literature. Our analysis estimates a minimum CO_2 avoidance cost of $107/t for SMR for capturing from furnace outlet, which is moderately higher than the $87/t reported in Collodi et al (3). For capturing CO_2 from PSA inlet and outlet, the minimum CO_2 avoidance costs are 94.8 and 77.6 $/t, respectively (see Table S33). Several reasons contribute to this discrepancy. First, our estimate incorporates the cost of CO_2 transportation and storage (approximately $11/t combined), along with CO_2 leakage from pipelines (ranging from 11.3 $ktCO_2/y$ to 173 $ktCO_2/y$), both of which are excluded in Collodi et al (3). Secondly, our CapEx assumptions are different. Following the approach of Gale et al (2), we include the cooling water and interconnecting sections for carbon capture unit. In Gale's analysis, these components together account for approximately 38% of total CapEx—$62,200k out of $163,000k. In our study, the interconnecting section alone accounts for 52% of total CapEx—$317,000k out of $614,000k. This is largely due to the higher CO_2 flow rate in our case (627 t/h vs. 102 t/h in Gale et al.), which significantly increases the scale and cost of shared infrastructure. Despite this, the difference in CO_2 avoidance cost between our study ($107/t) and Collodi et al. ($87/t) remains relatively modest. This is partly because Collodi et al. include the cost of building a standalone H_2 production facility, which accounts for 24% of their total CapEx—$99,700k out of $409,000k. This partially offsets the effect of our higher interconnection-related costs, resulting in broadly comparable cost estimates.

Our analysis further shows that CO_2 avoidance costs for SMR vary significantly across refinery configurations. While SMR is generally among the most cost-effective capture targets in deep conversion refineries, our absolute costs are somewhat elevated due to the inclusion of broader cost components. Importantly, SMR emissions vary with refinery H_2 demand, and in configurations with limited H_2 use, the poor economies of scale can render SMR a suboptimal capture target. For instance, in a medium conversion refinery with only FCC in PADD 4, the CO_2 avoidance cost for SMR is $442/t, the highest among all units in that configuration, with only 10.0 $ktCO_2/y$ avoided. In contrast, the same configuration in PADD 3 results in a much lower cost of $125/t, with 467 $ktCO_2/y$ avoided, making SMR the second most cost-effective unit for capture (FCC: $115/t). More broadly, configurations with high H_2 consumption—such as deep conversion or medium conversion

refineries with hydrocrackers—consistently rank SMR as one of the lowest-cost capture options. In PADDs 3 and 5, SMR represents the least-cost capture unit in deep conversion refineries with hydrocrackers, with avoidance costs ranging from $122–238/t. Conversely, in low-H_2 configurations (e.g., hydroskimming refineries, or medium and deep conversion refineries with only FCC), SMR can become the most expensive capture option. For instance, in hydroskimming refineries under a low fuel demand scenario in 2050, SMR capture costs reach $1,890/t in PADD 1, $311/t in PADD 3, and $2,800/t in PADD 4, with respective CO_2 avoidance potentials of 1.34 kt/y, 11.3 kt/y, and 0.540 kt/y.

These additional factors contribute to a higher overall cost of avoidance compared to values typically reported in literature.

Other process units also exhibit potential for achieving the lowest CO_2 avoidance costs in specific refinery configurations. For example, 14.0% (24 out of 172) of the minimum CO_2 avoidance costs, ranging from 159 – 493 $/tCO2, are achieved by hydrotreaters. All these cases occur in hydroskimming refineries, typically where other low-cost CO_2 avoidance units, such as FCC or SMR, are unavailable. Additionally, 3 instances of minimum CO_2 avoidance costs, ranging from 262 – 552 $/t$CO_2$ are achieved by ADU in hydroskimming refineries.

Contrastingly, the steam boiler is considered the most-costly process unit for CO_2 capture among various units in half of scenarios (86 out of 172). For instance, the CO_2 avoidance cost for a steam boiler is 156 $/tCO2, while it is 84.3 $/t$CO_2$ for FCC in MD35 scenario. The avoidance cost for capturing CO_2 from steam boilers ranges from 154$/t CO_2 (configuration 6 in PADD3, HD35) to 1313 $/t$CO_2$ (configuration 0 in PADD2, MD35). This high expense is primarily due to its lower GHG emissions, which are only 0.001 to 1.26 kt/y of CO_2 avoided. Detailed results are presented in Figure S17 and Figure S18.

From a location-wise perspective, the lowest average CO_2 avoidance cost is achieved by refineries in PADD 3, 160$/t$CO_2$, under a HD50 scenario, while the highest average CO_2 avoidance cost is achieved in PADD 1, 519 $/t$CO_2$, under a LD50 scenario. The location factor considered in this study as well as the total amount of GHG emissions are the two major reasons that contribute to the difference in CO_2 avoidance cost between PADDs. Detailed results are shown in Figure S19 and Figure S20.

A sensitivity analysis was also conducted to evaluate the impact of variability in CO_2 transport and storage costs on the overall avoidance cost. Figure S27 presents the distribution of CO_2 avoidance costs under low, baseline, and high cost assumptions for both transportation and storage. For transportation, the average avoidance cost increases modestly from 281 to 290 $/t$CO_2$ across the low to high cost range, corresponding to transport cost assumptions of 2.09 - 10.5 $/t$CO_2$, reflecting variation in pipeline distance and system conditions (59). Similarly, for storage, the average avoidance cost increases from 283 to 312 $/t$CO_2$ across a wider cost range of 1.49 - 30.9 $/t$CO_2$, representing the difference between onshore and offshore storage (60). The results indicate that

variations in CO_2 transport and storage costs lead to relatively limited changes in total avoidance cost, with increases of approximately 9 \$/tCO_2 for transportation and 29 \$/tCO_2 for storage between low and high scenarios. The influence of transport cost is relatively modest, whereas storage cost introduces a more noticeable shift. Overall, the magnitude and distribution of avoidance costs remain broadly consistent, suggesting that the main conclusions are preserved while confirming that transport and storage remain relevant factors that should be considered.

Supplementary Result 4 - Additional Results for Real Option Analysis

Additional Real Option Results for Two Different Carbon Price Trajectories

Under the A-CP scenario, the simulated mean carbon price reaches approximately 153 \$/tCO_2 by 2033, increases to 252 \$/tCO_2 by 2043, and approaches 270 \$/tCO_2 by 2050. In contrast, under the D-CP scenario, the mean carbon price remains at approximately 85 \$/tCO_2 until 2036, increases to 145 \$/tCO_2 by 2040, and reaches 259 \$/tCO_2 by 2050. For comparison, the mean carbon price under the R-CP scenario reaches 150 \$/tCO_2 by 2034, 206 \$/tCO_2 by 2042, and 284 \$/tCO_2 by 2050. Differences in the timing and shape of these carbon price trajectories result in modest but systematic changes in the modeled responses of refinery archetypes.

Figure S26 and Figure S27 shows that, across all carbon price scenarios, the hydroskimming refinery archetype continues to face a high likelihood of shutdown after the early 2030s, largely independent of the specific carbon price trajectory. Under the A-CP scenario, higher carbon prices at earlier stages do not incentivize CCS deployment for the hydroskimming archetype under less favorable market conditions (e.g., the medium fuel demand scenario). In contrast, medium and deep conversion refinery archetypes consistently exhibit a greater propensity to deploy CCS than the hydroskimming archetype across scenarios (See Figure S26).

Notably, the stepwise carbon price trajectory in the A-CP scenario—characterized by higher prices in earlier years—reduces the likelihood of shutdown for the medium conversion refinery archetype under the high fuel demand scenario relative to the other carbon price trajectories. Conversely, under the D-CP scenario, the medium conversion refinery archetype exhibits a higher shutdown risk under the high fuel demand scenario compared with the R-CP case, with the shutdown probability increasing from 5.08% to 16.2% by 2047, whereas it remains below 10% (8.45%) until 2049 under the R-CP scenario (see Figure S27). This increase in shutdown likelihood is driven by the sudden escalation of carbon prices in the late 2030s and early 2040s, which imposes additional financial pressure on refineries during a period of tightening crack spreads relative to the 2020s and early 2030s. A similar increase in shutdown risk associated with delayed carbon price escalation is also observed for the deep conversion refinery archetype under low and medium fuel demand scenarios.

Overall, these results indicate that the qualitative insights regarding refinery shutdown risk and CCS deployment are broadly robust to differences in the timing and stepwise structure of carbon price escalation, provided that carbon prices continue to rise and converge to comparable levels by 2050.

Expected Total GHG Emissions by Selected Refinery Configurations.

In 2019, the estimated total GHG emissions for hydroskimming (configuration 0), medium conversion: FCC & GO-HC (configuration 3), and deep conversion: FCC & GO-HC (configuration 6) refineries were 3.40, 22.1, and 101 Mt CO2eq/yr, respectively. The likelihood of deploying carbon capture varies significantly across different transportation fuel demand scenarios, as detailed in Dataset 5, influencing expected GHG emissions accordingly. For instance, under a low fuel demand scenario (i.e., BL to LD35/50), the probability of deploying carbon capture in a hydroskimming refinery remains extremely low, not exceeding 0.005% until 2028, peaking at 0.08% in 2033, and diminishing by 2048. Concurrently, the probability of refinery shutdown increases from 0% to 20.6% between 2027 and 2028, and then sharply to 99.9% by 2029. This shift leads to a marked reduction in total GHG emissions from the hydroskimming refinery, which initially increase from 3.40 to 6.03 Mt CO_2/year from 2019 to 2027 due to capacity expansion (from 507 to 862 MMbbl/d), before abruptly dropping to 5.06 Mt CO_2/year by 2028 and then to 0.008 Mt CO_2/year by 2029. Conversely, in a medium fuel demand scenario (i.e., BL to MD35/50), the percentage of deploying carbon capture at a hydroskimming refinery surged from 0% to 19.2% between 2021 and 2022, gradually reaching a peak of 53.7% by 2030, and then declining to negligible levels by 2046. Simultaneously, the likelihood of shutting down the refinery escalated from 0% to 97.1% by 2031 and continued to increase until it reached 100% by 2047. Despite a significant expansion in refinery capacity from 507 to 1093 MMbbl/d, these changes lead to a gradual increase in GHG emissions, from 3.40 Mt CO_2/year in 2019 to 5.35 Mt CO_2/year by 2030. Given the higher probability of deploying carbon capture, a less steep reduction in GHG emissions is noted as fuel demand increases. Notably, between 2028 and 2030, there is a marked increase in the likelihood of transitioning from continued operation to shutting down the refinery for hydroskimming refinery. This suggests that the window for considering carbon capture as a viable option is preferable before this critical time point.

A two-stage reduction in GHG emissions is observed in the deep conversion refinery as well. GHG emissions remained above 105 Mt CO_2/year from 2019 to 2021. In 2022, a sudden drop to 93.8, 95.9, and 99.2 Mt CO_2/year for the low, medium, and high fuel demand scenarios, respectively, was induced by an increased probability of deploying carbon capture. Subsequently, emissions further declined to 2.97 Mt CO_2/year in 2031, 6.07 Mt CO_2/year in 2032, and 24.3 Mt CO_2/year in 2036 for low, medium and high fuel demand scenario, respectively. This trend precedes another steep decline triggered by a rising probability of refinery shutdown.

As indicated in Figure S21, it is projected that the expected GHG emissions will likely align with the sectoral net-zero goal by 2050, and by 2100 for all configurations. However, achieving the net-zero

goal by 2035 appears unlikely. Although expected GHG emissions do not directly represent the actual future changes in emissions by the U.S. refining sector—owing to the dynamic market affected by individual refinery decisions—we can still derive insights into future emission trends. The GHG emissions from the U.S. refining sector are expected to decline, facilitated by increased utilization of low-carbon technologies within a specific timeframe. As market conditions evolve, operating refineries may become unprofitable, prompting refiners to consider shutting down or exiting the market. A massive closure of refineries will lead to a shortage in the supply of refinery products in the U.S. market (see Figure S22). This gap between supply and demand will likely be bridged by increasing imports and boosting domestic production. Consequently, the sectoral GHG emissions are unlikely to reach net-zero before 2050, indicating that additional efforts are necessary to achieve net-zero emissions within the refining sector.

It is also worth noting that while this study assesses how a refinery could react to economic signals (i.e., the expected present value changes driven by carbon pricing) reflecting how aggressively society may pursue decarbonization, the specific mechanism by which this signal is delivered can vary. Rather than a simple carbon price, it could be implemented via tax incentives (e.g., under the Inflation Reduction Act), other carbon pricing mechanisms, or performance-based regulations.

Figures

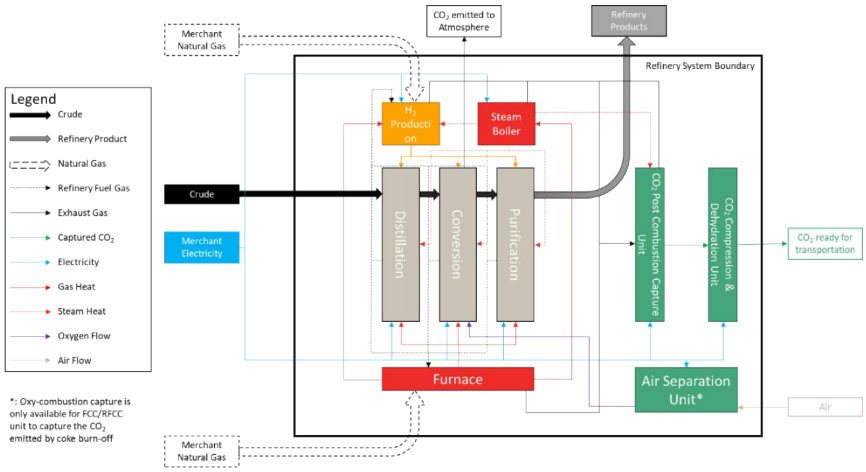


Figure S1. PRELIM and Carbon Capture Module

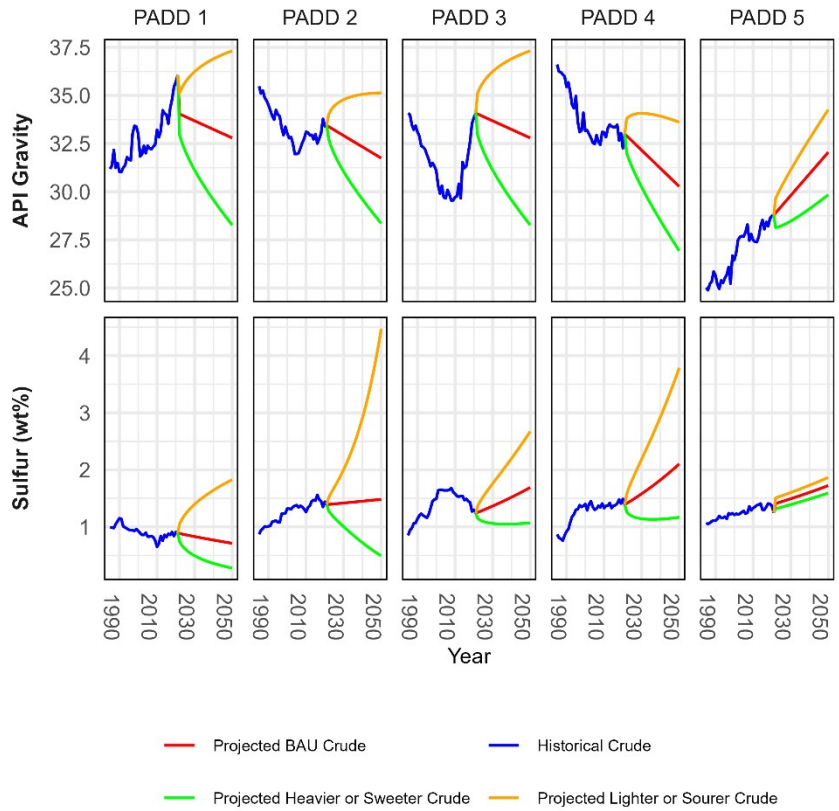


Figure S2. Projected Crude Qualities Using ETS

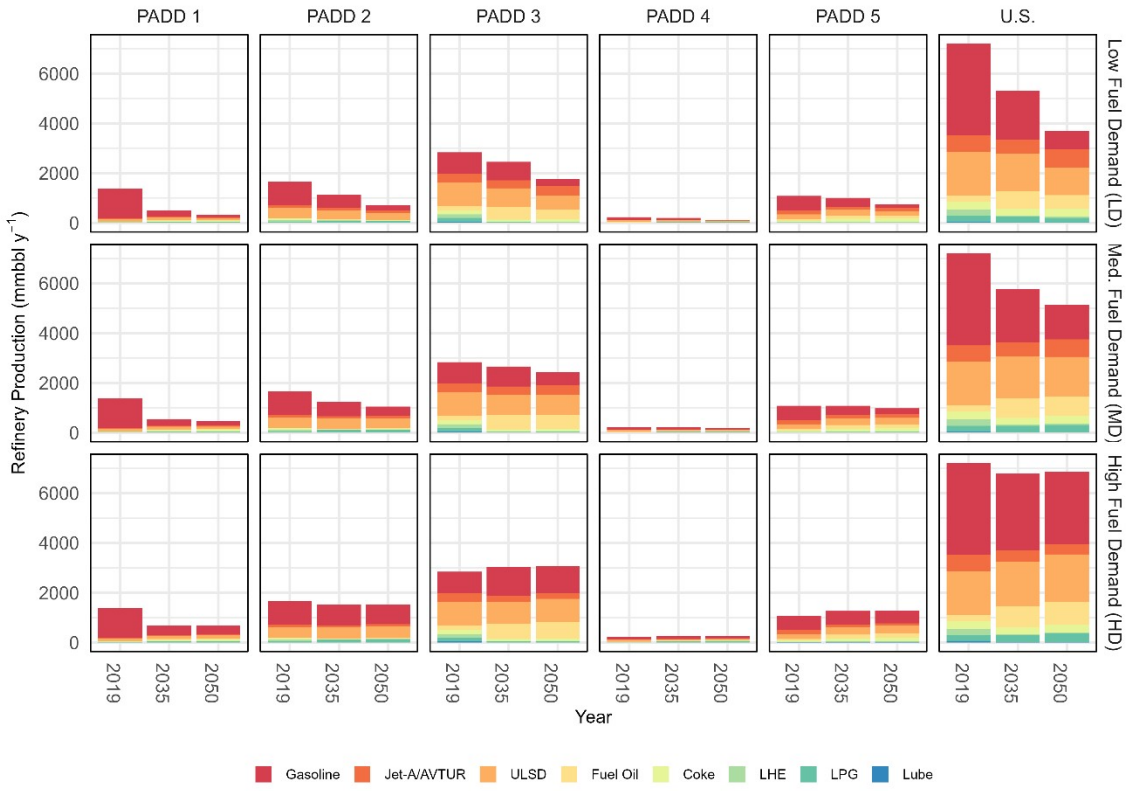


Figure S3. Projected U.S. Refinery Product Slates under Different Transportation Fuel Demand Scenarios.

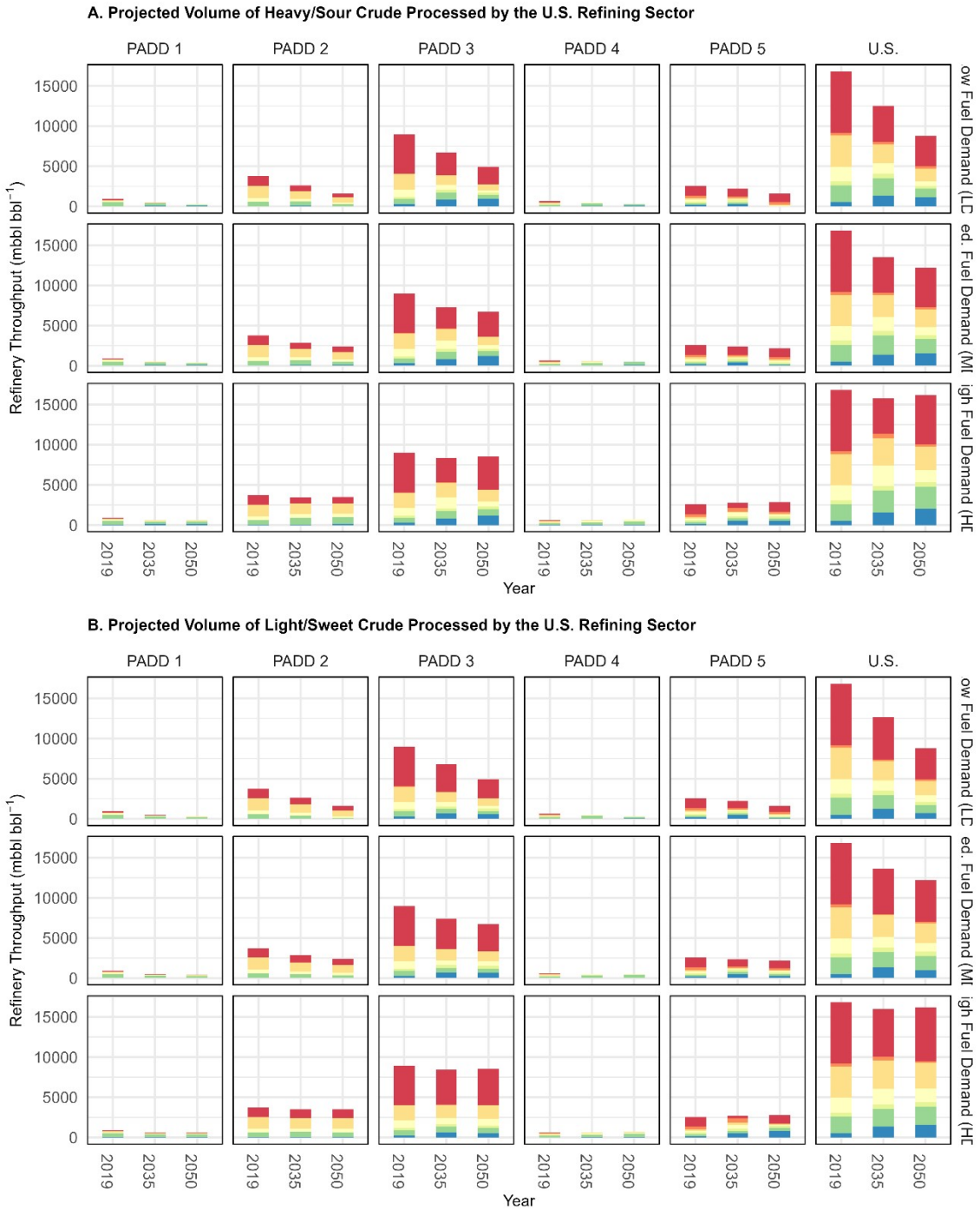
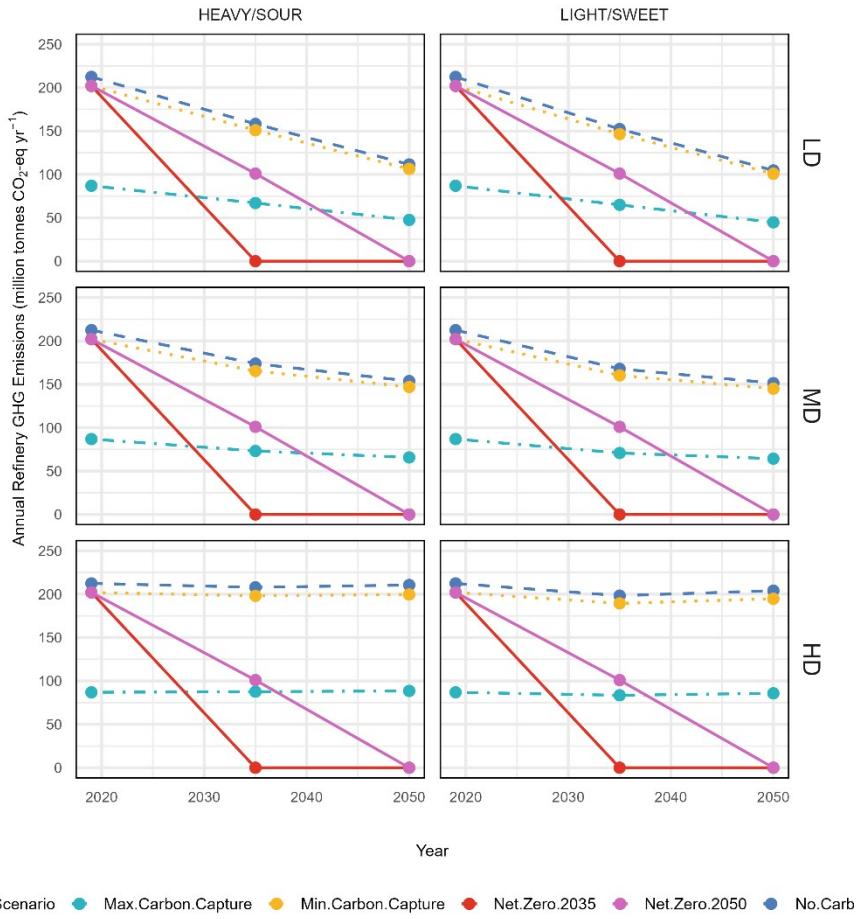


Figure S4. Estimated Future U.S. Refineries Crude input Changes. a) Light/Sweet Crude; b) Heavy/Sour Crude.



983

984 Figure S5. Estimated Annual U.S. Refineries GHG Emissions based on Different Decarbonization Goals.

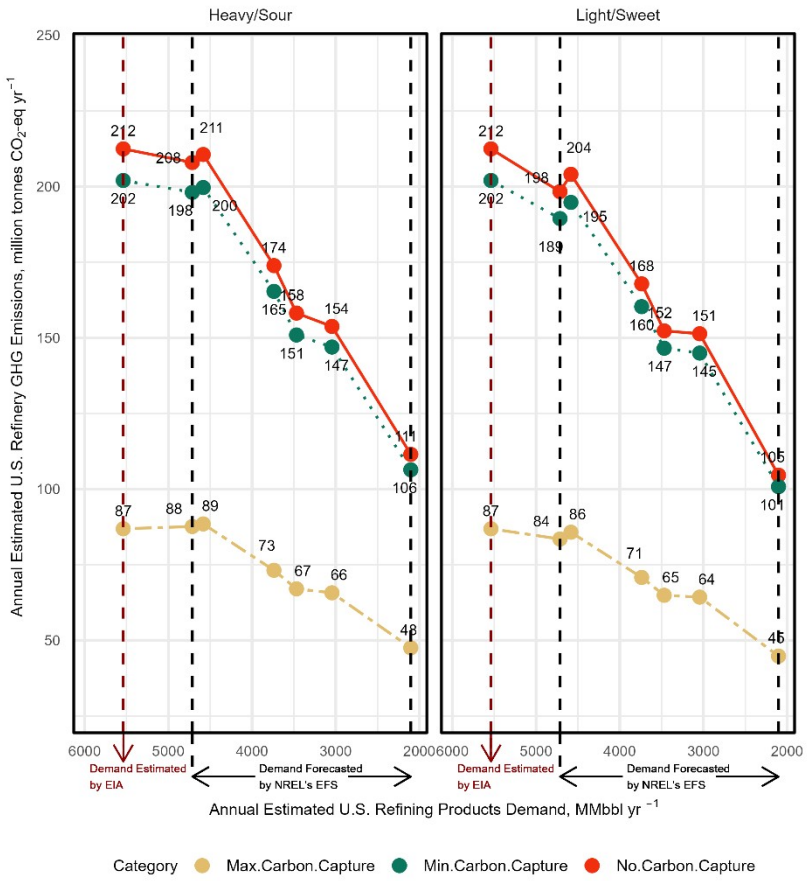


Figure S6. Annual Estimated GHG Emissions for U.S. Refining Sector under Different Refining Product Demands.

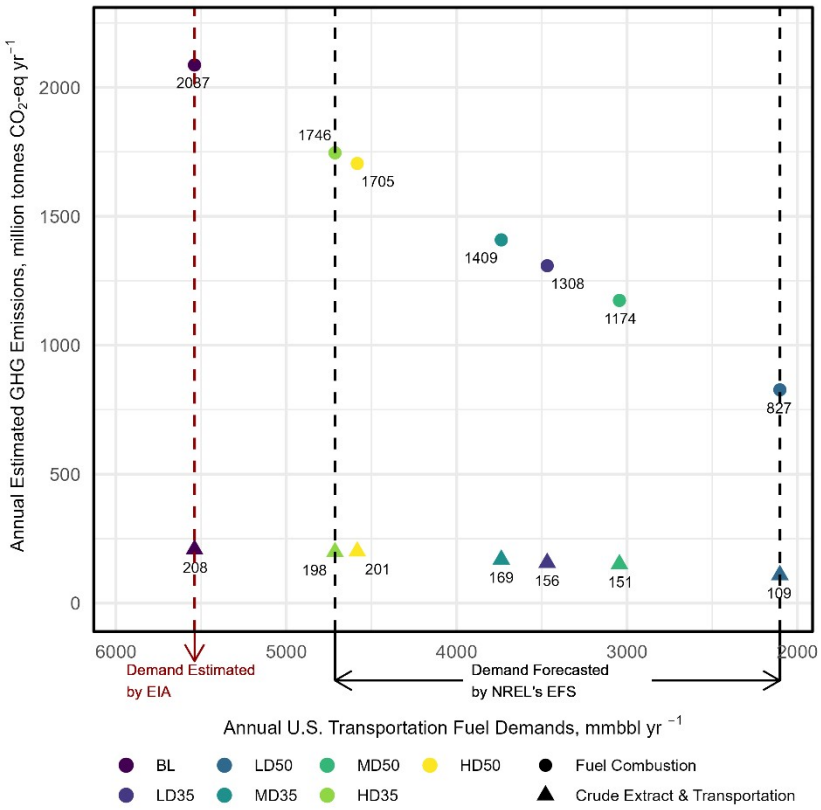


Figure S7. Estimated Annual GHG Emissions by Direct Combustion of Transportation Fuels and Crude Extraction and Transportation under Different Fuel Demand Scenarios.

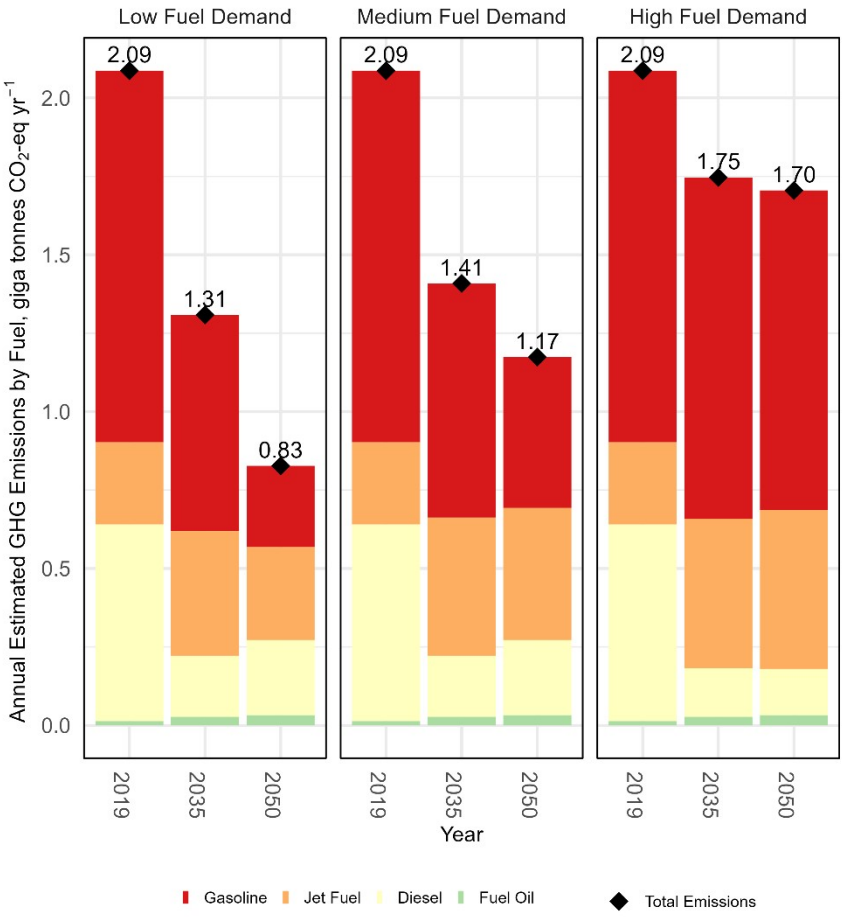


Figure S8. Breakdown of Estimated Annual GHG Emissions by Direct Combustion of Transportation Fuels under Different Fuel Demand Scenarios.

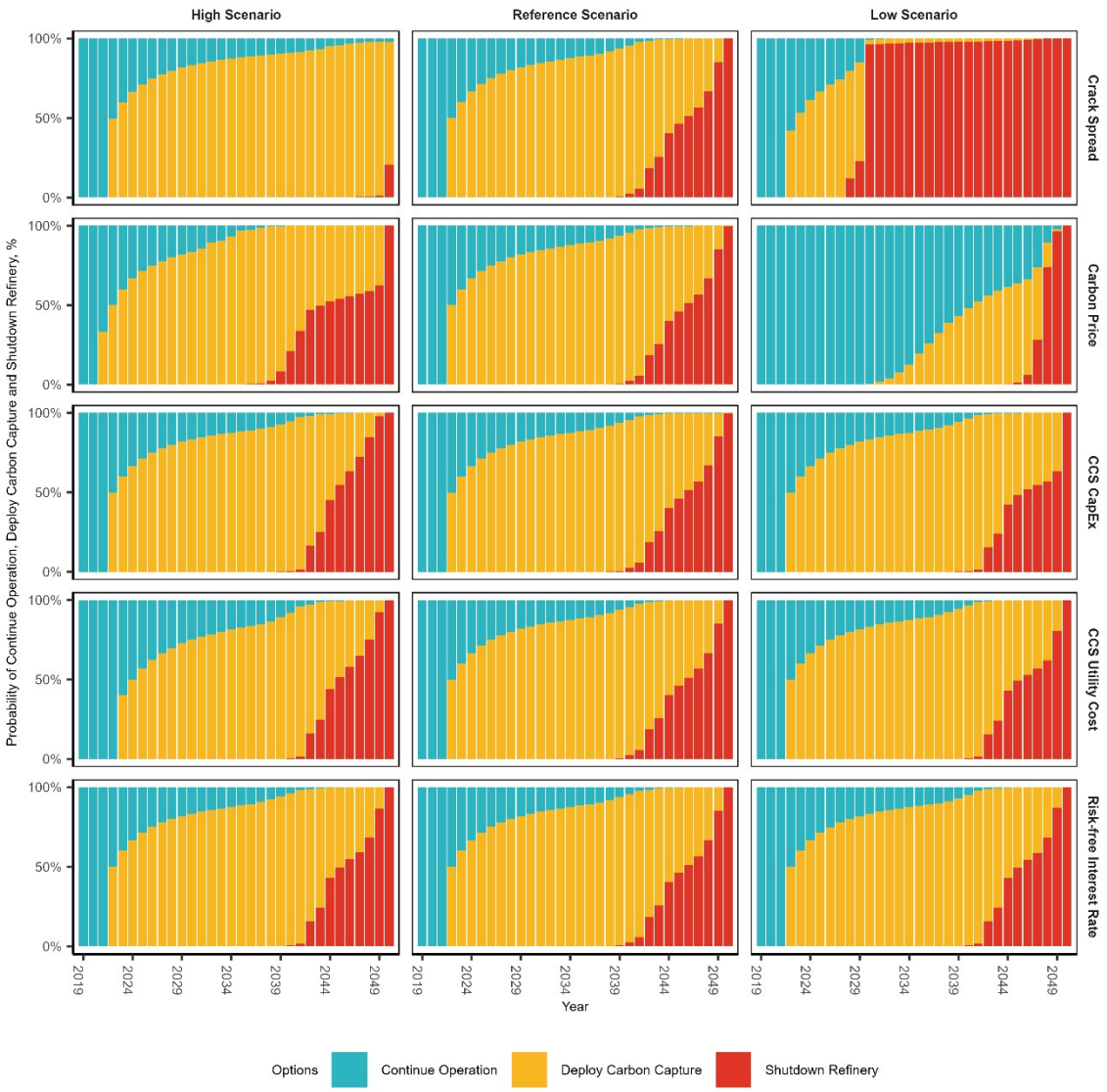


Figure S9. Sensitivity Analysis on Real Option

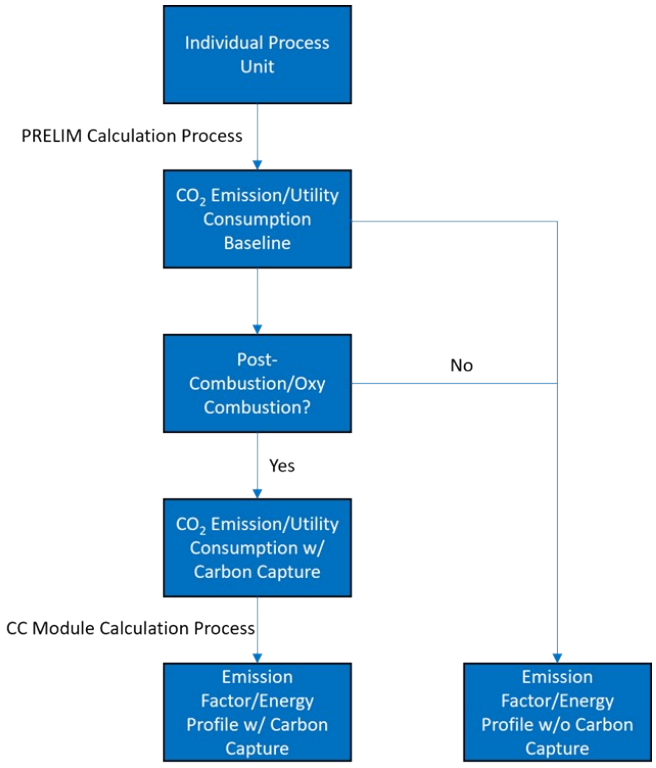


Figure S10. The General Computing Process for Carbon Capture Module

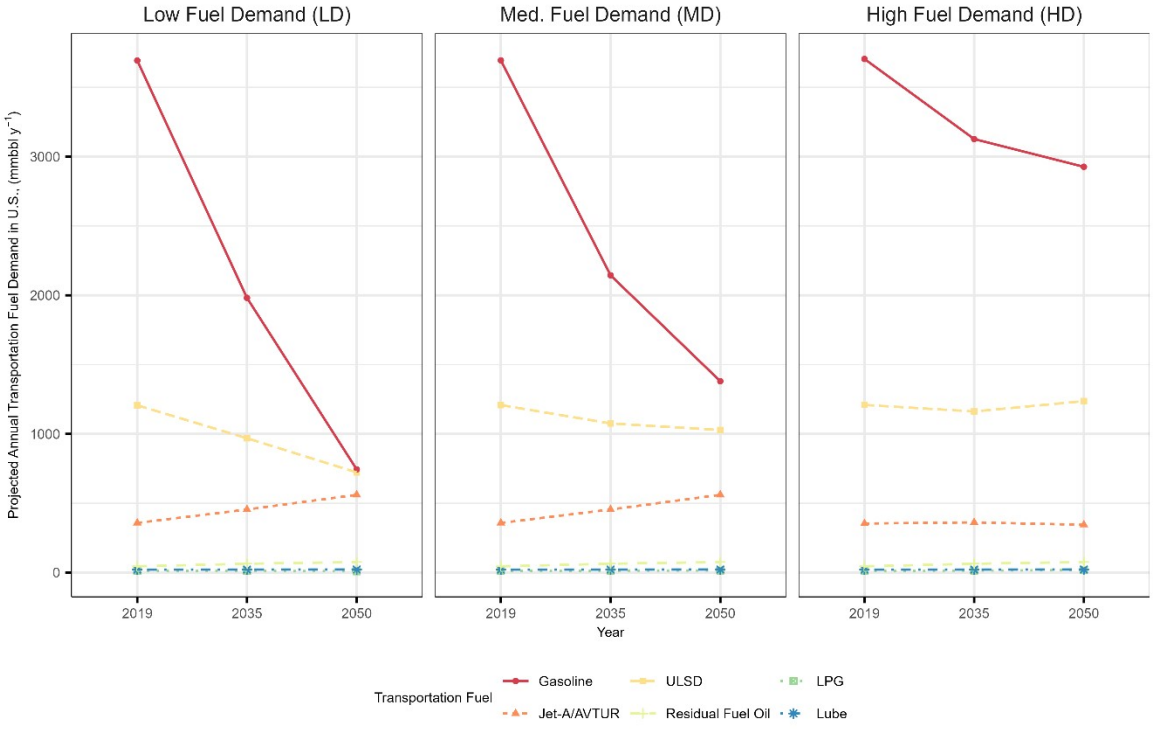


Figure S11. Projected Annual Transportation Fuel Demand in the U.S. from 2019 to 2050.

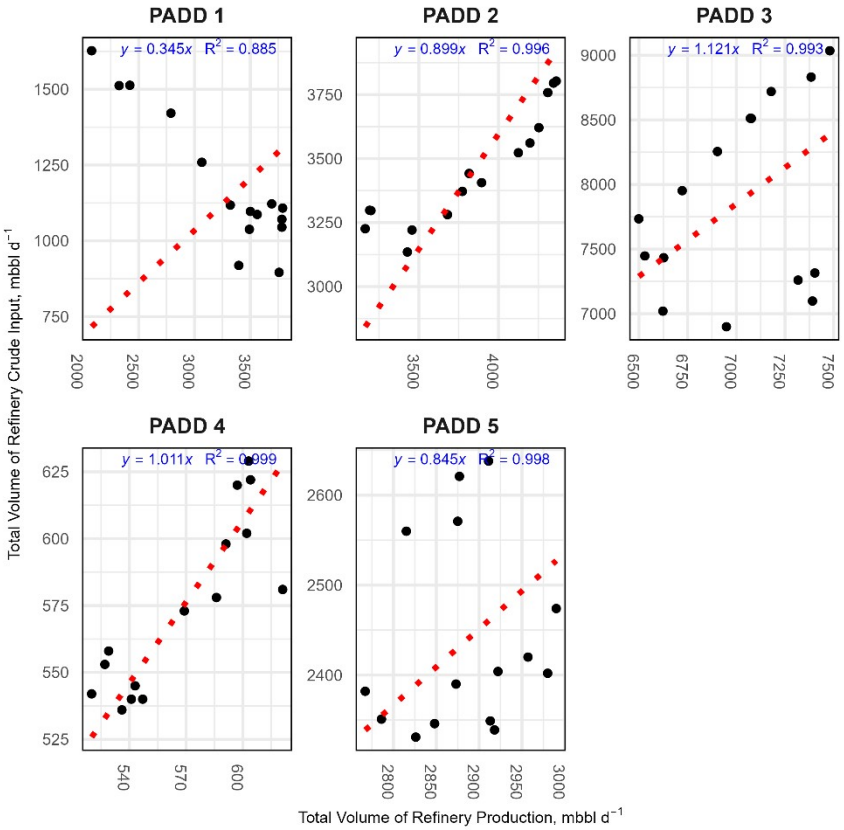


Figure S12. Linear Correlations between Refinery Historical Input and Output

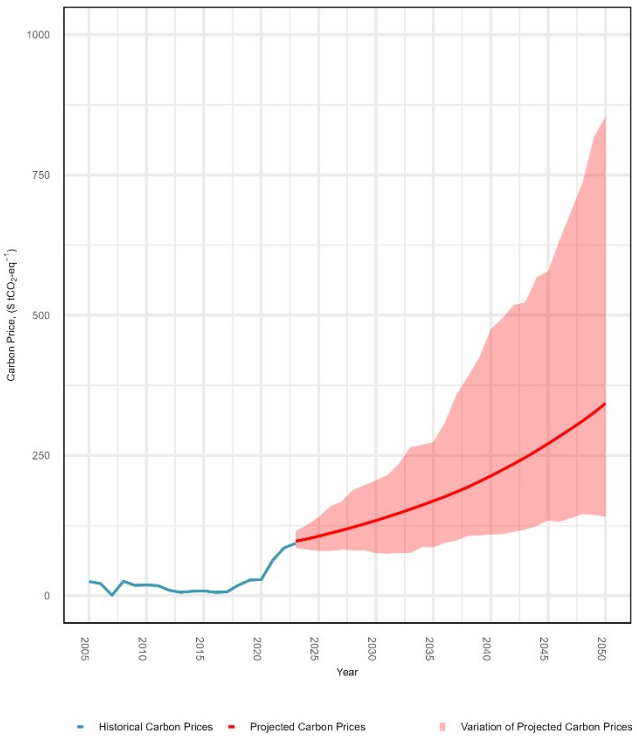


Figure S13. Simulated Carbon Price Following a GBM Process

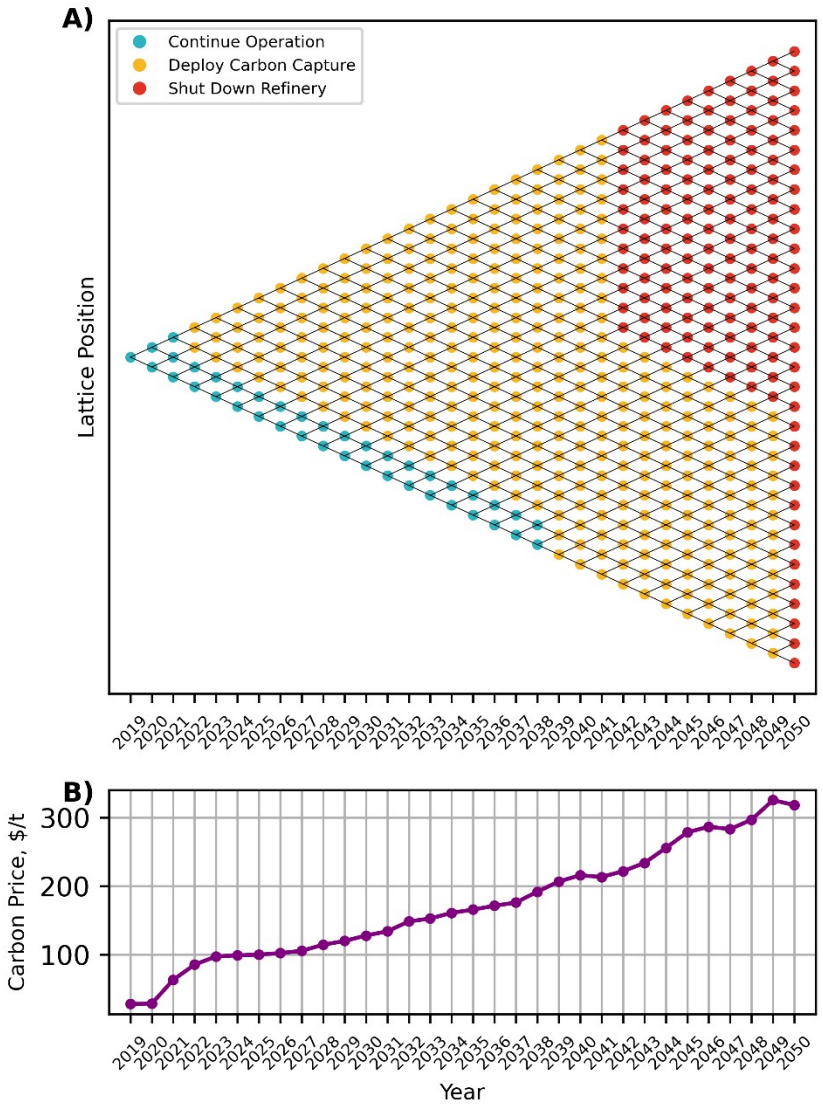


Figure S14. An Established Binomial Decision Tree for a Specific Carbon Price Trajectory.

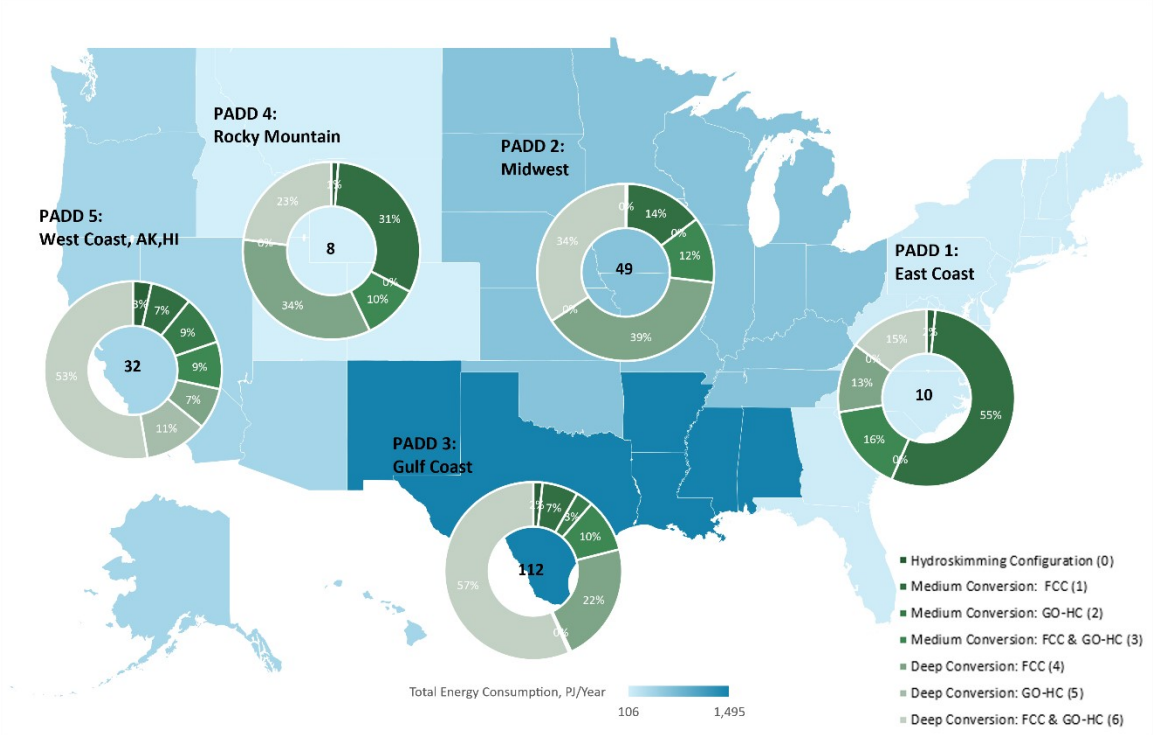


Figure S15. 2019 U.S Refining Sector Emission & Energy Consumption Profile. The opacity of the map indicates the annual refinery energy consumption for each region, PJ/y. The pie chart indicates the share of GHG emissions by each type of refinery configuration in the region. The numbers in the circle indicate daily GHG emissions from refineries in each region, MMtCO2eq/d.

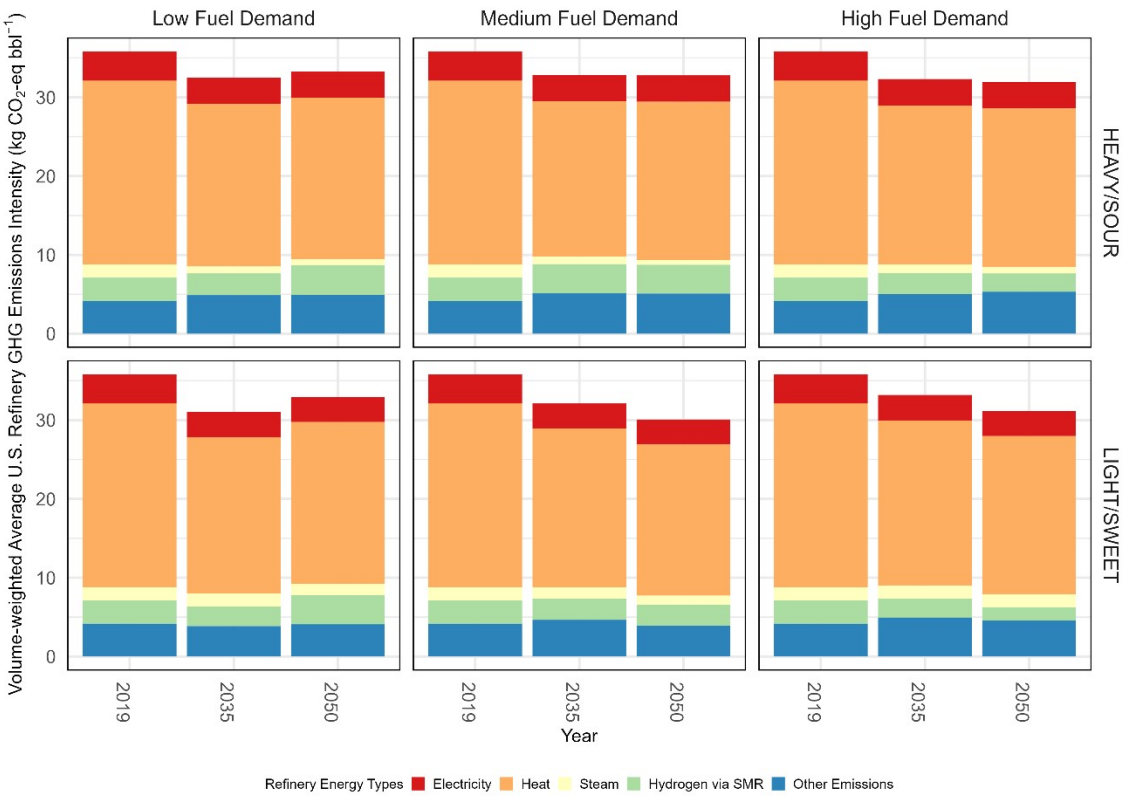


Figure S16. Projection on Refinery Emissions Intensity under Various Transportation Fuel Demand and Different Crude Qualities.

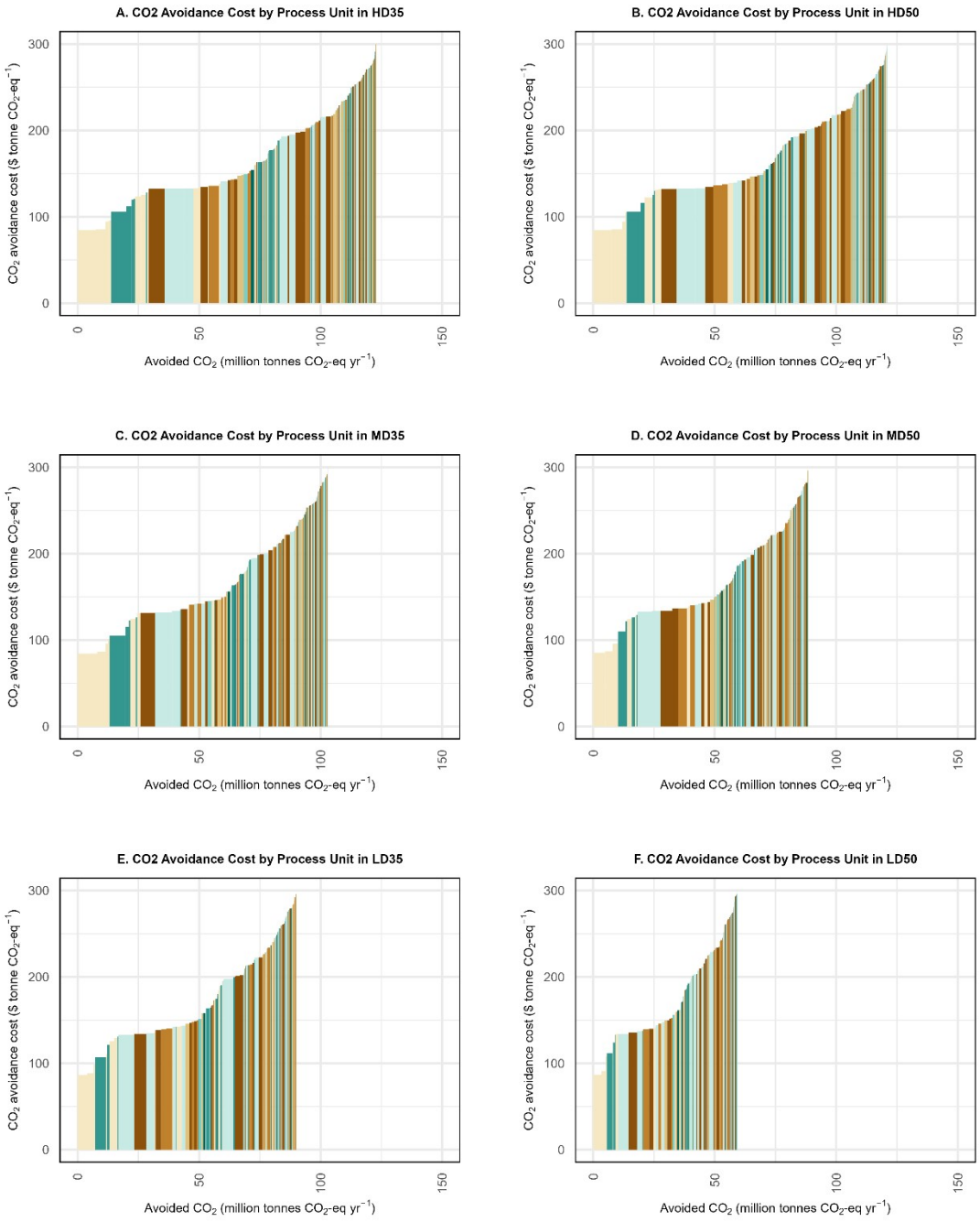


Figure S17. CO_2 Avoidance Cost Curve for Heavy/Sour Crudes

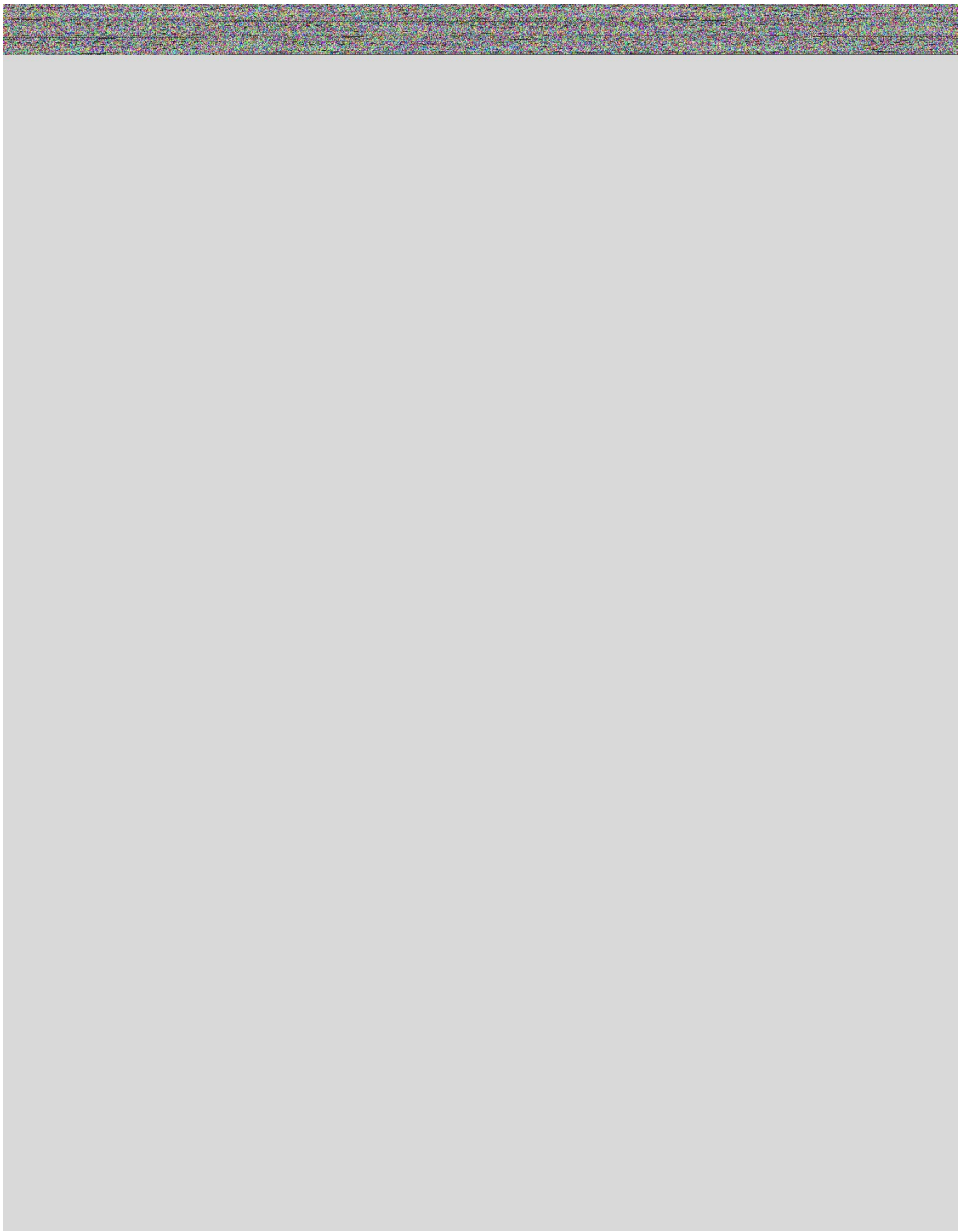


Figure S18. CO_2 Avoidance Cost Curve for Light/Sweet Crudes

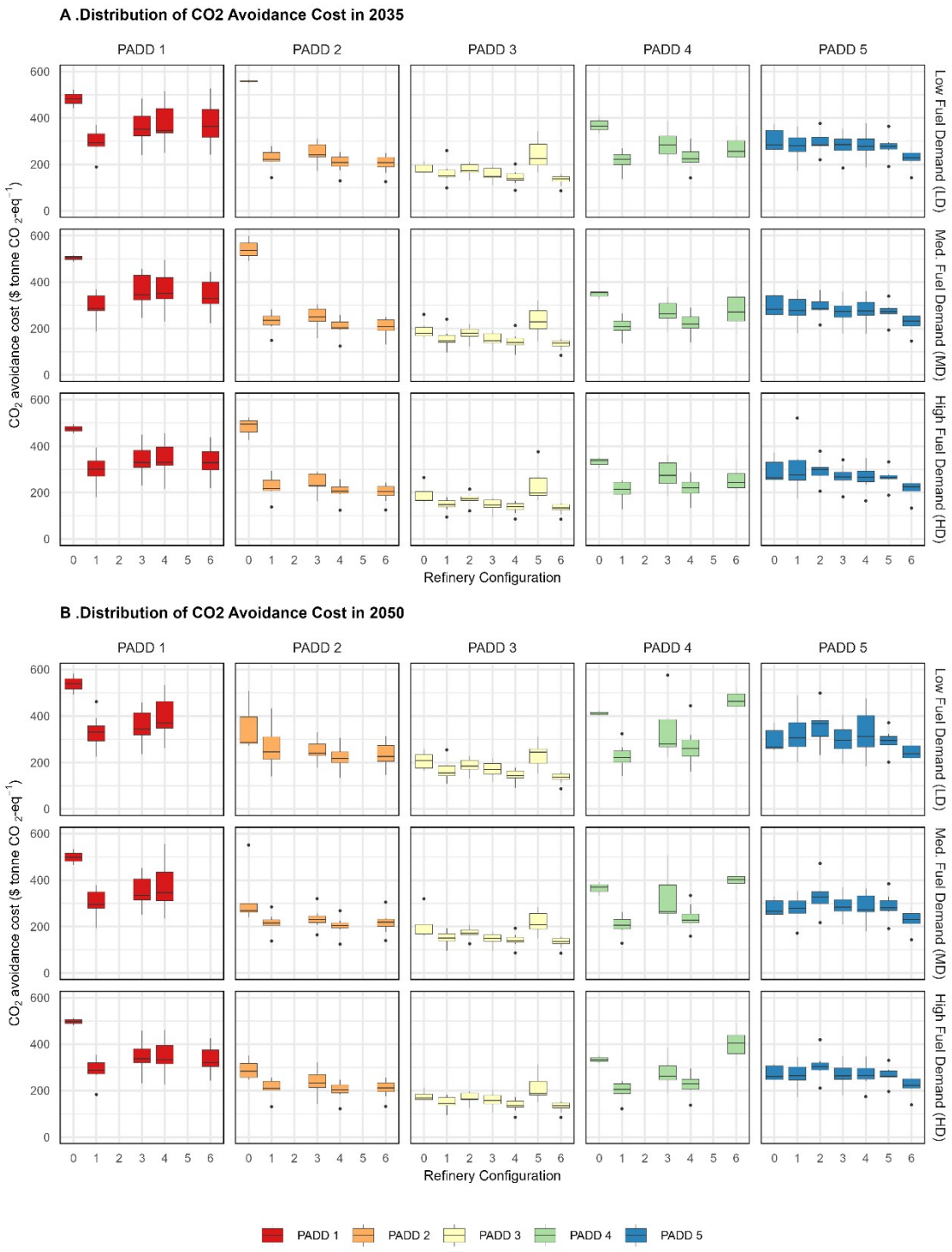


Figure S19. Distribution of CO_2 Avoidance Cost across Refinery Configurations for Heavy/Sour Crudes

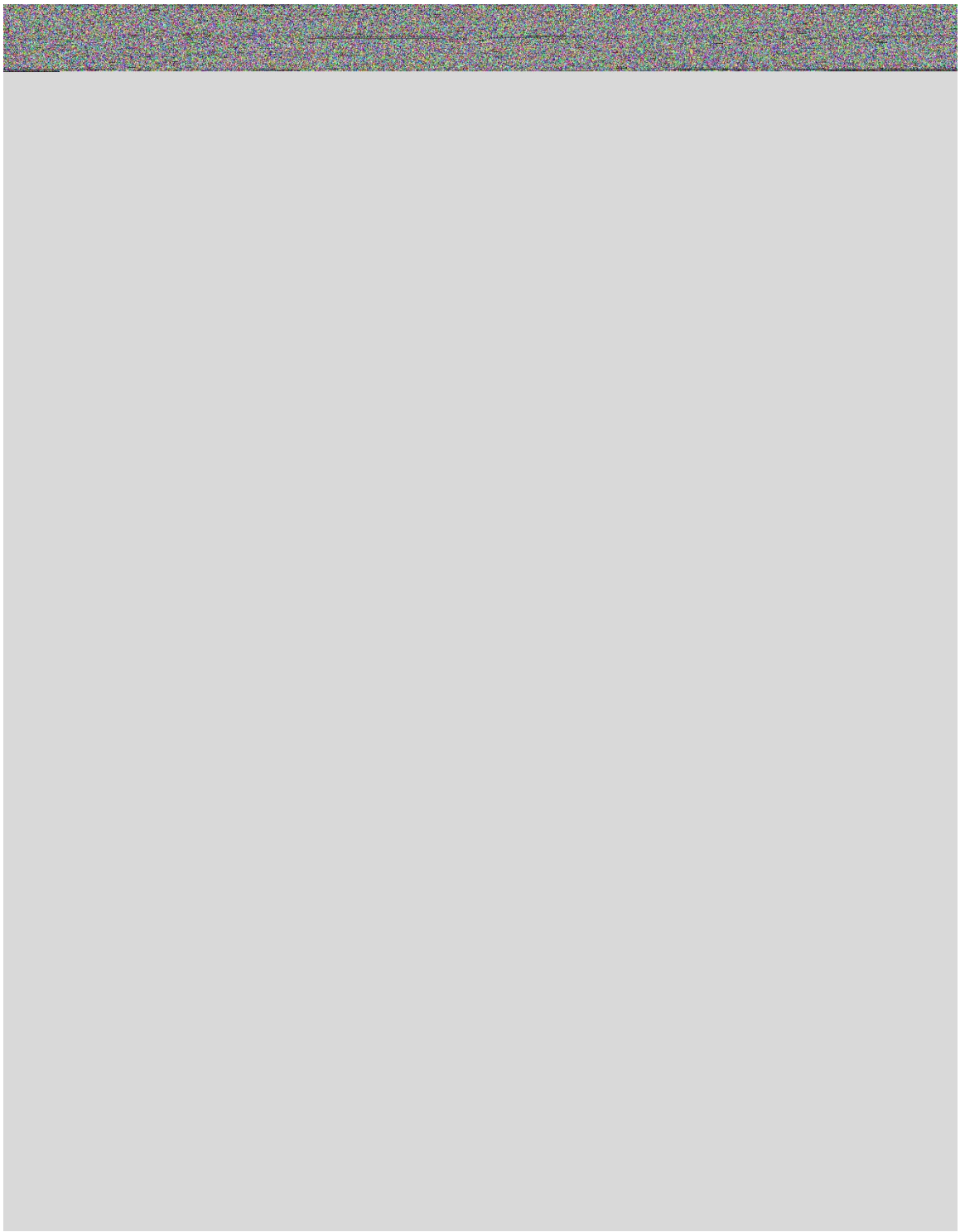


Figure S20. Distribution of CO_2 Avoidance Cost across Refinery Configurations for Light/Sweet Crudes

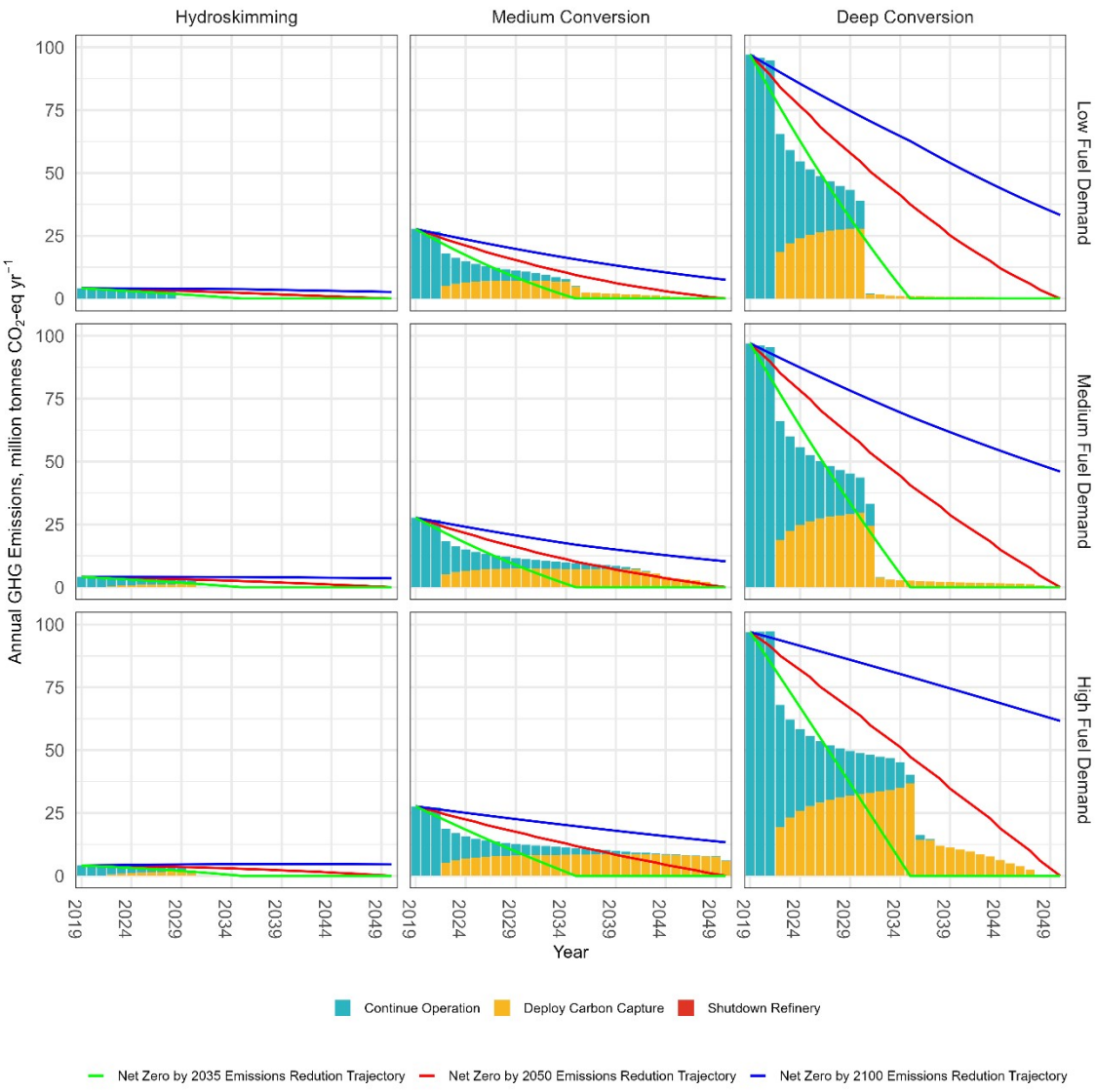


Figure S21. Expected U.S. GHG Emissions by Selected Refinery Configurations based on Likelihoods of Different Strategic Options.

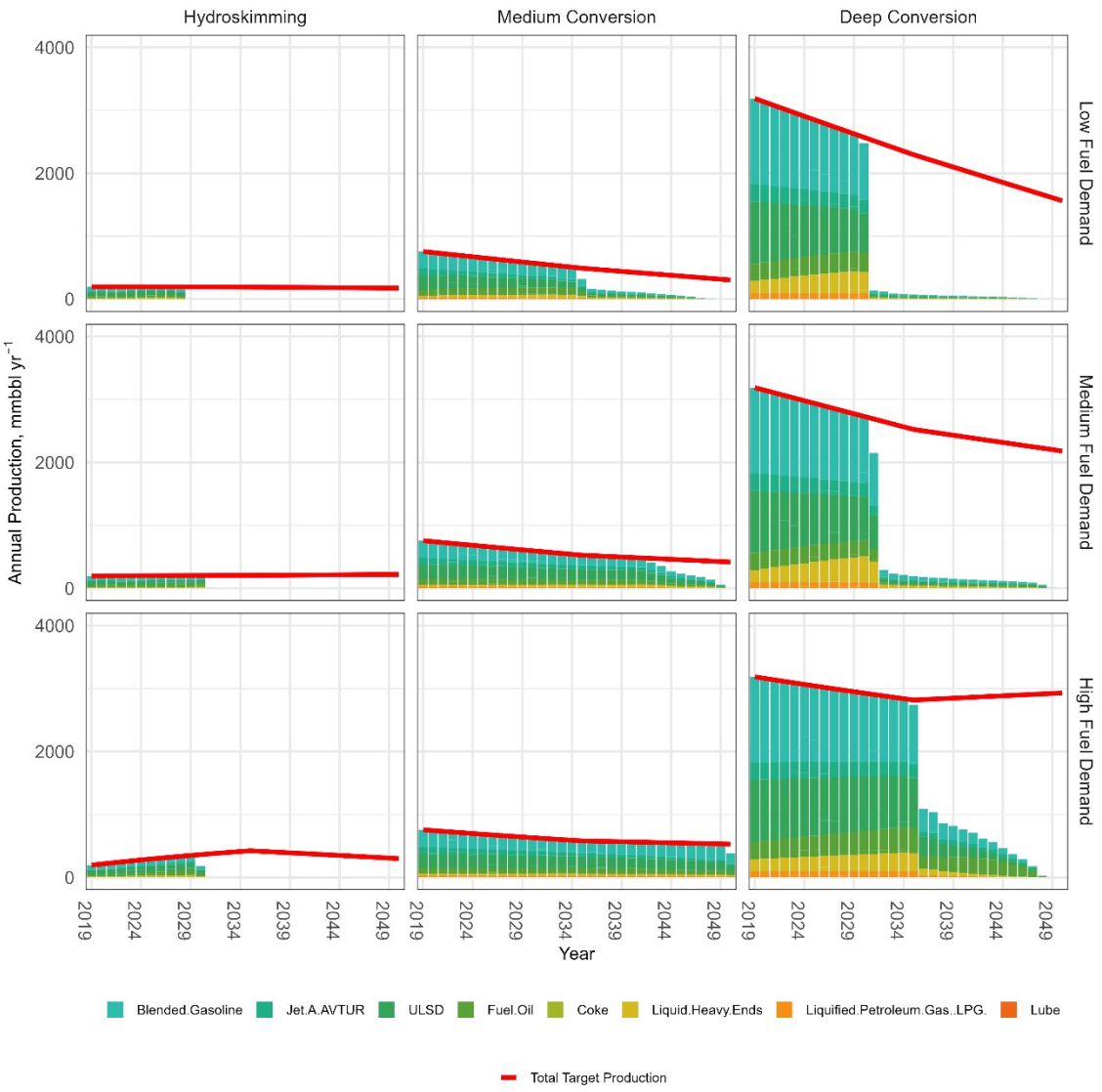


Figure S22. Expected U.S Refinery Production by Selected Refinery Configurations based on Likelihoods of Different Options.

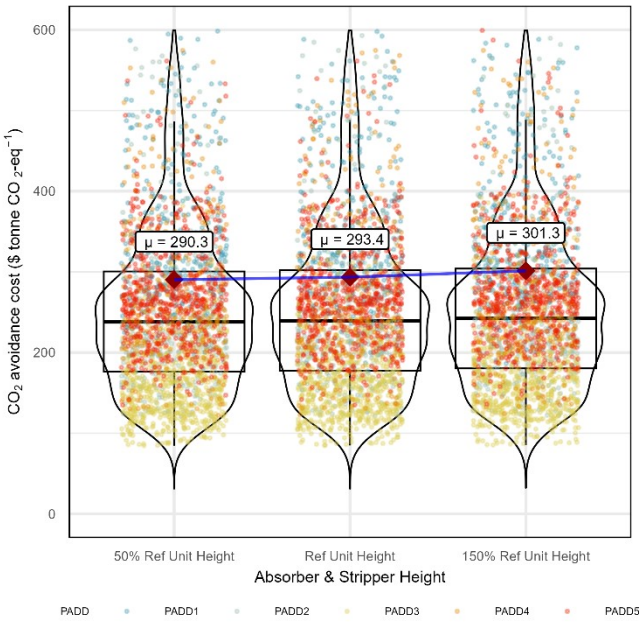


Figure S23. Sensitivity of CO_2 Avoidance Costs to Variations in Absorber and Stripper Heights (50%, 100%, and 150% of the Reference Case Dimensions)

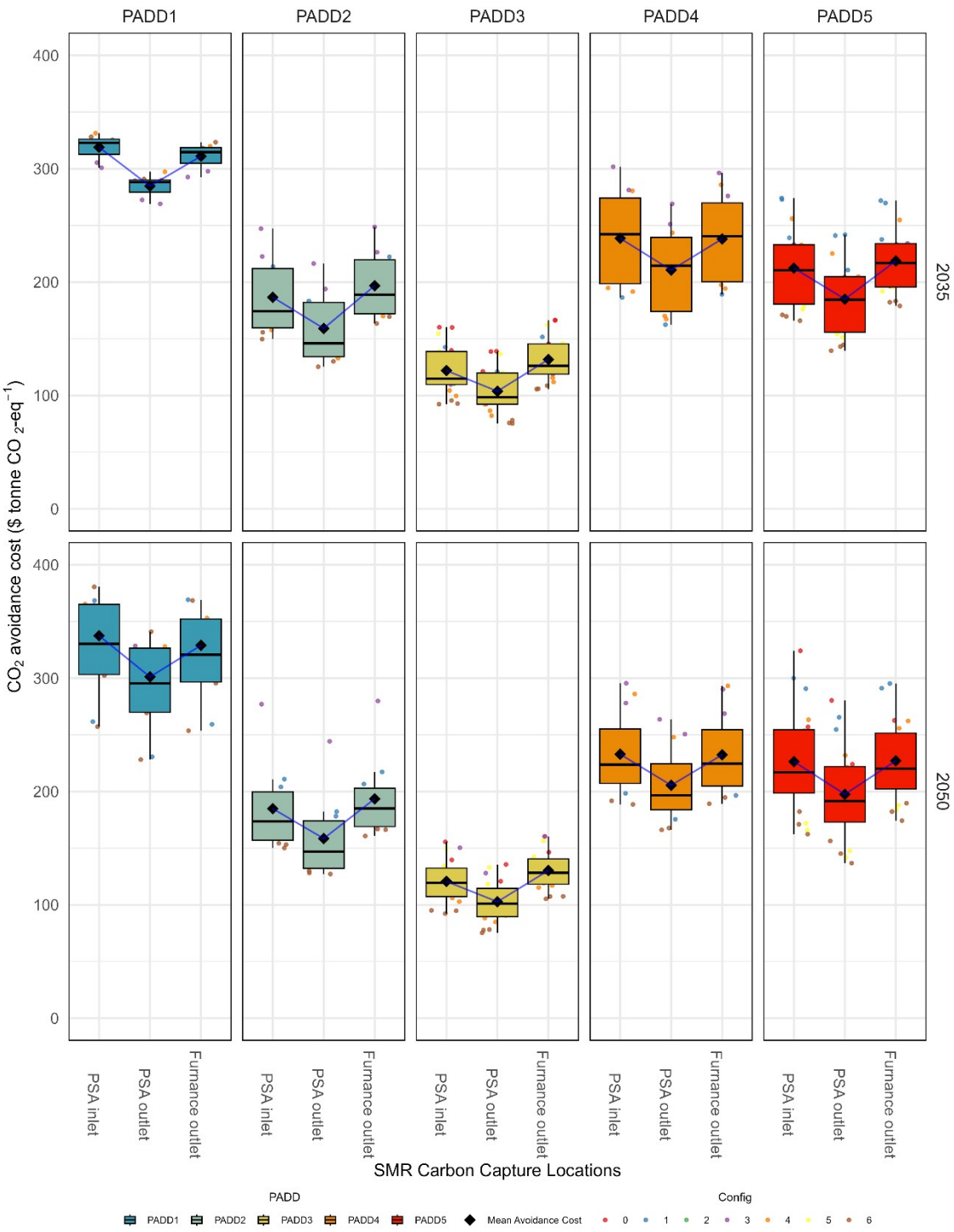


Figure S24. Distribution of CO_2 Avoidance Cost for Carbon Capture on Different SMR Streams for U.S. Refineries.

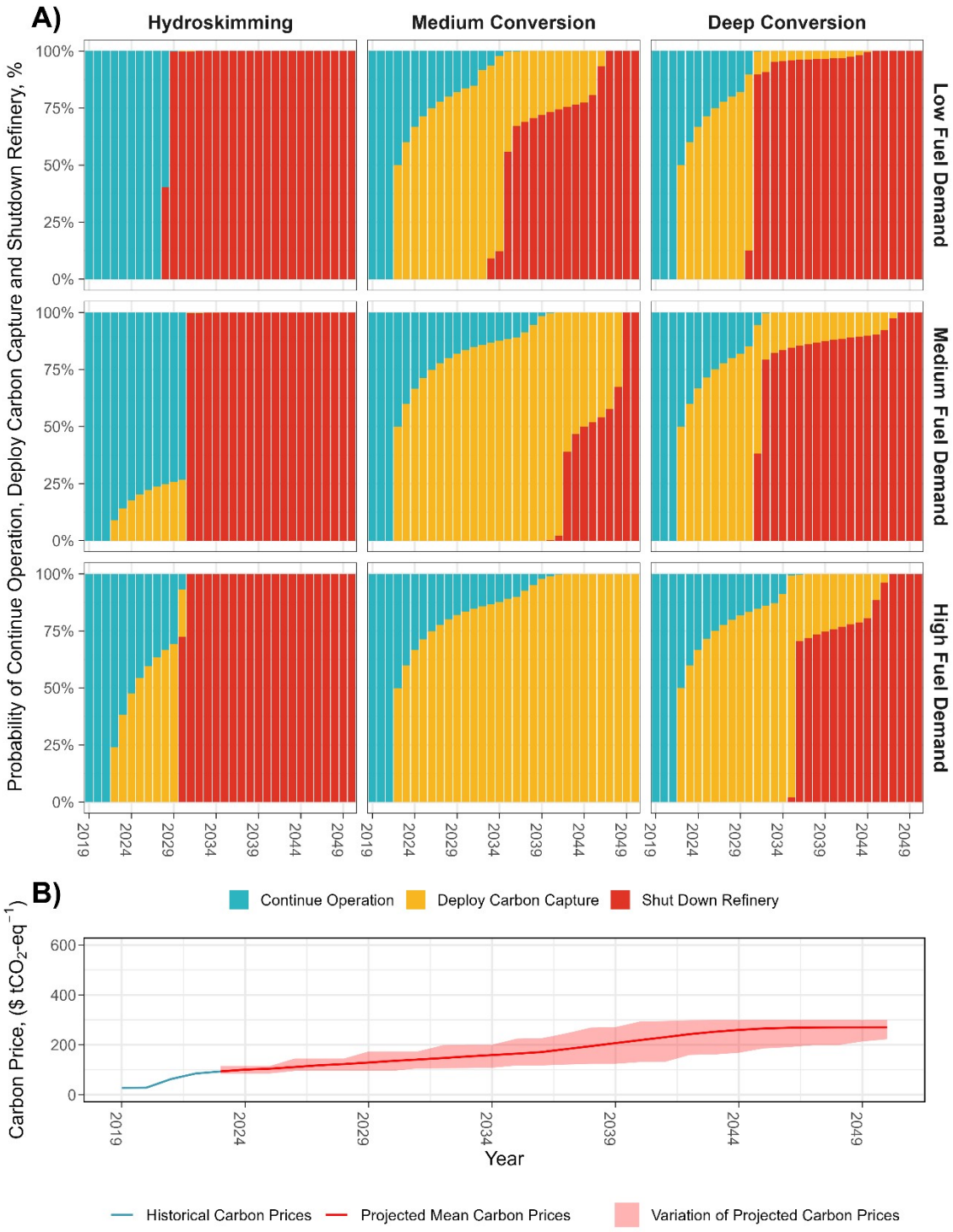


Figure S25. Likelihood of Three Choices Using Real Options Analysis for Three Refinery Types under Three Transportation Fuel Demand Scenarios for Anticipated Carbon Price Scenario.

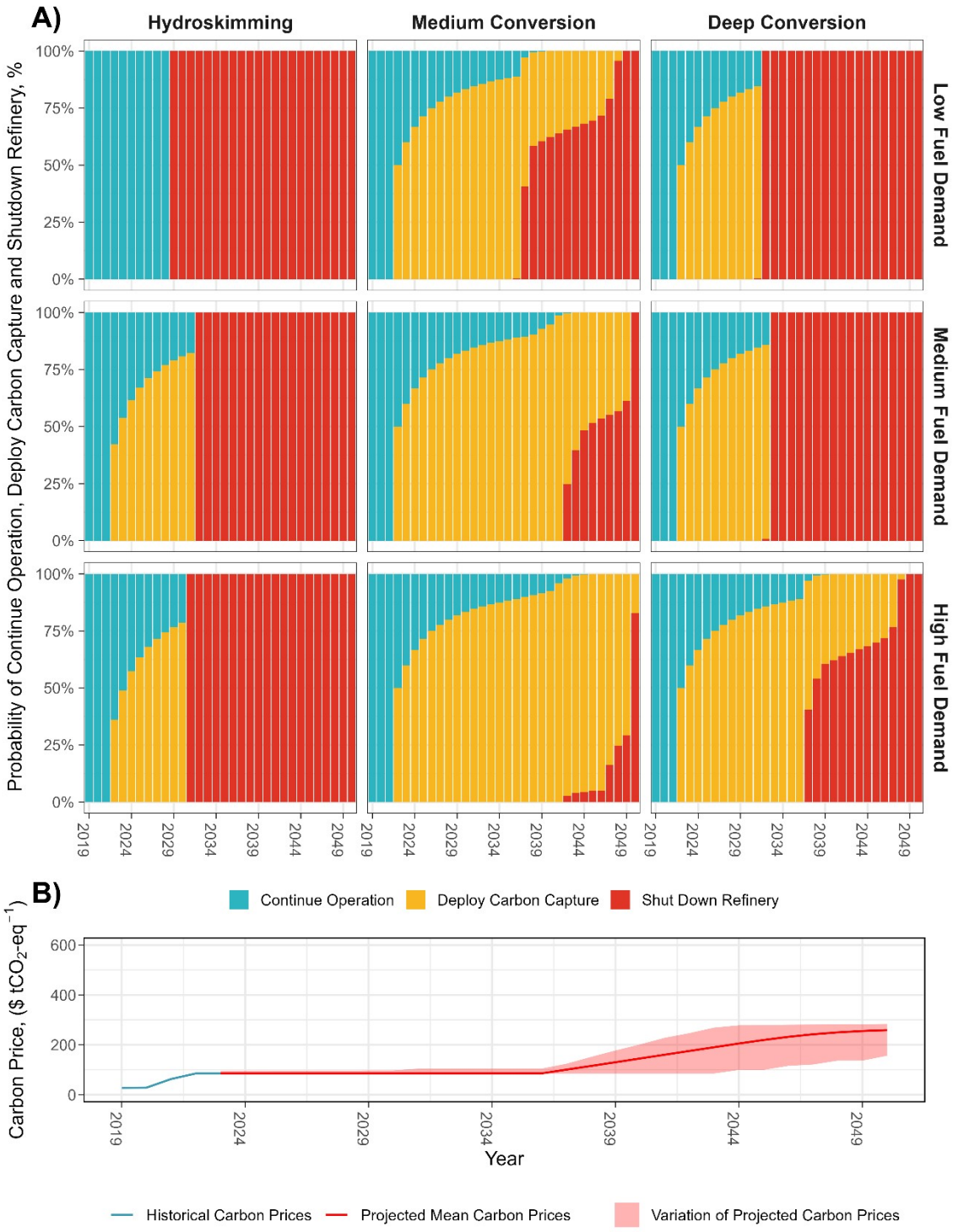


Figure S26. Likelihood of Three Choices Using Real Options Analysis for Three Refinery Types under Three Transportation Fuel Demand Scenarios for Deferred Carbon Price Scenario.

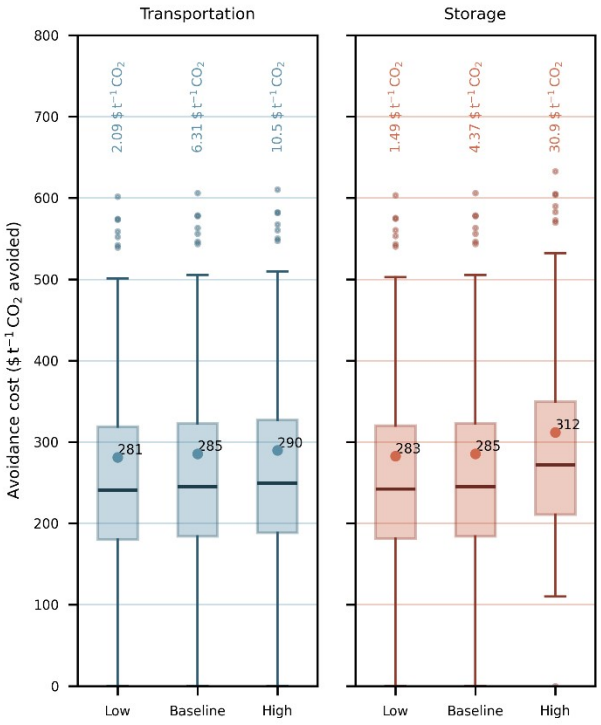


Figure S27. Sensitivity of CO_2 avoidance cost to transport and storage cost variations

Tables

Table S1. PRELIM Specified Refinery Configurations

Refinery Configuration	Configuration #
Hydroskimming Configuration (0)	0
Medium Conversion: FCC (1)	1
Medium Conversion: GO-HC (2)	2
Medium Conversion: FCC & GO-HC (3)	3
Deep Conversion: FCC (4)	4
Deep Conversion: GO-HC (5)	5
Deep Conversion: FCC & GO-HC (6)	6

Table S2. EIA (13) Reported Historical Contribution of Fuel to Demand.

1	The gasoline consumed in the transportation sector consists of 96 %volume of total gasoline supplied.
2	The jet fuel consumed in the transportation sector consists of 100 %volume of total jet fuel supplied.
3	The diesel consumed in the transportation sector consists of 77 %volume of total diesel supplied.
4	The heating oil consumed in the transportation sector consists of 23 %volume of total diesel supplied.
5	The residual fuel oil consumed in the transportation sector consists of 80%volume of total residual fuel oil supplied
6	The refinery-produced coke usually takes 1.4%volume of the total supplied refinery products, while the residual fuel oil takes about 1.2%.

Table S3. The proportion of locally produced transportation fuel to total fuel supplies

PADD	Gasoline	ULSD	Jet Fuel	Fuel Oil	Coke	LHE	LPG	Lube
1	35%	27%	27%	42%	131%	0%	107%	46%
2	89%	91%	94%	16%	121%	100%	192%	53%
3	279%	325%	453%	1004%	462%	100%	112%	181%

4	103%	101%	59%	107%	157%	57%	715%	0%
5	91%	107%	90%	247%	493%	43%	64%	226%

Table S4. Target Crude Input Qualities for Each PADD

PADD	1		2		3		4		5	
Sulfur & Year	Sour	Sweet	Sour	Sweet	Sour	Sweet	Sour	Sweet	Sour	Sweet
2035	1.88	0.34	2.82	0.73	2.18	0.95	2.89	1.00	1.70	1.41
2050	2.46	0.21	6.33	0.35	3.09	0.93	4.56	0.97	1.91	1.55
API & Year	Heavy	Light	Heavy	Light	Heavy	Light	Heavy	Light	Heavy	Light
2035	29.3	37.6	29.6	35.7	29.3	37.6	28.6	34.8	28.3	32.5
2050	26.9	38.7	27.3	36.2	26.9	38.7	25.9	34.7	29.1	35.0

Table S5. SRCI for Individual Refinery Configuration

Refinery Configuration	SRCI for Individual Refinery Configuration
0	3.13
1	7.92
2	6.14
3	9.70
4	10.3
5	9.42
6	10.6

Table S6. Constants and Coefficients for the Multi-linear Regression Model

PADD *n*	Intercept Coefficient	API Coefficient	Gasoline Coefficient	Diesel Coefficient	Jet Fuel Coefficient

1	9.26	0.03	-0.38	-7.91	-2.63
2	-45.97	-0.07	61.63	72.54	56.23
3	2.36	-0.09	10.41	14.57	11.01
4	-24.13	0.04	22.29	48.84	40.37
5	-6.86	0.24	9.81	10.91	20.88

1081

1082 Table S7. Assumed Key Performance Indicator (KPI) for CapEx

Item	KPI	Unit
Discount Rate	8	%
Project Lifetime	25	Years
Construction Periods	3	Years
Project Contingency	15	%
Ratio of Construction Cost to Absorber Equipment Cost	58.7	%
Ratio of EPC Cost to Absorber Equipment Cost	40.5	%
Ratio of Construction Cost to Stripper Equipment Cost	58.8	%
Ratio of EPC Cost to Stripper Equipment Cost	40.6	%
Ratio of Construction Cost to Compressor Equipment Cost	76.2	%
Ratio of EPC Cost to Compressor Equipment Cost	46.8	%
Ratio of Construction Cost to Interconnection Equipment Cost	115	%
Ratio of EPC Cost to Interconnection Equipment Cost	45.7	%
Ratio of Construction Cost to Cooling Tower Equipment Cost	72.1	%
Ratio of EPC Cost to Cooling Tower Equipment Cost	45.0	%
Ratio of Construction Cost to Wastewater Treatment Equipment Cost	58.8	%
Ratio of EPC Cost to Wastewater Treatment Equipment Cost	40.1	%
Ratio of Other Cost to CapEx	10.0	%

1083

1084 Table S8. Assumed KPI for OpEx

Item	KPI	Unit
Operating Hours	8,500	hours/y
MEA	2,000	$/ton MEA
MEA Sludge Disposal Cost	225	$/ton MEA
Natural Gas	4.99	USD/MMBtu
Electrical Energy Cost	0.07	per kWh
Calcium Carbonate Cost	44.0	$/ton CaCO3
Molecular Sieve Adsorbent Cost	7,200	$/ton Adsorbent
Raw Process Water Make-up Cost	0.10	$/ton Water
Waste treatment Cost	5.00%	of TPC
Miscellaneous Operating Materials	2.00%	of VOC
Local Tax	1.60%	of TPC

Insurance	1.00%	of TPC
Material Maintenance (MM)	6.00%	of TPC
Share Annual Material Maintenance Cost in the Overall Annual Maintenance Cost	60.0%	of M
Total Number of Employee	0.09	ppl/t CO_2
Average Annual Salary	80,000	USD/y
Operating Labor (OL)	5308	per Hour
Laboratory costs	20.0%	of OL
Supervision	20.0%	of OL
Plant Overheads	80.0%	of OL
Operating Supplies	15.0%	of M
Administrative Cost	15.0%	of OL
Distribution and Marketing	0.50%	of TOC
R&D cost	5.00%	of TOC

1085

1086 Table S9. Crude Input and Product Yields for Selected Individual Refineries in 2019, 2035 and 2050

Fuel Demand	Year	Config.	Total Cap. US (mil bbl/d)	Avg Cap. (MMbbl/d)	Total Prod. (MMbbl/d)	Gas Yield (%)	Jet Yield (%)	Diesel Yield (%)	Fuel Oil Yield (%)
Low	2019	0	507	28,450	28,776	34	5	30	14
Low	2035	0	574	28,450	28,776	22	4	25	23
Low	2050	0	525	28,450	28,776	23	3	25	22
Low	2019	3	1,847	158,147	158,291	34	14	29	12
Low	2035	3	1,299	111,225	111,326	31	10	28	16
Low	2050	3	808	69,184	69,247	43	4	27	11
Low	2019	6	7,614	288,262	282,040	42	8	30	8
Low	2035	6	6,131	232,116	227,106	34	7	28	14
Low	2050	6	4,228	160,070	156,615	8	14	5	70
Medium	2019	0	507	28,450	28,776	34	5	30	14
Medium	2035	0	619	28,450	28,776	25	6	23	21
Medium	2050	0	703	28,450	28,776	16	10	13	28
Medium	2019	3	1,847	158,147	158,291	34	14	29	12
Medium	2035	3	1,407	120,473	120,582	31	11	30	15
Medium	2050	3	1,112	95,214	95,300	25	11	30	14
Medium	2019	6	7,614	288,262	282,040	42	8	30	8
Medium	2035	6	6,638	251,311	245,887	37	7	22	10

Medium	2050	6	5,850	221,478	216,697	26	11	30	19
High	2019	0	507	28,450	28,776	34	5	30	14
High	2035	0	719	28,450	28,776	18	7	24	23
High	2050	0	910	28,450	28,776	14	10	24	25
High	2019	3	1,847	158,147	158,291	34	14	29	12
High	2035	3	1,646	140,937	141,065	34	6	21	22
High	2050	3	1,444	123,641	123,753	17	11	28	19
High	2019	6	7,614	288,262	282,040	42	8	30	8
High	2035	6	7,744	288,262	282,040	32	8	22	14
High	2050	6	7,837	288,262	282,040	21	14	31	15

Table S10. Description of Different Fuel Demand Scenarios and the Corresponding Electrification Scenarios Taking from EFS

Fuel Demand Scenario	Electrification Scenario	Year	Scenario Name	Description of Changing in Vehicle Fleet
Baseline (BL)	Baseline	2019	BL	2019 vehicle fleet profile
Low Fuel Demand (LD)	High Electrification	2035 & 2050	LD35/LD50	PEV (Plug-in Electric Vehicle) sales shares from AEO2017 reference case; PEV adoption is largely restricted to LDVs (Light Duty Vehicles)
Medium Fuel Demand (MD)	Medium Electrification	2035 & 2050	MD35/ MD50	Growing PEV adoption for LDVs; MDVs (Medium Duty Vehicles), HDVs (Heavy Duty Vehicles) and passenger bus electrification are primarily limited to short distance uses only.
High Fuel Demand (HD)	Low Electrification	2035 & 2050	HD35/ HD50	High PEV adoption in LDVs and passenger buses; PEVs adoption for MDVs and HDVs expands to both short and long distance uses.

Table S11. Changes in Refinery Product Production and Supply by PADD Region (HD35 to HD50 Scenario)

PADD	Scenario	Product	Delta_Production	Delta_Demand	Surplus/Deficit	Contribution
1	HD35-HD50	Gasoline	-27.4	-78.8	51.4	0.515
3	HD35-HD50	ULSD	52.9	16.2	36.6	0.367
3	HD35-HD50	Fuel Oil	37.5	3.74	33.8	0.338
5	HD35-HD50	Coke	24.1	4.88	19.2	0.192
3	HD35-HD50	Coke	15.4	3.34	12.1	0.121
2	HD35-HD50	LPG	24.7	12.9	11.9	0.119

4	HD35-HD50	LPG	12.5	1.75	10.8	0.108
2	HD35-HD50	Gasoline	-54.7	-61.2	6.52	0.0652
5	HD35-HD50	Fuel Oil	10.7	4.37	6.37	0.0638
5	HD35-HD50	Gasoline	-34.6	-38	3.43	0.0344
1	HD35-HD50	Jet	-0.906	-3.4	2.49	0.0249
1	HD35-HD50	Coke	8.66	6.64	2.03	0.0203
5	HD35-HD50	ULSD	20.5	19	1.44	0.0144
3	HD35-HD50	LPG	8.39	7.5	0.894	0.00895
5	HD35-HD50	Jet	-7.3	-8.14	0.831	0.00832
1	HD35-HD50	LPG	11.8	11	0.736	0.00737
2	HD35-HD50	Coke	2.89	2.39	0.497	0.00497
4	HD35-HD50	Jet	-0.359	-0.609	0.249	0.0025
4	HD35-HD50	Coke	0.326	0.207	0.119	0.00119
5	HD35-HD50	Lube	0.175	0.0771	0.0975	0.000976
2	HD35-HD50	Jet	-1.08	-1.15	0.0688	0.000689
4	HD35-HD50	Fuel Oil	0.798	0.754	0.0441	0.000442
3	HD35-HD50	Lube	0.088	0.0484	0.0396	0.000396
4	HD35-HD50	ULSD	3.3	3.28	0.0255	0.000255
2	HD35-HD50	LHE	2.05	2.05	0	0
3	HD35-HD50	LHE	2.86	2.86	0	0
4	HD35-HD50	Lube	0	0.0142	-0.0142	-0.00014
4	HD35-HD50	LHE	0.152	0.178	-0.0254	-0.00025
2	HD35-HD50	Lube	0.0621	0.118	-0.056	-0.00056
1	HD35-HD50	Lube	0.0709	0.154	-0.0828	-0.00083
4	HD35-HD50	Gasoline	-7.02	-6.81	-0.212	-0.00213
2	HD35-HD50	ULSD	23.7	26.1	-2.39	-0.0239
5	HD35-HD50	LHE	1.79	4.19	-2.39	-0.024
5	HD35-HD50	LPG	4.77	7.5	-2.73	-0.0273
1	HD35-HD50	Fuel Oil	3.11	7.43	-4.31	-0.0432
2	HD35-HD50	Fuel Oil	1.46	5.99	-4.54	-0.0454
1	HD35-HD50	LHE	0	5.69	-5.69	-0.0569
3	HD35-HD50	Jet	-17	-3.75	-13.2	-0.133
1	HD35-HD50	ULSD	8.84	32.3	-23.5	-0.235
3	HD35-HD50	Gasoline	-66.3	-23.7	-42.6	-0.426

Table S12. PRELIM Estimated Volume-weighted Average Emissions Intensities of Different Process Unit in U.S. Refineries

Process Unit	Emissions Intensity, $kgCO_2eq$ per bbl of crude fraction,
AT	7.61
VDU	4.75
NHT	17.9

KHT	24.6
GOHC	44.7
DHT	31.8
CHT	15.7
COKER	15.3
FCC	31.6
CNR	14.5
ISO	34
RHT	35.6

Table S13. Estimated U.S. GHG Emissions with or without Carbon Capture under Different Fuel Demand Scenarios

Crude Quality	Demand, MMbbl/year	GHG, No CC, MMt/y	GHG, No CC for FCC Refineries, MMt/y	GHG, Max CC, MMt/y	GHG, Min CC, MMt/y
REF/REF	5540	212	198	86.9	202
HEAVY/SOUR	4710	208	185	87.6	198
HEAVY/SOUR	4580	211	190	88.5	200
HEAVY/SOUR	3740	174	155	73.2	165
HEAVY/SOUR	3040	154	137	65.7	147
HEAVY/SOUR	3470	158	139	67	151
HEAVY/SOUR	2100	111	98.5	47.6	106
REF/REF	5540	212	198	86.9	202
LIGHT/SWEET	4710	198	171	83.5	189
LIGHT/SWEET	4580	204	177	85.8	195
LIGHT/SWEET	3740	168	147	70.8	160
LIGHT/SWEET	3040	151	129	64.3	145
LIGHT/SWEET	3470	152	132	64.9	147
LIGHT/SWEET	2100	105	88.8	44.8	101

Table S14. Post-Combustion Capture Unit's Specific Reboiler Duties and Electricity Consumptions for Various Gas Flows in Carbon Capture Module

	Assumed CO_2 Content (mol%)	Specific Reboiler Duty (MJ Steam/kg CO_2 captured)	Specific Electricity Consumption (MJ/kg CO_2 captured)
Natural gas combusted flue gas	8.74%	4.13	0.32
Refinery fuel gas combusted flue gas	6~11.0%	4.12	0.17
Internal Liquid fuel combusted flue gas	10~12.5%	3.00	0.15
PSA tail gas combusted flue gas	8.94%	4.13	0.19
PSA tail gas	19.28%	2.00	0.11
FCC coke combusted flue gas	16.85%	3.00	0.12

Table S15. Location Factor for Each PADD

PADD	Location Factor
1	1.4
2	1.4
3	1.0
4	1.2
5	1.7

Table S16. Variables Considered for Real Option Analysis

Variable	High	Reference	Low
Risk-free interest rate	7%	3%	1%
Crack Spread	AEO (39) Low Oil Price	AEO (39) Reference	AEO(39) High Oil Price
CCS Utility Cost	150%	100%	50%
Carbon Price	150%	100%	50%
CCS CapEx	150%	100%	50%

Table S17. List of Captured Individual and Combination of Process Units

Index	Carbon Capture Scenarios
1	No Carbon Capture
2	ADU/VDU

Comment [FL]: STM: Why utility cost and not capital cost?

Comment [FL]: Added a new parameter for CapEx

3	CNR
4	FCC/RFCC
5	Coker
6	Hydrotreaters
7	Steam Boiler
8	SMR
9	Others
10	ADU/VDU+CNR
11	ADU/VDU+FCC/RFCC
12	ADU/VDU+SMR
13	CNR+FCC/RFCC
14	CNR+SMR
15	FCC/RFCC+SMR
16	ADU/VDU+CNR+FCC/RFCC
17	ADU/VDU+CNR+SMR
18	ADU/VDU+FCC/RFCC+SMR
19	CNR+FCC/RFCC+SMR
20	ADU/VDU+CNR+FCC/RFCC+SMR
21	ADU/VDU+CNR+FCC/RFCC+Coker+Hydrotreater+SMR
22	ADU/VDU+CNR+FCC/RFCC+Coker+Hydrotreater+SMR+Others
23	ADU/VDU+CNR+Steam Boiler
24	ADU/VDU+FCC/RFCC+Steam Boiler
25	ADU/VDU+SMR+Steam Boiler
26	CNR+FCC/RFCC+Steam Boiler
27	CNR+SMR+Steam Boiler
28	FCC/RFCC+SMR+Steam Boiler
29	ADU/VDU+CNR+FCC/RFCC+Steam Boiler
30	ADU/VDU+CNR+SMR+Steam Boiler
31	ADU/VDU+FCC/RFCC+SMR+Steam Boiler
32	CNR+FCC/RFCC+SMR+Steam Boiler
33	ADU/VDU+CNR+FCC/RFCC+SMR+Steam Boiler
34	ADU/VDU+CNR+FCC/RFCC+Coker+Hydrotreater+Steam Boiler+SMR
35	ADU/VDU+CNR+FCC/RFCC+Coker+Hydrotreater+Steam Boiler+SMR+Others

Table S18. CO_2 Compression Unit Utility Consumption Estimation

CO_2 Product	Stage 1	Stage 2	Stage 3	Stage 4	Stage 5	Stage 6
Ps, psia	23.5	52.0	113	248	545	1,200
Pd, psia	53.7	116	253	550	1,210	2,220
Ts, K	294	294	294	294	294	294
Td, K	361	359	361	362	365	347
Zs	1.00	0.98	0.98	0.95	0.90	0.80
Zd	0.98	0.98	0.95	0.90	0.80	0.59
n/n-1	0.23	0.25	0.22	0.18	0.10	-0.14
Polytropic compression head, kJ/kg	58.1	55.5	55.1	52.4	47.9	29.7

Table S19. SCF for Individual Process Units **(21)**

Process Unit	SCF	Process Unit	SCF
AT	1	Coke	0
VT	2	Sulfur	0
Coker	6	Asphalt	2
Visbreaker	3	POLY	10
FCC	6	Aromatics	15
Reformer	5	Isomerization	15
Hydrocracker	6	Lube	10
Hydrotreater	2	Oxygenates	10
Alkylation	10	Hydrogen	1

Table S20. Regression Coefficient for the Linear Regression Model between Regional Refinery Output and Input.

PADD n	Coefficient b_n
1	0.345
2	0.899
3	1.12
4	1.01
5	0.845

Table S21. Mean refinery crude input historical change for different refinery configurations

configuration	15-year max $(\Delta p_{i,n}^{ub})$	15-year min $(\Delta p_{i,n}^{lb})$	30-year max $(2 \times \Delta p_{i,n}^{ub})$	30-year min $(2 \times \Delta p_{i,n}^{lb})$
0	1.77	-0.387	3.53	-0.774
1	0.348	-0.0162	0.696	-0.0324
2	0.0339	-0.454	0.0678	-0.909
3	0.915	-0.443	1.83	-0.886

4	0.242	-0.127	0.483	-0.254
5	0.411	-0.521	0.823	-1
6	0.0394	-0.324	0.0788	-0.648

1122

1123 Table S22. Background Information on the Reference Carbon Capture Configuration

Item	KPI	Unit
Refinery Configuration	6	-
Refinery Capacity	350,000	bbl/d
Flue Gas Flowrate	286	t/h
CO_2 Flowrate	49.2	t/h
CO_2 vol%	11.3	%

1124

1125 Table S23. Equipment List for Reference Carbon Capture Configuration **(2)**

Section	Letter	Description	Type	Size Metric	Type of Size Metric	Unit for Size Metric	Design Pressure (barg)	Material	Equipment Cost (k$)
Absorption	T	Flue Gas Desulfurization	Vertical	33,094	Shell Mass	kg	1.9	SS304L	2,084
Absorption	C	Exhaust Fan	Blower	179,009	Flow	m^3/h	19	SS304L	1,146
Absorption	C	Exhaust Fan After FGD	Blower	219,483	Flow	m^3/h	20	SS304L	1,404
Absorption	E	Flue Gas Reheater	P&F	653	Area	m^2	2	SS304L	37
Absorption	E	DCC Cooler	P&F	2,056	Area	m^2	2	SS304L	106
Absorption	E	Amine Wash Cooler	P&F	38	Area	m^2	40	SS304L	8
Absorption	E	Intercooler	P&F	326	Area	m^2	34	SS304L	20
Absorption	P	DCC Circulating Pump	Centrifugal	174	Flow	L/s	4.4	SS304L	39
Absorption	P	Amine Water Wash Pump	Centrifugal	25	Flow	L/s	4	SS304L	12
Absorption	P	Rich Amine Pump	Centrifugal	206	Flow	L/s	6	SS304L	46
Absorption	P	Intercooler Pump	Centrifugal	203	Flow	L/s	3.4	SS304L	45
Absorption	T	Direct Contact Cooler	Vertical	18,324	Shell Mass	kg	2	SS304L	282

Absorption	T	Absorber	Vertical	33,094	Shell Mass	kg	1.9	SS304L	401
Compressions	C	CO_2 Compression Package	Compressor	31,950	Duty	kW	-	SS304L	2,426
Compressions	P	CO_2 Product Pump	Centrifugal	440	Flow	L/s	-	SS304L	101
Compressions	PK	Molecular Sieve Package for Dehydration	304SS Packing	19	Volume	m3	-	CS	62
Regeneration	E	Lean/Rich Heat Exchanger	P&F	14,500	Area	m^2	6	SS316L	715
Regeneration	E	Lean Amine Cooler	P&F	405	Area	m^2	6	SS316L	26
Regeneration	E	Reflux Condenser	S&T	2,765	Area	m^2	6	SS316L	342
Regeneration	E	Stripper Reboiler	Kettle	13,025	Area	m^2	5.3	SS316L	1,442
Regeneration	P	Lean Amine Make-up Pump	Centrifugal	0	Flow	L/s	8.6	SS304L	9
Regeneration	P	Lean Amine Pump	Centrifugal	1,326	Flow	L/s	4.8	SS316L	290
Regeneration	P	Stripper Reflux Pump	Centrifugal	46	Flow	L/s	8.3	SS304L	15
Regeneration	P	Condensate Return Pump	Centrifugal	199	Flow	L/s	10	SS316L	44
Regeneration	T	Stripper	Vertical	336,291	Shell Mass	kg	6	SS316L	2,498
Regeneration	TK	Amine Storage Tank	Vertical	669,658	Shell Mass	kg	1.8	CS	3,806
Regeneration	TK	CO_2 Reflux Accumulator	Vertical	74,177	Shell Mass	kg	2.8	SS316L	652
Regeneration	TK	LP Condensate Separator	Horizontal	2,107	Shell Mass	kg	5.1	CS	25
Regeneration	TK	LP Condensate Separator	Horizontal	13,148	Shell Mass	kg	2.7	CS	136
Cooling Tower	CT	Cooling Tower & Pumps	-	694	Flow	L/s	-	CS	357
Wastewater Treatment	TK	Wastewater Treatment	Cone Roof	10,603	Capacity	m^3	-	CS	472

Table S24. Selected Equipment Cost for Common Plant

Letter	Type	Material	Type of Size Metric	Unit for Size	D_L	D_H	a	b	r
T	Vertical	SS30	Shell Mass	kg	90	1242	-100	600	0.

		4L				00			6
T	Horizontal	SS304L	Shell Mass	kg	170	114000	-15000	560	0.6
C	Compressor	SS304L	Duty	kW	132	29000	8400	3100	0.6
C	Blower	SS304L	Flow	m^3/h	200	5000	400	27	0.8
E	P&F	SS304L	Area	m^2	1	180	1100	850	0.4
E	Kettle	SS304L	Area	m^2	10	500	14000	83	1
E	S&T	SS304L	Area	m^2	10	1000	10000	88	1
P	Centrifugal	SS304L	Flow	L/s	0.2	500	3300	48	1.2
PK	304SS Packing	SS304L	Volume	m^3	0	100000	0	3200	1
PK	Packaged Mechanical Refrigerator	SS304L	Duty	kW	50	1500	4900	720	0.9
A	-	SS304L	Area	m^2	10	500	13000	95	1
U	-	SS304L	Flow	L/s	100	10000	61000	650	0.9
TK	Cone Roof	SS304L	Capacity	m^3	10	40000	5700	700	0.7

Table S25. Equipment Cost Comparison between the Reference Configuration and Literature Data (2019)

Item	Reference Configuration	Base Case 01-01
Flue Gas Flow Rate, t/h	286	317
CO_2 Flow Rate, t/h	49	42.3
Sections	Calculated Cost, k$	Base Case 01-01 Cost, k$
Absorber	7,166	18,550
Regenerator	11,784	7,950
Compression	3,286	4,685
Cooling Tower	453	2,353
Wastewater Treatment	599	869

Table S26. Equipment Cost Estimation Parameters (2)

Process Unit	Capacity Parameter	Parameter Unit	Exponent
Absorber	Flue gas mass rate	t/h	1.8
Stripper	CO_2 flowrate	t/h	0.9
Compressor	CO_2 flowrate	t/h	0.75
Interconnection	CO_2 flowrate	t/h	1
Cooling tower	Number of cells	Unit	0.87
Wastewater treatment	Wastewater inlet	t/h	0.75

Table S27. Extrapolated Chemical Engineering Plant Cost Index (CEPCI) **(61)**

Year	CEPCI
2019	607.5
2035	769.7
2050	915.7

Table S28. Parameters for Carbon Prices Generation from McKeller **(45)**.

Item	KPI
Drift, μ	0.047
Volatility, σ	0.05

Table S29. Maximum, Minimum and Average Values for Simulated Carbon Price Trajectories

Date	Max_price	Iteration_Max	Min_price	Iteration_Min	Mean_price
2019	27.8	Run1	27.8	Run1	27.8
2020	28.4	Run1	28.4	Run1	28.4
2021	63	Run1	63	Run1	63
2022	85.3	Run1	85.3	Run1	85.3
2023	106	Run5560	89.6	Run5560	97
2024	118	Run8392	84.2	Run8392	101
2025	130	Run8392	84.9	Run8392	105
2026	146	Run407	82.9	Run407	109
2027	156	Run407	83.6	Run407	113
2028	166	Run407	83.7	Run407	118
2029	176	Run407	83.6	Run407	123
2030	184	Run7281	84.2	Run7281	128
2031	197	Run1752	84.4	Run1752	133
2032	215	Run1752	86.3	Run1752	138
2033	231	Run1504	86.8	Run1504	144
2034	256	Run1504	92.4	Run1504	150
2035	269	Run1504	91.3	Run1504	156
2036	276	Run1504	97.6	Run1504	162
2037	291	Run1504	99.2	Run1504	169
2038	327	Run9969	100	Run9969	176
2039	366	Run9969	103	Run9969	183
2040	380	Run9969	106	Run9969	190
2041	374	Run9969	107	Run9969	198
2042	380	Run9969	111	Run9969	206
2043	420	Run6425	110	Run6425	215
2044	433	Run6425	113	Run6425	223

2045	459	Run9969	112	Run9969	232
2046	509	Run8011	122	Run8011	242
2047	530	Run8011	125	Run8011	252
2048	559	Run8011	130	Run8011	262
2049	573	Run5400	131	Run5400	273
2050	584	Run9969	134	Run9969	284

Table S30. Summary of Individual Refinery Operational Changes Since 1999

Operation	Size					PADD					Configurations						
	<9999	10000< & <49999	50000< & <99999	99999< & <199999	<200000	1	2	3	4	5	0	1	2	3	4	5	6
Shutdown	10	17	10	4	0	8	5	12	1	15	26	6	2	1	2	2	2
No Change	121	490	499	509	468	183	407	786	240	471	399	457	90	206	425	77	433
New Build	0	3	3	1	0	0	1	3	0	3	3	0	2	0	0	1	1
Expand	8	68	107	120	153	27	106	195	49	79	47	95	13	53	100	9	139
Shrink	3	27	36	52	51	9	29	72	15	44	16	24	8	17	38	9	57

Table S31. Estimated Annual GHG Emissions by Direct Combustion of Transportation Fuels

Year	Fuel Demand Scenario	Gasoline	Jet Fuel	Diesel	Fuel Oil	Total Emissions	Total Demands
2019	BL	1180	262	627	13.4	2090	5540
2035	HD35	1090	476	155	27.3	1750	4710
2035	MD35	746	441	195	27.3	1410	3740
2035	LD35	690	397	195	27.3	1310	3470
2050	HD50	1020	506	147	32.4	1700	4580
2050	MD50	480	421	240	32.4	1170	3040
2050	LD50	259	296	240	32.4	827	2100

Table S32. CO_2 Avoidance Cost Range for Carbon Capture on Different Locations from SMR

Scenario	Year	SMR Capture Locations	Max Avoid Cost, $/t	Min Avoid Cost, $/t
Low Fuel Demand (LD)	2035	PSA inlet	3030	95.6
Low Fuel Demand (LD)	2035	PSA outlet	2784	78.1
Low Fuel Demand (LD)	2035	Furnace outlet	2925	109
Low Fuel Demand (LD)	2050	PSA inlet	2883	95.0
Low Fuel Demand (LD)	2050	PSA outlet	2656	78.2

Low Fuel Demand (LD)	2050	Furnace outlet	2800	107
Med. Fuel Demand (MD)	2035	PSA inlet	1850	92.8
Med. Fuel Demand (MD)	2035	PSA outlet	1598	75.7
Med. Fuel Demand (MD)	2035	Furnace outlet	1881	106
Med. Fuel Demand (MD)	2050	PSA inlet	881	94.8
Med. Fuel Demand (MD)	2050	PSA outlet	811	77.6
Med. Fuel Demand (MD)	2050	Furnace outlet	853	107
High Fuel Demand (HD)	2035	PSA inlet	1150	92.3
High Fuel Demand (HD)	2035	PSA outlet	1056	75.2
High Fuel Demand (HD)	2035	Furnace outlet	1110	105
High Fuel Demand (HD)	2050	PSA inlet	1122	92.3
High Fuel Demand (HD)	2050	PSA outlet	1030	75.3
High Fuel Demand (HD)	2050	Furnace outlet	1085	105

1144

Datasets

Dataset S1. CO_2 Avoidance Cost of Carbon Capture for Selected Individual/Grouped Process Units in Various Transportation Fuel Demand & Crude Quality Scenarios.

Dataset S2. CO_2 Avoidance Cost of Carbon Capture for Combined Stacks in Selected Deep Conversion Refineries under Selected Transportation Fuel Demand & Crude Quality Scenarios.

Dataset S3. Monte Carlo Results for Expected Present Value for Selected Individual Refineries under Different Fuel Demand Scenarios

Dataset S4. Blended Crude Proxies for Different Crude Quality Scenarios in PRELIM Format

Dataset S5. Monte Carlo Simulation Results for the Probability of Different Decisions Made by Individual Refineries.

SI References

1. E. T. A. R. G. University of Calgary, Petroleum Refinery Lifecycle Inventory Model (PRELIM) v1.6. (2022). Deposited 2022.

2. J. Gale, “Understanding the cost of retrofitting CO2 capture to an integrated oil refinery” (International Energy Agency, 2017).

3. G. Collodi, *et al.*, Techno-economic evaluation of SMR based standalone (merchant) hydrogen plant with CCS. *IT Report, Editor* (2017).

4. F. El-Mahallawy, S.-D. Habik, *Fundamentals and technology of combustion* (Elsevier, 2002).

5. R. James, Cost and Performance Baseline for Fossil Energy Plants Volume 1: Bituminous Coal and Natural Gas to Electricity. **1**.

6. A. J. Kidnay, W. R. Parrish, FUNDAMENTALS of NATURAL GAS PROCESSING. (2006).

7. CO2 abatement in oil refineries: Fired Heaters.

8. N. Maruoka, H. Purwanto, T. Akiyama, Exergy Analysis of Methane Steam Reformer Utilizing Steelmaking Waste Heat. *ISIJ Int.* **50**, 1311–1318 (2010).

9. M. Wang, H. Lee, J. Molburg, Allocation of energy use in petroleum refineries to petroleum products: Implications for Life-Cyde energy use and emission inventory of petroleum transportation fuels. *Int J LCA* **9**, 34–44 (2004).

10. L. Brickett, “Carbon Capture Program: Carbon Dioxide Capture Handbook” (US DOE/NETL, 2015).

11. T. T. Mai, *et al.*, “Electrification Futures Study: Scenarios of Electric Technology Adoption and Power Consumption for the United States” (2018).

12. H. Kim, “Refinery GHG Emissions of US, Canada and EU from 2019 to 2050: An Exploration of Decarbonization Scenarios,” University of Calgary, Calgary, Alberta, Canada. (2023).

13. Total Energy Monthly Data - U.S. Energy Information Administration (EIA). Available at: https://www.eia.gov/totalenergy/data/monthly/index.php [Accessed 27 April 2023].

14. G. Cooney, *et al.*, Updating the U.S. Life Cycle GHG Petroleum Baseline to 2014 with Projections to 2040 Using Open-Source Engineering-Based Models. *Environ. Sci. Technol.* **51**, 977–987 (2017).

15. U.S. Refinery Crude Oil Input Qualities. Available at: https://www.eia.gov/dnav/pet/pet_pnp_crq_dcu_nus_m.htm [Accessed 27 April 2023].

16. Company Level Imports. Available at: https://www.eia.gov/petroleum/imports/companylevel/ [Accessed 27 April 2023].

17. W. L. Nelson, *Guide to refinery operating cost (process costimating)* (1976).

18. W. L. Nelson, *How the Nelson refinery constraction cost indexes evolved* (1976).

19. W. L. Nelson, *Here's how operating cost indexes are computed* (1977).

20. 2021 Worldwide Refining Capacity Summary. *Oil & Gas Journal* (2022). Available at: https://www.ogj.com/ogj-survey-downloads/worldwide-refining/document/14233364/2021-worldwide-refining-capacity-summary [Accessed 27 April 2023].

21. M. J. Kaiser, A review of refinery complexity applications. *Pet. Sci.* **14**, 167–194 (2017).

22. U.S. Refinery Yield. Available at: https://www.eia.gov/dnav/pet/pet_pnp_pct_dc_nus_pct_m.htm [Accessed 27 April 2023].

23. U.S. Refinery & Blender Net Input. Available at: https://www.eia.gov/dnav/pet/pet_pnp_inpt_dc_nus_mbbl_m.htm [Accessed 27 April 2023].

24. U.S. Refinery and Blender Net Production. Available at: https://www.eia.gov/dnav/pet/pet_pnp_refp_dc_nus_mbbl_m.htm [Accessed 27 April 2023].

25. U.S. Refinery Utilization and Capacity. Available at: https://www.eia.gov/dnav/pet/pet_pnp_unc_dcu_nus_m.htm [Accessed 28 April 2023].

26. A. A. V. Julio, R. Castro-Amoedo, F. Maréchal, A. M. González, J. C. Escobar Palacio, Exergy and economic analysis of the trade-off for design of post-combustion CO2 capture plant by chemical absorption with MEA. *Energy* **280**, 128004 (2023).

27. D. Danaci, M. Bui, C. Petit, N. Mac Dowell, En Route to Zero Emissions for Power and Industry with Amine-Based Post-combustion Capture. *Environ. Sci. Technol.* **55**, 10619–10632 (2021).

28. R. A. Y. Sinnott, *Chemical engineering design* (Elsevier, 2014).

29. E. S. Rubin, Understanding the pitfalls of CCS cost estimates. *International Journal of Greenhouse Gas Control* **10**, 181–190 (2012).

30. U.S. Department of Energy, Enhanced Oil Recovery. (2026).

31. I. E. Agency, *Putting CO 2 to Use: Creating Value from Emissions* (International Energy Agency (IEA) Paris, 2019).

32. E. and C. C. Canada, The federal carbon pollution pricing benchmark. (2021). Available at: https://www.canada.ca/en/environment-climate-change/services/climate-change/pricing-pollution-how-it-will-work/carbon-pollution-pricing-federal-benchmark-information.html [Accessed 20 February 2026].

33. Alberta TIER prices jump 25% after MOU, but outlook remains uncertain. *S&P Global Energy* (2025). Available at: https://www.spglobal.com/energy/en/news-research/latest-news/electric-power/120325-alberta-tier-prices-jump-25-after-mou-but-outlook-remains-uncertain [Accessed 25 February 2026].

34. S. M. Ross, Variations on Brownian motion. *Introduction to Probability Models* 612–614 (2014).

35. J. C. Cox, S. A. Ross, M. Rubinstein, Option pricing: A simplified approach. *Journal of Financial Economics* **7**, 229–263 (1979).

36. Y. He, *Real options in the energy markets* (s.n.], 2007).

37. S. Khan, S. Khan, Bumpy road ahead for US gasoline demand and energy transition. *Reuters* (2023).

38. T. E. Copeland, V. Antikarov, *Real Options: A Practitioner's Guide* (Texere, 2001).

39. S. Nalley, A. LaRose, Annual energy outlook 2022 (AEO2022). *Energy Information Agency* **23** (2022).

40. M. Zhang, L. Liu, Q. Wang, D. Zhou, Valuing investment decisions of renewable energy projects considering changing volatility. *Energy Economics* **92**, 104954 (2020).

41. J. A. Schachter, P. Mancarella, A critical review of Real Options thinking for valuing investment flexibility in Smart Grids and low carbon energy systems. *Renewable and Sustainable Energy Reviews* **56**, 261–271 (2016).

42. World Bank, Carbon Pricing Dashboard. *Carbon Pricing Dashboard* (2023). Available at: https://carbonpricingdashboard.worldbank.org/compliance/price [Accessed 22 May 2024].

43. B. R. Cobb, J. M. Charnes, Real Options Volatility Estimation With Correlated Inputs. *The Engineering Economist* **49**, 119–137 (2004).

44. P. S. Reinelt, D. W. Keith, Carbon Capture Retrofits and the Cost of Regulatory Uncertainty. *The Energy Journal* **28**, 101–128 (2007).

45. J. M. McKellar, J. A. Bergerson, J. Kettunen, H. L. MacLean, Predicting Project Environmental Performance under Market Uncertainties: Case Study of Oil Sands Coke. *Environ. Sci. Technol.* **47**, 5979–5987 (2013).

46. L. E. Brandão, J. S. Dyer, Decision Analysis and Real Options: A Discrete Time Approach to Real Option Valuation. *Ann Oper Res* **135**, 21–39 (2005).

47. S. Jiang, Survival in the U.S. petroleum refining industry. *Journal of Applied Statistics* **39**, 1505–1530 (2012).

48. M.-Y. Chen, The survival and growth of US petroleum refineries in response to changes in market conditions. *The Journal of Energy and Development* **29**, 115–146 (2003).

49. D. Mullen, L. Herraiz, J. Gibbins, M. Lucquiaud, On the cost of zero carbon hydrogen: A techno-economic analysis of steam methane reforming with carbon capture and storage. *International Journal of Greenhouse Gas Control* **126**, 103904 (2023).

50. Y. Khojasteh Salkuyeh, B. A. Saville, H. L. MacLean, Techno-economic analysis and life cycle assessment of hydrogen production from natural gas using current and emerging technologies. *International Journal of Hydrogen Energy* **42**, 18894–18909 (2017).

51. D. Bonaquist, "Analysis of CO2 Emissions, Reductions, and Capture for Large-Scale Hydrogen Production Plants" (Praxair, 2010).

52. R. G. Lemus, J. M. Martínez Duart, Updated hydrogen production costs and parities for conventional and renewable technologies. *International Journal of Hydrogen Energy* **35**, 3929–3936 (2010).

53. J. C. Meerman, *et al.*, Techno-economic assessment of CO2 capture at steam methane reforming facilities using commercially available technology. *International Journal of Greenhouse Gas Control* **9**, 160–171 (2012).

54. I. Lindsay, C. Lowe, S. Reddy, M. Bhakta, S. Balkenende, Designing a climate friendly hydrogen plant. *Energy Procedia* **1**, 4095–4102 (2009).

55. F. Pruvost, S. Cloete, C. Arnaiz Del Pozo, A. Zaabout, Blue, green, and turquoise pathways for minimizing hydrogen production costs from steam methane reforming with CO2 capture. *Energy Conversion and Management* **274**, 116458 (2022).

56. M. H. Ali Khan, *et al.*, A framework for assessing economics of blue hydrogen production from steam methane reforming using carbon capture storage & utilisation. *International Journal of Hydrogen Energy* **46**, 22685–22706 (2021).

57. G. Lozza, P. Chiesa, Natural Gas Decarbonization to Reduce CO2 Emission From Combined Cycles—Part II: Steam-Methane Reforming. *Journal of Engineering for Gas Turbines and Power* **124**, 89–95 (2002).

58. P. L. Spath, M. K. Mann, "Life Cycle Assessment of Hydrogen Production via Natural Gas Steam Reforming" (2000).

59. National Energy Technology Laboratory, "Carbon Dioxide Transport and Storage Costs in NETL Studies" (U.S. Department of Energy, National Energy Technology Laboratory (NETL), 2019).

60. Global CCS Institute, "The Cost of CO2 Storage" (Global CCS Institute, 2025).

61. D. Leeson, N. Mac Dowell, N. Shah, C. Petit, P. S. Fennell, A Techno-economic analysis and systematic review of carbon capture and storage (CCS) applied to the iron and steel, cement, oil refining and pulp and paper industries, as well as other high purity sources. *International Journal of Greenhouse Gas Control* **61**, 71–84 (2017).