# Transition to Green Technology along the Supply Chain

Philippe Aghion
*Collège de France, INSEAD, LSE* and CEPR

Lint Barrage
*ETH Zurich and CEPR*

David Hémous
*U Zurich and CEPR*

Ernest Liu*
*Princeton and NBER*

March 2024

**Abstract**

We analyze a model of green technological transition along a supply chain. In each layer, a good is produced with a dirty technology, or, if the required "electrification" innovation has occurred, with a clean technology which uses the immediate upstream good. We show that the economy is characterized by a single equilibrium but multiple steady-states, and that even in the presence of Pigouvian environmental taxation, a targeted industrial policy is generally necessary to implement the social optimum. We also show that: (i) small, targeted, industrial policy may bring large welfare gains; (ii) a government which is constrained to focus its subsidies to electrification on one particular sector, should primarily target downstream sectors; (iii) when extending the model so as to allow for supply chains also for the dirty technology, overinvesting in electrification in the wrong upstream branch may derail the overall transition towards electrification downstream. Finally, we illustrate our model with a calibration to decarbonization of global iron and steel production via hydrogen direct reduction, and show that, absent industrial policy, the economy can get stuck in a "wrong" steady-state with CO2 emissions vastly above the social optimum even with a carbon price in place.

# 1 Introduction

There is a growing consensus worldwide of the need to speed up the transition away from fossil fuel energy so as to slow down and eventually curb global warming. However, there is all but a

*We thank John Hassler for helpful discussions. Philippe Aghion acknowledges the support he received from the Kurt Björklund Chair in Innovation and Growth at INSEAD, and also from LVMH and SCOR. David Hémous gratefully acknowledges the financial support of the European Commission (DG ECFIN).

unanimity among policy leaders when it comes to the best choice of climate policy instrument. In Europe, carbon price and carbon taxes appear to hold the upper hand, whereas in the US or China, priority has been given to industrial policy as illustrated by the Inflation Reduction Act.

However, until recently most contributions in the economics of climate change have emphasized the carbon tax as the way to solve the climate problem without paying much attention to other instruments. In particular, despite long-standing recognition that a broader portfolio of policies may be required to address market failures in technological development and adoption, formalizations of such policy portfolios have been comparatively limited.[1] One influential recent attempt at introducing a second leg in the design of climate policy design was Acemoglu et al. (2012), which developed a growth model with directed innovation where both, the carbon tax and subsidies to green innovation, are needed to minimize the social cost of redirecting firms' innovation towards clean technologies. While some authors might already consider green innovation subsidies as a particular instance of industrial policy, no serious attempt so far has been made at rationalizing the use of vertically targeted subsidies to fight climate change. By "vertically targeted subsidies" we mean subsidies that target particular sectors, and/or that treat different sectors differently.

In this paper we develop a model of technological transition along a supply chain to make the case for sector-specific industrial policy to best address the energy transition problem. In our model, a final consumption good (or service) can be produced either with a dirty or with a clean technology, where the dirty technology directly generates pollution. Think of a clean car versus a dirty car. However, (even) the clean technology uses an upstream input which itself can be produced using a clean or a dirty technology. Think again of a clean car having either clean or dirty components. And, even if the input production uses a clean technology, that technology itself requires a more upstream input which itself can be produced using a cleaner or a dirtier technology. And so, we keep moving upstream to increasingly more basic inputs, where at each stage the choice can be made between the dirty technologies which directly generate pollution, and clean technologies which do not directly generate pollution, but require a more basic input which itself may be produced using either cleaner or dirtier technologies.

More specifically, each layer in the supply chain produces a "good" which is a Cobb-Douglas aggregate of a mass of industrial processes or varieties. Each variety of that good can be produced either in a dirty way using labor only or, if that process has been "electrified", in a clean way using labor and the immediate upstream good. To move from dirty to clean, a variety in a given sector—i.e. a given layer in the supply chain—needs to undergo "electrification".[2] We assume

[1] For early examples see, e.g., Jaffe and Stavins (1995) or Fischer and Newell (2008).

[2] Electrification is a stand-in here for a process that allows replacing fossil fuels as a production inputs with other inputs that do not directly generate emissions. In practice, this often but not always involve replacing fossil fuels with electricity.

heterogeneous fixed costs of electrification across varieties within a sector, so that some varieties may be electrified at one point in time while others still have to electrify. We also assume that once the social cost of carbon has been taken into account, producing using the clean technology is cheaper than producing using the dirty technology, so that a variety producer will always choose the former once the variety has been electrified, provided that a Pigouvian carbon tax is in place.

Under these assumptions, the incentive to electrify, for any variety producer in any sector on the supply chain, depends both on the degree of electrification downstream (more electrification downstream, which uses the variety as inputs, increases the demand for electrifying the variety) and on electrification upstream (more electrification upstream, which supplies inputs to the variety, reduces the cost of producing the variety once it has been electrified). The resulting cross-sectoral strategic complementarities in electrification generate a coordination problem: namely, insufficient electrification in sectors downstream or upstream of a given sector reduces the private incentives for variety producers in that sector to electrify.

We characterize the equilibrium in the absence of industrial policy. We show that the equilibrium achieved at any particular date is unique for given initial conditions, but that the cross-sectoral strategic complementarities in electrification generally lead to a multiplicity of steady-states.

We then characterize the social optimum, and show that it generally differs from the decentralized solution even once emissions are optimally priced through a Pigouvian tax. Abstracting from the difference in the time horizon of the social planner and private agents, the key source of inefficiency is that the decentralized economy may be stuck in a local steady-state which is not the steady-state corresponding to the global optimum. As a result, the social optimum may differ from the equilibrium with Pigouvian emission taxes even though the private incentives to electrify are locally in line with the social incentives. Yet we show that the socially optimal steady-state is uniquely implemented through the combination of a Pigouvian tax with a whole set of time-varying sector specific subsidies.

Even though we have an environment that displays complementarities as in other models, e.g. the Big Push model of Murphy et al. (1989), the fact that our complementarities occur along the supply chain leads to drastically different insights.

First, our model generates the possibility that a small and temporary subsidy to electrification that targets key-sectors can be sufficient to achieve large welfare gains by moving the decentralized equilibrium just a little out of an inefficient steady-state and then relying on market forces to complete the transition towards the socially efficient steady-state: then, large, sustained interventions across all sectors are not needed.

Second, our framework has implications for how to prioritize public intervention between upstream and downstream. In particular, a government which is constrained to focus its subsidies to electrification on one particular sector, should primarily target downstream sectors, the reason being that electrification propagates more easily upstream (through the demand channel) than downstream (through the cost channel). This stems from the fact that, while increasing electrification in a downstream sector generates proportional gains in electrification incentives in more upstream sectors (the demand channel), increasing electrification in an upstream sector instead generates less than proportional gains in electrification incentives in more downstream sectors (the cost channel).[3] In particular, we develop an example showing that a one-shot subsidy to electrification that targets the most upstream sector in a supply chain with at least three layers, does not allow electrification to propagate to more downstream sectors.

Third, we rationalize the possibility of horizontally misdirected clean industrial policies, with resulting welfare losses. We illustrate this point in an extended version of our model with two layers, where, on top of using labor, the dirty technology for producing varieties in the downstream sector also uses inputs from another upstream sector that can be electrified. In that context, electrification in the two upstream sectors—for the dirty input and for the clean input to downstream varieties—are strategic substitutes. Over-investing in the electrification of the upstream sector associated with the dirty technology downstream may derail the overall transition towards electrification downstream.

Finally, we present a quantitative application of our model to global iron and steel production, a sector that accounts for 7-9% of global carbon dioxide emissions (Kim et al. 2022). One of the most promising options for deep decarbonization in this sector is hydrogen-based direct reduction-electric arc furnace ("H2-DR-EAF") production (Delvin et al. 2023; BloombergNEF 2021). Hydrogen, in turn, is currently produced mainly from fossil fuels (IEA 2021) but can also be produced from electrolysis powered by renewable energy. We thus model the interplay between upstream hydrogen and downstream iron and steel production. The benchmark calibration confirms the quantitative importance of the multiplicity of steady states: A uniform carbon price of $25/tCO2, for example, is consistent with three steady states: (0% clean H2, 0% clean steel), (59% clean H2, 54% clean steel), and (82% clean H2, 84% clean steel). The stakes of being stuck in the "wrong" steady state are enormous: even at current levels of global steel production, the difference in annual emissions between the (0%, 0%) and (82%, 84%) scenarios is upwards of 2.4 billion tons of CO2 per year, close to the entirety of the European Union's CO2 emissions from all sectors (2.8 billion tons in 2022, EDGAR 2023). The difference between the (59%, 54%) and (82%, 84%) scenarios is similarly large at around1 billion tons of CO2 per year.

Our paper relates to several strands of literature. First, on the macroeconomics of climate change, starting with the so-called "Integrated assessment models" literature (IAMs) initiated by

[3] This in turn is due to the fact that the clean technology at all layers in the supply chain does not only use the more upstream input but also labor which captures a positive share of the benefit from the cost reduction induced by the initial electrification shock.

Nordhaus (1994), and pursued more recently, e.g. by Golosov, Hassler, Krusell, and Tsyvinski (2014). However, these papers take technological change as given, and their emphasis is on the optimal design of carbon tax policies. More closely related to our analysis is the literature on directed technical change and the environment, in particular Acemoglu, Aghion, Bursztyn, and Hémous (2012), henceforth AABH, who show that the optimal climate policy in the presence of endogenous directed innovation, amounts to the combination between a carbon tax and a green research subsidy. The model in AABH relies on knowledge externalities, not on cross-sectoral strategic complementarities and coordination, as it does not model clean innovations along the supply chain.[4]

Second, a set of recent contributions study the positive properties of environmental policies in static production networks with carbon emissions, a setting where the optimal intervention remains a uniform carbon tax (e.g., see King, Tarbush and Teytelboym 2019, Devulder and Lisack 2020).[5] Even though our model also features production network with environmental externalities, our analysis differs fundamentally from these papers because our setting features endogenous technological change, and the key policy issue is to address the dynamic strategic complementarity of technology adoption along the supply chain, and not the static environmental externalities. In this regard, we relate to papers on strategic complementarities in technology adoption, including the seminal work of Murphy et al. (1989), recent work by Sturm (2023), and in an environmental setting, Greaker and Midttømme (2016) and Dugoua and Dumas (2021). While these papers study static models featuring multiple equilibria, which poses a difficulty for analyzing the impact of policy interventions, our dynamic model has multiple steady-states but a unique equilibrium path;[6] our model therefore has unambiguous predictions on industrial policy's equilibrium impact. The feature enables our model to maintain tractability—despite rich strategic interactions along the supply chain—and generate new insights on how industrial policy could target key sectors to aid the technological transition.

A third related strand is the literature on industrial policy.[7] A first rationale for industrial policy is the infant industry argument, recently modeled by Greenwald and Stiglitz (2006), which emphasizes "learning by doing" externalities between the nascent domestic industry and the agricultural sector. Another justification is the existence of cross-sectoral demand externalities: here, a classical reference is Murphy et al (1989) who model Rosenstein-Rodan (1943)'s Big Push idea, namely, investments in increasing returns technologies in some markets or sectors, generates

[4] Acemoglu et al. (2016) present another model of directed technical change in the environmental context, which is perhaps closer to ours. In their model, intermediates are aggregated in a Cobb-Douglas fashion and can be produced with a clean or a dirty input (as in our model), however, there is no supply chain or coordination issues.

[5] See Martin, Muuls and Stoerk (2023) for an empirical evaluation of carbon pricing's effect along supply chains.

[6] Also see Crouzet, Gupta and Filippo Mezzanotti (2023) for a dynamic model of strategic complementarity in the context of adopting electronic payments.

[7] For an illuminating recent survey on the industrial policy literature, see Juhasz, Lane and Rodrik (2023).

additional income which increases the incentives for industrialization in other sectors through a demand channel. However, none of these papers considers supply chains and the associated motives for sectoral policies.

Closer to our paper is Liu (2019), who analyzes the role of sector-specific policy interventions in a production network setting with market imperfections and finds that targeting upstream sectors—as was done in South Korea and China—could lead to aggregate welfare gains. Relatedly, Liu and Ma (2023) investigate the optimal cross-sector allocation of R&D in the context of an endogenous growth framework that incorporates an innovation network, featuring knowledge spillovers across technologies. Donald (2023) embeds the innovation network à la Liu and Ma (2023) into AABH's model of directed innovation between clean and dirty technologies and studies the policy implications. Contemporaneous work by Buera and Trachter (2024) studies the role of industrial policy in a static production network model with endogenous technology adoption. Compared to all of these papers, the unique feature of our environment is the strategic complementarity of technology adoption *across* different sectors along the supply chain. As we show, such strategic complementarity has strong implications for the efficacy and targeting of industrial policy, the analysis of which is the focus of this paper.

The remaining part of the paper is organized as follows. Section 2 outlays our baseline model of supply chain and electrification. Section 3 derives the equilibrium equations and establishes both, the uniqueness of an equilibrium for given initial conditions and the possibility of multiple steady-states. Section 4 derives the equations for the social optimum. It compares the social optimum with the steady-states of an economy without industrial policy. It then shows how the social optimum can be implemented by combining a Pigouvian carbon tax with suitably a chosen time-varying set of sector-specific subsidies. Section 5 analyzes the propagation of an exogenous one-sector electrification shock, and how this propagation depends upon whether the selected sector is more downstream or upstream. Section 6 analyses the possibility of horizontally misguided—and consequently backfiring—industrial policy, using an extension of our baseline model where the dirty technology for producing the consumption good also uses an upstream input that can be produced with a clean or a dirty technology. Section 7 presents the quantitative application. Section 8 concludes.

# 2 Model

In this section, we present our baseline model of a vertical supply chain in a green transition.

## 2.1 Preferences and production technology

Time is discrete and denoted by $t$. The consumer demand side of the economy consists of a continuum of mass one of agents with the same intertemporal utility

$$\sum_{t=0}^{\infty} \beta^t \left(\ln c_t - \ell_t - a_t\right)$$

where $c_t$ denotes the consumption flow, $\ell_t$ denotes the representative individual's labor supply, $a_t$ denotes the disutility of pollution, at time $t$, and $\beta$ is the discount factor.

The production side is a vertical supply chain consisting of $N$ layers or "sectors" which we rank from the most upstream, namely $i =$1, to the most downstream, $i = N$, which corresponds to the consumption good. Production $y_{it}$ in each sector $i$ at time $t$ is the aggregate outcome of a continuum of mass one of sector-specific varieties ($\nu$), according to:

$$\ln y_{it} = \int_0^1 \ln y_{it}\left(\nu\right) \, \mathrm{d}\nu.$$

In turn each variety $\nu$ in any sector $i$ can be produced using either a dirty or a clean technology. Three features distinguish clean and dirty technologies. First, the dirty technology is associated with pollution while the clean one is not. Second, the dirty technology is always available, while the clean technology is only available for a variety $\nu$ in sector $i$ that is "electrified"—electrification occurs through a process described below. And third, the dirty technology only uses labor (one for one), while the clean technology uses good $i - 1$ as an intermediate input together with labor in a Cobb-Douglas fashion. More specifically, we assume that the most upstream input is produced according to:

$$y_{1t}\left(\nu\right) = \ell_{d1t}\left(\nu\right) + \gamma\left(\nu\right) e^z \ell_{cit}\left(\nu\right),$$

and that for all $i > 1$ :

$$y_{it}\left(\nu\right) = \ell_{dit}\left(\nu\right) + \gamma\left(\nu\right) \left(\frac{e^z \ell_{cit}\left(\nu\right)}{\alpha_i}\right)^{\alpha_i} \left(\frac{m_{it}\left(\nu\right)}{1-\alpha_i}\right)^{1-\alpha_i}, \tag{1}$$

where: (i) $\gamma(\nu)$ is an indicator function, equal to $1$ for electrified varieties and to $0$ for non-electrified varieties; (ii) $\alpha_i \in (0, 1)$ for all $i > 1$; (iii) for all $i$, $\ell_{dit}\left(\nu\right)$ denotes the labor input used by variety $\nu$ in sector $i$ using the dirty technology; (iv) for all $i$, $\ell_{cit}\left(\nu\right)$ and $m_{it}\left(\nu\right)$ denote respectively the labor input and the amount of intermediate input from sector $i - 1$ used for producing variety $\nu$ with the clean technology; (v) $e^z$ is the relative (labor-augmenting) productivity of using the clean rather than the dirty technology. Our assumption that only the clean technology uses a supply chain is less stringent than it may appear at first sight. Our analysis can directly accommodate a supply chain for the dirty input, as long as no layers on that chain can be electrified: in that case, the different dirty sectors can simply be aggregated along the value chain and this economy is de facto identical to one where only labor is used for the dirty technology.

Section 6 extends the model and considers the case where electrification is also possible for dirty technologies.

We assume that producing each unit of output using the dirty technology generates $\xi$ units of disutility due to pollution. The total disutility of pollution is thus $a_t = \xi \ell_{dt}$, where $\ell_{dt}$ is the total labor input used for dirty production in the economy, and $\xi$ is the social unit cost of pollution.[8]

For simplicity, we have assumed here that relative productivities $z$ and emission rates $\xi$ are the same for all sectors. This assumption can easily be relax and we do so in our calibration exercise in Section 7 below.

## 2.2 Market clearing

Labor market clearing requires that, at any time $t$ :

$$\ell_{ct} = \sum_{i=1}^{N} \ell_{cit} = \sum_{i=1}^{N} \int_0^1 \ell_{cit}(\nu)\, \mathrm{d}\nu, \quad \ell_{dt} = \sum_{i=1}^{N} \ell_{dit} = \sum_{i=1}^{N} \int_0^1 \ell_{dit}(\nu)\, \mathrm{d}\nu, \quad \ell_t = \ell_{ct} + \ell_{dt} + \ell_{et},$$

where: $\ell_{ct}$ and is the total labor used for clean production (summing across all sectors $i = 1, \ldots N$ and integrating across varieties within each sector); $\ell_{dt}$ is the total labor used for dirty production; and $\ell_{et}$ is the total labor employed for electrification (we specify the electrification technology below); and $\ell_{cit}$ and $\ell_{dit}$ are the total labor inputs used for clean and dirty usage in sector $i$.

Market clearing in the upstream sectors $i = 1, \ldots, N-1$, requires that for each $i$ the total output of sector $i$ be equal to the total use of it as intermediate inputs across varieties in sector $i+1$:

$$y_{i,t} = \int_0^1 m_{i+1,t}(\nu)\, \mathrm{d}\nu$$

whereas market clearing for the downstream consumption good $N$ simply boils down to:

$$c_{Nt} = y_{Nt}.$$

## 2.3 Electrification and market power

To operate the clean technology, the variety must be "electrified". For any variety $\nu$ in any sector $i$, the producer of that variety must incur a one-time sunk cost which is sector-variety specific to electrify production of the variety. We order varieties in any sector $i$ by increasing cost of electrification and therefore we let $\phi_i(s)$ denote the cost of electrification associated with the

[8] This formulation is equivalent to one where production with the dirty technology uses a free and inexhaustible fossil fuel resource together with labor in a Leontief fashion. Alternatively, the model can accommodate extraction costs for the resource if $\xi$ includes both environmental damages and the extraction costs and the carbon tax $\tau$ introduces below includes both the tax itself and the extraction cost.

variety quantile $s$ in sector $i$, where that cost is expressed in labor units.[9]

We let $F_i(\cdot)$ denote the CDF cost distribution, that is, for any cost $\phi$, $F_i(\phi)$ is the measure of the set of variety quantiles $s$ in sector $i$ with costs $\phi_i(s)$ less than $\phi$. A special case we shall consider is the case where the distribution of electrification costs across varieties in sector $i$ is a mass point at some cost level $\phi_i$. Let $\chi_{it}$ denote the fraction of electrified varieties at time $t$ in sector $i$, which also corresponds to the cut-off quantile $s$ beyond which varieties cease to be electrified. The sum of all electrification fixed costs up to $\chi_{it}$ is given by $\mathcal{F}_i(\chi_{it}) \equiv \int_0^{\chi_{it}} \phi_i(s)\, \mathrm{d}s$.[10] The collection of $\chi_{it}$'s across sectors form the key state variables of the economy. Starting from an initial condition $\{\chi_{i0}\}_{i=1}^N$ at $t = 0$, electrification raises $\chi_{it}$'s monotonically over time, until $\chi_{it}$'s converge to a steady-state.

We assume that the dirty technology is operated competitively, but that, if she electrifies her variety, the producer of that variety earns a one-period monopoly profit upon electrification, with fringe producers operating the dirty technology as competitors. Starting from the subsequent period, the clean technology becomes freely available to all producers so that the variety is again competitively produced.

Finally, note that we refer to the innovation that enables the use of the clean technology as "electrification" because, in practice, such an innovation often means replacing fossil fuels with electricity as an energy source. However, this need not always be the case, in our example below fossil fuels will be replaced by hydrogen, and we do not explicitly model a separate electricity sector.

## 2.4 Policy instruments

We assume that the government can: (i) impose a carbon tax $\tau$ on any variety which uses the dirty technology; (ii) impose a cap-and-trade limit $\bar{\ell}_d$ on the amount of dirty input used by any variety (leading to a carbon price $\tau$); (iii) (industrial policy) implement a set of sector-specific electrification subsidies $\kappa_i$. All of these policy instruments are allowed to be time-varying, and our analysis will focus on showing that the carbon tax and/or the cap-and-trade alone do not implement the optimum, whereas suitable sector-specific electrification subsidies $\kappa_i$ can implement the optimum.

[9] Our model of electrification is directly inspired by the automation literature where after an innovation or payment of a fixed cost, labor can be replaced by capital in the production process (see e.g., Zeira 1998).

[10] There is an obvious relationship between $F_i$ and $\mathcal{F}_i$. Namely, for any $s : F_i(\phi_i(s)) = s$ or equivalently $\phi_i(s) = F_i^{-1}(s)$. This directly leads to $\mathcal{F}_i(\chi_{it}) = \int_0^{\chi_{it}} F_i^{-1}(s)\, \mathrm{d}s$.

# 3 Equilibrium

In this section we characterize the equilibrium of the economy if it only resorts to a carbon tax.

## 3.1 Equilibrium price

In each period, the representative consumer solves

$$\max_{c_t,\ell_t} \ln c_t - \ell_t - a_t \qquad \text{s.t. } p_t c_t = w_t \ell_t + \pi_t$$

where $p_t$ is the price index of the consumption good, $\pi_t$ is the net transfer from monopolistic producers and the government, and $w_t$ is the wage rate with we normalize to one. Consumer optimization implies that the total expenditure on the consumption good is equal to one, that is:

$$p_t c_t = 1.$$

Given that dirty production uses only labor one-for-one, then, absent policy interventions the marginal cost of the dirty technology is equal to the wage, namely 1. If instead the government taxes pollution (or imposes a binding cap-and-trade limit), the marginal cost of dirty production is $1+\tau$, with $\tau > 0$. We assume that $\tau$ and/or $z$ are large enough to ensure that $1+\tau > e^{-z}$. This condition ensures that a producer always uses the clean technology once her variety has been electrified. For notational simplicity, we define $Z \equiv \ln(1+\tau) + z$, the tax-adjusted relative productivity of clean versus dirty technology, and we assume that $Z > 0$.

The equilibrium price index $p_{it}$ for good $i$ satisfies:

$$\ln p_{it} \equiv \int_0^1 \ln p_{it}(\nu)\, \mathrm{d}\nu, \tag{2}$$

where, for each variety $\nu$ in sector $i$, $p_{it}(\nu)$ is determined as follows. If variety $\nu$ has been electrified in the previous period, then it is priced at the marginal cost associated with the clean technology . If variety $\nu$ has not been electrified in the previous period, then it is priced at the marginal cost of a dirty producer. This is trivial for producers using the dirty technology, but it is also true for a newly electrified variety. The producer of a newly electrified variety is a monopolist for the clean technology facing a fringe that uses the dirty technology. Given the unit demand elasticity, she will charge a price equal to the marginal cost of the fringe.

The market structure implies that in the most upstream sector (sector 1), the price of a variety that has been electrified in the previous period is equal to $p_{1t}(\nu) = e^{-z}$, whereas the price of a non electrified or of a newly electrified variety is $p_{1t}(\nu) = 1+\tau$.

Next, in sectors $i > 1$, we have:

$$p_{it}(\nu) = \begin{cases} e^{-\alpha_i z} p_{i-1,t}^{1-\alpha_i} & \text{if the variety has been electrified by time } t-1, \\ 1+\tau & \text{otherwise,} \end{cases} \tag{3}$$

as the marginal cost of producing dirty is $1+\tau$, while the marginal cost of producing with the clean technology is $e^{-\alpha_i z} p_{i-1,t}^{1-\alpha_i}$.

To get more explicit expressions for these equilibrium prices, it will be helpful to consider, for all $i$, the network adjusted share $\mu_{it}$ of electrified content in the production of any electrified variety in sector $i$. In the most upstream sector 1, we obviously have $\mu_{1t} = 1$. In more downstream sectors $i > 1$, $\mu_{it}$ is recursively determined by:

$$\mu_{it} = \alpha_i + (1-\alpha_i)\, \chi_{i-1,t} \mu_{i-1,t} \tag{4}$$

In words, the electrified content of an electrified variety in sector $i$, is equal to the direct share of clean labor input $\alpha_i$ plus $(1-\alpha_i)$ times the aggregate electrified content of input $i-1$, which in turn is equal to the electrified content $\mu_{i-1,t}$ of each electrified variety in sector $i-1$ times the fraction $\chi_{i-1,t}$ of electrified varieties for good $i-1$.

We can now solve for the price index in each sector. For the most upstream sector 1, we can use (2) and (3) to get:

$$p_{1t} = (1+\tau)\, e^{-\chi_{1,t-1} Z}.$$

Then moving downstream, we get by induction:

$$p_{it} = (1+\tau)\, e^{-\chi_{i,t-1} \mu_{it-1} Z}. \tag{5}$$

## 3.2 Equilibrium profits

The incentive to electrify depends on the profits that a newly electrified variety obtain. We now derive these profits. The producer of a newly electrified variety charges a price $1+\tau$ but faces marginal costs equal to $e^{-\alpha_i z} p_{i-1,t}^{1-\alpha_i}$, therefore she charges a mark-up $\theta_{i,t}$ given by

$$\theta_{i,t} = \frac{1+\tau}{e^{-\alpha_i z} p_{i-1,t}^{1-\alpha_i},} = e^{Z \mu_{it-1}},$$

where the second equality uses equations (5) to substitute for $p_{i-1t}$ and (4). The share of revenues going to profits is $1-\theta_{i,t}^{-1}$ and the share going to input costs is $\theta_{i,t}^{-1}$.

The next step is to derive the equilibrium revenue $r_{it}$ of a producer in sector $i$ at time $t$. In the most downstream sector $N$, which produces the final consumption good, we already know that:

$$r_{Nt} = p_t c_t = 1.$$

From there we move upstream, as revenues trickle up from the most downstream to the upstream sectors. Take as given the revenues $r_{i+1,t}$ of a producer in sector $i+1$ at time $t$. Then, sector $i$'s good is only used as an input by the electrified varieties in sector $i+1$. For any variety in sector $i$,the revenue $r_{it}$ includes both sales to previously electrified varieties and sales to newly electrified varieties. There is a mass $\chi_{i+1,t-1}$ of previous electrified varieties. These are produced competitively, so that a share $1-\alpha_{i+1}$ of their revenues go to sector $i$. There is a mass $\chi_{i+1,t} - \chi_{i+1,t-1}$ of newly electrified varieties, only a share $\theta_{i+1,t}^{-1} = e^{-Z\mu_{i+1.t-1}}$ of their revenues go to the payment of inputs, out of which a share $1-\alpha_{i+1}$ go to sector $i$. We then obtain:

$$\begin{aligned} r_{it} &= \underbrace{\chi_{i+1,t-1} r_{i+1,t} \left(1-\alpha_{i+1}\right)}_{\text{sales to previously electrified varieties}} + \underbrace{\left(\chi_{i+1,t} - \chi_{i+1,t-1}\right) r_{i+1,t} e^{-Z\mu_{i+1,t-1}} \left(1-\alpha_{i+1}\right)}_{\text{sales to newly electrified varieties}} \\ &= \widetilde{\chi}_{i+1,t} r_{i+1,t} \left(1-\alpha_{i+1}\right) \end{aligned}$$

where we define the share of revenues of sector $i+1$ spent on inputs for the clean technology:

$$\widetilde{\chi}_{i+1,t} \equiv \chi_{i+1,t-1} + \left(\chi_{i+1,t} - \chi_{i+1,t-1}\right) e^{-Z\mu_{i+1,t-1}}. \tag{6}$$

This in turn immediately yields the following expression for the equilibrium revenue accruing from downstream to good $i$ production:

$$r_{it} = \prod_{j=i+1}^{N} \left(\widetilde{\chi}_{jt}(1-\alpha_j)\right). \tag{7}$$

The corresponding profit rent from electrification for any variety producer in sector $i$, is then simply equal to the profit share times revenues:

$$\pi_{it} = (1 - e^{-Z\mu_{it-1}}) r_{it} \tag{8}$$

## 3.3 Electrification

A variety producer in sector $i$ electrifies at time $t$ if and only if $\pi_{it}$ is bigger than the cost of electrification of that variety. It then immediately follows that the equilibrium share of electrifying varieties $\chi_{it}$ in sector $i$ at time $t$ satisfies:

$$\chi_{it} = F_i(\pi_{it}) = F_i\left(\left(1 - e^{-Z\mu_{it-1}}\right) \prod_{j=i+1}^{N} \left(\widetilde{\chi}_{jt}(1-\alpha_j)\right)\right) \tag{9}$$

whenever this is greater than $\chi_{it-1}$ and to $\chi_{it-1}$ otherwise, as there is no dis-electrification.

As equation (9) shows very clearly, the incentive of a variety producer in sector $i$ to electrify depends positively upon both upstream and downstream electrification. On the one hand, (past and contemporaneous) downstream electrification increases total revenue to good $r_{it}$ accruing to good $i$ production, and on the other hand (past) upstream electrification increases $\mu_{i,t-1}$ and therefore the rent share $\left[1 - e^{-Z\mu_{it-1}}\right]$ to a newly electrified producer in sector $i$.

Two further remarks can be made at this point. First, there are complementarities in electrification across sectors which will be a source of multiplicity of steady-states. As we shall see below, absent adequate industrial policy nothing guarantees that the economy will converge to the equilibrium with maximum electrification.

Second, note that the incentives to electrify travel upstream contemporaneously whereas they travel downstream with a one period lag. This features ensures that, given any initial condition $\{\chi_{i0}\}$ the model features a unique equilibrium path $\{\chi_{it}\}_{t\geq 0}$. We now explain the timing of how incentives to electrify travel along the supply chain. Moving upstream: as downstream sectors $i+1$ to $N$ electrify they increases the revenue accruing to the immediate upstream sector $i$ which in turn also increases the incentive to electrify in that sector: this is captured by the $\prod_{j=i+1}^{N}\widetilde{\chi}_{jt}$ factor in $r_{it}$. Moving downstream: as variety producers in sector $i$ electrify, the cost of production of good $i$ declines, which in turn leads to the decline in the price of that good next period once the monopoly power of the electrifying producers in sector $i$ has expired. This in turn increases the profit margin of electrifying producers in the more downstream sector $i+1$, thereby inducing more electrification in that sector. The logic extends to sectors which are further downstream.

### 3.4 Equilibrium equations

Our above analysis leads us to the following simple characterization of an equilibrium. Given initial electrification shares $\{\chi_{i0}\}$ at time zero, an equilibrium with a sequence of carbon taxes $\{\tau_t\}$ is a sequence of electrification shares $\{\chi_{it}\}_{t>0}$ such that

$$\mu_{1t}=1, \quad \mu_{it}=\alpha_i+\chi_{i-1,t}\mu_{i-1,t}(1-\alpha_i), \tag{10}$$

$$\widetilde{\chi}_{i+1,t}\equiv \chi_{i+1,t-1}+\left(\chi_{i+1,t}-\chi_{i+1,t-1}\right)\left(e^{z}\left(1+\tau_t\right)\right)^{-\mu_{i+1,t-1}} \tag{11}$$

$$\chi_{it}=\max\{\chi_{i,t-1},F_i\left(\left(1-\left[e^{z}\left(1+\tau_t\right)\right]^{-\mu_{it-1}}\right)\prod_{j=i+1}^{N}\left(\widetilde{\chi}_{jt}(1-\alpha_j)\right)\right). \tag{12}$$

Despite the complementarities in electrification, the market structure ensures that the equilibrium is unique:

**Proposition 1.** *Given initial condition $\{\chi_{i0}\}_{i=1}^{N}$, the economy with carbon taxes $\{\tau_t\}$ features a unique equilibrium path $\{\chi_{it}\}_{t>0}$.*

*Proof.* We proceed by induction. Suppose we know the sequence of electrification shares $\{\chi_{it-1}\}_{i=1}^{N}$ at time $t-1$. Then, note first that we can compute the sequence $\{\mu_{it}\}_{i=1}^{N}$ recursively from upstream moving downstream using the equations:

$$\mu_{1t}=1, \quad \mu_{it}=\alpha_i+\chi_{i-1,t}\mu_{i-1,t}\sigma_i$$

Next, given $\chi_{N,t-1}$ and the $\mu$'s at time $t-1$, one can compute $\chi_{Nt}$ using

$$\chi_{Nt}=\max\left\{\chi_{N,t-1},F_N\left(\left(1-\left[e^{z}\left(1+\tau_t\right)\right]^{-\mu_{Nt-1}}\right)\right)\right\}$$

meaning that the measure of electrified varieties at time $t$ cannot decrease, and new varieties are electrified if the cost of doing so is lower than the contemporary profit $\left(1-\left[e^{z}\left(1+\tau_t\right)\right]^{-\mu_{Nt-1}}\right)$.

Next, for given electrification shares $\chi_{jt}$ for $j \in \{j+1, N\}$ one can compute the equilibrium electrification share $\chi_{it}$, hence moving from downstream to upstream, using the equations

$$\widetilde{\chi}_{j,t} \equiv \chi_{j,t-1} + (\chi_{j,t} - \chi_{j,t-1}) \left(e^z (1+\tau_t)\right)^{-\mu_{j,t-1}} \text{ for all } j \in \{j+1, N\}$$

and then

$$\chi_{it} = \max\left\{\chi_{i,t-1}, F_i\left(\left(1 - \left[e^z (1+\tau_t)\right]^{-\mu_{it-1}}\right) \prod_{j=i+1}^{N} (\widetilde{\chi}_{jt}(1-\alpha_j)\right)\right\}.$$

By induction the equilibrium sequence of shares at date $t$, $\{\chi_{it}\}_{i=1}^{N}$, is uniquely pinned down. □

## 3.5 Multiplicity of Steady-States

We next characterize the steady-state(s) of the economy. In a steady-state, there are no newly electrified varieties, so that $\widetilde{\chi}_i = \chi_i$ for all $i$. It then follows that a steady-state with carbon tax $\tau$ is a set of electrification shares $\{\chi_i\}$ such that:

$$\mu_1 = 1, \quad \mu_i = \alpha_i + \chi_{i-1}\mu_{i-1}(1-\alpha_i) \text{ and} \tag{13}$$

$$\chi_i \geq F_i\left(\left(1 - \left[e^z (1+\tau)\right]^{-\mu_i}\right) \prod_{j=i+1}^{N} (\chi_j(1-\alpha_j)\right). \tag{14}$$

Condition (14) is an inequality for the same reason as the "max" notation in (12): technically, because electrification cannot decrease, any $\{\chi_i^{ss}\}$ for which (14) holds as a strict inequality is also a steady-state providing that the economy starts with $\chi_{i0} = \chi_i^{ss}$. These are, however, not very interesting steady-states since starting from $\chi_{i0} < \chi_i^{ss}$, there is no path that the economy can follow to reach these steady-states (without a direct government intervention). We therefore generally ignore them in our analysis, and we focus only on steady-states in which condition (14) holds as an equality:

$$\chi_i = F_i\left(\left(1 - \left[e^z (1+\tau)\right]^{-\mu_i}\right) \prod_{j=i+1}^{N} (\chi_j(1-\alpha_j)\right). \tag{15}$$

While the equilibrium is unique, complementarities in electrification ensure that the economy generally features more than 1 steady-state.

**Proposition 2.** *For given carbon tax $\tau$, there may exist multiple steady-states over a non-empty open set of parameters whenever $N \geq 2$. There exists a unique steady-state when $N = 1$.*

*Proof.* When $N = 1$, the steady-state equilibrium electrification share $\chi_1$ satisfies the equation:

$$\chi_1 = F_1\left(1 - e^{-Z}\right),$$

which has a unique solution.

Assume now that $N \geq 2$, to establish the result, we simply need to build an example with multiple steady state equilibria. We do so by assuming that the distribution of electrification costs is a mass point at some value $\phi$. We further assume that all the $\alpha_i$ 's are equal to the same $\alpha$ (except for $i = 1$). We derive conditions under which there exist a steady-state where all sectors fully electrify and another steady-states where no sector electrifies.

Consider first that no firm electrifies in any sector. In this case, the profit from electrification in all sectors $i < N$ is zero due to zero demand. The rent from electrifying in the most downstream sector $N$ is $\pi_N = 1 - e^{-\alpha Z}$. Thus there will be no electrification in sector $N$ whenever:

$$1 - e^{-\alpha Z} < \phi.$$

Next, suppose that all but a measure zero of firms electrify in all sectors. In this case, the profit from electrification in sector $i$ is $\pi_i = (1-\alpha)^{N-i}\left[1 - e^{-Z}\right]$. For all firms to have an incentive to electrify, we need that:

$$(1-\alpha)^{N-1}\left[1 - e^{-Z}\right] > \phi$$

Hence, in order for both, full and no electrification to be steady state equilibria, $\phi$ must satisfy:

$$1 - e^{-\alpha Z} < \phi < (1-\alpha)^{N-1}\left[1 - e^{-Z}\right].$$

This is possible as soon as there exist values of $z, \tau$, and $\alpha$ such that:

$$1 - e^{-\alpha Z} < (1-\alpha)^{N-1}\left[1 - e^{-Z}\right].$$

For instance, we note that for small $Z$ ($e^Z \approx 1 + z$) this inequality boils down hat the above inequality boils down to $\alpha < (1-\alpha)^{N-1}$which is satisfied for $\alpha$ sufficiently small. This completes the proof. □

Naturally, the same logic extends to the case where the government implements a cap-and-trade system instead of a carbon tax. While the equilibrium is always unique, multiple steady-states are possible: for instance, a steady-state with a high level of electrification and a low price of carbon can exist together with a steady-state with a low-level of electrification and a high price for carbon. Both steady-states may achieve the same level of emissions (if the cap binds) but the low electrification steady-state does it with lower output (see an example in Appendix 9.1).

# 4 Social optimum

## 4.1 Characterizing the optimum

We now characterize the social optimum, while in Section 5 below we shall consider a constrained optimum where only a limited number of sectors can undergo electrification each period. Electrification, once the fixed costs are paid, weakly pushes out the production possibility frontier; and, with linear labor disutility, the social costs of electrifying a given varieties are independent of the share of already electrified varieties. As a result, all electrification happens immediately in the optimum: the optimum features instantaneous electrification of (a fraction of) varieties in all sectors in the initial period, and no further electrification in subsequent periods.

Starting from initial condition $\{\chi_{i0}\}_i$, the social planner seeks to maximize the intertemporal utility of consumption minus the labor costs (including for labor hired in electrification) and minus the pollution costs associated with the use of the dirty technology. Keeping in mind that $c_t = y_{Nt}$ and that $\xi$ is the social cost of pollution, the social planner will solve:

$$\max_{c_t, \ell_{dit}, \ell_{cit}, \chi_i} \sum_{t=0}^{\infty} \beta^t \left( \ln y_{Nt} - (1+\xi) \sum_i \ell_{dit} - \sum_i \ell_{cit} \right) - \sum_i \left( \mathcal{F}_i \left( \chi_i \right) - \mathcal{F}_i \left( \chi_{i,0} \right) \right).$$

We assume that $(1+\xi) e^z > 1$, which ensures that the social planner uses the clean technology whenever it is available. Then, the social planner treats all electrified varieties in a given sector symmetrically and all non-electrified varieties symmetrically as well, which ensures that $\ell_{it}(\nu) = \ell_{cit}(\nu) = \ell_{cit}/\chi_{it}$ and $m_{cit}(\nu) = y_{i-1,t}/\chi_{it}$ if the variety is electrified, and $\ell_{it}(\nu) = \ell_{dit}(\nu) = \ell_{dit}/(1-\chi_{it})$ if the variety is not electrified. We can then write output in sector $i$ as:

$$\ln y_{it} = \chi_{it} \ln \left[ \left( \frac{e^z \ell_{cit}}{\chi_{it} \alpha_i} \right)^{\alpha_i} \left( \frac{y_{i-1,t}}{\chi_{it}(1-\alpha_i)} \right)^{1-\alpha_i} \right] + (1-\chi_{it}) \ln \frac{\ell_{dit}}{1-\chi_{it}}. \tag{16}$$

We can then characterize the social planner problem as follows:

**Proposition 3.** *The social planner's problem can be rewritten as:*

$$\max_{\{\chi_i \geq \chi_{i0}\}} \ln \left( (1+\xi) e^z \sum_{i=1}^{N} \chi_i \alpha_i \prod_{j=i+1}^{N} \left[ \chi_j (1-\alpha_j) \right] \right) - (1-\beta) \sum_i \mathcal{F}_i (\chi_i). \tag{17}$$

*If the solution $\{\chi_i\}$ is interior, it must satisfy the first-order conditions:*

$$\chi_i = F_i \left( \frac{\ln((1+\xi) e^z)}{1-\beta} \mu_i \prod_{j=i+1}^{N} (\chi_j (1-\alpha_j)) \right) \tag{18}$$

*Proof.* See Appendix 9.2. □

## 4.2 Steady-states versus the social optimum

At this stage it is worth comparing the steady-state decentralized equilibrium with the social optimum. Obviously, for given electrification shares $\chi's$, the decentralized economy in steady-state follows the optimal allocation provided that the carbon tax is set at its Pigouvian level: $\tau = \xi$.

More interestingly, we can compare electrification in the two economies by comparing equation (15) and the first-order condition for the social optimum (18). There are two differences. First, the factor $\frac{1}{1-\beta}$ on the LHS of (18) is absent from the LHS of (15). This simply captures an intertemporal spillover effect whereby the social gain from electrification carries over to the whole future, whereas private producers benefit from electrification for one period only, given our assumption that patents expire after one period. This inefficiency could potentially be corrected with a uniform subsidy to electrification at rate $\beta$.

A second difference comes from the term $1-[e^z(1+\xi)]^{-\mu_i}$ on the LHS of (15) (assuming that the social planner implements a Pigouvian tax: $\tau = \xi$) versus $\mu_i \ln((1+\xi)e^z)$ on the LHS of (18). This difference reflects the fact that the social planner maximizes total surplus including consumer surplus whereas producers maximize their private rent from electrification.[11] The social surplus is generally larger than the private surplus, so that ceteris paribus, there is too little electrification in equilibrium (for $e^z(1+\xi) > 1$, we get that $1-[e^z(1+\xi)]^{-\mu_i} < \mu_i \ln((1+\xi)e^z)$). Nevertheless, if $(1+\xi)e^z$ is close to one, the difference is negligible: $1-[e^z(1+\tau)]^{-\mu_i} \approx \mu_i \ln((1+\xi)e^z)$ so that abstracting from $\frac{1}{1-\beta}$ factor on the LHS of (18), the two equations (15) and (18) become identical.

Is it the case then, that a uniform subsidy $\beta$ to electrification coupled with a Pigouvian carbon tax is enough to decentralize the optimum in steady-state when the overall advantage of the clean technology is small (i.e. $(1+\xi)e^z$ is close to 1)? The answer is no, which reveals the fundamental reason why industrial policy is warranted in our set-up: cross-sectoral strategic complementarities in electrification. That is, insufficient electrification in sectors downstream and/or upstream to sector $i$, reduces private incentives to electrify in sector $i$ (see (**??**)), typically leading to multiple steady-state equilibria as showed above.[12] Then, starting from initially low levels of electrification in all sectors, the economy may end up being stuck in a steady-state which also features low-level of electrification, even though the optimum might involve a high level of electrification in all sectors. The following example with Pigouvian taxation on emissions illustrates this point.

**Example 1.** We consider a two sector supply chain, hence $N = 2$. We set $z = 0$ for simplicity and assume that $\xi$ is small. In addition we assume that electrification is uniformly subsidized at rate $\beta$ , so that firms only face the electrification cost $(1-\beta)\phi$. Under these conditions, both the steady-state equilibrium equations and the first order conditions for the social optimum can be

[11]To see this, suppose there is just one sector, and that the carbon tax is set equal to the social cost of pollution. The model is then equivalent to having a (dirty) technology with productivity $\frac{1}{1+\xi}$ and a clean technology with productivity $e^z$. Under Cobb-Douglas preferences, the demand curve is $p = 1/q$. Under competitive production, the consumer surplus of improving productivity of a variety from $\frac{1}{1+\xi}$ to $e^z$ is $\int_{\frac{1}{1+\xi}}^{e^z} \frac{1}{q}\,\mathrm{d}q = \ln(1+\xi)e^z$, which differs from the private rent from such productivity improvement.

[12]This also implies that the first order conditions (18) are generally not sufficient to identify the global optimum.

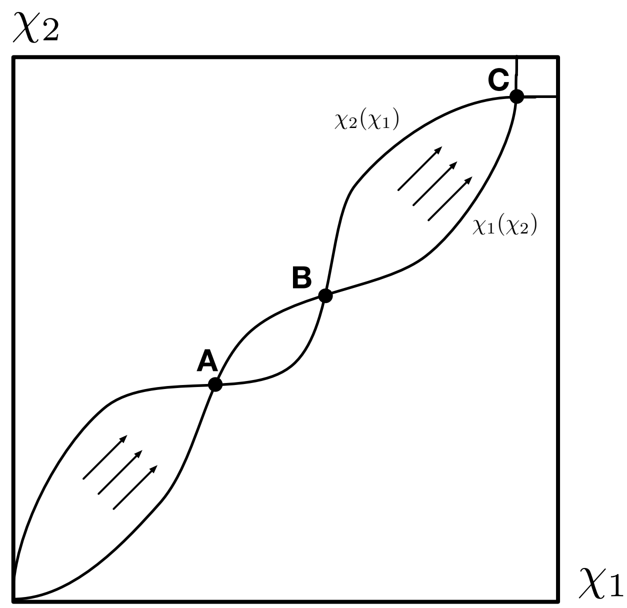


**Figure 1.** First order conditions of the social planner problem and the decentralized economy with a subsidy $\beta$

written as:

$$\chi_1 = F_1\left(\frac{\xi\chi_2(1-\alpha_2)}{1-\beta}\right) \text{ and} \tag{19}$$

$$\chi_2 = F_2\left(\frac{\xi(\alpha_2+\chi_1(1-\alpha_2))}{1-\beta}\right). \tag{20}$$

where $F_1$ and $F_2$ are chosen so that the two curves 1 and 2 in Figure 1 intersect three times, at A, B, and C, all of which correspond to a steady-state of the decentralized economy. As Figure 1 shows, the two steady-states A and C are both stable whereas B is unstable. For $\beta$ sufficiently large, the social optimum will correspond to point C, whereas a decentralized economy starting from initially low or no electrification will end up being stuck at the low-electrification steady-state A.

Without the uniform subsidy $\beta$, the decentralized steady-states satisfy:

$$\chi_1 = F_1(\xi\chi_2(1-\alpha_2)) \text{ and} \chi_2 = F_2(\xi(\alpha_2+\chi_1(1-\alpha_2))).$$

These two equations correspond to the two dash lines in Figure 2, which also intersect three times at the steady-state equilibria A', B' and C'. Then an economy with low or no initial electrification will end up being stuck at the—even lower electrification—steady-state A'. The uniform subsidy $\beta$ allows an economy with initially very low or no electrification to converge to A instead of A' but yet sector specific subsidies are required on top of the uniform subsidy to make the economy converge to C.

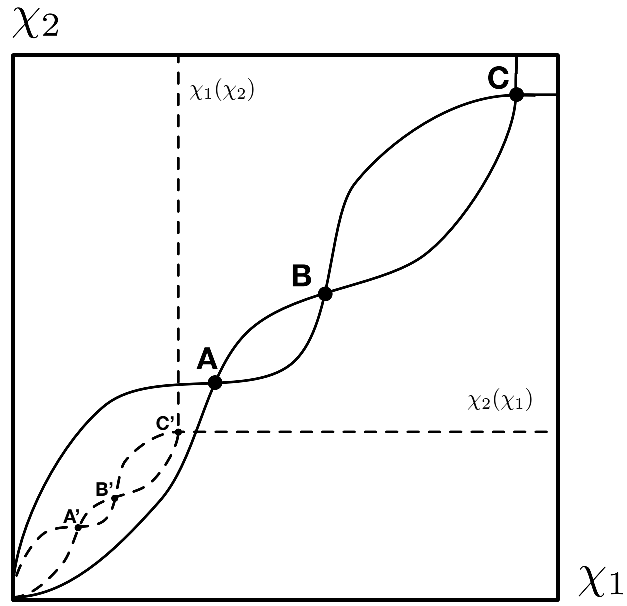


**Figure 2.** First order conditions of the social planner problem and the decentralized economy without a subsidy $\beta$

## 4.3 Implementing the social optimum

Yet, starting from the same initial conditions with low levels of electrification in all sectors, the social planner can achieve the optimum through a set of temporary subsidies to electrification—given that the equilibrium for any given policy and given initial conditions, is unique as shown in the previous section. More precisely, we can easily establish:

**Proposition 4.** *The optimal steady-state can be implemented through a carbon price together with a whole set of time-varying sector specific subsidies.*

*Proof.* Consider any initial allocation $\{\chi_{i,0}\}$ and assume that the social planner imposes a Pigouvian tax $\tau_i = \xi$, and a set of sector specific subsidies $\{q_{i,t}\}$. Then, the equilibrium level of electrification at time $t$ is given by

$$\chi_{i,t} = F_i \left( \frac{\left(1 - \left[e^z \left(1 + \xi\right)\right]^{-\mu_{i,t-1}}\right)}{1 - q_{i,t}} \prod_{j=i+1}^{N} \left(\widetilde{\chi}_{j,t}(1 - \alpha_j)\right) \right).$$

For sector $N$ at time $1$, we can always set $q_{N,1}$ such that

$$\chi_N^{SP} = F_N \left( \frac{\left(1 - \left[e^z \left(1 + \xi\right)\right]^{-\mu_{N,0}}\right)}{1 - q_{N,1}} \right),$$

where $\chi_N^{SP}$ is the social planner level of electrification in sector $N$ and $\mu_{N,0}$ is predetermined. Assume now that the social planner uses a set of sector specific subsidies $\{q_{j,1}\}$ for $j > i$, in order to implement the social planner level of electrification $\chi_j^{SP}$ for $j > i$ at $t = 1$. Then for sector $i$, the social planner can choose $q_{i,1}$ such that

$$\chi_i^{SP} = F_i\left(\frac{\left(1-\left[e^z\left(1+\xi\right)\right]^{-\mu_{i,0}}\right)}{1-q_{i,1}}\prod_{j=i+1}^{N}\left(\widetilde{\chi}_{j,1}(1-\alpha_j)\right)\right)$$

since $\mu_{i,0}$ is again pre-determined and $\widetilde{\chi}_{j,1} = \chi_{j,0} + \left(\chi_j^{SP} - \chi_{j,0}\right)\left(e^z\left(1+\xi\right)\right)^{-\mu_{j,0}}$ is also given. Therefore, by induction, the social planner can implement the socially optimal levels of electrification $\chi_i^{SP}$ in all sectors from the most downstream to the most upstream at time $t = 1$.

At time $t = 2$, there is no more incentives to electrify when $q_{i,1} = 0$, because if $\chi_{j,2} = \chi_j^{SP}$, we get that:

$$\chi_i^{SP} = F_i\left(\frac{\ln\left(\left(1+\xi\right)e^z\right)}{1-\beta}\mu_i^{SP}\prod_{j=i+1}^{N}\left(\chi_j^{SP}(1-\alpha_j)\right)\right) > F_i\left(\left(1-\left[e^z\left(1+\xi\right)\right]^{-\mu_{i,1}}\right)\prod_{j=i+1}^{N}\left(\widetilde{\chi}_{j,2}(1-\alpha_j)\right)\right),$$

where the equality stems from the fact that $\chi_i^{SP}$ is the optimum, while the inequality uses that $\mu_{i,1} = \mu_i^{SP}$, $\widetilde{\chi}_{j,2} = \chi_j^{SP}$ if there is no further electrification, and that $\frac{\ln((1+\xi)e^z)}{1-\beta}\mu_i^{SP} > \left(1-\left[e^z\left(1+\xi\right)\right]^{-\mu_i^{SP}}\right)$. The overall inequality implies that there is no further incentive to electrify in the decentralized economy. Since, there is no electrification, there are no profits either and the social optimum is implemented from $t = 2$ onward.[13] This completes the proof. □

## 4.4 Small subsidies can make a big difference

Are large subsidies—big-push policies—always necessary to move the economy away from a low-electrification steady-state? To conclude this section, we provide an example where small sector-specific subsidies are enough to move the economy from an initial low electrification steady-state towards a high electrification steady-state, which dominates it in terms of welfare (and may be the optimum).

We consider again a two-sector supply chain, but now assume mass point distributions of electrification costs across varieties in the two sectors. We build the example such that there are three steady-state equilibria, respectively with no electrification, full electrification and in-between an interior, unstable steady-state. Provided that consumers are sufficiently patient, the full electrification steady-state dominates the other two and may even correspond to the optimum. Temporary subsidies can then ensure that the economy moves away from the no electrification steady-state to a little beyond the interior one. Afterwards, the economy will move to the full electrification steady-state on its own without the use of any subsidy. We show that if the downstream cost of electrification is just high enough that the no electrification steady-state exists but not much higher, then a very small electrification subsidy is enough to make the economy eventually con-

[13] At $t = 1$, the optimal environmental tax and the optimal electrification levels are implemented, but newly electrified varieties are priced monopolistically and therefore inefficiently under-supplied.

verge to the full adoption steady-state.

**Example 2.** We assume that the distributions of fixed costs $F_1$ and $F_2$ have mass points at $\phi_1$ and $\phi_2$ respectively. As before, we denote $Z = z + \ln(1+\tau)$. We first derive conditions under which there are three steady-states characterized by no electrification, full electrification and an interior level of electrification. We then derive conditions under which moving from the no electrification to the interior level of electrification only requires a small intervention.

First, assume that the economy features no electrification, then there are no incentive to electrify downstream (sector 2) provided that:

$$1 - e^{-\alpha_2 Z} < \phi_2.$$

n that case, there is no market for sector 1 and no electrification upstream either as long as $\phi_1 > 0$. Because these are inequalities, there are still no incentives to electrify for $\chi$'s slightly different from $0$, so that the steady-state is stable.

Second, full electrification is also a steady-state provided that:

$$\left(1 - e^{Z}\right)(1-\alpha_2) > \phi_1$$

which ensures full electrification upstream, and

$$\left(1 - e^{-Z}\right) > \phi_2,$$

which ensures full electrification downstream. For the same reason as before, this steady-state is also stable.

Therefore, no-electrification and full-electrification steady-states can coexist provided that:

$$1 - e^{-\alpha_2 Z} < \phi_2 < 1 - e^{-Z} \text{ and } 0 < \phi_1 < \left(1 - e^{-Z}\right)(1-\alpha_2).$$

Third, an interior steady-state equilibrium $(\chi_1^*, \chi_2^*)$ must satisfy:

$$\left(1 - e^{-Z}\right)\chi_2^*(1-\alpha_2) = \phi_1 \tag{21}$$

$$1 - e^{-\left(\alpha_2 + \chi_1^*(1-\alpha_2)\right)Z} = \phi_2 \tag{22}$$

Given that $1 - e^{-\alpha_2 Z} < \phi_2 < 1 - e^{-Z}$, there always exists a $\chi_1^* \in (0,1)$ which satisfies the second equation. Similarly, given that $\left(1 - e^{-Z}\right)(1-\alpha_2) > \phi_1 > 0$, there also always exists a $\chi_2^* \in (0,1)$ that satisfies the first equation. Since the left hand side of (21) is increasing in $\chi_2$ and the left-hand side of (22) is increasing in $\chi_1$ while the right-hand sides are fixed at $\phi_1$ and $\phi_2$, this interior steady-state is necessarily unstable. Therefore, a small increase in $\chi_1$ and/or $\chi_2$ starting from $(\chi_1^*, \chi_2^*)$ will lead to further electrification.

Next, we derive a set of subsidies sufficient to ensure that the economy move from the no electrification to the interior steady-states at time $1$. For electrification in sector 1 to be interior, we need a subsidy $q_{1,1}$ which satisfies:

$$\frac{1-e^{-Z}}{1-q_{1,1}}\widetilde{\chi}_{2,1}\left(1-\alpha_2\right)=\phi_1 \text{ with } \widetilde{\chi}_{2,1}=\chi_2^* e^{-\alpha_2 Z},$$

which, using (21), yields $q_{1,1}=1-e^{-\alpha_2 Z}$. Similarly, for electrification in sector 2 to be interior, we need a subsidy $q_{2,1}$ such that:

$$\frac{1-e^{-\alpha_2 Z}}{1-q_{2,1}}=\phi_2 \Longrightarrow q_{2,1}=1-\frac{1-e^{-\alpha_2 Z}}{\phi_2}$$

Overall, the total amount of subsidies to move the economy to the interior unstable steady-state, is given by

$$Q_1=\chi_1^* q_{1,1}+\chi_2^* q_{2,1}=\chi_1^*\left(1-e^{-\alpha_2 Z}\right)+\chi_2^*\left(1-\frac{1-e^{-\alpha_2 Z}}{\phi_2}\right)$$

If $\phi_2$ is above but close to $1-e^{-\alpha_2 Z}$, then $q_{2,1}$ is small. In addition, in that case, (22) implies that $\chi_1^*$ is small too, which ensure that the total amount spent $Q_1$ is small. This establishes the result described above: if electrification costs are just a little too high downstream, then a small intervention is enough to push the economy toward the interior steady-state and eventually the full electrification one.[14]

# 5 Propagation and appropriate industrial policy

Our analysis so far has considered a government that could and generally would intervene in several sectors at the same time – since implementing the optimum generically require such a multi-faceted intervention. In practice, however, it may be that a government is constrained to focus on one or a few key sectors at the time. This raises a number of questions that our framework can shed light upon: What can be achieved when the government is constrained to electrify at most one sector at a time? How do electrification incentives propagate along the supply chain starting from a particular sector? In other words, what are the effects of an exogenous change in the electrification share $\chi_k$ in sector $k$ at a given time—for instance due to targeted policy—on the equilibrium electrification incentive in all other sectors in the supply chain?

We first analyze these questions in our general framework before providing more definitive answers on which sectors should be targeted by the government in an example.

[14]This logic does not extend to the case where electrification costs are just a little too high upstream: If $\phi_1$ is positive but close to 0, then $\chi_2^*$ is close to 0 and subsidies spent for sector 2 are small. However, there is no guarantee that subsidies spent for electrification in sector 1 are small without additional assumptions.

## 5.1 Basic intuition

Recall that a steady-state satisfies:

$$\chi_i = F_i\left(\pi_i\right) \text{ with } \pi_i \equiv \underbrace{\left(1 - e^{-\mu_i Z}\right)}_{\text{productivity gain}} \underbrace{\prod_{j=i+1}^{N} \left(\chi_j(1-\alpha_j\right)}_{\text{demand from downstream}},$$

$$\text{and } \mu_1 = 1, \quad \mu_i = \alpha_i + \chi_{i-1}\mu_{i-1}(1-\alpha_i) \text{ for } i \in \{2, ...N\},$$

where with Pigouvian taxes $e^Z = e^z(1+\xi)$. The above expression for $\pi_i$ suggests that, a priori, the incentives to electrify could propagate in both directions: more electrification in a sector $k$ which is downstream from $i$ ($k > i$) affects the demand component of $\pi_i$, whereas more electrification in a sector $k$ which is upstream relative to $i$ ($k < i$) affects the productivity gain component of $\pi_i$.

Yet we argue below that the same perturbation in the electrification rate $\chi_k$ in sector $k$, always generates more electrification incentives in sectors $i < k$, i.e., those that are more upstream than $k$. That is, the impact of electrification in $k$ working through the demand components, are always stronger than the impacts working through the productivity components. In fact, the latter is close to zero in a steady-state with low electrification rates $\chi$'s. In other words, starting from a steady-state with low electrification across all sectors, the bang-for-the-buck is generically higher if policy can induce electrification first in more downstream sectors.

To see this more formally, we show in Appendix 9.3 that

$$\frac{\partial \ln \pi_i}{\partial \ln \chi_k} = \begin{cases} 1 & k > i \\ \frac{\mu_i Z e^{-\mu_i Z}}{1-e^{-\mu_i Z}} \frac{\left(\prod_{j=0}^{i-k-1}(1-\alpha_{i-j})\chi_{i-j-1}\right)\mu_k}{\mu_i} < 1 & k < i \end{cases}. \tag{23}$$

This in turn leads us to make the following two observations. First, a 1% increase in electrification in a sector $k$ which is downstream to sector $i$ (i.e., with $k > i$), always induces a 1% proportional increase in the steady-state electrification incentive in sector $i$. Intuitively, electrification downstream increases the market for electrified varieties upstream 1 for 1. Second, a 1% increase in electrification in a sector $k$ which is upstream to sector $i$ (i.e. with $k < i$) always provides lower marginal incentives for electrification in sector $i$ than a 1% increase in electrification downstream from sector $i$. This is because both $\frac{\mu_i Z e^{-\mu_i Z}}{1-e^{-\mu_i Z}} < 1$ and $\frac{\left(\prod_{j=0}^{i-k-1}(1-\alpha_{i-j})\chi_{i-j-1}\right)\mu_k}{\mu_i} < 1$, where in fact the latter expression decreases as $k$ becomes smaller (that is as $k$ is further upstream from $i$). Intuitively, electrification upstream reduces the cost of electrified varieties downstream but less than proportionately, because at each stage of the production process, the costs of producing a given intermediate depends on the costs of the more upstream intermediates – but only in proportion to how much electrification has occurred – and also on labor costs.

To summarize, exogenously more electrification in downstream sectors generate proportional gains in electrification incentives in all upstream sectors. However exogenously more electrification upstream generates less than proportional gains in electrification incentives in downstream sectors, and the incentives gets smaller and smaller the further downstream we go.

This in turn suggests that starting from a steady-state with low electrification in all sectors ($\chi_i \approx 0$ for all $i$), the downstream propagation of incentives from targeting upstream sectors is close to zero, so that industrial policy should first target downstream sectors and leverage the equilibrium upstream propagation of incentives.

Note that our intuitive reasoning in this subsection is incomplete in two respects. First, it considers the change in incentives in sector $i$ while perturbing $\chi_k$, holding all other $\chi_j$'s constant. However, in equilibrium, $\chi_j$ would respond for all $j$, generating further changes in incentives. Second, it focuses on perturbations in a steady-state, while ignoring the sequence of changes in incentives along the transition. But the intuitions continue to extend when the path of $\chi_{jt}$'s respond to the exogenous shock in $\chi_k$. We shall now focus on a special case where a more formal statement can be made.

## 5.2 A Proposition

In this subsection, we now consider a special case of an economy starting in a no-electrification steady-state. Should the government only be able to intervene in one sector, then electrification will only propagate contemporaneously if it targets the most downstream sector and with a lag if it targets the sector immediately above. Formally, we prove the following proposition:

**Proposition 5.** *Consider a supply chain with $N \geq 3$ sectors, and suppose that the basic parameters and the $F_i$'s are such that no electrification in all sectors is a steady-state. Suppose also that initially the economy is stuck in this no-electrification steady-state and that the government can directly electrify a positive mass of varieties in one sector only. Then, provided that electrification costs are not too high, (i) the one-sector exogenous electrification will start to propagate immediately—creating incentives for other sectors to electrify—only if the government intervenes in the most downstream sector $N$; (ii) the one-sector electrification will start propagating with a one-period delay if the government intervenes in sector $N-1$; (iii) the one-sector electrification will not propagate at all if the government intervenes in a sector which is more upstream than sector $N-1$.*

*Proof.* Part (i):

Suppose the government starts electrifying in sector $N$ at the beginning of time $t$, so that $\chi_{N,t} > 0$. Then, at the end of time $t$, electrification in sector $N-1$ is given by

$$\chi_{N-1.t} = F_{N-1}\left(\left(1 - e^{-\alpha_{N-1}Z}\right)\chi_{N,t}e^{-\alpha_{N-1}Z}\left(1 - \alpha_{N-1}\right)\right),$$

With $\chi_{N,t} > 0$, we get $\chi_{N-1,t} > 0$ (as long as electrification costs at $N-1$ are not too high:

$F_{N-1}\left(\left(1-e^{-\alpha_{N-1}Z}\right)\chi_{N,t}e^{-\alpha_{N-1}Z}\left(1-\alpha_{N-1}\right)\right) > 0$). In contrast at $t-1$, pre-intervention, we had $\chi_{N,t-1}=\chi_{N-1,t-1}=0$.

Now suppose electrification did propagate to sectors $j>i$ (i.e. $\chi_{j,t}>0$). Then it also propagates to sector $i$ since:

$$\chi_{it}=F_i\left(\left(1-e^{-\alpha_i Z}\right)\prod_{j=i+1}^{N}\chi_{j,t}e^{-\alpha_j Z}\left(1-\alpha_j\right)\right)\text{ for } i<N,$$

which is positive if electrification costs are not too high. Hence electrification propagates all the way from the most downstream sector $N$ to the most upstream sector 1 at time $t$.

At time $t+1$, we have:

$$\chi_{i,t+1}=F_i\left(\left(1-e^{-\mu_{i,t}Z}\right)\prod_{j=i+1}^{N}\left(\widetilde{\chi}_{j,t+1}\left(1-\alpha_j\right)\right)\right)$$

with $\widetilde{\chi}_{j,t+1}$ given by (6) and $\mu_{it}$ by (4). Then, $\mu_{it}\geq\alpha_i$ and $\widetilde{\chi}_{j,t+1}\geq\chi_{j,t}>\chi_{j,t}e^{-\alpha_j Z}$, which implies that further electrification occurs in all sectors at time $t+1$ (as long as $F_i$ has positive mass around the relevant range) and this continues in subsequent periods until we reach a steady-state with positive electrification in all sectors.

Part (ii):

Suppose now that the government starts electrifying in sector $N-1$. Then we get that $\mu_{N,t-1}=\alpha_N$ (i.e. the pre-intervention value) so that $\chi_{Nt}$ must satisfy:

$$\chi_{Nt}=F_N\left(1-e^{-\alpha_N Z}\right)=\chi_{N,t-1}=0.$$

In other words, electrifying first in sector $N-1$ at time $t$ does not immediately propagate to the most downstream sector $N$.

Consider now sector $N-2$. We have:

$$\begin{aligned}\chi_{N-2.t} &= F_{N-2}\left(\left(1-e^{-\alpha_{N-2}Z}\right)\widetilde{\chi}_{N-1.t}\left(1-\alpha_{N-1}\right)\widetilde{\chi}_{Nt}\left(1-\alpha_N\right)\right)\\ &= F_{N-2}\left(0\right)=\chi_{N-2,t-1}=0\end{aligned}$$

And given that $\widetilde{\chi}_{Nt}=0$, electrification incentives in sector $N$ are the same as at time $t-1$, i.e. there is no electrification in sector $N$ at time $t$. The same logic applies to any sector $i<N-1$. In other words, there is no propagation of the exogenous electrification in sector $N-1$ to other sectors at time $t$.

Consider now time $t+1$. In sector $N$, we get: $\mu_{N,t}=\alpha_N+\chi_{N-1,t}\alpha_{N-1}\left(1-\alpha_N\right)>\alpha_N$ and electrification incentives in sector $N$ obey:

$$\chi_{N,t+1}=F_N\left(1-e^{-\mu_{N,t}Z}\right).$$

Since $\mu_{N,t}>\alpha_N$, then provided that electrification costs are not too high, we can have $F_N\left(1-e^{-\mu_{N,t}Z}\right)>F_N\left(1-e^{-\alpha_N Z}\right)=0$, such that $\chi_{N,t+1}>0$ in which case electrification propagates to sector $N$.

Moving back to sector $N-1$, we have:

$$\chi_{N-1,t+1} = \max\left\{\chi_{N-1,t}, F_{N-1}\left(\left(1-e^{-\alpha_{N-1}Z}\right)\widetilde{\chi}_{N,t+1}\left(1-\alpha_N\right)\right)\right\}$$

with $\widetilde{\chi}_{N,t+1} = \chi_{N,t+1}\left(e^z\left(1+\tau_t\right)\right)^{-\mu_{N,t}}$.

Now moving to sector $N-2$, we get that:

$$\chi_{N-2,t+1} = F_i\left(\left(1-e^{-\alpha_{N-2Z}Z}\right)\widetilde{\chi}_{N-1,t+1}\left(1-\alpha_{N-1}\right)\widetilde{\chi}_{N,t+1}\left(1-\alpha_N\right)\right),$$

with $\widetilde{\chi}_{N-1,t+1} = \chi_{N-1,t} + \left(\chi_{N-1,t+1} - \chi_{N-1,t}\right)e^{-\mu_{N-1,t}Z} \geq \chi_{N-1.t}$ and $\widetilde{\chi}_{N,t+1} > 0$ if $\chi_{N,t+1} > 0$. This in turn implies that electrification will also propagates to sector $N-2$ – provided that the distribution of fixed costs $F_i$ has positive mass in the relevant range. The logic extends to all sectors $j > N-1$ so that electrification propagates to all sectors at time $t+1$.

And electrification intensifies in subsequent periods until we reach a steady-state with positive electrification in all sectors.

Part (iii):

Now suppose that the government starts electrifying in a sector $j$ which more upstream than $N-1$,i.e. $j < N-1$, at time $t$. Consider first sector $N$ at time $t$. Given that electrification incentives in that sector only depend upon $\mu_{N,t-1}$, they are the same as pre-intervention so that $\chi_{N,t} = 0$. Next, consider any sector $k \neq N, j$ : electrification incentives in sector $k$ depend multiplicatively upon on $\widetilde{\chi}_{N,t} = 0$, thus profits from electrifying in sector $j$ are equal to zero as it was pre-intervention. We then have that $\chi_{k,t} = 0$ for all $k \neq j$.

Consider now sector $N$ at time $t+1$. Incentives to electrify in that sector hinge upon:

$$\mu_{N,t} = \alpha_N + \chi_{N-1,t}\mu_{N-1,t}\left(1-\alpha_N\right) = \alpha_N = \mu_{N,t-1}$$

since $\chi_{N-1,t} = 0$. Therefore the incentives to electrify in sector $N$ remain the same as in period $t$, i.e. $\chi_{N,t+1} = 0$. It follows that we also have $\chi_{k,t+1} = 0$ for all $k \neq j$. And the same reasoning carries over to all periods $t+h$ with $h > 1$. Hence, the initial electrification push in sector $j < N-1$ never propagates to the other sectors of the supply chain. This establishes the proposition. □

The finding that industrial policy should first target the most downstream sector, hinges heavily, first on industrial policy being one-shot and focused on one sector, and second on having a single vertical supply chain. In particular, to induce electrification in production structures involving a several number of supply chains with the same most upstream sector, a government with the ability to intervene in several sectors, might find it optimal to also intervene in the most upstream sector. This leads us to consider the case of parallel supply chains in the next section.

# 6 Horizontal misallocation

In Sections 4 and 6, we showed that industrial policy can bring large welfare gains. Yet, misdirected green industrial policies can also bring welfare losses if it "picks the wrong winner". We already saw in the previous section that in supply chains with more than two sectors, a one-time-one-sector intervention should focus on the most downstream sectors. In this section, we entertain the possibility of "horizontal" misallocations of public electrification investment. More specifically, we consider an extended version of our basic model with two layers, where on top of using labor, the downstream dirty industrial process also uses inputs from an upstream sector that can be also electrified.[15] In that context, electrification in the two upstream sectors are strategic substitutes, and over-investing in the electrification of the upstream sector associated with the dirty downstream process may derail the overall transition towards a clean economy.

## 6.1 A simple two-legs model

We consider a simple two-layer network, with a single downstream sector and two upstream sectors denoted by $1a$ and $1b$. The downstream sector uses input $1a$ when producing using the dirty technology and it uses input $1b$ when producing using the clean technology. More formally, a non-electrified downstream variety is produced according to:

$$y_2(v) = \left(\frac{l_{2t}^d(v)}{\alpha}\right)^{\alpha}\left(\frac{m_{1at}(\nu)}{1-\alpha}\right)^{1-\alpha}$$

whereas an electrified downstream variety is produced according to:

$$y_2(v) = \left(\frac{l_{2t}^d(v)}{\alpha}\right)^{\alpha}\left(\frac{m_{1at}(\nu)}{1-\alpha}\right)^{1-\alpha} + \left(\frac{e^z l_{2t}^c(v)}{\alpha}\right)^{\alpha}\left(\frac{m_{1bt}(\nu)}{1-\alpha}\right)^{1-\alpha}$$

Varieties in the upstream sectors are produced according to:

$$y_{1j}(v) = \begin{cases} l_{1j}^d(v) & \text{if not electrified} \\ l_{1j}^d(v) + e^z l_{1j}^c(v) & \text{if electrified} \end{cases} \quad \text{for } j \in \{a, b\}.$$

Electrification involves the fixed cost distributions $F_2$, $F_{1a}$ and $F_{1b}$.[16]

[15] For instance, fossil fuel engines for vehicles can be produced in a cleaner way or become more fuel efficient.

[16] A generalization would involve the clean and dirty industrial processes using both upstream goods with Cobb-Douglas shares: $\sigma_a^d$ and $\sigma_b^d$ denoting respectively the use of sector $1a$ and $1b$ in the dirty downstream process, and $\sigma_a^c$ and $\sigma_b^c$ denoting the corresponding shares for the clean process. Then the case with $\sigma_a^d = \sigma_b^d = 0$ corresponds to the original model and its results based on strategic complementarity extend directly for $\sigma_a^d < \sigma_a^c$ and $\sigma_b^d < \sigma_b^c$. If $\sigma_a^d = \sigma_a^c$ and $\sigma_b^d = \sigma_b^c$, then there will be no strategic complementarity. If $\sigma_a^d > \sigma_a^c$ and $\sigma_b^d < \sigma_b^c$ (as here), there is strategic substitutability.

## 6.2 Equilibrium

We solve for the equilibrium following similar steps to Section 3.

**Prices.** As before, in the upstream sectors, the price of a variety electrified in the previous period is $p_{1jt}(\nu) = e^{-z}$ for $j \in \{a,b\}$,whereas the price of a newly electrified variety or of a non-electrified variety is $p_{1jt}(\nu) = (1+\tau_t)$. We then get that in the downstream sector 2, the price of a variety is given by:

$$p_{2t}(\nu) = \begin{cases} \min\left[(1+\tau)^{\alpha} p_{1,a,t}^{1-\alpha}, e^{-\alpha z} p_{1,b,t}^{1-\alpha}\right] & \text{if sector 2 is electrified by time } t-1, \\ (1+\tau)^{\alpha} p_{1,a,t}^{1-\alpha} & \text{otherwise.} \end{cases} \tag{24}$$

We still assume that $e^{-z} < 1+\tau$, so that the upstream sectors use the clean production process when electrified. Therefore, the price of the upstream input $1j$ at date $t$ follows the same formula as in the baseline model, namely: $p_{1jt} = (1+\tau)\, e^{-Z\chi_{1j,t-1}}$ with $Z = z + \ln(1+\tau) > 0$.

The assumption that $e^{-z} < 1+\tau$ is no longer sufficient to ensure that the downstream sector always uses the clean production process since the dirty input may be cheaper than the clean one if it is itself sufficiently electrified. Instead, the downstream sector will use the clean production process if and only if

$$e^{-\alpha z} p_{1,b,t}^{1-\alpha} < (1+\tau)^{\alpha} p_{1,a,t}^{1-\alpha}$$

or equivalently if and only if

$$\mu_{2,t-1} \equiv \alpha - \chi_{1a,t-1}(1-\alpha) + \chi_{1b,t-1}(1-\alpha) > 0, \tag{25}$$

where we adjusted the definition of $\mu_{2,t}$ relative to the baseline model. For expositional purposes we henceforth assume that $\alpha > 1/2$ so that the above condition is always satisfied.

**Equilibrium revenues and profits.** As in our baseline model, the profit margin is equal to $1 - e^{-Z}$ in the upstream sectors $1a$ and $1b$. In the downstream sector 2, the profit margin is given by:

$$1 - \frac{e^{-\alpha z} p_{1,b,t}^{1-\alpha}}{(1+\tau_t)^{\alpha} p_{1,a,t}^{1-\alpha}} = 1 - e^{-Z\mu_{2,t-1}},$$

which is positive since by assumption, we always have $\mu_{2,t} > 0$. This is again the same expression as in the baseline model.

We now derive the revenue accruing to each variety. In the downstream sector 2, we have, as before that revenues obey $r_{2t} = p_t c_t = 1$.

Consider now the upstream sector $1b$. Note that $r_{1bt}$ follows exactly the same logic as the upstream sector in the baseline model, namely:

$$r_{1bt} = \underbrace{\chi_{2,t-1} r_{2,t} (1-\alpha)}_{\text{sales to previously electrified varieties}} + \underbrace{(\chi_{2,t} - \chi_{2,t-1}) r_{2,t} e^{-Z\mu_{2,t-1}} (1-\alpha)}_{\text{sales to newly electrified varieties}}$$

which we again rewrite as $r_{1bt} = \widetilde{\chi}_{2,t} (1-\alpha)$ with $\widetilde{\chi}_{2,t} \equiv \chi_{2,t-1} + (\chi_{2,t} - \chi_{2,t-1}) e^{-Z\mu_{2,t-1}}$ as before.

Consider now the upstream sector $1a$ : all the revenues accruing to electrified varieties in that sector come from still non-electrified varieties in sector 2, so that:

$$r_{1at} = (1-\chi_{2,t}) r_{2,t} (1-\alpha) = (1-\chi_{2,t})(1-\alpha).$$

The electrification rents in the three sectors $2, 1a, 1b$ are then given by the profit margins times the revenues, namely, they follow (8) for $i \in \{2, 1a, 1b\}$ ,with $\mu_{1a,t} = \mu_{1b,t} = 1$ as in the baseline model.

**Electrification levels.** As in the baseline model, the producer of a variety in sector $i$ will electrify at time $t$ if and only if the rents from electrification $\pi_{it}$ are bigger than her fixed cost of electrification. We then obtain the equilibrium conditions for the electrification levels as:[17]

$$\chi_{1at} = F_{1a}\left(\left(1-e^{-Z}\right)(1-\chi_{2,t})(1-\alpha)\right), \tag{26}$$

$$\chi_{1bt} = F_{1b}\left(\left(1-e^{-Z}\right)\widetilde{\chi}_{2,t}(1-\alpha)\right), \tag{27}$$

$$\chi_{2t} = F_{2b}\left(1-e^{-Z\mu_{2,t-1}}\right). \tag{28}$$

This system is recursive: given that $\mu_{2,t-1}$ is predetermined, the third equation determines $\chi_{2t}$, from which we uniquely determine $\chi_{1at}$ and $\chi_{1bt}$ using the first two equations. Thus, as in the baseline model, the equilibrium is unique given initial electrification shares in the three sectors.

**Steady-states.** In a steady-state the electrification levels are constant over time, so that steady-states are now defined by:[18]

$$\chi_{1a} = F_{1a}\left(\left(1-e^{-Z}\right)(1-\chi_2)(1-\alpha)\right)$$

$$\chi_{1b} = F_{1b}\left(\left(1-e^{-Z}\right)\chi_2(1-\alpha)\right)$$

$$\chi_2 = F_{2b}\left(1-e^{-Z\mu_2}\right)$$

[17]Technically, electrification levels can never decrease so they are given by the minimum of the right-hand side and the levels in the previous period exactly as in the baseline model.

[18]Again we focus on strict steady-states where the three conditions below hold as equalities; the actual set of steady-states is larger since whenever the left-hand side is greater than the right-hand side we also have a steady-state as electrification levels cannot decrease.

with $\mu_2 = \alpha - (1-\alpha)\chi_{1a} + \chi_{1b}(1-\alpha)$ . The first equation establishes that $\chi_{1a}$ is a decreasing function of $\chi_2$. The second equation establishes that $\chi_{1b}$ is increasing $\chi_2$ and the third equation establishes that $\chi_2$ is increasing in $\chi_{1b}$, but decreasing in $\chi_{1a}$. Therefore electrifications in $1b$ and $2$ are strategic complement but electrifications in $1a$ versus $1b$ and $2$ are strategic substitutes. In other words, electrifying sector $1a$ diverts efforts away from electrifying sectors $2$ and $1b$, since it reduces the cost of the dirty industrial process in the downstream sector compared to the clean industrial process.

The strategic substitutability between $1a$ and $2$ plus $1b$ will have two important implications: First, an industrial policy that favors $1a$ may backfire and halt the electrification that would have otherwise happened in $2$ and $1b$ without government intervention — thereby reducing long-term welfare compared to no intervention. Second, the laissez-faire equilibrium may involve too much electrification of $1a$ compared to the ex-ante optimum. But, given the history dependence in the dynamics of electrification shares over time, delaying electrification in $2$ plus $1b$ in turn leads to irreversible long-term consequences. We illustrate these two possibilities in the remaining part of this section.

## 6.3 Backfiring industrial policy

Here we build an example where an industrial policy initially focused on the upstream sector $1a$ backfires: namely, in this example, given initial conditions and the electrification cost functions, the laissez-faire is associated with full electrification in sectors $2$ and $1b$, whereas a (misguided) industrial policy focusing on sector $1a$ in the initial period, reduces long-run welfare by preventing full electrification in sectors $2$ and $1b$. Key in our example is the fact that electrification in sector $1a$ reduces the cost-advantage of electrified varieties in sector $2$, which reduces the incentives to electrify both sector 2 and its upstream sector, 1b.

**Example 3.** We assume that the economy initially features no-electrification in all three sectors ($\chi_{i,0} = 0$ for $i \in \{2, 1a, 1b\}$). We consider two potential policies: In the first one, the government does not subsidize electrification. In the second one, the government subsidizes electrification in sector $1a$ at time $t = 1$,such that $\chi_{1a,1} = 1$. In both cases, the government implements a Pigouvian carbon tax: $\tau = \xi$.

Consider first the case with no government intervention. Then electrification at time $1$ in sector 2 follows from (28): with $\mu_{2,0} = \alpha$, $\chi_{2,1}$ is given by $\chi_{2,1} = F_2\left(1 - e^{-Z\alpha}\right)$. In addition, following (6), $\widetilde{\chi}_{2,1} = \chi_{2,1} e^{-Z\alpha}$,so that, using (27), $\chi_{1b,1}$ satisfies:

$$\chi_{1b,1} = F_{1b}\left(\left(1 - e^{-Z}\right)\chi_{2,1} e^{-Z\alpha}(1-\alpha)\right).$$

We assume that this is positive, i.e. the smallest fixed cost of electrification in sector $1b$ lies below

$(1-e^{-Z})\chi_{2,1}e^{-Z\alpha}\,(1-\alpha)$. For sector $1a$, we then get:

$$\chi_{1a,1} = F_{1a}\left(\left(1-e^{-Z}\right)(1-\chi_{2,1})\,(1-\alpha)\right),$$

which we also assume to be equal to zero. This in turn will be the case if the smallest fixed cost of electrification in sector $1a$ is greater than $\left(1-e^{-Z}\right)(1-\alpha)$.

We now consider period $t=2$. Since $\chi_{2,t}$ is non-decreasing, then incentives to electrify sector 1a are weakly smaller, so that we still have $\chi_{1a,2}=0$. Using (25), we get $\mu_{2,1}=\alpha+\chi_{1b,1}\,(1-\alpha)$. Therefore, following (28), we get that electrification in sector 2 is given by:

$$\chi_{2,2} = F_{2b}\left(1-e^{-Z\mu_{2,1}}\right),$$

which we take to be equal to 1. That is all electrification costs in sector 2 are below $1-e^{-Z\mu_{2,1}}$. We now get $\widetilde{\chi}_{2,2}=\chi_{2,1}+(1-\chi_{2,1})\,e^{-Z\mu_{2,1}}$, so that using (27), $\chi_{1b,2}$ satisfies:

$$\chi_{1b,2} = F_{1b}\left(\left(1-e^{-Z}\right)\left(\chi_{2,1}+(1-\chi_{2,1})\,e^{-Z\mu_{2,1}}\right)(1-\alpha)\right).$$

Again, we take this to be equal to $1$, namely all electrification costs in sector $1b$ are below $\left(1-e^{-Z}\right)\left(\chi_{2,1}+(1-\chi\right.$ With full electrification in sectors 1b and 2, the economy has reached a steady-state with $\chi^*_{1a}=0$, $\chi^*_{1b}=\chi^*_2=1$. Not electrifying sector $1a$ in this context comes at no cost, since the input from that sector is used only for a dirty sector 2 which disappears from $t=3$ onwards. The corresponding utility flow is:[19]

$$\ln y_2 - 1 = z - 1.$$

Now, suppose instead that starting from no-electrification in all sectors, the government decides to fully electrify the upstream sector $1a$ at time $t=1$, i.e. sets $\chi^{\dagger}_{1a,1}=1$ (we add $^{\dagger}$ to denote variables under this alternative scenario). Since electrifications incentives only move downstream with a lag, this does not change $\chi_{2,1}$ and $\chi_{1b,1}$, which remain the same as without the policy: $\chi^{\dagger}_{2,1}=\chi_{2,1}$ and $\chi^{\dagger}_{1b,1}=\chi_{1b,1}=0$.

Consider now time $t=2$. Using (25), we get that: $\mu^{\dagger}_{2,1}=\alpha-(1-\chi_{1b,t-1})\,(1-\alpha)<\mu_{2,0}$. As a result, there is no further electrification in sector 2, namely: $\chi^{\dagger}_{2,2}=\widetilde{\chi}^{\dagger}_{2,2}=\chi^{\dagger}_{2,1}=\chi_{2,1}$. In sector $1b$, using (27), we get:

$$\chi^{\dagger}_{1b,2} = F_{1b}\left(\left(1-e^{-Z}\right)\chi_{2,1}\,(1-\alpha)\right).$$

We assume that this is still equal to $\chi_{1b,1}$: That is the distribution of electrification costs in sector $1b$ is such that, a positive mass of varieties have fixed costs below $\left(1-e^{-Z}\right)\chi_{2,1}e^{-Z\alpha}\,(1-\alpha)$, no varieties has fixed costs in the interval $\left(1-e^{-Z}\right)(1-\alpha)\times\left(\chi_{2,1}e^{-Z\alpha},\chi_{2,1}\right)$, and all remaining vari-

[19] In Appendix 9.4, we show that in steady-state, the utility flow is given by

$$\ln y_2 - \sum_{i\in\{1a,1b,2\}}\left((1+\xi)\,\ell_{di}+\ell_{ci}\right) = \left[\chi_2\alpha+\chi_2\,(1-\alpha)\,\chi_{1b}+(1-\chi_2)\,(1-\alpha)\,\chi_{1a}\right]Z-\ln(1+\xi)-1.$$

eties have their fixed costs in the interval $\left(1-e^{-Z}\right)(1-\alpha)\times\left(\chi_{2,1},\left(\chi_{2,1}+\left(1-\chi_{2,1}\right)e^{-Z\mu_{2,1}}\right)\right)$.[20] At this point, the economy has reached a steady-state and no further electrification occurs. This gives a flow utility

$$\ln y_2 - 1 = \left[\chi_{2,1}\alpha + \chi_{2,1}(1-\alpha)\chi_{1b,1} + (1-\chi_{2,1})(1-\alpha)\right]Z - \ln(1+\xi) - 1.$$

As long as $\chi_{2,1}$ is sufficiently small – which is certainly feasible –, this is strictly less than the flow utility $z-1$ under laissez-faire. This establishes the result described at the beginning of the subsection.

Note however that industrial policy cannot backfire when the decentralized economy is initially already stuck in a steady-state. Since the steady-state utility flow is weakly increasing in all the $\chi' s$, the only cost associated with industrial policy in that case are the costs of the subsidies themselves (i.e. the costs associated with the corresponding increase in the $\mathcal{F}_i(\chi_i)$'s).

## 6.4 Excessive electrification and path dependence

We now build an example where, given initial conditions and the electrification cost functions, the laissez-faire is associated with positive electrification in sector $1a$ and no electrification in sectors $2$ and $1b$, whereas thee optimal policy would involve less electrification in sector $1a$ but electrification in sectors $2$ and $1b$. In other words: the laissez-faire economy exhibits excessive electrification in sector $1a$.

**Example 4.** We still assume that a Pigouvian carbon tax is in place. We choose the cost functions so that there exists a steady-state equilibrium $(\chi_{1a}^*, \chi_{1b}^*, \chi_2^*)$ with $\chi_{1a}^* > 0$ and $\chi_{1b}^* = \chi_2^* = 0$. That is:

$$\chi_{1a}^* = F_{1a}\left(\left(1-e^{-Z}\right)(1-\alpha)\right)$$
$$\chi_2^* = 0 = F_2\left(1-e^{-Z\left(\alpha-(1-\alpha)\chi_{1a}^*\right)}\right)$$

(the latter condition requires that the lowest fixed cost in sector 2 lie above $1-e^{-Z\left(\alpha-(1-\alpha)\chi_{1a}^*\right)}$.

Suppose now that initially at time $t=0$ the economy starts with electrification shares $\chi_{1a,0} = \chi_{1a}^* - \varepsilon$ (with $\varepsilon > 0$ but small) and $\chi_{1b,0} = \chi_{2,0} = 0$. Then, we get that $\chi_{1a,1}$ and $\chi_{2,1}$ solve:

$$\chi_{1a,1} = F_{1a}\left(\left(1-e^{-Z}\right)(1-\alpha)\right) = \chi_{1a}^* \text{ and } \chi_{2,1} = F_2\left(1-e^{-Z(\alpha-(1-\alpha)\chi_{1a,0})}\right) = 0,$$

if $\varepsilon$ is sufficiently small and the smallest fixed electrification cost in sector 2 lies significantly above $1-e^{-Z\left(\alpha-(1-\alpha)\chi_{1a}^*\right)}$ relative to $\varepsilon$. Then the laissez-faire economy will reach the steady-state $(\chi_{1a}^*, \chi_{1b}^* = 0, \chi_2^* = 0)$.

[20] Of course, this is a very specific case, but our goal here is simply to build an example and that one certainly exists.

Let us compare this laissez-faire equilibrium with the social optimum. We note first that the social planner always wants to implement the socially optimal steady-state immediately (the Pigouvian tax $\tau = \xi$). Then, provided that electrification costs are bounded above, a sufficiently patient social planner will seek to maximize steady-state utility flow, which is maximized for $\chi_{1b} = 1$ and $\chi_2 = 1$ and is independent of $\chi_{1a}$ – see footnote 19. Therefore, the social planner will immediately set $(\chi_{1a,0}, 1, 1)$. The corresponding $\chi_{1a}$ is lower than under laissez-faire, which in turn establishes that there is excessive electrification of $1a$ in the laissez-faire equilibrium compared to the social optimum.

# 7 Quantitative Application

Our quantitative application focuses on decarbonization of global iron and steel production and the role of hydrogen supply chain in this process. The iron and steel industry is one of the largest contributors to global greenhouse gas emissions, accounting for an estimated 7-9% of total anthropogenic CO2 emissions (Kim et al. 2022). This sector's emissions intensity is high both because of the high temperatures required to convert iron ore into iron and steel and because carbon is typically used as a reductant in the chemical process converting iron oxide into iron. Conventional steel production using this integrated blast furnace-basic oxygen furnace ("BF-BOF") technology currently accounts for around 70% of the global total (Benavides et al. 2024; WSA 2021). While several alternative technologies can reduce the emissions intensity of steel—such as scrap-based electric arc furnace production—our calibration focuses on hydrogen-based direct reduction and electric arc furnace ("H2-DR-EAF") technology as clean alternative to the BF-BOF process. This focus is motivated by the facts that H2-DR-EAF technology (i) can produce high quality steel (which scrap-based production cannot necessarily match, Jamarollo et al. 2023), (ii) can potentially reduce CO2 emissions to near zero, and (iii) has been projected to produce steel at competitive or even lowest levelized costs globally after initial innovation investments (BloombergNEF 2021). Of course, the economic viability and emissions intensity of H2-DR-EAF steel production depends on upstream hydrogen production. Here, we consider both traditional fossil fuels-based (mainly steam-methane reforming or "SMR") hydrogen production possibilities and renewable energy electrolysis-based production as clean alternative. Similar to clean steel production, clean hydrogen production is technologically understood but faces initial high cost hurdles and is projected to become cost competitive after initial investments have been incurred (e.g., BloombergNEF 2023). Our benchmark calibration thus features 2 sectors: hydrogen as upstream sector 1 and steel as downstream sector 2. The quantitative model generalizes slightly the benchmark theoretical framework to allow for heterogeneity in the relative input efficiency parameter $z_i$ and the emission rate $\xi_i$ across sectors and in adding total factor productivity (TFP) parameters $A_i$ to each sector $i$'s production function (see Appendix 9.5 for details).

## 7.1 Calibration

By assumption, the hydrogen sector uses only clean or dirty labor as inputs, and hence $\alpha_1 = 1$. Intuitively, this assumption means that we take upstream inputs from the electricity sector and its decarbonization incentives as exogenous to the hydrogen and steel industries – as the production of clean hydrogen would only represent a small share of the total use of electricity. We calibrate the relative input efficiency parameter $z_1$ based on estimates of the ratio of production costs for clean and dirty hydrogen from BloombergNEF (2023), which provides such estimates for 28 countries (accounting for over 80% of world GDP in 2022) from 2023-2050. While the costs of fossil fuels-based hydrogen production are projected to remain relatively stable over time, the costs of electrolysis-based production are estimated to fall significantly due to learning-by-doing effects and improvements in electrolyzer technologies. Conceptually, we consider the excess cost of initial vs. "n-th of a kind" clean hydrogen production as initial innovation costs. In order to quantify relative efficiencies of clean and dirty hydrogen production after incurment of these one-time costs, we thus compare the average levelized cost of hydrogen (LCOH2) for fossil fuel-based (SMR, \$2.14/kgH2 in \$2022) vs. future clean hydrogen (\$1.00/kgH2),[21] yielding $z_1 = 0.7608$. We then quantify the distribution of one-time innovation costs $\phi_1(\chi_1)$ based on the distribution of excess initial vs. n-th of a kind clean LCOH2's across countries, with several adjustments to map cost estimates into the model structure and units (such as removing innovation costs related to renewable electricity; see Appendix 9.5 for further details). In dollar terms, these costs range from \$0.78/kgH2 in China, where Chinese alkaline electrolyzers are already significantly cheaper than Western models, to more than \$4/kgH2 in higher production cost countries. Finally, the TFP parameter $A_1$ is set so that the model units of hydrogen production costs match those of the data used to estimate levelized costs of steel production.

For the steel sector, we first quantify the hydrogen cost share in clean steel production, $1 - \alpha_2$. Using BloombergNEF (2021) estimates of the (n-th of a kind) levelized cost of clean steel (LCOS) across China, the U.S., and Germany, we infer an average hydrogen cost share of around 14%, implying $\alpha_2 = 0.86$. We note that using future clean hydrogen production cost estimates from Devlin et al. (2023) across 17 countries instead yields an almost identical hydrogen cost share estimate of 13%.[22] Next, we again infer the relative production efficiency parameter $z_2$ based on the relative (post-innovation) costs of clean (\$489/tS) and dirty (\$544/tS) steel production,

[21]Our calculation assumes that western alkaline electrolyzers are used everywhere except China where Chinese alkaline electrolyzers are used, in line with BloombergNEF assumptions that the latter technology - though cheaper - is not yet widely available outside of China.

[22]Here we infer the hydrogen cost share based on the share of the LCOS due to electrolyzer plus the relevant (cheapest) costs of wind turbines or solar panels, respectively, for 2050.

yielding $z_2 = 0.1239$.[23] For the distribution of fixed innovation costs $\phi_2(\chi_2)$ , we also use estimates of the distribution of first vs. n-th of a kind clean LCOS's across countries, again with several adjustments (described in the Appendix). Our preferred calibration uses BloombergNEF (2021) estimates across three major steel producers (China, the United States, and Germany) as these map most precisely into the model, but we also consider estimates from Devlin et al. (2023) for robustness. Finally, the TFP parameter is set to match the model normalization that steel sector revenues equal 1 given base year steel industry revenues of $1.034 trillion ($2021).[24]

One important difference between the two-sector model and reality is that, in the model, clean steel production is the only potential downstream revenue source for the hydrogen sector. In practice, however, there are already other sectors using hydrogen at scale (e.g., ammonia production), and in future decarbonization scenarios, additional applications may add further to demand (e.g., aviation). The two-sector model thus likely overestimates fixed H2 innovation costs relative to revenue benefits. Our preferred specification thus scales the H2 innovation costs based on the projected hydrogen demand share of iron and steel production by 2050 across a range of decarbonization scenarios and modeling groups, which we take to be 23%.[25] Of course we also explore robustness to omission of this scaling.

Finally, we set the emissions intensity of dirty steel production to 2.2tCO2/ton steel (BloombergNEF 2021), and the emissions intensity of dirty hydrogen production to its 2021 global average of 12.5kgCO2/kgH2 (IEA 2023).

## 7.2 Results

This section presents quantitative results for potential steady-states of the 2-sector hydrogen/steel economy in decentralized equilibrium across different carbon price scenarios. All scenarios assume uniform carbon pricing across sectors (implying heterogeneous values of $\tau_1$and $\tau_2$).[26] The main results are as follows.

First, neither current estimated levels of global average effective carbon prices ($6/tCO2 in $2022) nor projected 2050 carbon prices in the "business-as-usual" scenario from the DICE-2023

[23]In the relevant expression, $\frac{mc_{c2}}{mc_{d2}} = \frac{e^{-\alpha_2 z_2} p_1^{1-\alpha_2}/A_2}{1/A_2}$, we use $p_1 = \$1/kgH2$ as input price as this matches both the underlying assumptions in the BloombergNEF LCOS estimation and the model calibration of the hydrogen sector.

[24]Calculated as the product of 2022 global steel production (1.9 billion tons, WSA 2023) and the average price assumed to equal the levelized cost of dirty steel production ($544/tS in $2021, BloombergNEF 2021) since $\chi_{2,0} \sim 0$ and we assume $\tau_{2,0} = 0$.

[25]We first calculate the average hydrogen demand share of industry across 64 decarbonization scenarios (from the European Commission, BP, Deloitte, etc.) collected by the European Hydrogen Observatory, which is 58%. We then scale this figure by the estimated steel sector industrial hydrogen demand share of 40% in 2050 from BP (2023, same share for both the accelerated and net-zero scenarios) to arrive at an overall 23% demand share.

[26]All calculations also assume that subsidies correcting investor myopia are in place. That is, fixed electrification costs and revenue benefits are compared over the same time horizon.

model ($12.5/tCO2, both Barrage and Nordhaus 2024) are sufficient to induce decarbonization of hydrogen or steel production. That is, the only stable steady-state associated with such low carbon prices is $(\chi_1 = 0, \chi_2 = 0)$, in line with the observed state of the world today where H2-DR-EAF steelmaking has been limited to demonstration projects and less than 0.7% of global hydrogen production is clean (i.e., the low-carbon share including fossil fuels-based production with carbon capture is 0.7%, IEA 2019). Second, while doubling the baseline 2050 carbon price (to $25/tCO2) can potentially induce substantial decarbonization, the multiplicity of steady states demonstrated in the theoretical results is quantitatively apparent and important here: there are three stable steady-states with decarbonization rates of (0%, 0%), (59%, 54%), and (82%, 84%). Figure 3 showcases these results. The stakes of being in the "wrong" steady-state are large: even at current quantities of annual global BF-BOF steel production, the difference in annual emissions between the (0%, 0%) and (82%, 84%) scenarios would be upwards of 2.4 billion tons of CO2 per year, close to the entirety of the European Union's CO2 emissions from all sectors (2.8 billion tons in 2022, EDGAR 2022). The difference between the (59%, 54%) and (82%, 84%) scenarios is around 1 billion tons of CO2 per year.

Third, with sufficiently high carbon prices, the low decarbonization steady states can likely be avoided. For example, with a carbon price of $100/tCO2, in the benchmark calibration the only steady-state equilibrium is the one with full decarbonization (100%, 100%). While such a carbon price would be broadly in line with recent estimates of the social cost of carbon (e.g., $78/tCO2 for 2025 in DICE-2023, $185/tCO2 in Rennert et al. 2022, etc.), it remains far from the global policy reality. In China, for example - which currently produces 54% of the world's steel and 30% of the world's hydrogen - the OECD estimates an average net effective carbon price for industry of only EUR 1.07/tCO2 (OECD 2023).

Finally, we consider the robustness of these results. On the one hand, using Devlin et al. (2023) estimates instead of BloombergNEF data to quantify clean steel innovation costs yields broadly similar results. While slightly higher carbon prices are required to induce any decarbonization, we again see consequential multiplicity of steady states. For example, a $40/tCO2 tax is consistent with both (0%, 0%) and (82%, 81%) as stable steady states. On the other hand, however, without the downscaling of H2 innovation costs (to account for additional downstream hydrogen markets besides steel) even very high carbon prices (e.g., $500/tCO2) fall short of inducing full decarbonization (resulting in a stable steady-state at 30%, 100%). Though not (yet) formalized in the model, this finding arguably illustrates again the broader point of the importance of accounting for supply chains linkages in modeling decarbonization incentives and policy impacts.

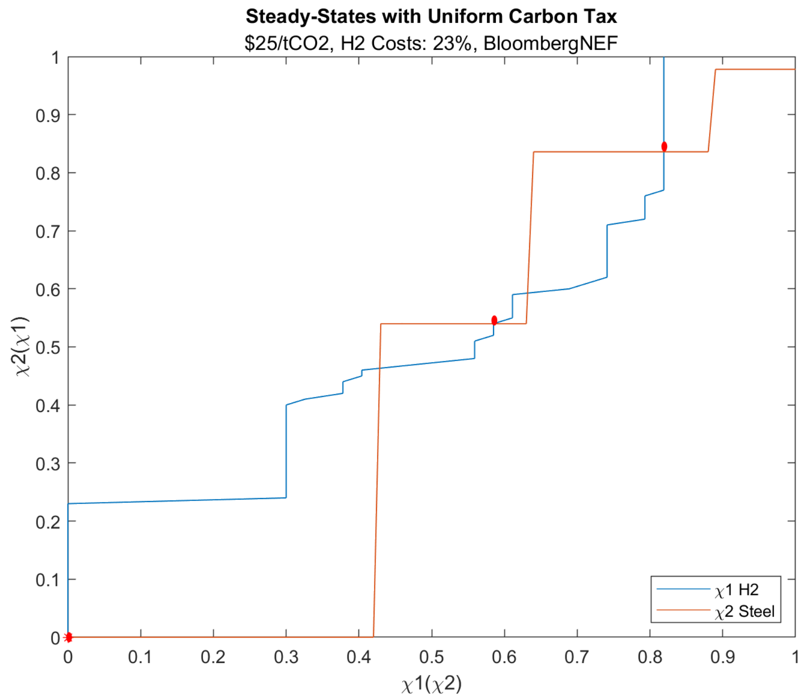


**Figure 3.** Steady-States with Uniform \$25/tCO2 Carbon Tax

# 8 Conclusion

In this paper we analyzed a model of green technological transition along a supply chain, in which each layer produces a good which is an aggregate of varieties, each of which can be produced using either a dirty technology which uses only labor, or a clean, "electrified" technology which uses labor and the good produced by the next upstream layer in the chain. We assumed heterogeneous fixed costs of electrification across varieties within any layer, and that producing a variety using the clean technology once the variety has been electrified, is cheaper than producing the variety using the dirty technology.

Under these assumptions, we first showed that the resulting cross-sectoral strategic complementarities in electrification across layers, generally lead to a multiplicity of steady states, and that the social optimum generally differs from the decentralized solution even when allowing for a Pigouvian tax. Second, we showed that our model generates the possibility that a small and temporary subsidy to electrification that targets key-sectors can be sufficient to achieve large welfare gains by moving the decentralized equilibrium just a little out of an inefficient steady-state. Third, we argued that a government which is constrained to focus its subsidies to electrification on one particular sector, should primarily target downstream sectors. Fourth, in an extension of our model where the dirty technology also uses inputs from another upstream sector that can also be electrified, we showed overinvesting in electrification in the wrong upstream branch may derail the overall transition towards electrification downstream. Finally, we applied a two-layer version of our basic model to iron and steel production, and showed that, despite using a uniform carbon price, the economy could get stuck in a "wrong" steady-state with CO2 emissions way above the social optimum.

Our model and analysis in this paper could be extended in several interesting directions. A first extension would be to develop a quantitative macroeconomic model of green transition of the overall economy seen as a grand network comprising multiple parallel supply chains. Another extension would be to look at coordination and multiple steady states not only on across sectors and layers within a country but also along international value chains. A third extension would be to use our framework to compute the overall elasticities of substitution between clean and dirty inputs to produce final goods once the supply chains involved in these technologies are fully taken into account. We know from previous work (e.g. see Acemoglu et. al., 2012, 2023) that these elasticities play a major role in the design of optimal policies, yet rigorous methodologies to compute these elasticities remain to be found. These and other extensions are left for future research.

# References


**Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn, and David HÃ©mous**, "The Environment and Directed Technical Change," *The American Economic Review*, 2012, *102 (1)*, 131–166.

_ , _ , **Lint Barrage, and David Hémous**, "Climate Change, Directed Innovation and Energy Transition: The Long-run Consequences of the Shale Gas Revolution," 2023. NBER working paper 31657.

_ , **Ufuk Akcigit, Douglas Hanley, and William Kerr**, "The Transition to Clean Technology," *Journal of Political Economy*, 2016, *124* (1), 52–104.

**Agency, International Energy**, "The Future of Hydrogen," Technical Report 2019.

_ , "Opportunities for Hydrogen Production with CCUS in China," Technical Report 2022.

_ , "Towards hydrogen definitions based on their emissions intensities," Technical Report 2023.

**Association, World Steel**, "Sustainability Indicators 2023 Report," Technical Report 2021.

_ , "World Steel in Figures 2023," Technical Report 2023.

**Barrage, Lint and William D Nordhaus**, "Policies, projections, and the social cost of carbon: results from the DICE-2023 model," *Proceedings of the National Academy of Sciences*, 2024, *Forthcoming.*

**Benavides, Kali, Angelo Gurgel, Jennifer Morris, Bryan Mignone, Bryan Chapman, Haroon Kheshgi, Howard Herzog, and Sergey Paltsev**, "Mitigating emissions in the global steel industry: Representing CCS and hydrogen technologies in integrated assessment modeling," *International Journal of Greenhouse Gas Control*, 2024, *131*, 103963.

**Bhashyam, Adithya**, "2023 Hydrogen Levelized Cost Update," *BloombergNEF*, 2023.

**BloombergNEF**, "Decarbonizing Steel: Technologies and Costs," *BloombergNEF*, 2021.

**BP**, "Energy Outlook 2023 Edition," Technical Report 2023.

**Buera, Francisco and Nicholas Trachter**, "Sectoral Development Multipliers," Technical Report 2024.

**Crouzet, Nicolas, Apoorv Gupta, and Filippo Mezzanotti**, "Shocks and Technology Adoption: Evidence from Electronic Payment Systems," *Journal of Political Economy*, 2023, *131* (11), 3003–3065.

**Devlin, Alexandra, Jannik Kossen, Haulwen Goldie-Jones, and Aidong Yang**, "Global green hydrogen-based steel opportunities surrounding high quality renewable energy and iron ore deposits," *Nature Communications*, 2023, *14* (1), 2578.

**Devulder, Antoine and Noemie Lisack**, "Carbon Tax in a Production Network: Propagation and Sectoral Incidence," Working papers 760, Banque de France 2020.

**Donald, Eric**, "Spillovers and the Direction of Innovation: An Application to the Clean Energy Transition," Technical Report 2023.

**Dugoua, Eugenie and Marion Dumas**, "Green product innovation in industrial networks: A theoretical model," *Journal of Environmental Economics and Management*, 2021, *107*, 102420.

**Fischer, Carolyn and Richard G Newell**, "Environmental and technology policies for climate mitigation," *Journal of environmental economics and management*, 2008, *55* (2), 142–162.

**for Global Atmospheric Research, EDGAR Emissions Database**, "CO2 Emissions of All World Countries," Technical Report 2022.

**Golosov, Mikhail, John Hassler, Per Krusell, and Aleh Tsyvinski**, "Optimal Taxes on Fossil Fuel in General Equilibrium," *Econometrica*, 2014, *82*, 41–88.

**Greaker, Mads and Kristoffer Midttomme**, "Network effects and environmental externalities: Do clean technologies suffer from excess inertia?," *Journal of Public Economics*, 2016, *143*, 27–38.

**Greenwald, Bruce and Joseph E. Stiglitz**, "Helping Infant Economies Grow: Foundations of Trade Policies for Developing Countries," *American Economic Review*, May 2006, *96* (2), 141–146.

**Jaffe, Adam B and Robert N Stavins**, "Dynamic incentives of environmental regulations: The effects of alternative policy instruments on technology diffusion," *Journal of environmental economics and management*, 1995, *29* (3), S43–S63.

**Jaramillo, Paulina, Valerie J. Karplus, P. Chris Pistorius, and Edson Severini**, "The Costs and Distributional Impacts of Decarbonizing the Iron and Steel Industry in the United States," Technical Report 2023.

**Kim, Jinsoo, Benjamin K Sovacool, Morgan Bazilian, Steve Griffiths, Junghwan Lee, Minyoung Yang, and Jordy Lee**, "Decarbonizing the iron and steel industry: A systematic review of sociotechnical systems, technological innovations, and policy options," *Energy Research & Social Science*, 2022, *89*, 102565.

**King, Maia, Bassel Tarbush, and Alexander Teytelboym**, "Targeted carbon tax reforms," *European Economic Review*, 2019, *119*, 526–547.

**Liu, Ernest**, "Industrial Policies in Production Networks," *The Quarterly Journal of Economics*, 08 2019, *134* (4), 1883–1948.

_ **and Song Ma**, "Innovation Networks and R&D Allocation," Working Paper 29607, National Bureau of Economic Research December 2021.

**Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny**, "Industrialization and the Big Push," *Journal of Political Economy*, 1989, *97* (5), 1003–1026.

**Nordhaus, William D.**, *Managing the Global Commons: The Economics of Climate Change*, MIT Press. Cambridge, Massachusetts., 1994.

**OECD**, "Net Effective Carbon Rates," Technical Report 2023.

**Rennert, Kevin, Frank Errickson, Brian C Prest, Lisa Rennels, Richard G Newell, William Pizer, Cora Kingdon, Jordan Wingenroth, Roger Cooke, Bryan Parthum et al.**, "Comprehensive evidence implies a higher social cost of CO2," *Nature*, 2022, *610* (7933), 687–692.

**Rosenstein-Rodan, P. N.**, “Problems of Industrialisation of Eastern and South-Eastern Europe,” *The Economic Journal*, 1943, *53* (210/211), 202–211.

**Sturm, John**, “How to Fix a Coordination Failure: A Šuper-PigouvianÄpproach,” Technical Report 2023.

**Zeira, Joseph**, “Workers, Machines, and Economic Growth,” *Quarterly Journal of Economics*, 1998, *113* (4), 1091–1117.

# 9 Appendix

## 9.1 Example with multiple steady-states in the cap-and-trade case

In this section, we show that in the presence of a cap-and-trade system with a cap $\bar{\ell}_d$, there exist multiple steady-states over a non-empty open set of parameters whenever $N \geq 2$, while there is a unique steady-state when $N = 1$.

Revenues of the dirty production process get allocated to the payment of labor in these processes and emission permits. Given a price on emissions $\tau_t$, and using that revenues of each sector are given by (7), we then get that

$$(1+\tau_t)\,\ell_{dt} = \sum_{i=1}^{N} (1-\chi_{it}) \prod_{j=i+1}^{N} \widetilde{\chi}_{jt}\,(1-\alpha_j)\,, \tag{29}$$

under the maintained assumption that $(1+\tau_t)\,e^z > 1$. If the cap does not bind, then $\tau_t = 0$ and if it binds, $\ell_{dt} = \bar{\ell}_d$ and the previous equation determines uniquely the price of emissions for given technology levels (noting that $\widetilde{\chi}_{jt}$ decreases in $\tau_t$).

With $N = 1$, a steady-state is then characterized by

$$\chi_1 = F_1\left(1 - \frac{e^{-z}}{1+\tau}\right) \text{ and } (1+\tau_t)\,\ell_{dt} = 1 - \chi_1,$$

which defines the steady-state uniquely.

To show that there can be multiple steady-states for $N \geq 2$, we build an example with 2 sectors. We consider parameter values for which the cap always binds. A steady-state is then a pair $\{\chi_1, \chi_2\}$ and a price on emissions $\tau$, which satisfy (13), (15) and (29) such that

$$(1+\tau)\,\bar{\ell}_d = (1-\chi_1)\,\chi_2\,(1-\alpha_2) + (1-\chi_2)\,, \tag{30}$$

$$\chi_1 = F_1\left[\left(1 - \frac{e^{-z}}{1+\tau}\right)\chi_2\,(1-\alpha_2)\right] \tag{31}$$

$$\chi_2 = F_2\left(1 - \left[\frac{e^{-z}}{1+\tau}\right]^{\mu_2}\right) \text{ with } \mu_2 = \alpha_2 + \chi_1\,(1-\alpha_2)\,. \tag{32}$$

We construct non-knife edge examples where one steady-state features $\chi_1 = 0$, $\chi_2 > 0$ and the other features $\chi_1^\dagger = 1$, $\chi_2^\dagger > \chi_2$.

In the first steady-state, $\mu_2 = \alpha_2$ and given (30), $1+\tau = (1-\alpha_2\chi_2)\,/\bar{\ell}_d$, hence (31) and (32) give:

$$0 = F_1\left[\left(1 - \frac{e^{-z}\bar{\ell}_d}{1-\alpha_2\chi_2}\right)\chi_2\,(1-\alpha_2)\right], \tag{33}$$

$$\chi_2 = F_2\left(1 - \left[\frac{e^{-z}\bar{\ell}_d}{1-\alpha_2\chi_2}\right]^{\alpha_2}\right). \tag{34}$$

In the second steady-state, $\mu_2^\dagger = 1$ and given (30), $1+\tau = \left(1-\chi_2^\dagger\right)/\bar{\ell}_d$, hence (31) and (32) give:

$$1 = F_1\left[\left(1-\frac{e^{-z}\bar{\ell}_d}{1-\chi_2^\dagger}\right)\chi_2^\dagger\left(1-\alpha_2\right)\right] \tag{35}$$

$$\chi_2^\dagger = F_2\left(1-\frac{e^{-z}\bar{\ell}_d}{1-\chi_2^\dagger}\right). \tag{36}$$

The right-hand sides of (34) and (36) are decreasing in $\chi_2$ and $\chi_2^\dagger$ respectively. For a sufficiently low cap $\bar{\ell}_d$, we necessarily have that $1-\left[\frac{e^{-z}\bar{\ell}_d}{1-\alpha_2\chi_2}\right]^{\alpha_2} < 1-\frac{e^{-z}\bar{\ell}_d}{1-\chi_2^\dagger}$, in that case, (34) and (36) imply that $\chi_2^\dagger > \chi_2$, and one can build $F_2$ such that there is a large gap between $\chi_2^\dagger$ and $\chi_2$. Again, for a sufficiently low cap, we can then obtain that $\left(1-\frac{e^{-z}\bar{\ell}_d}{1-\alpha_2\chi_2}\right)\chi_2 < \left(1-\frac{e^{-z}\bar{\ell}_d}{1-\chi_2^\dagger}\right)\chi_2^\dagger$. Building $F_1$ such that all the mass of the distribution is between these two values, we can satisfy both (33) and (35). This shows that multiple steady-states are possible.

## 9.2 Proof of Proposition 3

The proof proceeds in several steps: 1) We write down output as a function of labor allocation and technology, 2) we derive the optimal labor allocation, 3) we establish that disutility from labor and pollution is a constant, 4) we derived output as a function of technology only and the simplified social planner problem (equation (17)), and 5) we derive the first order conditions.

Step 1: Under the assumption that $(1+\xi)\,e^z > 1$, the social planner will use the electrified production process whenever it is available. As argued in the text, all electrification in the optimum happens immediately, so that the share of electrified varieties $\{\chi_{it}\}$ is constant over time. In that case, the allocation of labor to production is also constant over time, which means that we can drop time subscript. We then get that for any $k < N$,

$$y_{k+1} = \left(\frac{\ell_{d(k+1)}}{1-\chi_{k+1}}\right)^{1-\chi_{k+1}}\frac{1}{\chi_{k+1}^{\chi_{k+1}}}\left(\frac{e^z\ell_{c(k+1)}}{\alpha_{k+1}}\right)^{\chi_{k+1}\alpha_{k+1}}\left(\frac{y_k}{1-\alpha_{k+1}}\right)^{\chi_{k+1}(1-\alpha_{k+1})}. \tag{37}$$

Assume that for sector $k$, we have:

$$\ln y_k = \sum_{i=1}^{k}\prod_{j=i+1}^{k}\chi_j\left(1-\alpha_j\right)\left[\begin{array}{c}-\left(\chi_i\ln\chi_i+\left(1-\chi_i\right)\ln\left(1-\chi_i\right)\right)+\chi_i\alpha_i\ln\left(\frac{e^z\ell_{ci}}{\alpha_i}\right)\\ -\chi_i\left(1-\alpha_i\right)\ln\left(1-\alpha_i\right)+\left(1-\chi_i\right)\ln\ell_{di}\end{array}\right].$$

Then taking the log of (37), it is immediate that sector $k+1$ follows the same formula. Therefore, denoting by $\omega_i \equiv \prod_{j=i+1}^{N}\chi_j\left(1-\alpha_j\right)$, we can rewrite output as:

$$\ln y_{Nt} = \sum_{i=1}^{N}\omega_i\left[-\left(\chi_{it}\ln\chi_{it}+\left(1-\chi_{it}\right)\ln\left(1-\chi_{it}\right)\right)+\chi_{it}\alpha_i\ln\left(\frac{e^z\ell_{ci}}{\alpha_i}\right)-\chi_{it}\left(1-\alpha_i\right)\ln\left(1-\alpha_i\right)+\left(1-\chi_{it}\right)\right. \tag{38}$$

where, to economize notations, we keep $(1-\alpha_1)\ln(1-\alpha_1)$ in the sum above but treat it as a zero (since this term is not there for sector 1 which features $\alpha_1 = 1$).

Step 2: The planner's problem can thus be re-written as

$$\max_{\{\ell_{di},\ell_{ci}\}} \frac{1}{1-\beta}\left(\ln y_N - (1+\xi)\sum_i \ell_{di} - \sum_i \ell_{ci}\right) - \sum_i \mathcal{F}_i(\chi_i) \tag{39}$$

Taking first order condition with respect to clean and dirty labor inputs, we immediately get that

$$\ell_{ci} = \chi_i \alpha_i \omega_i, \qquad \ell_{di} = \frac{(1-\chi_i)\,\omega_i}{1+\xi} \tag{40}$$

Step 3: Note that the total disutility of labor in production + pollution is always equal to one: Indeed, using (40), we get:

$$(1+\xi)\,\ell_{d1} + \ell_{c1} = \omega_1.$$

Assume that

$$\sum_{j=1}^{i} \left[(1+\xi)\,\ell_{dj} + \ell_{cj}\right] = \omega_i.$$

Then, we get for $i < N$ :

$$\begin{aligned}
\sum_{j=1}^{i+1} \left[(1+\xi)\,\ell_{dj} + \ell_{cj}\right] &= \chi_{i+1}\alpha_{i+1}\omega_{i+1} + (1-\chi_i)\,\omega_{i+1} + \omega_i \\
&= \left[\chi_{i+1}\alpha_{i+1} + (1-\chi_i) + \chi_{i+1}(1-\alpha_{i+1})\right]\omega_{i+1} \\
&= \omega_{i+1}.
\end{aligned}$$

This implies the stated result, namely:

$$\sum_{j=1}^{N} \left[(1+\xi)\,\ell_{dj} + \ell_{cj}\right] = \omega_N = 1.$$

Step 4: Plugging the labor allocation equations, (40), into (40) we can rewrite output as:

$$\begin{aligned}
\ln y_N &= \sum_{i=1}^{N} \omega_i \left[-(1-\alpha_i)\,\chi_i \ln \chi_i + \chi_i\alpha_i \ln(e^z\omega_i) - \chi_i(1-\alpha_i)\ln(1-\alpha_i) + (1-\chi_i)\ln\frac{\omega_i}{1+\xi}\right]. \\
&= \left(\sum_{i=1}^{N} \omega_i\chi_i\alpha_i\right) z + \left(\sum_{i=1}^{N} \omega_i(1-\chi_i)\right)\ln(1+\xi) \qquad (41) \\
&\quad + \sum_{i=1}^{N} \omega_i\left[-(1-\alpha_i)\,\chi_i \ln(\chi_i(1-\alpha_i)) + (1-\chi_i+\chi_i\alpha_i)\ln\omega_i\right]
\end{aligned}$$

We note that

$$\begin{aligned}\sum_{i=1}^{N}(1-\chi_i)\,\omega_i+\sum_{i=1}^{N}\omega_i\chi_i\alpha_i &= \sum_{i=1}^{N}\omega_i-\sum_{i=1}^{N}\omega_i\chi_i\,(1-\alpha_i)\\ &= \sum_{i=1}^{N}\omega_i-\sum_{i=2}^{N}\omega_{i-1}\\ &= \omega_N=1.\end{aligned}$$

Moreover,

$$\begin{aligned}&\sum_{i=1}^{N}\omega_i\left[-(1-\alpha_i)\,\chi_i\ln\left(\chi_i\,(1-\alpha_i)\right)+(1-\chi_i+\chi_i\alpha_i)\ln\omega_i\right]\\ =\;&\sum_{i=1}^{N}\omega_i\ln\omega_i-\sum_{i=1}^{N}\omega_i\,(1-\alpha_i)\,\chi_i\ln\left(\chi_i\,(1-\alpha_i)\,\omega_i\right)\\ =\;&\sum_{i=1}^{N}\omega_i\ln\omega_i-\sum_{i=1}^{N-1}\omega_i\ln\omega_i=0.\end{aligned}$$

Using both relationships in (41), we can rewrite

$$\ln y_N=\left(\sum_{i=1}^{N}\omega_i\chi_i\alpha_i\right)(z+\ln(1+\xi))-1. \tag{42}$$

Dropping constants and multiplying by $1-\beta$ the original problem, we then get that the social planner solves the simplified social planner problem given equation (17).

Step 5: If the solution $\{\chi_i\}$ is interior, it must satisfy the first-order conditions given by:

$$\frac{\ln\left((1+\xi)\,e^z\right)}{1-\beta}\sum_{j=1}^{i}\frac{\omega_j\alpha_j\chi_j}{\chi_i}=\mathcal{F}_i'(\chi_i)\,.$$

We note that $\sum_{j=1}^{i}\frac{\omega_j\alpha_j\chi_j}{\chi_i}=\mu_i\omega_i$, which then delivers (18). To see this, note that for $i>1$,

$$\begin{aligned}\chi_i\omega_i\mu_i &= \chi_i\omega_i\left(\alpha_i+(1-\alpha_i)\,\chi_{i-1}\mu_{i-1}\right)\\ &= \chi_i\omega_i\alpha_i+\chi_{i-1}\mu_{i-1}\omega_{i-1}\\ &= \sum_{j=1}^{i}\omega_j\alpha_j\chi_j.\end{aligned}$$

This establishes Step 5 and thus completes the proof of Proposition 3.

## 9.3 Derivation of equation (23)

To derive the result, we first derive $\frac{\partial\mu_i}{\partial\chi_k}$. We immediately note that $\frac{\partial\mu_i}{\partial\chi_k}=0$ for $k>i$. Further, we have that $\frac{\partial\mu_i}{\partial\chi_{i-1}}=(1-\alpha_i)\,\mu_{i-1}$, while for $k<i-1$, we get: $\frac{\partial\mu_i}{\partial\chi_k}=\chi_{i-1}\,(1-\alpha_i)\,\frac{\partial\mu_{i-1}}{\partial\chi_k}$. Iterating

on $j$ such that $i-j$ goes down to $k+1$, we then obtain:

$$\frac{\partial \mu_i}{\partial \chi_k} = \begin{cases} \frac{\left(\prod_{j=0}^{i-k-1}(1-\alpha_{i-j})\chi_{i-j-1}\right)\mu_k}{\chi_k} & k<i \\ 0 & \text{otherwise} \end{cases}. \tag{43}$$

For $k>i$, it is then immediate that $\frac{\partial \ln \pi_i}{\partial \ln \chi_k}=1$. While for $k<i$, we get

$$\frac{\partial \ln \pi_i}{\partial \ln \chi_k} = \frac{\mu_i Z e^{-\mu_i Z}}{1-e^{-\mu_i Z}} \frac{\partial \ln \mu_i}{\partial \ln \chi_k},$$

which combined with (43) gives (23).

We note that (for $Z>0$)

$$\begin{aligned} \frac{\mu_i Z e^{-\mu_i Z}}{1-e^{-\mu_i Z}} < 1 \quad &\Longleftrightarrow \quad \mu_i Z e^{-\mu_i Z} < 1 - e^{-\mu_i Z} \\ &\Longleftrightarrow \quad 1 + \mu_i Z < e^{\mu_i Z}, \end{aligned}$$

which is true for $Z>0$. Next, for $k=i-1$, we get using (4) that:

$$\begin{aligned} \frac{\left(\prod_{j=0}^{i-k-1}(1-\alpha_{i-j})\chi_{i-j-1}\right)\mu_k}{\mu_i} &= \frac{(1-\alpha_i)\chi_{i-1}\mu_{i-1}}{\mu_i} \\ &= \frac{\mu_i - \alpha_i}{\mu_i} < 1. \end{aligned}$$

In addition, for any $k<i-1$, then, using again (4), we get:

$$\begin{aligned} \frac{\left(\prod_{j=0}^{i-k-1}(1-\alpha_{i-j})\chi_{i-j-1}\right)\mu_k}{\mu_i} &= \frac{\left(\prod_{j=0}^{i-(k+1)-1}(1-\alpha_{i-j})\chi_{i-j-1}\right)(\mu_{k+1}-\alpha_{k+1})}{\mu_i} \\ &< \frac{\left(\prod_{j=0}^{i-(k+1)-1}(1-\alpha_{i-j})\chi_{i-j-1}\right)\mu_{k+1}}{\mu_i}. \end{aligned}$$

Therefore we get that $\frac{\left(\prod_{j=0}^{i-k-1}(1-\alpha_{i-j})\chi_{i-j-1}\right)\mu_k}{\mu_i}$ is increasing in $k$ for $k \leq i-1$ and by induction $\frac{\left(\prod_{j=0}^{i-k-1}(1-\alpha_{i-j})\chi_{i-j-1}\right)\mu_k}{\mu_i} < 1$ for all $k \leq i-1$.

## 9.4 Utility flow in the extended model

The utility flow is given by $\ln y_{2t} - \sum_{i\in\{1a,1b,2\}}\left((1+\xi)\,\ell_{dit} + \ell_{cit}\right)$. In steady-state, all inputs are priced at marginal costs. Since downstream production is Cobb-Douglas between the clean and dirty production process, we get that in steady-state output in sector 2 is given by:

$$y_2 = \frac{\left(e^{\alpha z}\ell_{c2}^{\alpha} y_{1b}^{1-\alpha}\right)^{\chi_2}\left(\ell_{d2}^{\alpha} y_{1a}^{1-\alpha}\right)^{1-\chi_2}}{\alpha^{\alpha}(1-\alpha)^{1-\alpha}\chi_2^{\chi_2}(1-\chi_2)^{1-\chi_2}}, \tag{44}$$

where we kept the assumption that downstream electrified producers prefer to use the clean production process. In sector $1a$ and $1b$, we similarly have

$$y_{1k}=\frac{\left(e^{z}\ell_{c1k}\right)^{\chi_{1k}}\ell_{d1k}^{1-\chi_{1k}}}{\chi_{1k}^{\chi_{1k}}\left(1-\chi_{1k}\right)^{1-\chi_{1k}}}\text{ for }k\in\{a,b\}\,. \tag{45}$$

Since wages and revenues are both equal to $1$, then clean labor commands income shares equal to its overall factor share in production, namely we have:

$$\ell_{c2}=\chi_2\alpha,\ \ell_{c1b}=\chi_2\left(1-\alpha\right)\chi_{1b}\text{ and }\ell_{c1a}=\left(1-\chi_2\right)\left(1-\alpha\right)\chi_{1a}. \tag{46}$$

With Pigouvian taxation, the cost of using dirty labor is $1+\xi$, such that we obtain:

$$\left(1+\xi\right)\ell_{d2}=\left(1-\chi_2\right)\alpha,\ \left(1+\xi\right)\ell_{d1b}=\chi_2\left(1-\alpha\right)\left(1-\chi_{1b}\right)\text{ and }\left(1+\xi\right)\ell_{d1a}=\left(1-\chi_2\right)\left(1-\alpha\right)\left(1-\chi_{1a}\right). \tag{47}$$

Using these expressions, we get that the total disutility of labor in production plus pollution is always equal to one:

$$\left(1+\xi\right)\sum_i\ell_{di}+\sum_i\ell_{ci}=1.$$

Plugging (45) in (44) and taking logs we can express log output as:

$$\begin{aligned}\ln y_2 =& -\left(\chi_2\ln\chi_2+\left(1-\chi_2\right)\ln\left(1-\chi_2\right)\right)-\alpha\ln\alpha-\left(1-\alpha\right)\ln\left(1-\alpha\right)\\ &+\chi_2\alpha z+\chi_2\alpha\ln\ell_{c2}+\left(1-\chi_2\right)\alpha\ln\ell_{d2}\\ &+\chi_2\left(1-\alpha\right)\left[\chi_{1b}z+\chi_{1b}\ln\ell_{c1b}-\chi_{1b}\ln\chi_{1b}+\left(1-\chi_{1b}\right)\ln\ell_{d1bt}-\left(1-\chi_{1b}\right)\ln\left(1-\chi_{1b}\right)\right]\\ &+\left(1-\chi_2\right)\left(1-\alpha\right)\left[\chi_{1a}z+\chi_{1a}\ln\ell_{c1a}-\chi_{1a}\ln\chi_{1a}+\left(1-\chi_{1a}\right)\ln\ell_{d1a}-\left(1-\chi_{1a}\right)\ln\left(1-\chi_{1a}\right)\right]\end{aligned}$$

Substituting the labor allocations, (46) and (47), into the previous expression gives:

$$\begin{aligned}\ln y_2 =& -\left(\chi_2\ln\chi_2+\left(1-\chi_2\right)\ln\left(1-\chi_2\right)\right)-\alpha\ln\alpha-\left(1-\alpha\right)\ln\left(1-\alpha\right)\\ &+\chi_2\alpha z+\chi_2\alpha\ln\left(\chi_2\alpha\right)+\left(1-\chi_2\right)\alpha\ln\frac{\left(1-\chi_2\right)\alpha}{\left(1+\xi\right)}\\ &+\chi_2\left(1-\alpha\right)\left[\chi_{1b}z+\ln\chi_2+\ln\left(1-\alpha\right)-\left(1-\chi_{1b}\right)\ln\left(1+\xi\right)\right]\\ &+\left(1-\chi_2\right)\left(1-\alpha\right)\left[\chi_{1a}z+\ln\left(1-\chi_2\right)+\ln\left(1-\alpha\right)-\left(1-\chi_{1a}\right)\ln\left(1+\xi\right)\right].\end{aligned}$$

Rearranging terms, we then obtain:

$$\begin{aligned}\ln y_2 =& \chi_2\alpha z-\left(1-\chi_2\right)\alpha\ln\left(1+\xi\right)\\ &+\chi_2\left(1-\alpha\right)\left[\chi_{1b}z-\left(1-\chi_{1b}\right)\ln\left(1+\xi\right)\right]\\ &+\left(1-\chi_2\right)\left(1-\alpha\right)\left[\chi_{1a}z-\left(1-\chi_{1a}\right)\ln\left(1+\xi\right)\right].\end{aligned}$$

We then get that the utility flow in steady state is given by

$$\begin{aligned}&\ln y_2-\textstyle\sum_{i\in\{1a,1b,2\}}\left(\left(1+\xi\right)\ell_{di}+\ell_{ci}\right)\\ =&\left[\chi_2\alpha+\chi_2\left(1-\alpha\right)\chi_{1b}+\left(1-\chi_2\right)\left(1-\alpha\right)\chi_{1a}\right]Z-\ln\left(1+\xi\right)-1.\end{aligned}$$

## 9.5 Details on the calibration

### 9.5.1 Calibrated model

We briefly present the model that we calibrate in Section 7. We modify the model such that each sector may feature heterogeneity in the relative productivity of clean and dirty technologies $z_i$, the emission rate associated with the use of the dirty production process $\xi_i$ and a TFP parameter $A_i$. We present the model for $N \geq 2$ – though we will have $N = 2$ in the calibration. Therefore a variety $\nu$ in sector $i$ is now produced according to

$$y_{it}(\nu) = A_i \left[ \ell_{dit}(\nu) + 1_{elec}(v,i) \left( \frac{e^{z_i} \ell_{cit}(\nu)}{\alpha_i} \right)^{\alpha_i} \left( \frac{m_{it}(\nu)}{1-\alpha_i} \right)^{1-\alpha_i} \right],$$

where $1_{elec}(v,i)$ is an index function which takes value 1 if and only if production process $\nu$ in sector $i$ has been electrified.

Following the same logic as in the baseline model, the price of production process $\nu$ in sector $i$ is given by:

$$p_{it}(\nu) = \begin{cases} \min\left(e^{-\alpha_i z_i} p_{i-1,t}^{1-\alpha_i}, 1+\tau_{it}\right)/A_i & \text{if electrified by time } t-1, \\ (1+\tau_{it})/A_i & \text{otherwise,} \end{cases} \tag{48}$$

where for sector 1, $\alpha_1 = 1$ and the term $p_{i-1,t}^{1-\alpha_i}$ drops from the previous expression.

Empirically, we get that, even when $\tau_{it} = 0$, the clean production process is cheaper than the dirty one (i.e. $e^{-\alpha_2 z_2} p_{1,t}^{1-\alpha_2} < 1$ and $z_1 > 0$). Assuming that this is the case, we can then solve for the price index in each sector as:

$$p_{1t} = \frac{1}{A_1} (1+\tau_{1t})^{1-\chi_{1,t-1}} e^{-\chi_{1,t-1} z_1} \tag{49}$$

$$\text{and} p_{it} = \frac{1}{A_i} (1+\tau_{it})^{1-\chi_{i,t-1}} e^{-\chi_{i,t-1}\alpha_i z_i} p_{i-1,t}^{(1-\alpha_i)\chi_{i,t-1}} \text{for } i > 1. \tag{50}$$

Because sectors are heterogeneous in $z$ and $\tau$, we cannot introduce the variable $\mu_i$ that allowed for closed form solutions as before. However, this does not affect the logic of the model.

Following the same steps as in the baseline model, we get that if a variety from sector $i$ is electrified at time $t$, the innovator obtains a profit margin given by $1 - \frac{e^{-\alpha_i z_i} p_{i-1,t}^{(1-\alpha_i)}}{1+\tau_{it}}$ (or $1 - \frac{e^{-z_1}}{1+\tau_{1t}}$ for sector 1). That is, an innovator obtains profits given by

$$\pi_{it}(\nu) = r_{it} \left[ 1 - \frac{e^{-\alpha_i z_i} p_{i-1,t}^{(1-\alpha_i)}}{1+\tau_{it}} \right],$$

where, as before, $r_{it}$ denote the revenue of a variety in sector $i$ at time $t$. We still get that in sector

$N$, $r_{Nt} = p_t c_t = 1$, while we can obtain revenues for sector $i < N$ recursively from:

$$r_{it} = \prod_{j=i+1}^{N} \widetilde{\chi}_{jt} \left(1 - \alpha_j\right),$$

where $\widetilde{\chi}_{j,t}$ is now defined as

$$\widetilde{\chi}_{j,t} \equiv \chi_{j,t-1} + \left(\chi_{j,t} - \chi_{j,t-1}\right) \frac{e^{-\alpha_j z_j} p_{j-1,t}^{(1-\alpha_j)}}{1 + \tau_{jt}} \tag{51}$$

. For completeness, we introduce the possibility that the government subsidizes electrification in sector $i$ at rate $s_{it}$ (for the calibration, we assume that $s_{it} = \beta$). Then get that an equilibrium is defined as follows:

**Definition.** Given initial electrification shares $\{\chi_{i0}\}$, an equilibrium with a sequence of carbon taxes $\{\tau_{it}\}$ and electrification subsidies $\{s_{it}\}$ is the sequence $\{\chi_{it}, \widetilde{\chi}_{it}, p_{it}\}_{t>0}$ such that $p_{1t}$ and $p_{it}$ obey (49) and (50), $\widetilde{\chi}_{it}$ is given by (51), and $\chi_{it}$ is given by

$$\chi_{it} = \max\left\{ F_i\left(\left[1 - \frac{e^{-\alpha_i z_i} p_{i-1,t}^{(1-\alpha_i)}}{1 + \tau_{it}}\right] \frac{\prod_{j=i+1}^{N} \widetilde{\chi}_{jt}\left(1 - \alpha_j\right)}{1 - s_{it}}\right), \chi_{i,t-1} \right\}. \tag{52}$$

As before, the equilibrium is unique. In turn, a steady-state is characterized by:

**Definition.** For given carbon taxes $\{\tau_i\}$ and electrification subsidies $\{s_i\}$, a steady-state is a vector of electrification shares and prices $\{\chi_i, p_i\}$ such that

$$p_1 = \frac{1}{A_1}\left(1 + \tau_1\right)^{1-\chi_1} e^{-\chi_1 z_1}, \quad p_i = \frac{\left(1 + \tau_i\right)^{1-\chi_i} e^{-\chi_i \alpha_i z_i} p_{i-1}^{(1-\alpha_i)\chi_i}}{A_i}, \tag{53}$$

$$\chi_i \geq F_i\left(\left(1 - \frac{e^{-\alpha_i z_i} p_{i-1}^{(1-\alpha_i)}}{1 + \tau_i}\right) \frac{\prod_{j=i+1}^{N} \chi_j\left(1 - \alpha_j\right)}{1 - s_i}\right). \tag{54}$$

As before, the inequality in (54) results from the fact that electrification can never decrease and we focus on interesting steady-states where (54) holds with equality.

To derive output in each sector, we start from the output for each variety, which using the Cobb-Douglas aggregator must obey: $y_{it}(\nu) = r_{it}/p_{it}(\nu)$. We then get that output of a non-electrified or newly electrified variety is $y_{it}(\nu) = A_i \left(1 + \tau_{it}\right)^{-1} r_{it}$, while the employment of dirty labor for non-electrified varieties is $\left(1 + \tau_{it}\right)^{-1} r_{it}$. The output of a previously electrified variety is $y_{it}(\nu) = \frac{A_i r_{it}}{e^{-\alpha_i z_i} p_{i-1,t}^{1-\alpha_i}}$. We then get that the output of sectoral good $i$—combining all varieties—is

$$y_{it} = A_i \left(e^{-\alpha_i z_i} p_{i-1,t}^{1-\alpha_i}\right)^{-\chi_{i,t-1}} \left(1 + \tau_{it}\right)^{-(1-\chi_{i,t-1})} r_{it},$$

with the disutility from emissions in sector $i$ is given by:

$$a_{it} = \xi_i \frac{1 - \chi_{it}}{1 + \tau_{it}} r_{it}.$$

### 9.5.2 Data appendix

This section describes in more detail three adjustments we make to the data for the calibration of the distribution of fixed innovation costs in the model. Broadly speaking, we quantify these one-time costs based on the differences between first and n-th of a kind costs of producing clean H2 or steel, respectively, where the differences are due to learning by doing and improvements technology. For renewable energy electrolysis-based H2 production we use data from BloombergNEF (2023) and for clean (H2-DRI-EAF) steel production we use data from BloombergNEF (2021) in the benchmark calibration. The adjustments are as follows. First, we keep only *relevant* innovation costs. That is, for clean H2 production, we remove projected cost changes due to renewable electricity innovation as we take take these to be exogenous to the hydrogen sector.[27] For clean steel production, we similarly remove hydrogen input costs from the LCOS calculations as hydrogen innovation costs are already accounted for in the model.[28] Second, we map per-unit excess costs into total fixed costs in model units by multiplying by the total number of units produced in sector $i$ in each region $j$ and dividing by regional revenues in period $t$:

$$\frac{\int \phi(\chi_i) d\chi_i}{p_{i,t} C_{i,t,j}} = \frac{TotalFixedCosts_{i,j}}{Revenues_{i,t,j}} = \frac{(FixedCosts/unit)_j \cdot Units_{j,t}}{p_{i,t} \cdot Units_{j,t}}. \tag{55}$$

If producers' myopia is corrected by appropriate subsidies, quantity units cancel out of this calculation so that the model results are robust to different time horizon assumptions. Third, in assigning global production shares $\chi_i$ to each country's fixed costs, we face two challenges. One is that our cost estimates do not cover all producers in the world, so we must assign un-modeled producers' output to those for which we observe costs. This problem is especially relevant for steel production in the BloombergNEF data used in the benchmark calibration, which reports cost estimates only for three major steel producing countries (China, the United States, and Germany). Two, for hydrogen it is not clear how useful present-day production shares across countries are for predicting potential future shares given that today's hydrogen production primarily serves ammonia and oil refining industries, whereas our model is about new hydrogen production as an input for iron and steel. We thus proceed as follows. First, given its outsized role in the modeled industies, we keep China's production shares as they are in (fully global) output data in the base year, namely 54% of global steel production (WSA 2023) and 30% of global hydrogen production in 2022 (IEA 2022). For hydrogen, we then distribute the remainder of potential global production evenly across the other countries in the data, attributing an output share of 2.6% to each. For

[27]We specifically remove the average cost share of renewable electricity - which is 56% across the years and technologies for which the breakdown is reported in BloombergNEF (2023) - from both present and future clean LCOH2s before calculating excess initial technology costs.

[28]We use the projected *future* hydrogen efficiency (66kgH2/tSteel) in this calculation, thus maintaining the initial hydrogen consumption requirement's (73kgH2/tSteel) excess cost as part of the innovation costs to achieve long-run n-th of a kind costs.

steel, we split the remaining potential output aross the U.S. and Germany based on their relative current global steel production shares of 4% and 1.9%, respectively (WSA 2023), net of an assumed additional global output share of 0.3% going to a synthetic "high cost" producer country assumed to have 20% higher fixed costs than Germany (motivated by the fact that estimates with broader country coverage - such as Devlin et al. (2023) discussed below - suggest that countries such as Brazil or Chile may face especially high clean steel production and innovation costs).

Given the limited country coverage of the BloombergNEF (2021) LCOS estimates, we also use estimates from Devlin et al. (2023) for robustness, which cover production across muliple regions within each of 17 countries. The downsides to these estimates for our purposes are that they (i) begin in 2030 as first model year rather than 2023, and (ii) focus exclusively on integrated on-site green electricity, hydrogen, and steel production rather than allowing for traded hydrogen and storage. We remove the costs of electrolyzers, wind turbines and/or solar panels from their clean LCOS estimates for each country so as to focus on the H2-DR-EAF cost components, but note that this is an imperfect correction since, on the one hand, O&M and labor costs may still reflect hydrogen and electricity production components and, on the other hand, hydrogen transportation and storage costs - which are included in the BloombergNEF data - are not covered here in the relevant way. For production shares, we again assume a 54% share for China in line with current data, and redistribute the non-modeled global production to the other modeled countries proportional to their within-sample output shares (computed by Devlin et al. (2023) based on a 5-year moving average).