Optimal Carbon tax and Energy Substitution

Emmanuel Murray Leclair

Department of Economics, Western University

August 2021

1 Model

1.1 Framework

The model to study energy input substitution is very similar to Acemoglu and Restrepo (2021) which I adapt to the energy market and augment it with costly switching. It features the assignment of a mix of energy inputs to energy tasks which allows for flexible variation in input usage, both at the extensive and intensive margin which is characteristic of data on fuel consumption the plant-level. I first present the structure of production for a given plant in a static setting, and then consider what happens when a new fuel can be added to the mix in a dynamic setting. There are two levels to production which correspond to two CES nests. The outer nest is standard and features typical TFP z, labor (l), capital (k), intermediates (m) and energy (E) inputs:

$$y = z\left(\alpha_k k^{\frac{\sigma-1}{\sigma}} + \alpha_l l^{\frac{\sigma-1}{\sigma}} + \alpha_m m^{\frac{\sigma-1}{\sigma}} + \alpha_e E^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \tag{1}$$

$$\alpha_k + \alpha_l + \alpha_m + \alpha_e = 1$$

Where $\sigma \geq 0$ is the elasticity of substitution between inputs. The inner nest that composes E features a mass of M energy tasks in the $\omega \in (0,1)$ interval:

$$E = \left(\frac{1}{M} \int_0^1 (M\tau(\omega))^{(\lambda-1)/\lambda} d\omega\right)^{\lambda/(\lambda-1)} \tag{2}$$

I assume that tasks are gross complements, $\lambda \in (0,1)$, which will be motivated with specific examples. Each energy task can $\tau(\omega)$ and can be performed with a set of fuels $\{e_f\}_{f=1}^F$ (oil, gas, coal, electricity, etc.) which are in principle perfect substitutes:

$$\tau(\omega) = \sum_f \psi_f(\omega) e_f \tag{3}$$

1

Where $\psi_f(\omega)$ are fuel-by-task specific productivity terms, which are an important distinction from standard task-based production models and will allow for very flexible input usage. to motivate this framework one should be thinking of tasks such as the steps required to produce crude steel: preparation of raw material, conversion of iron ore into iron, and conversion of iron into crude steel (Luh et al. 2020). The preparation of raw materials typically requires coal whereas the two subsequent steps can be done with different fuels and the $\psi_f(\omega)$ terms can reflect these particularities. Moreover, these steps are complementary and require high amounts of energy.

Going back to the model, the inner nest problem can be solved in two steps:

1. Find the cheapest fuel to perform each task (section 1.2 and 1.3)

2. Aggregate across tasks into fuel categories to get fuel demand (section 1.4).

1.2 Task choices:

Minimize the cost of producing one unit of energy E given task prices $p(\omega)$

$$\min_{\tau(\omega)} \left\{ \int p(\omega)\tau(\omega)d\omega \right\}$$
$$s.t. \ E = \left(\frac{1}{M} \int_0^1 (M\tau(\omega))^{(\lambda-1)/\lambda}d\omega \right)^{\lambda/(\lambda-1)}$$

This leads to demand for task ω:

$$\tau(\omega)^* = \frac{E}{M} \left(\frac{p_\tau}{p(\omega)} \right)^\lambda \tag{4}$$

Where p_τ is the price index of energy tasks:

$$p_\tau = \left(\int_0^1 p(\omega)^{1-\lambda} M d\omega \right)^{1/(1-\lambda)} \tag{5}$$

1.3 Assignment of fuels to energy tasks:

To study the assignment of fuels to task with costly switching, I need to define a reference set of fuels available to the plant (which will the set of fuels used last period in the dynamic model) $\mathcal{F}$ which defines the set of tasks that were performed by fuel f in the reference distribution. For now, I shall present the assignment of fuels to energy tasks given $\mathcal{F}$. A plant finds the fuel that minimize the cost of performing task ω given (exogenous) fuel prices:

$$\min_{\{e_f(\omega)\}_{f \in \mathcal{F}}} \sum_{f \in \mathcal{F}} p_f e_f(\omega)$$

$$s.t. \sum_{f \in \mathcal{F}} \psi_f(\omega) e_f(\omega) = \tau(\omega)$$

By the linearity of this problem implied by perfect substitution, the plant chooses the fuel that has the lowest unit cost to produce the task. Hence, task prices follow:

$$p(\omega) = \min \left\{ \frac{p_1}{\psi_1(\omega)}, ..., \frac{p_F}{\psi_F(\omega)} \right\} \tag{6}$$

This set-up has one important implications. The fuel-by-task productivity terms will play an important role to determine the energy efficiency of different fuels. For example, since natural gas is consistently more expensive than coal and oil in India, only plants with large productivity draws for natural gas will use it, and if tasks are gross complements ($\lambda < 1$) the quantity consumed will be decreasing in its productivity, making natural gas more energy efficient as observed in the data[1].

1.4 Aggregation from tasks to fuels

From the problem before, I can define the set of tasks that are performed by each fuel:

$$\mathcal{T}_f = \left\{ \omega : \frac{p_f}{\psi_f(\omega)} \le \frac{p_j}{\psi_j(\omega)} \forall j \neq f \right\} \tag{7}$$

Then, I can redefine the aggregate price index of energy tasks in (5):

$$p_\tau = \left(\sum_{f \in \mathcal{F}} p_f^{1-\lambda} M \int_{\mathcal{T}_f} \psi_f(\omega)^{\lambda-1} d\omega \right)^{1/(1-\lambda)} \tag{8}$$

Notice that the price index of energy tasks is the same as the price of the aggregate energy input E:

---

[1] proof in appendix

$$p_e E = \int_0^1 \tau(\omega)p(\omega)d\omega$$

$$p_e E = E \int_0^1 \left(\frac{p_\tau}{p(\omega)}\right)^\lambda p(\omega)d\omega$$

$$p_e = p_\tau^\lambda \int_0^1 p(\omega)^{1-\lambda}d\omega$$

$$p_e = p_\tau^\lambda p_\tau^{1-\lambda}$$

$$p_e = p_\tau$$

This has important implications. It means fuel choices are such that they minimizes the effective price of energy p_e. Hence, expanding $\mathcal{F}$ to $\mathcal{F}'$ by adding a new fuel will serve to expand the choice set in the minimization problem of (6) and allow the plant to face a weakly lower price of energy. Moreover, it means that all the heterogeneity from the fuel-by-task productivity terms is captured by a single term, p_e. I can now define the demand for a task performed by fuel f from (4):

$$\tau_f(\omega) = \frac{E}{M}\left(\frac{\psi_f(\omega)}{p_f}p_e\right)^\lambda \tag{9}$$

Using the task production function from section 1.1, I can find the conditional demand for fuel f by equalizing (9) with (3) and integrating over all tasks performed by fuel f ($\mathcal{T}_f$):

$$\psi_f(\omega)e_f(\omega) = \frac{E}{M}\left(\frac{\psi_f(\omega)}{p_f}p_e\right)^\lambda$$

$$e_f(\omega) = \frac{E}{M}\left(\frac{p_e}{p_f}\right)^\lambda (\psi_f(\omega))^{\lambda-1}$$

Then,

$$e_f = E\left(\frac{p_e}{p_f}\right)^\lambda \underbrace{\frac{1}{M}\int_{\mathcal{T}_f}\psi_f(\omega)^{\lambda-1}d\omega}_{\Gamma_f} \tag{10}$$

$$= E\left(\frac{p_e}{p_f}\right)^\lambda \Gamma_F \tag{11}$$

Where Γ_f is referred to the task-share of energy input (fuel) f.

I can now use this conditional fuel demand and the energy price index in (8) to back out the production function for E as a function of energy inputs:

4

1. Rearrange (10):

$$p_f = \frac{E^{1/\lambda} P_e \Gamma_f^{1/\lambda}}{e_f^{1/\lambda}} \tag{12}$$

2. plug (11) into (8) and solve for E:

$$E = \left(\sum_{f \in \mathcal{F}} \Gamma_f^{1/\lambda} \left(e_f \right)^{(\lambda-1)/\lambda} \right)^{\lambda/(\lambda-1)} \tag{13}$$

The production function for energy in (12) looks like a standard CES in different energy inputs with a key distinction: the presence of Γ_f, the endogenous share of tasks performed by fuel f. This term can be 0 if a fuel is always dominated by at least one other fuel, or it can be 1 if this fuel always dominates every other fuels. This implies that the production function in (12) is very flexible and allows for corner solutions, a key feature of data on fossil fuel consumption.

1.5 Solving for Optimal Energy Input Choices given $\mathcal{F}$

Plants indexed by i are heterogeneous in fuel-by-task specific productivity $(\psi_{if}(\omega))$. Without loss of generality, I remove fuel-augmenting productivity terms (A_f) from the analysis as they are endogenously determined by the distribution of ψ's across all tasks. In the first year of production, I assume that plants draw fuel-by-task productivity terms for each fuels and tasks from a fuel-specific Pareto distribution such that:

$$P(\psi_{if}(\omega) \geq x) = \left(\frac{x}{\gamma_f} \right)^{-\alpha_f}$$

Then, the probability fuel f is chosen over fuel j is:

$$
\begin{aligned}
P_r\left(\frac{p_f}{\psi_{if}(\omega)} \leq \frac{p_j}{\psi_{if}(\omega)} \right) &= \left(\frac{p_f}{p_j \gamma_f} \right)^{-\alpha_f} \int_{\gamma_j}^{\infty} \psi_j^{-\alpha_f} f(\psi_j(\omega)) d\psi_j \\
&= \left(\frac{p_f}{p_j \gamma_f} \right)^{-\alpha_f} \int_{\gamma_j}^{\infty} \psi_j^{-\alpha_f - \alpha_j - 1} \alpha_j \gamma_j^{\alpha_j} d\psi_j \\
&= \left(\frac{p_f}{p_j \gamma_f} \right)^{-\alpha_f} \left[\frac{\psi_j^{-\alpha_f - \alpha_j}}{-\alpha_f - \alpha_j} \alpha_j \gamma_j^{\alpha_j} \right]_{\gamma_j}^{\infty} \\
&= \left(\frac{p_f}{p_j} \right)^{-\alpha_f} \left(\frac{\gamma_j}{\gamma_f} \right)^{-\alpha_f} \frac{\alpha_j}{\alpha_f + \alpha_j}
\end{aligned}
$$

Then, if there are F number of fuels, the probability that fuel f is chosen over all other fuels to perform task ω is going to be:

5

$$\tau_f(\omega) = \prod_{j \neq f} P_r \left(\frac{p_f}{\psi_{if}(\omega)} \leq \frac{p_j}{\psi_{ij}(\omega)} \right)$$

$$= \left(\frac{\gamma_f}{p_f} \right)^{(F-1)\alpha_f} \prod_{j \neq f} \left(\frac{\gamma_j}{p_j} \right)^{-\alpha_f} \frac{\alpha_j}{\alpha_f + \alpha_j} = \tau_f \tag{14}$$

If there are M tasks, I can define the probability that a given fuel is used m times which follows a binomial distribution. Let $M\mathcal{T}_f$ be the number of times fuel f is chosen:

$$P_r(M\mathcal{T}_f = m) = \binom{M}{m} \tau_f^m (1 - \tau_f)^{M-m} \tag{15}$$

Where $\binom{M}{m} = \frac{M!}{m!(M-m)!}$. Notice that by the Poisson Limit Theorem, the limiting distribution (as $M \to \infty$) only exist for $m < M$. This assumption is used extensively in the literature on technological choices (**Kortum 1997, Buera and Oberfield 2020, Boehm and Oberfield 2020**). However, in my context it means there can not be any corner solution where a fuel is used for all tasks if the number of tasks goes to infinity. For this reasons, I keep the assumption of a finite number of tasks for the remainder of the paper.

With (15) I can now characterize the distribution of fuel shares, which is going to be continuous because productivity draws affect how much fuel is consumed. For example, when tasks are gross complement ($\lambda < 1$) a lower quantity of fuel is going to be needed if it is very productive and vice versa. For a given plant, let fuel shares be:

$$s_f = \frac{e_f}{\sum_f e_f}$$

$$= \frac{p_f^{-\lambda} \Gamma_f}{\sum_f p_f^{-\lambda} \Gamma_f} \tag{16}$$

First moment (mean)

Notice that with M tasks and a probability of success τ_f the average number of times fuel f is chosen in the data is going to be:

$$E(M\mathcal{T}_f) = M\tau_f$$
$$E(\mathcal{T}_f) = \tau_f$$

Then, I can use that to characterize $E(\Gamma_f)$ and $E(s_f)$:'

$$E(\Gamma_f) = E\left(\frac{1}{M} \int_{\mathcal{T}_f} \psi_f(\omega)^{\lambda-1} d\omega \right)$$

$$= ???$$

6

1.6 Full model and expanding fuel set $\mathcal{F}$

Each period there is a representative consumer with exogenous income I_t who consumes a final good Y_t with price P_t:

$$P_t Y_t = I_t$$
$$Y_t = \frac{I_t}{P_t}$$

To form the final good Y, there is a continuum of plants who produce differentiated goods index by φ that can be substituted at rate $\rho > 1$, which can be extended to include multiple industries later on:

$$Y_t = \left(\int_0^1 y_t(\varphi)^{\frac{\rho}{\rho-1}} dF(\varphi) \right)^{\frac{\rho}{\rho-1}} \tag{17}$$

Each differentiated good (plant) faces then faces inverse demand:

$$p_t(\varphi) = \left(\frac{I_t}{y_t(\varphi)} \right)^{\frac{1}{\rho}} P_t^{\frac{\rho}{\rho-1}} \tag{18}$$

At the beginning of each period a given plant observes the set of fuels available today $\mathcal{F}_t$, the set of fuel-by-task productivity $\Psi_t \in \mathbb{R}^{Mx\mathcal{F}_t}$ and fuel prices p_{ft}, and it solves the static problem of allocating fuels to tasks, which gives each plant its price of energy $p_{et}(\Psi_t, \mathcal{F}_t)$ from (8) and its pollution intensity of energy, which will be useful later for characterizing optimal policy. The plant then observes its total factor productivity z_t, the fixed cost of expanding $\mathcal{F}_t$ by one fuel κ, capital k_t and input prices w_t, p_{mt}, r_t. It then chooses flexible inputs today l_t, m_t, E_t, capital next year k_{t+1} and the set of fuels available for next year to maximize the following value function:

$$V_t(k_t, z_t, \Psi_t, \mathcal{F}_t) = \max_{k_{t+1}, l_t, m_t, E_t, \mathcal{F}_{t+1}} I_t^{\frac{1}{\rho}} (P_t y_t(\varphi))^{\frac{\rho-1}{\rho}} - w_t l_t - p_{mt} m_t - r_t k_{t+1} - p_e(\Psi_t, \mathcal{F}_t)$$
$$- \mathbb{I}(\mathcal{F}_{t+1} > \mathcal{F}_t) + \beta E_t V_{t+1}(k_{t+1}, z_{t+1}, \Psi_{t+1}, \mathcal{F}_{t+1}) \tag{19}$$

$$\text{s.t. } y_t(\varphi) = z_t \left(\alpha_k k_t^{\frac{\sigma-1}{\sigma}} + \alpha_l l_t^{\frac{\sigma-1}{\sigma}} + \alpha_m m_t^{\frac{\sigma-1}{\sigma}} + \alpha_e E_t^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

I also assume that fuel-by-task productivity terms follow a unit root process, which means plants expected to get the same productivity draws tomorrow as the ones they observe today for fuels in $\mathcal{F}_t$, whereas they take expectation over the cross-sectional distribution of productivity for fuels that are not yet in $\mathcal{F}_t$. This has the implication that the allocation of fuels to tasks carries no dynamic implications. For example, if a plant doesn't use a fuel that

is in $\mathcal{F}_t$ because it got bad productivity draws for that fuel, it drops from $\mathcal{F}_{t+1}$. However, if the plant expects to get the same productivity next period, it also expect not to use that fuel next period. **Need to show this formally (I want $E(p_e(\Psi_{t+1}, \mathcal{F}_t)|\Psi_t) = E(p_e(\Psi_t, \mathcal{F}_t))$.**

$$\psi_{f,t+1}(\omega) = \psi_{ft}(\omega) + \epsilon_{f,t+1}(\omega)$$
$$\psi_{f,t+1} = \psi_{ft} + \epsilon_{f,t+1} \text{ by symmetry of tasks}$$

Where $E(\epsilon_{f,t+1}) = 0$. For now, I also assume that plants can only increase the set of fuels by one fuel. To get some intuition, it is useful to see this problem with a timeline:
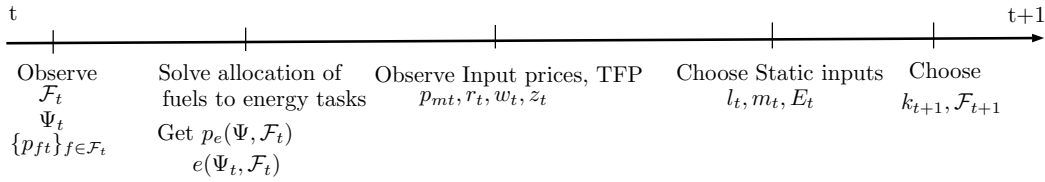

Figure 1: timeline of production decisions

Proposition 1. *Larger (more productive) plants are more likely to add a new fuel to the mix*

1.7 Adding a new fuel in a dynamic setting (preliminary)

Let us now consider the problem of a firm that wants to add a new fuel to its fuel mix $\mathcal{F}$. It pays a fixed cost κ today and have access to this new fuel next period and all subsequent periods. For example, one can think of a firm located away from gas pipelines considering the inclusion of natural gas to its mix. It may need to transport liquefied natural gas (LNG) and install its own compressors/vaporizers to turn the liquid back into a gas which is a significant investment. The value function of a firm takes capital today, a vector of technology/productivity terms $(\mathcal{Z} = \{A_f\}, \{\psi_f(\omega)\}, z)$ which are assumed to evolve stochastically and the set of energy inputs/fuels available to the firm $\mathcal{F}$:

$$\mathcal{V}(k, \mathcal{Z}, \mathcal{F}) = Pz\left(l^{\frac{\sigma-1}{\sigma}} + k^{\frac{\sigma-1}{\sigma}} + m^{\frac{\sigma-1}{\sigma}} + E^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$$
$$- wl - p_m m - r_k k - p_e(\mathcal{F}, \mathcal{Z})E - \sum_f \kappa_f \mathbb{I}(e_f \in \mathcal{F}' \setminus \mathcal{F}) + \beta \mathbb{E}_{x \in \mathcal{Z}} \mathcal{V}(k', \mathcal{Z}', \mathcal{F}')$$

Where $p_e(\mathcal{F}, \mathcal{Z})$ is given when solving:

$$\min_{\{e_f\}_{f \in \mathcal{F}}} \sum_{f \in \mathcal{F}} p_f e_f$$

$$s.t. \ E = \left(\sum_{f \in F} \Gamma_f(\mathcal{Z})^{1/\lambda} \left(A_f e_f\right)^{(\lambda-1)/\lambda}\right)^{\lambda/(\lambda-1)}$$

1.8 Solving for Optimal Energy Input Choices - Initial year Static

Plants indexed by i are heterogeneous in fuel-by-task specific productivity $(\psi_{if}(\omega))$. Without loss of generality, I remove fuel-augmenting productivity terms (A_f) from the analysis as they are endogenously determined by the distribution of ψ's across all tasks. In the first year of production, I assume that plants draw fuel-by-task productivity terms for each fuels and tasks from a fuel-specific Pareto distribution such that:

$$P(\psi_{if}(\omega) \geq x) = \left(\frac{x}{\gamma_f}\right)^{-\alpha_f}$$

Then, the probability fuel f is chosen over fuel j is:

$$\begin{aligned}
P_r\left(\frac{p_f}{\psi_{if}(\omega)} \leq \frac{p_j}{\psi_{if}(\omega)}\right) &= \left(\frac{p_f}{p_j\gamma_f}\right)^{-\alpha_f} \int_{\gamma_j}^{\infty} \psi_j^{-\alpha_f} f(\psi_j(\omega)) d\psi_j \\
&= \left(\frac{p_f}{p_j\gamma_f}\right)^{-\alpha_f} \int_{\gamma_j}^{\infty} \psi_j^{-\alpha_f-\alpha_j-1} \alpha_j \gamma_j^{\alpha_j} d\psi_j \\
&= \left(\frac{p_f}{p_j\gamma_f}\right)^{-\alpha_f} \left[\frac{\psi_j^{-\alpha_f-\alpha_j}}{-\alpha_f-\alpha_j} \alpha_j \gamma_j^{\alpha_j}\right]_{\gamma_j}^{\infty} \\
&= \left(\frac{p_f}{p_j}\right)^{-\alpha_f} \left(\frac{\gamma_j}{\gamma_f}\right)^{-\alpha_f} \frac{\alpha_j}{\alpha_f+\alpha_j}
\end{aligned}$$

Then, if there are F number of fuels, the probability that fuel f is chosen over all other fuels to perform task ω is going to be:

$$\begin{aligned}
\tau_f(\omega) &= \prod_{j \neq f} P_r\left(\frac{p_f}{\psi_{if}(\omega)} \leq \frac{p_j}{\psi_{ij}(\omega)}\right) \\
&= \left(\frac{\gamma_f}{p_f}\right)^{(F-1)\alpha_f} \prod_{j \neq f} \left(\frac{\gamma_j}{p_j}\right)^{-\alpha_f} \frac{\alpha_j}{\alpha_f+\alpha_j} = \tau_f
\end{aligned} \tag{20}$$

If there are M tasks, I can define the probability that a given fuel is used m times which follows a binomial distribution. Let $M\mathcal{T}_f$ be the number of times fuel f is chosen:

$$P_r(M\mathcal{T}_f = m) = \left(\frac{M}{m}\right) \tau_f^m (1-\tau_f)^{M-m} \tag{21}$$

Where $\left(\frac{M}{m}\right) = \frac{M!}{m!(M-m)!}$. Notice that by the Poisson Limit Theorem, the limiting distribution (as $M \to \infty$) only exist for $m < M$. This assumption is used extensively in the literature on technological choices (**Kortum 1997, Buera and Oberfield 2020, Boehm and Oberfield 2020**). However, in my context it means there can not be any corner solution where a fuel is used for all tasks if the number of tasks goes to infinity. For this reasons, I keep the assumption of a finite number of tasks for the remainder of the paper.

With (15) I can now characterize the distribution of fuel shares, which is going to be continuous because productivity draws affect how much fuel is consumed. For example, when tasks are gross complement ($\lambda < 1$) a lower quantity of fuel is going to be needed if it is very productive and vice versa. For a given plant, let fuel shares be:

$$s_f = \frac{e_f}{\sum_f e_f}$$

$$= \frac{p_f^{-\lambda} \Gamma_f}{\sum_f p_f^{-\lambda} \Gamma_f} \tag{22}$$

First moment (mean)

Notice that with M tasks and a probability of success τ_f the average number of times fuel f is chosen in the data is going to be:

$$E(M\mathcal{T}_f) = M\tau_f$$
$$E(\mathcal{T}_f) = \tau_f$$

Then, I can use that to characterize $E(\Gamma_f)$ and $E(s_f)$:'

$$E(\Gamma_f) = E\left(\frac{1}{M} \int_{\mathcal{T}_f} \psi_f(\omega)^{\lambda-1} d\omega\right)$$
$$= ???$$

10