Spillovers and the Direction of Innovation: An Application to the Clean Energy Transition*

Eric Donald[†]
Boston University

September 15, 2023
Click here for latest version

Abstract

This paper studies the role of policy in the transition to clean technology. To do so, I develop a general endogenous growth model with clean and dirty technologies and a rich network of cross-technology knowledge spillovers. I show that the size and speed of technological redirection, following a policy reform, depend on two forces: (i) substitution patterns in production and (ii) the network of cross-technology knowledge spillovers, each summarized by a sufficient statistic matrix. Calibrating my spillover network with patent data, I apply my theory to transportation and electricity generation in the US. I find that a carbon price and clean innovation subsidy create large long-run increases in clean technology for both sectors, but that the transition is slow, with half-lives of convergence of 121 years and 127 years, respectively. Using spectral analysis, I unpack the connections in the spillover network that drive each sector's convergence speed. I then characterize the optimal mix of carbon prices and innovation subsidies along the transition, showing that innovation subsidies should reflect a technology's centrality in the spillover network, rather than its associated pollution externality. Simulating the optimal policy path, I find that clean technologies create few knowledge spillovers and receive correspondingly small innovation subsidies. However, if carbon prices are incomplete, innovation subsidies should adjust to reflect the social cost of carbon, with clean technologies rewarded for their ability to reduce emissions in equilibrium.

*I am deeply grateful to my advisors Pascual Restrepo, David Lagakos, and Yuhei Miyauchi for their guidance and support. I thank Stephen Terry, Masao Fukui, Bob King, Adam Guren, Kevin Kuruc, and John Firth as well as seminar participants at Boston University and Global Priorities Institute for their helpful comments.

[†]Email: ericdon@bu.edu

1 Introduction

Decarbonizing the economy will require significant investments in clean innovation, so policymakers have set out to redirect the path of innovation toward clean technology. The main challenge for this clean technological transition is that innovation choices today depend not only on current policies but also on the history of invention. Technological transitions induced by policy reforms can be slow, or nonexistent, when established technologies face *increasing returns to innovation*: where more advanced technologies offer greater rewards for innovation, incentivizing yet more innovation that further deepens their lead. This raises two questions. On the descriptive side, how much, and how quickly, can policy reforms such as carbon pricing redirect the path of innovation? On the normative side, what path of carbon prices and innovation subsidies would generate an optimal clean technological transition?

To answer these questions, I develop a general endogenous growth model with clean and dirty technologies whose key feature is a network of cross-technology knowledge spillovers. These spillovers allow researchers in one technology to build on the productive knowledge of other technologies; a process that has featured prominently in the history of clean innovation. For instance, the first Tesla prototype – the Mule 1 – was a gasoline-powered car that engineers at Tesla reconfigured by tearing out the combustion engine and filling the engine compartment with batteries. Similarly, when researchers at Bell Labs invented the modern solar cell, they made use of conductive properties of silicon already known from previous research on semiconductors. The inventors of these clean technologies did not have to start from scratch. Instead, they could build on existing knowledge from other technologies.[1]

My paper's first contribution is to show that these cross-technology knowledge spillovers, together with substitution patterns in production, shape the impact of policy on the direction of innovation. I provide precise formulas for the size and speed of technology's transition, following a policy reform, that depend on two matrices: one summarizing substitution patterns in production and another summarizing the network of knowledge spillovers. These two matrices are sufficient statistics, providing a clear mapping from the model to the data for an arbitrary number of technologies and a general class of production and spillover structures. That is, estimates of these matrices allow a first-order characterization of policy's impact on technology, in both the steady-state and transition, without needing to know the entire underlying structure of spillovers and production.

Cross-technology knowledge spillovers prevent increasing returns to innovation by enhancing the research productivity of less advanced technologies, while substitutability creates increasing returns by increasing the market size of more advanced technologies. By combining the two sufficient statistic matrices, I quantify the degree of increasing returns to innovation in terms of a spectral radius. When this spectral radius is above one, technology is path dependent and only very large policy reforms can redirect innovation. Conversely, a spectral radius below one implies that policy reforms will always influence the long-run path of innovation, but the transition may be slow if the spectral radius is near one.

My formulas show that the long-run impact of policy and the speed of technology's transition are inversely related because increasing returns to innovation increase the overall size of transition while at the same time slowing it down. By favoring more advanced technologies, increasing returns allow technologies favored by policy to establish larger leads over their peers in the eventual steady-state. However, the same mechanism slows down the transition by preventing less advanced technologies from

[1] Figures C.1 and C.2 provide illustrations of these two examples of cross-technology knowledge spillovers.

achieving catchup growth. Thus, when policy reforms are transformative in the long-run, they will tend to have little impact in the short-run, and vice versa. For example, cross-technology knowledge spillovers speed up transitions by stimulating catchup innovation, but those same spillovers prevent any one technology from gaining too large a lead in the long run.

My paper's second contribution is to characterize the path of carbon prices and innovation subsidies that generate an optimal clean technological transition. I show that policy should separately correct the pollution externality and knowledge spillover externality, in keeping with the Pigou Principle. In particular, carbon prices are set so production internalizes the social cost of carbon, but carbon prices should not explicitly influence the direction of innovation. Instead, with the pollution externality corrected, innovators simply need to internalize the knowledge spillovers they create. To this end, I derive a general, recursive formula for innovation subsidies that reflects the social value of knowledge spillovers sent through the spillover network. These innovation subsidies consider both the value of technology in producing goods, as summarized by Domar weights, and the value of technology in producing innovations, as summarized by centrality in the spillover network, with the weight between the two determined by the Planner's level of patience. Just as carbon prices do not explicitly target innovation, innovation subsidies do not explicitly target the pollution externalities associated with each technology.

Given the political difficulty of carbon pricing, I consider second-best innovation subsidies where carbon prices are set to some external, potentially sub-optimal, level. I show that the same recursive formula for innovation subsidies holds as before but with a simple adjustment that adds the social cost of carbon, net of the external carbon price, multiplied by the impact of innovation on equilibrium emissions. This provides a general formula for innovation policy in the presence of improperly priced externalities. As a result, clean technologies, that reduce equilibrium emissions, are further subsidized, whereas dirty technologies, that increase equilibrium emissions, are punished. In this way, the Planner modifies innovation subsidies to prevent "immiserizing growth", where innovation reduces welfare in inefficient economies by exacerbating distortions (Bhagwati, 1958).

Knowledge spillovers play a fundamental role in both describing policy's impact on the direction of innovation and prescribing optimal policy to steer the direction of innovation. However, the role of spillovers is slightly different in each case. As described above, lagging technologies enjoy greater research productivity when they *receive* spillovers from their more advanced peers, reducing increasing returns to innovation. However, on the normative side, the increase in research productivity enjoyed by the recipients of spillovers is already internalized by innovators. Optimal innovation policy instead focuses on the spillovers that a technology *sends* to its peers, as summarized by a technology's centrality in the spillover network.

My paper's third contribution is a quantitative application of my model to the transportation and electricity generation sectors in the United States. These are two sectors where production can be clearly delineated into "clean" and "dirty", and they together account for a little more than half of US emissions.[2] Using the citation network of granted US patents as a proxy for the network of knowledge spillovers, I find, for both sectors, that increasing returns are large enough to create sluggish transitions, but not so large as to induce path dependence. Cross-technology knowledge spillovers are strong enough to overcome the substitutability of clean and dirty goods, thereby preventing destabilizing increasing

[2] According to the EPA, transportation and electricity generation accounted for 28.5% and 25% of total US greenhouse gas emissions in 2021, respectively.

returns to innovation.

I apply my model to the data with four quantitative exercises. First, to validate my model, I examine the progress made in clean transportation and electricity generation during the 2010s.[3] Despite being an untargeted moment, my model can explain this episode of clean technological progress while also accounting for US climate policy in this period. However, when I shut down the spillover network, my model predicts a counterfactual outcome: the past decade would have seen dirty technologies become *more* entrenched in both sectors, not less.

Second, I examine the impact on the direction of innovation of introducing a \$51 carbon price (the Biden Administration's current estimate of the social cost of carbon) and 30% clean innovation subsidy (consistent with the subsidies in the Inflation Reduction Act). I find a large long-run effect of this policy reform, with the steady-state prevalence of clean technology increasing by 116.6% and 120.3% for transportation and electricity generation, respectively. However, the transition to this new steady-state is slow, with half-lives of convergence of 121 years and 127 years, respectively. Underpinning both the sizable long-run impact and sluggish transition is a spectral radius near one. Furthermore, using spectral analysis, I unpack the connections in the spillover network that drive convergence speeds for each sector.

Third, I simulate the optimal path of carbon prices and innovation subsidies. This exercise highlights the distinction between the spillovers a technology *receives* and the spillovers it *sends*. When carbon pollution is priced at its social marginal cost, clean innovation subsidies play only a limited role in the clean transition. This is because clean technologies are beneficial for their ability to produce goods without pollution, not necessarily for their ability to produce knowledge spillovers. Indeed, the patent data indicates that clean technologies do not have high centrality in the spillover network. As a result, in the first century of policy, clean transportation and electricity generation receive average innovation subsidies of only 39.8% and 59.8% of the baseline innovation wedge, respectively.[4] As the Planner becomes more patient, with a lower discount rate, they implement higher carbon prices and *lower* clean innovation subsidies because their greater concern for the future increases the weight they place on a technology's centrality in the spillover network.

Finally, I simulate the path of second-best clean innovation subsidies, where carbon prices are incomplete. I consider two cases, one where the carbon price is ten percent of the social cost of carbon and another where it is set to zero, and show that these two cases differ substantially. In the case with a small carbon price, clean innovation subsidies shift upward by about 20-30% of the baseline innovation wedge, and little welfare is lost relative to the first-best. However, if pollution is unpriced, then decarbonization requires shutting down dirty innovation and relying exclusively on clean innovation. This leads to slow emission reductions, due to a rebound effect of clean innovation, and slow economic growth, due to the loss of spillovers from dirty technology, which together lead to far greater welfare losses.

1.1 Related Literature

This paper contributes to several strands of literature. First is the extensive literature on endogenous innovation and the climate, such as Gans (2012), Acemoglu et al. (2016, 2023), Fried (2018), and Hassler et al. (2021). The most notable point of comparison in this literature is the seminal work by Acemoglu

[3]From 2010 to 2021, the percentage of new light vehicles that were hybrid or electric rose from 2.4% to 9.8%, while the percentage of electricity generated from non-fossil sources rose from 30.1% to 39%.

[4]The baseline innovation wedge is the steady-state innovation subsidy all technologies would receive in the absence of a spillover network.

et al. (2012) who show that the substitutability of clean and dirty goods implies increasing returns to innovation. My model generalizes this line of reasoning by quantifying the degree of increasing returns to innovation in terms of two sufficient statistic matrices that reflect both substitutability and cross-technology knowledge spillovers.[5] I also show that increasing returns to innovation has different implications at different time horizons.

My paper also relates to the topic of technological path dependence.[6] Several papers in this literature argue for the role of cross-technology knowledge spillovers in stabilizing technology's steady-state, including Acemoglu (2002, 2023) and Fried (2018). My formulas generalize these results by providing a test for technological path dependence in terms of a spectral radius, applicable for an arbitrary number of technologies and a broad class of production and spillover structures.

More generally, there is extensive literature, going back to Hicks (1932), that considers the role of input prices in determining the direction of innovation. Examples of empirical research looking at the influence of dirty input prices on clean innovation include Newell et al. (1999), Popp (2002), Aghion et al. (2016, 2023), and Känzig (2023).[7] Examples of theoretical contributions from the skill-biased technological change literature include Acemoglu (1998, 2002) and Acemoglu and Restrepo (2018, 2022a).[8] While my paper focuses on climate policy, my model provides a general framework that explains the key forces determining the impact of input prices on the direction of innovation, while also providing guidance for the measurement of these forces.[9]

My paper also relates to the vast literature on integrated assessment models (IAMs) that, starting with the seminal contribution of Nordhaus (1992), seeks to quantify the social harm of carbon pollution. Examples from this literature include Stern (2007), Golosov et al. (2014), Nordhaus (2017), Krusell and Smith (2022), and Barrage and Nordhaus (2023).[10] My paper provides a rich model of the response of clean innovation to climate policy and shows that innovation subsidies should include the social cost of carbon whenever pollution prices are limited. In doing so, I provide policymakers in the innovation space a way to make use of the information provided by IAMs.

By focusing on innovation policy in the context of a spillover network, my work is closely related to the recent work of Liu and Ma (2021).[11] Specific to my context is the capacity for some technologies to enable externalities in production, and I show that innovation subsidies should treat improperly priced externalities in the same manner as a Domar weight, reflecting the contemporaneous impact of innovation on welfare. Finally, the methodological tools I use to characterize the transition path of technology following a policy reform are similar to those used by Kleinman et al. (2023).[12] Their paper considers

[5] Fried (2018) also argues that cross-technology knowledge spillovers mediate the impact of carbon taxation.

[6] For research on the topic of path dependence in spatial economies, see Krugman (1991), Bleakley and Lin (2012), and Allen and Donaldson (2020).

[7] For empirical work on the direction of innovation in other settings, see Acemoglu and Linn (2004) and Acemoglu and Finkelstein (2008) for the context of healthcare and Hanlon (2015) for a historical examination of the textile industry.

[8] Other examples of research on the determinants of the direction of innovation more broadly include Budish et al. (2015) on the choice of healthcare firms to invest in preventive medicine and Moscona and Sastry (2022) on the crop-specific pests and pathogens that receive attention in agricultural research.

[9] My focus on the role of policy in steering the direction of innovation is also related to the recent AEA Distinguished Lecture by Acemoglu (2023).

[10] For IAMs with endogenous innovation, see Nordhaus (2002) and Popp (2004).

[11] In general, knowledge spillovers have long been a central consideration of innovation policy (Arrow, 1962; Romer, 1990; Bloom et al., 2019; Akcigit et al., 2020; Bryan and Williams, 2021). Closely related is a body of empirical work focused on estimating the spillover benefits of innovation (Jones and Williams, 1998; Bloom et al., 2013; Jones and Summers, 2021; Myers and Lanahan, 2022).

[12] The technique of characterizing macroeconomic dynamics using the spectral properties of a matrix is discussed in Blanchard and Kahn (1980), Uhlig (1999), and Galor (2007).

the transition dynamics of the distribution of capital and labor across space following a shock to regional fundamentals. I take a complementary approach to describe transition dynamics in the context of directed technological change.

The remainder of the paper is organized as follows. In Section 2, I describe the endogenous growth model used throughout the paper. In Section 3, I consider a once-and-for-all policy reform, characterize the change in technology's steady-state, and derive technology's transition path to its new steady-state. In doing so, I quantify the degree of increasing returns to innovation. In Section 4, I derive optimal innovation subsidies and carbon prices. I also consider a second-best case where the Planner cannot control the carbon price. In Section 5, I describe the calibration of the spillover network and other structural parameters of my model. I also showcase my model's ability to match the 2010s advance of clean technology in the US. In Section 6, I perform quantitative exercises that simulate the impact of a policy reform on the direction of innovation as well as optimal climate policy in the first- and second-best. Section 7 concludes.

2 Model

This section lays out the endogenous growth model used throughout the paper. Time is discrete and indexed by t. Production in this economy consists of j-many technologies, which can produce varying levels of pollution in production. The innovation process for each of these technologies is endogenously determined by the incentives to produce a step-ladder innovation which entitles its owner to a single period of monopoly rents.

The economic environment has two externalities, which are the focus of policy. First, inputs specific to the dirty forms of production, e.g. fossil fuels, generate carbon emissions, and those emissions damage future productive capacity. Second, innovation creates knowledge spillovers via a spillover network, so innovators do not consider the benefit they bestow on the future production of knowledge. To correct these externalities, I will consider a Planner with access to (i) a carbon price and (ii) an array of technology-specific innovation subsidies.

I will start by describing the production and innovation technology of this economy, followed by a description of how the direction of innovation is determined in competitive equilibrium. Throughout the paper, I will use bold notation to denote matrices. I will use i or j to index technologies, with the final technology denoted by J.

2.1 Production

Each technology in this economy produces a distinct good, and these technology-specific goods aggregate into final output. Final output is produced according to

$$\mathcal{Y}_t = \Omega_t F\big(\{Y_{jt}\}\big), \tag{1}$$

where Ω_t is the climate damage function, and $F(.)$ is a constant returns aggregator of technology-specific goods. Climate damages Ω_t are a function of the past sequence of emissions $\{\mathcal{E}_{\hat{t}}\}_{\hat{t}\leq t}$. An important assumption embedded in this specification is that climate damages are Hicks-neutral, so they do not influence relative marginal products.

The good specific to technology j is produced according to

$$\ln(Y_{jt}) = \alpha \ln(\Lambda_{jt}) + (1-\alpha) \int_0^1 \ln(y_{j\iota t}) d\iota, \tag{2}$$

where Λ_{jt} is an input and $y_{j\iota t}$ is an intermediate, both for technology j. That is, technology-specific goods are produced via a Cobb-Douglas with share parameter α that combines an input and a unit interval of intermediates.

Using one unit of input Λ_{jt} produces ω_j units of carbon emissions, so I will call technologies with $\omega_j = 0$ "clean" and those with $\omega_j > 0$ "dirty". Thus, total carbon emissions follow

$$\mathcal{E}_t = \sum_j \omega_j \Lambda_{jt}. \tag{3}$$

Dirty inputs can be thought of as fossil fuels whose use generates carbon emissions. Carbon intensities ω_j may differ across dirty technologies due to the different types of fossil fuels they use. For instance, the main fuels used to produce electricity – natural gas and coal – have different carbon intensities than that used in transportation: gasoline.[13] Denote the overall emissions intensity of the economy by $\bar{\omega}_t \equiv \mathcal{E}_t / \mathcal{Y}_t$.

Inputs Λ_{jt} are produced using r_j units of the final good, so the economy features "round-about" production. For concreteness, one can imagine that dirty input costs correspond to the resources required to extract fossil fuels, but more generally, r_j should be thought of as representing the intrinsic ability of technology j to convert resources into a useful output. This is relevant in the case of the clean energy transition because fossil fuel-based machinery is relatively inefficient at converting primary energy into useful energy (BP, 2022).

Intermediate production is linear in labor, so

$$y_{j\iota t} = a_{j\iota t} \ell_{j\iota t}, \tag{4}$$

where $a_{j\iota t}$ is the labor productivity of intermediate ι for technology j. I define the technology stock of j as the geometric average of its labor productivities:

$$A_{jt} \equiv \exp\Big(\int_0^1 \ln(a_{j\iota t}) d\iota \Big). \tag{5}$$

The economy has a fixed endowment of labor that is supplied inelastically, so the total demand for labor must satisfy

$$\sum_j \int_0^1 \ell_{j\iota t} d\iota \leq L, \tag{6}$$

where L is the aggregate supply of labor.[14]

Finally, the resource constraint requires that

$$\mathcal{Y}_t = c_t + \sum_j r_j \Lambda_{jt}, \tag{7}$$

[13] According to the EIA, the carbon intensity of natural gas, coal, and gasoline is 52.9, 96.1, and 70.7 kilograms of CO_2 per million Btu, respectively.

[14] I abstract from capital accumulation, which is a standard simplifying assumption in the directed technological change literature (Acemoglu, 2002; Acemoglu et al., 2012, 2016; Fried, 2018).

where c_t is household consumption.

2.2 Innovation

In this section, I describe the innovation production function. As is typical in the endogenous growth literature, new ideas are produced by combining scientists and old ideas. Because there are multiple technologies, I specify a spillover function that defines the network of cross-technology knowledge spillovers.

Innovation follows a step-ladder process à la Grossman and Helpman (1991) and Aghion and Howitt (1992), so if intermediate ι for technology j receives an innovation, we have that $a_{j\iota t}$ increases to $\gamma a_{j\iota t}$, where $\gamma > 1$ is the step size of innovation. Define z_{jt} as the mass of intermediates for technology j that receives an innovation. Plugging this into Equation (5), we have that technology evolves according to

$$A_{jt} = \gamma^{z_{jt}} A_{jt-1}. \tag{8}$$

This formulation implies that technology-specific growth rates follow

$$g_{jt} = \ln(\gamma) z_{jt}. \tag{9}$$

The production of innovation follows

$$z_{jt} = \chi_j s_{jt}^{\eta} \phi_j\big(\{A_{it-1}\}\big), \tag{10}$$

where s_{jt} is the number of scientists devoted to technology j, and $\phi_j(.)$ is a spillover function that governs how research productivity is affected by the state of technology in the economy. I assume that spillover functions are homogeneous of degree zero, so spillovers remain constant if all technologies scale together. The parameter $\eta \in (0, 1)$ determines the degree of diminishing returns in research, and χ_j is a general research productivity term specific to each technology.

To describe the network of knowledge spillovers in this economy, consider the matrix of spillover elasticities:

Definition 1 (Spillover Network $\boldsymbol{\varphi}$)**.** *The spillover network is a $J \times J$ matrix with elements*

$$\varphi_{ijt} \equiv \frac{\partial \ln(\phi_{it})}{\partial \ln(A_{jt-1})}. \tag{11}$$

The spillover network $\boldsymbol{\varphi}_t$ describes the response of spillovers to changes in the knowledge stock of each technology. I will assume that $\boldsymbol{\varphi}_t$ is weakly positive off of its diagonal: $\varphi_{ijt}|_{i\neq j} \geq 0$. This guarantees that technologies receive higher spillovers when they are relatively less advanced.[15] By assuming that spillover functions are homogeneous of degree zero, i.e. $\sum_j \varphi_{ijt} = 0$, I ensure balanced growth in the presence of a fixed supply of scientists.[16] It will occasionally be helpful to refer to the *gross spillover network* $\tilde{\boldsymbol{\varphi}}_t$,

[15]To see this, note that because ϕ_{it} is homogeneous of degree zero, $\varphi_{ijt}|_{i\neq j} \geq 0$ implies $\frac{\partial \phi_i(\{A_{qt-1}/A_{it-1}\})}{\partial (A_{jt-1}/A_{it-1})}|_{i\neq j} \geq 0$. That is, technology i's spillovers increase as the other technologies become relatively more advanced. This implies $\varphi_{iit} \leq 0$.

[16]Allowing for $\sum_j \varphi_{ijt} < 0$ would be consistent with the fishing out feature of semi-endogenous growth models (Jones, 1995, 2022). In that case, balanced growth would require an exponential increase in the supply of scientists. The possibility that $\varphi_{iit} < 0$ introduces some notion of fishing out because it implies that additional innovations are more difficult for more advanced technologies (Bloom et al., 2020). However, if all technologies advance concurrently, then reductions in research productivity can be avoided, and balanced growth can still be achieved with a fixed supply of scientists.

which is defined in relation to the (net) spillover network according to $\boldsymbol{\varphi}_t = \tilde{\boldsymbol{\varphi}}_t - \mathbf{I}$. Thus, the rows of the gross spillover network describe an array of spillovers that sum to one, rather than zero.

The spillover network is a central factor in determining both the marginal cost and marginal benefit of innovation. The rows of the spillover network determine the marginal cost of innovation as they specify how the research productivity of each technology depends the overall state of technology in the economy. This is particularly important for transition dynamics as less advanced technologies can receive a catchup-inducing boost to their research productivity. However, the rows of the spillover network are *not* relevant for policy because researchers fully internalize the productivity benefits they receive from spillovers.

On the other hand, the columns of the spillover matrix determine the marginal benefit of innovation as they specify how the efforts of future researchers will be enhanced by research that takes place today. Therefore, the columns of the spillover matrix are the focus of policy because the positive spillovers they mediate are not internalized by today's researchers. However, the columns of the spillover network are *not* relevant for laissez-faire transition dynamics because the external nature of spillovers implies they do not influence equilibrium behavior.

As an example, an economy with no cross-technology knowledge spillovers would have a spillover network of all zeros: $\boldsymbol{\varphi}_t = \mathbf{0}$. In that case, spillovers ϕ_{it} would reduce to a positive constant for each technology, so the spillover network would cease to play any role. Note that there would still be spillovers *within* technologies as Equation (8) states that, for a given technology, innovation builds multiplicatively on the existing technology stock. The absence of cross-technology knowledge spillovers is the implicit assumption in much of the climate innovation literature (e.g. Acemoglu et al. (2012)), but as argued throughout this paper, cross-technology knowledge spillovers play a critical role in shaping both transition dynamics and optimal policy.

Finally, there is a fixed endowment of scientists supplied inelastically, so the allocation of scientists must satisfy

$$\sum_j s_{jt} \leq \mathcal{S}, \tag{12}$$

where $\mathcal{S}$ is the aggregate supply of scientists. The fact that the aggregate supply of scientists is fixed implies that the Planner's concern is with the *composition* of scientific effort, not the total level. The starting point of most innovation policies is that innovators do not internalize the full social marginal benefit of their efforts (Arrow, 1962; Bryan and Williams, 2021). In a setting where scientists have an outside option, e.g. production work or leisure, equilibrium scientific effort is likely to be below the optimum. This has been the traditional justification of innovation subsidies (Romer, 1990; Bloom et al., 2019). Instead, when the supply of scientists is fixed and there are multiple technologies, the fact that differences in spillover creation across technologies may not track private profit incentives implies that the equilibrium composition of scientific effort may not be optimal. Consequently, innovation subsidies in this context close the technology-specific wedge between spillover creation and profit.

2.3 Preferences & Policy Instruments

The Planner seeks to maximize the utility of a representative household with two main policy instruments: a carbon price, denoted by τ_t, and technology-specific innovation subsidies, denoted by ξ_{jt}. With these instruments (and two others discussed below), the Planner can implement the first-best as a competitive equilibrium by correcting the economy's externalities: carbon pollution and knowledge spillovers.

The social objective is household utility

$$\sum_{t\geq 0} \frac{1}{(1+\rho)^t} u(c_t), \tag{13}$$

where $u(.)$ is a concave function of consumption, and ρ is the rate of pure time preference. I will refer to the rate of pure time preference ρ as the (social) discount rate.[17]

As discussed below, producers of intermediates have market power, so I will allow the Planner to subsidize intermediates to correct the monopolist markup. This allows the Planner to correct every market failure and achieve the first-best, but given that market power is not the focus of this paper, this policy will play little role in the analysis. For the sake of simplicity, I will assume that all subsidies on intermediates are constant, which would be the case in both laissez-faire and the first-best. Denote by Υ the subsidy on intermediates. Finally, the Planner may levy a lump-sum tax D_t to fund(rebate) any deficit(surplus) from the use of corrective instruments.

2.4 Equilibrium

I make the typical assumption in endogenous growth models that innovation for a given intermediate grants exclusive ownership of the most advanced form of production for that intermediate. This exclusive ownership creates a rent which drives the incentive to innovate (Romer, 1990; Aghion and Howitt, 1992).

The final good is the numeraire in each period. The producers of intermediates are the only firms in the economy with market power, so prices of all other goods are set competitively. The final good producer solves

$$\max_{\{Y_{jt}\}} \mathcal{Y}_t - \sum_j p_{jt} Y_{jt}, \tag{14}$$

where p_{jt} is the price of the technology j good. This yields the condition

$$\Omega_t \frac{\partial F_t}{\partial Y_{jt}} = p_{jt}. \tag{15}$$

The producers of technology-specific goods solve

$$\max_{\Lambda_{jt}, \{y_{j\iota t}\}} p_{jt} Y_{jt} - (r_j + \omega_j \tau_t)\Lambda_{jt} - \int_0^1 p_{j\iota t} y_{j\iota t} d\iota \tag{16}$$

where $p_{j\iota t}$ is the price of intermediate ι for technology j. This yields the conditions

$$\alpha p_{jt} \frac{Y_{jt}}{\Lambda_{jt}} = r_j + \omega_j \tau_t \tag{17}$$

$$(1-\alpha) p_{jt} \frac{Y_{jt}}{y_{j\iota t}} = p_{j\iota t}. \tag{18}$$

[17] A distinction can be drawn between the household's (private) discount rate and the Planner's (social) discount rate, but this difference only matters in contexts where the household is making an intertemporal decision, e.g. savings, that the Planner can influence with policy (e.g. Farhi and Werning (2007)). Because the choice to innovate depends on a one-period monopoly and I have abstracted from capital accumulation, none of the equilibrium outcomes of this economy depend on the household's private discount rate. For this reason, I can ignore the distinction between private and social discount rates and set ρ in accordance with the social discount rate.

One can see from Equation (17) that carbon pricing reduces demand for dirty inputs by making those dirty inputs more expensive. This, in turn, reduces the demand for all goods and intermediates associated with dirty production as well.

Given their market power, an intermediate producer internalizes their demand curve from Equation (18), so they solve

$$\max_{p_{j\iota t}, y_{j\iota t}} \Upsilon p_{j\iota t} y_{j\iota t} - \frac{w_{\ell t}}{a_{j\iota t}} y_{j\iota t} \quad s.t. \tag{19}$$
$$(1-\alpha) p_{jt} \frac{Y_{jt}}{y_{j\iota t}} = p_{j\iota t},$$

where $w_{\ell t}$ is the wage paid to production workers. The optimal markup is infinite with a Cobb-Douglas demand system, so to pin down the markup, I will assume intermediate producers limit price one step down the productivity ladder to the competitive fringe. This yields

$$p_{j\iota t} = \frac{\gamma}{\Upsilon} \frac{w_{\ell t}}{a_{j\iota t}}. \tag{20}$$

That is, intermediate prices equal the subsidy-inclusive break-even price of the competitive fringe. This allows intermediate producers to achieve profit

$$\pi_{j\iota t} = \Upsilon \frac{\gamma - 1}{\gamma} (1-\alpha) p_{jt} Y_{jt}. \tag{21}$$

We can see from Equation (21) that relative profits depend on relative prices of technology-specific goods, so I show in Appendix A.1 that relative prices follow

$$\frac{p_{jt}}{p_{Jt}} = \left(\frac{r_j + \omega_j \tau_t}{r_J + \omega_J \tau_t}\right)^{\alpha} \left(\frac{A_{jt}}{A_{Jt}}\right)^{\alpha - 1}. \tag{22}$$

That is, relative prices are increasing in relative input costs – inclusive of the carbon price – and decreasing in relative technology.

Turning to the innovation side of the economy, I assume there is a competitive research firm that produces innovations for every technology. The purchase of an innovation in technology j allows its owner to be the productivity leader of a random intermediate for one period, so by no-arbitrage, the equilibrium price for such an innovation is equal to expected profit: $\Pi_{jt} = \int_0^1 \pi_{j\iota t} d\iota$.

The research firm solves

$$\max_{\{s_{jt}\}} \sum_j \xi_{jt} z_{jt} \Pi_{jt} - w_{st} \sum_j s_{jt} \tag{23}$$

where w_{st} is the wage paid to scientists. That is, the research firm produces mass z_{jt} of innovations for technology j and sells them at price Π_{jt}. Considering the no-arbitrage condition, we have that the price of innovation for technology j is proportional to market size:

$$\Pi_{jt} = \Upsilon \frac{\gamma - 1}{\gamma} (1-\alpha) p_{jt} Y_{jt} \propto S_{jt} \mathcal{Y}_t, \tag{24}$$

where S_{jt} is the income share of technology j. That is, the rent one can achieve by innovating in technology j scales in accordance with total spending on the technology j good. This is a standard result in the directed technological change literature (Acemoglu, 2002; Acemoglu et al., 2012).

The research optimality condition yields

$$\left(\frac{s_{jt}}{s_{Jt}}\right)^{1-\eta} = \left(\frac{\chi_j}{\chi_J}\right)\left(\frac{\xi_{jt}}{\xi_{Jt}}\right)\left(\frac{\phi_{jt}}{\phi_{Jt}}\right)\left(\frac{\Pi_{jt}}{\Pi_{Jt}}\right), \tag{25}$$

which, together with the fixed supply of scientists (12), pins down the innovation equilibrium. Thus, the allocation of scientists, and therefore the direction of innovation, is determined by three forces: subsidies, spillovers, and market size. First, innovation subsidies have a natural influence; as policy favors one technology with higher subsidies, more scientists are devoted to that technology. Second, spillovers influence the allocation of scientists by affecting the productivity of research. As discussed in Section 2.2, technologies that are relatively less advanced receive a boost in research productivity.

Third, technologies with larger markets, i.e. greater income shares, have larger innovation rents, and therefore, stronger incentives to innovate. This is the channel through which the substitutability of technology-specific goods influences the direction of innovation. If technologies are substitutes(complements), then the less advanced technology will obtain a smaller(larger) innovation rent as it accounts for a smaller(larger) share of final output. Market size is also the channel through which carbon prices influence the direction of innovation. In the case of substitutes(complements), a price on carbon reduces(increases) relative market size for dirty forms of production as it increases dirty input costs.

Finally, the representative household owns all of the factors and firms in the economy and consumes their income in each period. Both the household and government must satisfy their budget constraints:

$$c_t = w_{\ell t} L + w_{st}\mathcal{S} + \Pi_t - D_t \tag{26}$$

$$D_t + \tau_t \mathcal{E}_t = \sum_j (\xi_{jt} - 1) z_{jt} \Pi_{jt} + \sum_j \int_0^1 (\Upsilon - 1) p_{j\iota t} y_{j\iota t} d\iota, \tag{27}$$

where Π_t is a dividend equal to all of the profits in the economy. Because the household owns all of the factors and firms in the economy, and net government expenditure is financed lump-sum, equilibrium consumption is equal to final output net of input costs.

Given these conditions, the competitive equilibrium is defined as follows:

Definition 2 (Equilibrium)**.** *Given an initial condition for technology $\{A_{j,-1}\}$ and an array of input costs $\{r_j\}$, an equilibrium consists of a sequence of carbon prices $\{\tau_t\}$, innovation subsidies $\{\xi_{jt}\}$, intermediate subsidies Υ, lump-sum taxes $\{D_t\}$, final output $\{\mathcal{Y}_t\}$, technology-specific goods $\{Y_{jt}\}$, technology-specific good prices $\{p_{jt}\}$, inputs $\{\Lambda_{jt}\}$, intermediates $\{y_{j\iota t}\}$, intermediate prices $\{p_{j\iota t}\}$, labor $\{\ell_{j\iota t}\}$, production wages $\{w_{\ell t}\}$, innovation $\{z_{jt}\}$, innovation rents $\{\Pi_{jt}\}$, scientists $\{s_{jt}\}$, scientist wages $\{w_{st}\}$, technology $\{A_{jt}\}$, emissions $\{\mathcal{E}_t\}$, and household consumption $\{c_t\}$ such that:*

(i) Prices $\{\{p_{jt}, \{p_{j\iota t}\}\}, w_{\ell t}\}$ and quantities $\{\mathcal{Y}_t, \{Y_{jt}, \Lambda_{jt}, \{y_{j\iota t}, \ell_{j\iota t}\}\}\}$ on the production side of the economy solve the profit maximization problems (14), (16), and (19);

(ii) The labor market clears (6);

(iii) Prices $\{\{\Pi_{jt}\}, w_{st}\}$ and quantities $\{z_{jt}, s_{jt}\}$ on the innovation side of the economy satisfy the no-arbitrage condition (24) and solve the profit maximization problem (23);

(iv) The scientist market clears (12);

(v) Technology $\{A_{jt}\}$ evolves according to (8);

(vi) Emissions $\{\mathcal{E}_t\}$ follow (3);

(vii) The resource constraint (7) and budget constraints (26) and (27) hold.

Given this paper's focus on innovation, one can think about the equilibrium evolution of technology as a dynamic process with technology $\{A_{jt}\}$ as the state and scientists $\{s_{jt}\}$ as the control. In each period, the inherited state $\{A_{jt-1}\}$ pins down the control $\{s_{jt}\}$ via a spillover effect and a market size effect. The state then updates to $\{A_{jt}\}$, and the process starts over in the next period.

2.5 Example Production & Spillover Structure

The production and spillover structure I have described thus far involves minimal assumptions, so I will now sketch an example economy where I take a stance on the functional forms of production and spillovers. I will assume this structure when I calibrate and simulate my model in Sections 5 and 6, but the theoretical results of Sections 3 and 4 apply to the more general setup described above.

Production is divided into Θ-many sectors, each of which has a clean and dirty form of production. Final output follows

$$\mathcal{Y}_t = \Omega_t \Big(\sum_\theta \nu_\theta^{\frac{1}{\lambda}} E_{\theta t}^{\frac{\lambda-1}{\lambda}} \Big)^{\frac{\lambda}{\lambda-1}}, \tag{28}$$

where $E_{\theta t}$ is the output of sector θ. That is, final output is a complements CES over sectors with elasticity of substitution $\lambda < 1$ and sector shares ν_θ.

Output at the sector level follows

$$E_{\theta t} = \Big(Y_{\theta ct}^{\frac{\sigma-1}{\sigma}} + Y_{\theta dt}^{\frac{\sigma-1}{\sigma}} \Big)^{\frac{\sigma}{\sigma-1}}, \tag{29}$$

where $Y_{\theta c}$ and $Y_{\theta d}$ are the clean and dirty forms of production for sector θ, respectively. That is, sectoral output is a substitutes CES over clean and dirty production with elasticity of substitution $\sigma > 1$. This production structure nests the setup of Acemoglu et al. (2012) when there is a single sector.

I will assume the final sector of the economy Θ only has a clean form of production. My application considers the transportation and electricity generation sectors, but because these sectors make up only a small portion of the economy, the final sector, which I will refer to as the "general" sector, closes the model while abstracting from emissions that originate outside of transportation or electricity generation.[18]

On the innovation side of the economy, I will assume that research productivity follows $\chi_j = \chi \nu_{\theta(j)}^{-\eta}$, where $\theta(j)$ is the sector associated with technology j. Thus, the production of innovation follows

$$z_{jt} = \chi \Big(\frac{s_{jt}}{\nu_{\theta(j)}} \Big)^\eta \phi_{jt}. \tag{30}$$

Following Acemoglu and Restrepo (2022b), I show in Appendix A.2 that one can interpret the sectoral CES shares ν_θ as the share of tasks in the economy performed by sector θ. Thus, the relevant notion of effective scientific effort in this case is the number of scientists per economic task associated with a technology.

[18] Motor vehicle output and electricity generation accounted for 2.6% and 1.8% of US GDP in 2021, respectively.

The spillover functions follow

$$\phi_{it} = \frac{\prod_j A_{jt-1}^{\tilde{\varphi}_{ij}}}{A_{it-1}}, \tag{31}$$

which implies that the spillover network is constant. The numerator represents an idiosyncratic Cobb-Douglas aggregator of all knowledge stocks in the economy, and the denominator represents the recipient's own knowledge stock. The Cobb-Douglas shares define the gross spillover network $\tilde{\boldsymbol{\varphi}}$, so as discussed above, the (net) spillover network follows mechanically from $\boldsymbol{\varphi} = \tilde{\boldsymbol{\varphi}} - \mathbf{I}$. For example, the case of no cross-technology knowledge spillovers corresponds to a gross spillover network equal to the identity matrix: $\tilde{\boldsymbol{\varphi}} = \mathbf{I}$. This specification for spillovers was studied in the recent work by Liu and Ma (2021).

Finally, household utility follows the common CRRA specification

$$u(c_t) = \frac{c_t^{1-\vartheta} - 1}{1 - \vartheta}, \tag{32}$$

where $1/\vartheta$ is the intertemporal elasticity of substitution.[19]

3 Policy's Impact on the Direction of Innovation

This section characterizes the dynamic response of the direction of innovation to a policy reform. Specifically, I will imagine there is a once-and-for-all policy reform $\{d\tau, \{d\xi_j\}\}$.[20] I first characterize how technology's steady-state changes in response to the policy reform using what I call the *amplification matrix*. Next, I characterize the transition path of technology towards – or away from – the new steady-state using what I call the *transition matrix*.[21]

Both my amplification matrix and transition matrix are composed of the same two sufficient statistic matrices: one describing substitution patterns in production and another describing the network of knowledge spillovers. These two matrices show how the two forces, substitutability and spillovers, determine both the degree to which the policy reform will change technology's steady-state and the speed at which technology will converge to its new steady-state. These matrices are sufficient statistics in the sense that, conditional on estimating each matrix, one does not need to know the underlying structure of production and spillovers to first-order characterize the impact of policy on the direction of innovation.

Taking an eigendecomposition of the transition matrix, I quantify the degree of increasing returns to innovation in terms of the spectral radius of the transition matrix. I show that increasing returns determine both the long-run impact of policy and the speed of technology's transition, but that the two are inversely related. Finally, the spectral radius of the transition matrix provides me with a condition for technological path dependence, where the path of technology depends discontinuously on its initial condition. I conclude the section with a simple, two-technology example that illustrates the arguments made throughout the section.

[19] This specification implies a consumption discount rate of $\rho + \vartheta g_{ct}$ via the Ramsey equation, where g_{ct} is the growth rate of consumption.

[20] As discussed in Section 4, time-invariant policy is not necessarily optimal. However, I focus on time-invariant policy in this section to analytically characterize how the direction of innovation would respond to a given policy regime, given unlimited time.

[21] My analysis does not require technology to be in a previous steady-state at the time of the policy reform.

3.1 Sufficient Statistic Matrices

I start by defining the two sufficient statistic matrices that will be relevant throughout this section. These matrices summarize the two main forces of my model: (i) substitution patterns in production and (ii) the network of cross-technology knowledge spillovers.

First, consider the substitution matrix:

Definition 3 (Substitution Matrix $\mathbf{\Sigma}$)**.** *The substitution matrix is a $J-1 \times J-1$ matrix with elements*

$$\Sigma_{ijt} \equiv \frac{\partial \ln \left(Y_{it}/Y_{Jt}\right)}{\partial \ln \left(p_{Jt}/p_{jt}\right)}. \tag{33}$$

The substitution matrix $\mathbf{\Sigma}_t$ describes how relative demand responds to changes in relative prices. As the relative price of good j changes, this induces a change in relative equilibrium demand for good i, as described by Σ_{ijt}. For both relative prices and quantities, good J is the base good.

The elements of the substitution matrix $\mathbf{\Sigma}_t$ are similar to elasticities of substitution. Indeed, the diagonal elements of $\mathbf{\Sigma}_t$ are elasticities of substitution between goods i and J, but for off-diagonal elements, we are allowing the change in relative price and change in relative quantity to refer to different goods. We then have the usual interpretation that diagonal elements above one describe substitutes, while diagonal elements below one describe complements. As an illustration, if the output aggregator were a standard CES with elasticity of substitution σ, we would have $\mathbf{\Sigma}_t = \sigma \mathbf{I}$. Therefore, the demand responses described by the substitution matrix $\mathbf{\Sigma}_t$ provide a general characterization of substitution patterns in production.

In the case of the nested-CES production structure described in Section 2.5, I show in Appendix A.3 that $\mathbf{\Sigma}_t$ is a block-diagonal matrix that depends on the elasticities of substitution of both nests, as well as income shares within sectors. When I calibrate and simulate my model in Sections 5 and 6, I will use this closed-form solution for the substitution matrix to estimate $\mathbf{\Sigma}_t$ using existing estimates of the relevant elasticities of substitution.

Second, consider the spillover matrix:

Definition 4 (Spillover Matrix $\mathbf{\Phi}$)**.** *The spillover matrix is a $J-1 \times J-1$ matrix with elements*

$$\Phi_{ijt} \equiv \frac{\partial \ln \left(\phi_{Jt}/\phi_{it}\right)}{\partial \ln \left(A_{jt-1}/A_{Jt-1}\right)}. \tag{34}$$

The spillover matrix $\mathbf{\Phi_t}$ describes how relative spillovers change in response to changes in relative technology stocks. As before, the base technology is J. The fact that spillover functions are homogeneous of degree zero implies that one can map from the spillover network $\boldsymbol{\varphi}_t$ to the spillover matrix $\mathbf{\Phi}_t$ using

$$\Phi_{ijt} = \varphi_{Jjt} - \varphi_{ijt}. \tag{35}$$

Therefore, we can interpret the spillover matrix $\mathbf{\Phi}_t$ as a matrix of relative spillover elasticities.[22] That is, a positive value for Φ_{ijt} implies that the elasticity of ϕ_{Jt} with respect to technology j is greater than

[22]To see this, note that if ϕ_{it} is homogeneous of degree zero, then its derivatives are homogeneous of degree negative one. Thus, we have $\frac{\partial \phi_{it}(\{A_{jt-1}/A_{Jt-1}\})}{\partial (A_{jt-1}/A_{Jt-1})} \frac{A_{jt-1}/A_{Jt-1}}{\phi_{it}(\{A_{jt-1}/A_{Jt-1}\})} = \frac{\partial \phi_{it}(\{A_{jt-1}\})}{\partial A_{jt-1}} \frac{A_{jt-1}}{\phi_{it}(\{A_{jt-1}\})}$.

that of ϕ_{it}, and a negative value implies the reverse. Moreover, the mapping of Equation (35) implies that any estimate of the spillover network provides an estimate of the spillover matrix as well.

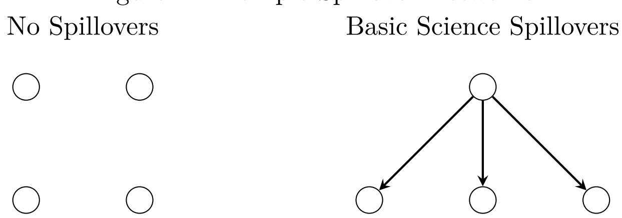
The spillover matrix $\mathbf{\Phi}_t$ provides a general description of an economy's spillovers for a wide class of spillover functions. In particular, the diagonal of $\mathbf{\Phi}_t$ provides a measure of the degree to which technologies receive spillovers from the *other* technologies in the economy. To see this, note that

$$\Phi_{iit} = \varphi_{Jit} - \varphi_{iit} = \sum_{j \neq i} \left(\varphi_{ijt} - \varphi_{Jjt}\right), \tag{36}$$

so the diagonal elements of $\mathbf{\Phi}_t$ will be large when technologies receive high spillovers from their peers, relative to the spillovers received by the base technology. When technologies receive minimal spillovers via the spillover network, i.e. $\sum_{j \neq i} \varphi_{ijt}$ approaches 0, less advanced technologies will experience only a minor enhancement in their research productivity. Therefore, the spillover matrix $\mathbf{\Phi}_t$ is informative of the extent to which cross-technology knowledge spillovers will stimulate catch-up growth.

For illustration, consider the two spillover networks of Figure 1.

Figure 1: Example Spillover Networks



In the first network, the "No Spillovers" network, there are no cross-technology knowledge spillovers, so $\boldsymbol{\varphi}_t = \mathbf{0}$. In this case, the spillover matrix is also zero $\mathbf{\Phi}_t = \mathbf{0}$ to reflect the irrelevance of cross-technology spillovers in this economy. In the second network, the "Basic Science Spillovers" network, there is one technology that creates all of the spillovers in the economy. Assuming these spillovers are symmetric, the spillover matrix follows $\mathbf{\Phi}_t \propto \mathbf{I}$, reflecting the predominance of cross-technology spillovers in this economy.[23]

3.2 Steady-State Impact of Policy

I will now characterize how the balanced growth steady-state of this economy changes in response to a policy reform.

Definition 5 (Balanced Growth Steady-State). *A balanced growth steady-state is an equilibrium in which every technology experiences uniform growth at rate g.*

I will describe the steady-state of technology in terms of relative technology, rather than technology levels, because relative technology is constant along a balanced growth path. Denote technology j's relative technology stock by

$$\bar{A}_{jt} \equiv \frac{A_{jt}}{A_{Jt}}, \tag{37}$$

[23]In these two stylized cases, the value of $\mathbf{\Phi}_t$ is independent of the choice of base technology, but this is not true in general.

where the base technology is J. Therefore, we can characterize the steady-state of this economy as a fixed point in a dynamic process for relative technology.

The following proposition characterizes the steady-state impact of a policy reform in terms of an amplification matrix.

Proposition 1. *A policy reform* $\{d\tau, \{d\xi_j\}\}$ *induces first-order changes in the steady-state level of relative technology according to*

$$d\ln(\bar{A}_{ss}) = \eta\boldsymbol{\mathcal{M}}\big[d\ln(\Xi) - \alpha\big(\boldsymbol{\Sigma} - \boldsymbol{I}\big)d\ln(\mathcal{R})\big], \tag{38}$$

where $d\ln(\Xi) \equiv d\ln\big(\xi_j/\xi_J\big)$ *and* $d\ln(\mathcal{R}) \equiv d\ln\big((r_j + \omega_j\tau)/(r_J + \omega_J\tau_t)\big)$ *are the changes to relative innovation subsidies and relative input costs induced by the policy reform, respectively. The amplification matrix* $\boldsymbol{\mathcal{M}}$*, evaluated at the steady-state, follows*

$$\boldsymbol{\mathcal{M}} = \big[\boldsymbol{\Phi} - \eta(1-\alpha)\big(\boldsymbol{\Sigma} - \boldsymbol{I}\big)\big]^{-1}. \tag{39}$$

The proof of Proposition 1 can be found in Appendix A.4. Proposition 1 defines the amplification matrix $\boldsymbol{\mathcal{M}}$, which describes the role of substitutability and spillovers in shaping the steady-state impact of a policy reform. I will assume that the amplification matrix $\boldsymbol{\mathcal{M}}$ does not flip the sign of policy impacts, which I show to be the empirically relevant case in Section 6.1. Furthermore, I provide an economic interpretation for the case when the amplification matrix *does* flip the sign of policy impacts in Section 3.4.

First, increases in relative subsidies Ξ increase the prevalence of favored technologies in steady-state. Next, carbon pricing reduces the prevalence of dirty technologies in steady-state, insofar as goods are substitutes, by increasing input costs for dirty technologies. Both of these effects are mediated by substitutability and spillovers via the amplification matrix $\boldsymbol{\mathcal{M}}$. When goods are substitutes, technologies that are favored by policy are able to increase their market size, further increasing the equilibrium incentive to innovate and amplifying the effect of policy. The opposite is true when goods are complements because technologies favored by policy have their market size reduced. This channel shows up through the substitution matrix net of the identity matrix ($\boldsymbol{\Sigma} - \mathbf{I}$) because one is the boundary point between complements and substitutes. Conversely, cross-technology knowledge spillovers dampen the impact of policy by increasing the research productivity of technologies disfavored by policy. If some technologies become relatively advanced due to policy, they will lose part of their lead as their peers experience catchup growth. Intuitively, the substitution matrix increases the "size" of the amplification matrix as it enters with a negative in the "denominator", while the spillover matrix decreases the "size" of the amplification matrix as it enters with a positive in the "denominator".

3.3 Technology's Transition Path

Having examined how policy affects technology in the economy's long-run steady-state, I can now describe the transition path technology takes to its new steady-state. To determine the forces that are of first-order importance in governing this transition, I will linearize the equilibrium conditions around the steady-state to derive a transition matrix.

I will describe the transition path of technology in terms of the log deviation of relative technology from steady-state

$$\bar{\mathcal{A}}_t \equiv \ln(\bar{A}_t) - \ln(\bar{A}_{ss}). \tag{40}$$

That is, relative technology converges to its steady-state as $\bar{\mathcal{A}}_t$ converges to zero.

The following proposition provides a first-order characterization of the transition path of technology in terms of a transition matrix.

Proposition 2. *To a first-order, the transition path of technology follows the linear process*

$$\bar{\mathcal{A}}_t \approx \boldsymbol{\mathcal{J}} \bar{\mathcal{A}}_{t-1}, \tag{41}$$

where the transition matrix $\boldsymbol{\mathcal{J}}$, evaluated at the steady-state, follows

$$\boldsymbol{\mathcal{J}} = \big[(1-\eta)\boldsymbol{I} - g\eta(1-\alpha)\big(\boldsymbol{\Sigma} - \boldsymbol{I}\big)\big]^{-1}\big[(1-\eta)\boldsymbol{I} - g\boldsymbol{\Phi}\big]. \tag{42}$$

The proof of Proposition 2 can be found in Appendix A.6. The main object of interest in Proposition 2 is the transition matrix $\boldsymbol{\mathcal{J}}$. This matrix shows that the transition path of technology is also governed by the two forces emphasized throughout this paper: Substitution patterns in production and the network of knowledge spillovers. Recall from Equation (25) in Section 2.4 that the allocation of scientists is determined by subsidies, spillovers, and market size. Subsidies determine the location of the steady-state, but we can now see that they do not directly determine technology's transition matrix. Instead, it is the latter two forces that govern the transition matrix. Spillovers influence the transition matrix by enhancing the research productivity of less advanced technologies. As stated above, the diagonal of the spillover matrix $\boldsymbol{\Phi}$ provides a measure of the degree to which less advanced technologies can use spillovers to achieve catchup growth.

Substitutability influences the transition matrix through a market size effect that reduces the profits of technologies that are less advanced than their substitutes. Conversely, when technologies are less advanced than their complements, they achieve higher profits. These effects are summarized by the substitution matrix net of the identity matrix $(\mathbf{\Sigma} - \mathbf{I})$. As before, the identity matrix is subtracted from the substitution matrix because one is the boundary point between complements and substitutes. Furthermore, substitutability shows up through an inverse matrix because it follows the logic of a Leontief inverse. A small change in the inherited technological state $\{\bar{A}_{jt-1}\}$ induces a change in the realized technological state $\{\bar{A}_{jt}\}$. The realized technological state is what governs market size, so the change in market sizes induces yet another change in the realized technological state, and so on. This sequence of ripple effects generates a geometric series that culminates in an inverse matrix which describes the cumulative impact of the market size effect.[24]

[24]Formally, we have that

$$\Big[\mathbf{I} - \frac{g\eta(1-\alpha)}{1-\eta}(\mathbf{\Sigma}-\mathbf{I})\Big]^{-1} = \sum_{n\geq 0}\Big(\frac{g\eta(1-\alpha)}{1-\eta}(\mathbf{\Sigma}-\mathbf{I})\Big)^n = \mathbf{I} + \Big(\frac{g\eta(1-\alpha)}{1-\eta}(\mathbf{\Sigma}-\mathbf{I})\Big) + \Big(\frac{g\eta(1-\alpha)}{1-\eta}(\mathbf{\Sigma}-\mathbf{I})\Big)^2 + \dots, \tag{43}$$

which is proportional to the inverse matrix contained in $\boldsymbol{\mathcal{J}}$.

3.4 Degree of Increasing Returns to Innovation

In this section, I eigendecompose the transition matrix and show that the degree of increasing returns to innovation can be quantified in terms of the transition matrix's spectral radius, which governs whether technology converges to its new steady-state quickly, slowly, or in an extreme case, not at all. Furthermore, I show that the amplification matrix can be written in terms of the eigenvalues of the transition matrix, with the implication that the steady-state impact of policy and the speed of technology's transition are inversely related.

The speed of convergence quantifies the degree of increasing returns to innovation. When increasing returns are high, convergence will be slow as technologies that are more advanced in the initial condition than in the steady-state will only slowly lose their advantage. Conversely, when increasing returns are low, convergence will be fast as any advantage enjoyed in the initial condition, relative to the steady-state, will quickly disappear. In what follows, I derive the speed of convergence from the transition matrix using spectral analysis.

First, use Equation (41) of Proposition 2 to iterate forward from technology's initial log deviation from steady-state

$$\bar{\mathcal{A}}_t \approx \boldsymbol{\mathcal{J}}^t \bar{\mathcal{A}}_0. \tag{44}$$

Intuitively, we can see that the speed of transition depends on how quickly the transition matrix $\boldsymbol{\mathcal{J}}$ shrinks technology's log deviation from steady-state $\bar{\mathcal{A}}_t$. If $\boldsymbol{\mathcal{J}}$ were a matrix of zeros, technology would converge immediately, whereas if $\boldsymbol{\mathcal{J}}$ were the identity matrix, technology would remain in its initial state forever. Thus, the speed at which less advanced technologies can catch up to their peers, and therefore the degree of increasing returns to innovation, depends on the "size" of the transition matrix. As we will see, the proper notion of a matrix's "size" is in terms of the magnitude of its eigenvalues, i.e. its spectral radius. I will focus on the case where the transition matrix $\boldsymbol{\mathcal{J}}$ has $J-1$ distinct real eigenvalues $\{\kappa_j\}$. I verify numerically that this is the empirically relevant case under the calibration described in Section 5. Consider the eigendecomposition of the transition matrix

$$\boldsymbol{\mathcal{J}} = \boldsymbol{Q}\boldsymbol{D}(\kappa)\boldsymbol{Q}^{-1}, \tag{45}$$

where $\boldsymbol{D}(\kappa)$ is a diagonal matrix whose diagonal elements are the eigenvalues of the transition matrix, and $\boldsymbol{Q}$ is an invertible matrix whose columns are the eigenvectors of the transition matrix. This eigendecomposition follows from the fact that any real square matrix has a Jordan decomposition (Galor, 2007).[25] Plugging the eigendecomposition into Equation (44), we have

$$\bar{\mathcal{A}}_t \approx \boldsymbol{Q}\boldsymbol{D}(\kappa)^t \boldsymbol{Q}^{-1} \bar{\mathcal{A}}_0. \tag{46}$$

From this, we can see that the speed of convergence is governed by the eigenvalues $\{\kappa_j\}$ of the transition matrix. As time progresses, technology's transition path is driven by the geometric decay of the eigenvalues. Eigenvalues slightly above zero will rapidly shrink to zero, signifying fast convergence, while eigenvalues slightly below one will gradually shrink to zero, signifying slow convergence.

[25] The basic arguments of this section hold for the more general case of repeated or complex eigenvalues because the main properties of Equation (45) hold for the Jordan decomposition of any real square matrix: $\boldsymbol{\mathcal{J}} = \boldsymbol{Q}\boldsymbol{D}\boldsymbol{Q}^{-1}$, where $\boldsymbol{D}$ is the Jordan normal form of $\boldsymbol{\mathcal{J}}$, and $\boldsymbol{Q}$ is some invertible matrix. Therefore, the convergence properties of technology always depend on the convergence properties of the Jordan normal form of $\boldsymbol{\mathcal{J}}$. See Lemma 2.8 of Galor (2007).

The eigenvalues of the transition matrix determine the speed of convergence, but which eigenvalues play the most important role quantitatively? To answer this question, consider the linear combination of eigenvectors which recovers the initial state of technology

$$\beta = \boldsymbol{Q}^{-1}\bar{\mathcal{A}}_0. \tag{47}$$

That is, β is a linear projection of the initial state $\bar{\mathcal{A}}_0$ onto the eigenvectors of the transition matrix. In other words, one can regress $\bar{\mathcal{A}}_0$ on $\boldsymbol{Q}$ to obtain β.[26] I will refer to the eigenvectors of the transition matrix as *eigenstates*. These eigenstates represent specific states of technology that correspond to the eigenvectors of the transition matrix.[27] Since $\boldsymbol{Q}$ forms a basis, any state of technology can be expressed as a linear combination of these eigenstates. Moreover, each eigenstate has a convergence rate determined by its corresponding eigenvalue. For instance, if the initial state of technology is proportional to the jth eigenstate, then β will load exclusively on the jth dimension, and technology will geometrically converge to the steady state with a decay rate of κ_j. More generally, the loading of the initial state on the eigenstates determines the speed of convergence as a weighted average of the eigenvalues. The following proposition provides a formal statement of how the spectral properties of the transition matrix determine technology's transition.

Proposition 3. *Suppose the transition matrix $\boldsymbol{\mathcal{J}}$ has $J-1$ distinct real eigenvalues $\{\kappa_j\}$. Then the transition path of technology follows*

$$\bar{\mathcal{A}}_t \approx \sum_j \kappa_j^t \beta_j Q_j, \tag{48}$$

where Q_j is the jth eigenstate of the transition matrix, and β is the linear combination of the eigenstates that recovers the initial state of technology $\beta = \boldsymbol{Q}^{-1}\bar{\mathcal{A}}_0$.

Proposition 3 is derived through matrix manipulation of the eigendecomposition (45) of the transition matrix. It demonstrates that technology's transition depends on three factors: eigenvalues, eigenstates, and the initial state's loading on the eigenstates. The initial state of technology is composed of a linear combination of the eigenstates, and as time progresses, each eigenstate component converges toward the steady-state based on the geometric decay of its respective eigenvalue. Consequently, the overall speed of convergence is determined by each eigenvalue according to the extent to which the initial state loads on its respective eigenstate.

In summary, the speed of convergence depends on a weighted average of the transition matrix's eigenvalues, with the weights determined by the initial state's loading on the eigenstates. The largest eigenvalue will tend to have an outsized influence on the speed of convergence, so the spectral radius provides a summary measure of increasing returns to innovation. Thus, the formula for the transition matrix explains the determinants of increasing returns to innovation. Combining Equations (42) and (45), I obtain the expression

$$\big[(1-\eta)\mathbf{I} - g\eta(1-\alpha)\big(\boldsymbol{\Sigma} - \mathbf{I}\big)\big]^{-1}\big[(1-\eta)\mathbf{I} - g\boldsymbol{\Phi}\big] = \boldsymbol{Q}\boldsymbol{D}(\kappa)\boldsymbol{Q}^{-1}. \tag{49}$$

[26] Using the standard regression formula, we have $\beta = (\boldsymbol{Q}'\boldsymbol{Q})^{-1}\boldsymbol{Q}'\bar{\mathcal{A}}_0$, but since $\boldsymbol{Q}$ forms a basis, this formula simplifies to $\beta = \boldsymbol{Q}^{-1}\bar{\mathcal{A}}_0$. Thus, such a regression would have zero mean squared error.

[27] The "eigenstates" of my paper are analogous to the "eigenshocks" of Kleinman et al. (2023). In both cases, these terms describe particular configurations of state variables that (i) correspond to eigenvectors of a transition matrix and (ii) determine the eigenvalues that are most relevant for convergence speeds.

Therefore, substitution patterns in production and cross-technology knowledge spillovers determine the degree of increasing returns to innovation through their influence on the eigenvalues of the transition matrix. Substitutability slows convergence by reducing profits for less advanced technologies. Conversely, knowledge spillovers accelerate convergence by stimulating catchup growth. Intuitively, the substitution matrix increases the ”size” of the transition matrix as it enters with a negative in the ”denominator”, while the spillover matrix decreases the ”size” of the transition matrix as it enters with a negative in the ”numerator”.

Measuring increasing returns with the spectral radius of the transition matrix allows me to derive a condition for technological path dependence, a central topic in the directed technical change literature (Acemoglu et al., 2012; Aghion et al., 2016, 2019). When the direction of innovation is path dependent, the path of technology depends discontinuously on its initial state. We can see this case play out by considering the eigenvalues of the transition matrix. Technological convergence slows down as the eigenvalues of the transition matrix approach unity, and once one of the eigenvalues goes outside of the unit circle, the steady-state becomes unstable. The following corollary states this point formally.

Corollary 1. *The direction of innovation is path dependent if the transition matrix's spectral radius exceeds one.*

Corollary 1 is an application of a well-known result in dynamical systems establishing the local instability of steady-states (Galor, 2007).[28] It provides a sufficient condition for technological path dependence, which hinges on the spectral radius of the transition matrix. By examining Equation (48), we can see that if the initial state of technology loads on an eigenstate with an eigenvalue exceeding one, then technology will geometrically diverge from the steady-state in the direction of this unstable eigenstate. Consequently, the dynamic path for technology will depend discontinuously on the initial state, as a flip in the sign of the loading on the unstable eigenstate would result in divergence along the same ray but in the opposite direction.

My second corollary establishes the relationship between the amplification matrix and the eigenvalues of the transition matrix.

Corollary 2. *The amplification matrix follows*

$$\boldsymbol{\mathcal{M}} = g\boldsymbol{Q}\boldsymbol{D}(1-\kappa)^{-1}\boldsymbol{Q}^{-1}\big[(1-\eta)\boldsymbol{I} - g\eta(1-\alpha)\big(\boldsymbol{\Sigma} - \boldsymbol{I}\big)\big]^{-1}. \tag{50}$$

The proof of Corollary 2 can be found in Appendix A.7. This result establishes that increasing returns to innovation increase the long-run impact of policy reforms through the term $\boldsymbol{D}(1-\kappa)^{-1}$. This implies that the long-run impact of policy and the speed of technology's transition are inversely related. When increasing returns are high ($\kappa_j \to 1$), the steady-state impact of policy will be large, but the transition will be slow. Conversely, when increasing returns are low ($\kappa_j \to 0$), the steady-state impact of policy will be small, but the transition will be fast. By favoring more advanced technologies, increasing returns allow technologies that are favored by policy to gain larger leads over their peers in the long run. However, the same mechanism leads to sluggish transitions by preventing less advanced technologies from catching up. For instance, high cross-technology knowledge spillovers reduce increasing returns and allow rapid

[28] In particular, see Theorem 4.8 of Galor (2007). The more general statement pertains to the modulus, rather than the absolute value, of the eigenvalues of the transition matrix, allowing for the possibility of complex eigenvalues.

transitions by facilitating catchup growth, but those same spillovers prevent technologies from gaining a significant advantage in the long run.

In describing the policy impacts of Proposition 1, I assumed the amplification matrix does not flip the sign of policy impacts. For example, clean innovation subsidies or carbon pricing reduce the prevalence of dirty technology in the steady-state. By combining Corollaries 1 and 2, we can interpret this assumption in terms of the degree of increasing returns to innovation. A positive-definite amplification matrix makes this assumption more likely to be correct, while a negative-definite amplification matrix is more likely to flip the sign of policy impacts. As we can see from Equation (50), eigenvalues that are much larger than one will tend to make the amplification matrix negative-definite. This is because, in the case of path dependence, the steady-state defines the boundary of the different basins of attraction. This is why the impact of policy can flip when technology is path dependent; policy reforms that favor clean technology now shrink the dirty basins of attraction. For instance, an increase in the carbon price now *increases* the prevalence of dirty technology in the steady-state, implying that fewer initial conditions will land in dirty basins of attraction.

Finally, to measure convergence speeds in units of time, I will consider half-lives of convergence. That is, the number of periods required to halve the log distance from steady-state. Each eigenstate has its own half-life, which follows

$$t_{ES_j}^{(1/2)} = \left\lceil \frac{\ln(1/2)}{\ln(\kappa_j)} \right\rceil, \tag{51}$$

where $\lceil . \rceil$ is the ceiling function. Thus, if the initial state of technology is proportional to the jth eigenstate, the half-life of convergence for each technology will equal $t_{ES_j}^{(1/2)}$. More generally, as stated in Proposition 3, each technology will have its own half-life governed by the initial state's loading on the eigenstates. Half-lives are scale independent, so the overall distance of technology's initial state from its steady-state will not affect this metric of convergence.[29] However, the location of the initial state relative to the steady-state matters in that it influences the loading on the eigenstates β.

3.5 Two Technology Example

To illustrate the arguments of this section, I will now consider a simple example where there are only two technologies. I will make the parametric assumptions of Section 2.5 and imagine that there is a single sector with clean and dirty technology, denoted by c and d. In this case, there is only one relative technology $\bar{A}_t = A_{ct}/A_{dt}$. This setup nests the canonical model of Acemoglu et al. (2012). Throughout this section, I will assume a mild regularity condition to ease exposition.[30]

When there are only two technologies, the various matrices reduce to scalars, and the two sufficient statistic matrices become

$$\Sigma = \sigma \tag{52}$$

$$\Phi = \varphi_{cd} + \varphi_{dc}. \tag{53}$$

That is, Σ is the elasticity of substitution between clean and dirty production, and Φ is the sum of

[29] To be exact, the half-life of convergence for each technology will be the same for initial states $\bar{\mathcal{A}}_0$ and $k\bar{\mathcal{A}}_0$, where $k > 0$.

[30] Specifically, I assume $1 - \eta > g\Phi$ and $1 - \eta > g\eta(1 - \alpha)(\Sigma - 1)$, both of which are easily satisfied under empirically plausible parameter choices (see Section 5). The main reason for the ease with which these conditions are satisfied is that empirically plausible growth rates are on the order of $g = 0.02$, so this shrinks the right-hand side of both inequalities.

cross-technology knowledge spillovers in the economy. Consistent with Proposition 1, the steady-state for relative technology follows

$$\ln(\bar{A}_{ss}) = \eta \frac{\ln(\Xi) - \alpha(\Sigma - 1)\ln(\mathcal{R})}{\Phi - \eta(1-\alpha)(\Sigma - 1)}, \tag{54}$$

where, as before, $\Xi \equiv \xi_c/\xi_d$ is the relative innovation subsidy and $\mathcal{R} \equiv r_c/(r_d + \omega_d \tau)$ is the relative input price (inclusive of the carbon price). From this, we can see that the amplification matrix follows

$$\mathcal{M} = \frac{1}{\Phi - \eta(1-\alpha)(\Sigma - 1)}. \tag{55}$$

As discussed in Section 3.2, the amplification matrix is increasing in the elasticity of substitution Σ and decreasing in the level of cross-technology spillovers Φ. Applying Proposition 2, we have that the transition matrix follows

$$\mathcal{J} = \frac{(1-\eta) - g\Phi}{(1-\eta) - g\eta(1-\alpha)(\Sigma - 1)}. \tag{56}$$

The transition matrix is also increasing in the elasticity of substitution Σ and decreasing in the level of cross-technology spillovers Φ, as argued in Section 3.4. In accordance with Proposition 3, technology's transition path is determined by the geometric decay, or expansion, of $\mathcal{J}^t$, so the speed of convergence increases as the transition matrix shrinks. In particular, the half-life of convergence follows

$$t^{(1/2)} = \left\lceil \frac{\ln(1/2)}{\ln(\mathcal{J})} \right\rceil. \tag{57}$$

To consider the possibility of path dependence, we can apply Corollary 1 and ask if $\mathcal{J}$ is outside the unit circle. This will be the case whenever

$$\Phi < \eta(1-\alpha)(\Sigma - 1). \tag{58}$$

That is, the direction of innovation is path dependent if the substitutability of clean and dirty technology is high relative to the size of cross-technology spillovers. This effect is also modulated by the curvature of research effort η and the importance of labor in production $(1-\alpha)$. As research effort in each period experiences steeper diminishing returns ($\eta \downarrow$), this reduces the incentive to allocate greater research resources to the more advanced technology. As labor becomes less important in production ($(1-\alpha) \downarrow$), this reduces the competitive advantage of the more advanced technology as the factor that benefits from the productivity differential is less important.

Finally, the two forces of substitutability and spillovers have an intuitive influence: if the two technologies are highly substitutable relative to catchup-inducing spillovers, then the less advanced technology will not be profitable enough to attract the innovation effort necessary to overcome their disadvantage. Instead, the more advanced technology will enjoy greater innovation and further solidify its lead. As this process continues, the more advanced technology will strengthen its lead until, in the limit, all production and innovation will be devoted to the more advanced technology. This is the path dependence scenario considered by Acemoglu et al. (2012), so Inequality (58) provides a test for their path dependence hypothesis.

Finally, by applying Corollary 2, we can see that the long-run impact of policy and the speed of technology's transition are inversely related as

$$\mathcal{M} = \frac{g(1-\mathcal{J})^{-1}}{(1-\eta) - g\eta(1-\alpha)(\Sigma - 1)}. \tag{59}$$

This shows that increasing returns to innovation, as measured by $\mathcal{J}$, slow the transition while at the same time increasing the long-run impact of policy. Equation (59) also shows that the steady-state is unstable if and only if the amplification matrix $\mathcal{M}$ is negative. As discussed in Section 3.4, in the case of path dependence, the steady-state defines the boundary of the two basins of attraction. If the economy starts with dirty technology relatively more advanced than the steady-state value (i.e. $\bar{A}_0 < \bar{A}_{ss}$), then innovation will continue to favor dirty technology. Thus, policies that favor clean technology now lower steady-state relative technology $\bar{A}_{ss}$ and shrink the dirty basin of attraction.

4 Optimal Policy

We have seen the response of the direction of innovation to an arbitrary policy regime, so I will now characterize optimal climate innovation policy. I start with a first-best problem where the Planner's instruments are unrestricted. In that case, the two externalities, carbon pollution and knowledge spillovers, are corrected separately, in keeping with the Pigou Principle. In particular, the fact that clean technology can produce goods without pollution does not warrant an additional innovation subsidy when carbon pollution is priced properly. Instead, innovation subsidies reward technologies for their ability to produce knowledge spillovers, and I derive a recursive formula for optimal innovation subsidies that corrects the knowledge spillover externality created via the spillover network. This analysis explains the differing role of knowledge spillovers in describing the impact of policy and prescribing efficient policy.

I then consider a realistic second-best problem where the Planner cannot control the price of carbon pollution. In that case, the Planner must choose innovation subsidies when there is some externally imposed carbon price. I show that the adjustment to optimal innovation subsidies is simple; the same recursive formula holds as in the first-best but with a simple adjustment for the wedge between the external and efficient carbon price. Now clean technologies *do* receive additional innovation subsidies for their ability to produce goods without pollution as the Planner must correct both externalities with a single instrument.

4.1 First-Best

I start with the first-best planning problem. The Planner has a complete set of instruments, so I will consider a first-best primal problem and back out the corrective instruments that implement the Planner's allocation as an equilibrium.

Definition 6 (First-Best Planning Problem)**.** *The Planner solves*

$$\max_{\{c_t,\{\Lambda_{jt},\{\ell_{j\iota t}\},A_{jt},s_{jt}\}\}} \sum_{t\geq 0} \frac{1}{(1+\rho)^t} u(c_t) \;\; s.t. \tag{60}$$

$$\mathcal{Y}_t = c_t + \sum_j r_j \Lambda_{jt}$$

$$L = \sum_j \int_0^1 \ell_{j\iota t} d\iota$$

$$A_{jt} = \gamma^{\chi_j s_{jt}^{\eta} \phi_{jt}} A_{jt-1}$$

$$\mathcal{S} = \sum_j s_{jt}.$$

That is, the Planner picks the technologically feasible allocation that will maximize household utility. Note that climate damage from past carbon emissions is embedded in $\mathcal{Y}_t$. It will be helpful to define the Planner's intertemporal marginal rate of substitution $R_{t+1} \equiv (1+\rho)u'_t/u'_{t+1}$. This is the Planner's discount factor over consumption goods. The following proposition describes first-best carbon prices and innovation subsidies.

Proposition 4. *To implement the first-best allocation as an equilibrium, the Planner separately corrects the externalities from carbon pollution and knowledge spillovers. They set the carbon price according to*

$$\tau_t = -\sum_{\hat{t}\geq t} \prod_{\hat{s}=1}^{\hat{t}-t} \frac{1}{R_{t+\hat{s}}} \mathcal{Y}_{\hat{t}} \frac{\partial \ln(\Omega_{\hat{t}})}{\partial \mathcal{E}_t}, \tag{61}$$

which is the social cost of carbon. Next, they set innovation subsidies according to the recursion

$$\xi_{jt}\Pi_{jt} = (1-\alpha)S_{jt}\mathcal{Y}_t + \frac{1}{R_{t+1}}\left[\xi_{jt+1}\Pi_{jt+1} + \sum_i \xi_{it+1}\Pi_{it+1} g_{it+1}\varphi_{ijt+1}\right]. \tag{62}$$

The proof of Proposition 4 can be found in Appendix A.8. Several comments are in order. First, the Planner corrects the two externalities in the economy with distinct instruments. This principle of separating externalities has been argued for elsewhere in the climate innovation literature (Acemoglu et al., 2012, 2016; Golosov et al., 2014). Clean innovators are blind to the climate benefit of their innovations in laissez-faire, but once the carbon price has been set equal to the social cost of carbon (61), innovation rents will be properly adjusted to reward clean technology. Innovation subsidies do not need to provide additional support to clean innovation as such.

Instead, innovation subsidies correct the knowledge spillover externalities mediated through the spillover network, and in doing so, they ignore the pollution externalities associated with each technology. Innovation subsidies set according to the recursive formula of Equation (62) guarantee that the private reward for innovation $\xi_{jt}\Pi_{jt}$ equals the social value of innovation. This social value includes two terms. First, there is a contemporaneous benefit to production from having more advanced technology. This effect is captured by the $(1-\alpha)S_{jt}\mathcal{Y}_t$ term, which reflects the Domar weight of technology à la Hulten's Theorem (Hulten, 1978). Next, the terms in the brackets represent the spillover benefits realized in the following period. Tomorrow's innovators will build on top of the knowledge stock they inherit, represented by the $\xi_{jt+1}\Pi_{jt+1}$ term, and their research productivity will be affected by knowledge stocks via the spillover network $\boldsymbol{\varphi}_{t+1}$, represented by the $\sum_i \xi_{it+1}\Pi_{it+1}g_{it+1}\varphi_{ijt+1}$ term. Therefore, the innovation subsidies of Proposition 4 provide a general formula for innovation policy in the presence of a spillover network as it relies on minimal parametric assumptions about the structure of production and spillovers. Finally, the Planner sets the intermediate subsidy Υ to close the monopoly markup γ and the lump-sum tax D_t to balance the government's budget (27) in each period.

As I argue throughout this paper, knowledge spillovers are critical for understanding both the impact and efficiency of policy in shaping the direction of innovation. However, we can see from Proposition 4 that the role of spillovers differs in these two cases. For describing the impact of policy, the spillovers technologies *receive* are what matter. These spillovers allow less advanced technologies to achieve catchup growth, and as argued in Section 3, this shapes both the steady-state impact of policy and the speed with which technology converges to its steady-state. Put differently, the rows of the spillover matrix $\boldsymbol{\varphi}_t$ are what matter for describing the impact of policy, and this can be seen from the fact that the diagonal of the spillover matrix $\boldsymbol{\Phi}_t$ (36) is composed of row sums. For prescribing optimal policy, the spillovers technologies *send* are what matter. This is because the recipients of spillovers already internalize the boost to their research productivity, so policy needs to reflect the spillover externalities that profit signals ignore. It is the columns of the spillover matrix $\boldsymbol{\varphi}_t$ that matter for efficiency, and this can be seen from the fact that the spillover benefit term of Equation (62) is composed of column sums.

To provide further intuition for the determinants of optimal innovation subsidies, I will consider steady-state innovation policy. My results mirror those of Liu and Ma (2021), but I will consider a more general structure for production and spillovers, albeit with a narrow focus on steady-state policy. I will describe the steady-state in terms of innovation subsidies multiplied by income shares $\tilde{\xi}_{jt} \equiv \xi_{jt} \cdot S_{jt}$ as this value converges to a constant even in cases where income shares of some technologies go to zero. For instance, ever-increasing carbon prices may drive the income share of dirty technologies asymptotically to zero.[31] It will be helpful to define the Planner's growth-adjusted intertemporal marginal rate of substitution $\tilde{R}_t \equiv R_t/(1+g_{yt})$, where g_{yt} is the growth rate of output. The following corollary characterizes steady-state innovation policy.

Corollary 3. *In steady-state, innovation subsidies multiplied by income shares satisfy*

$$\tilde{\xi}' = \frac{1}{\gamma - 1} S' \Big[(1 - \tilde{R}^{-1})\boldsymbol{I} - g\tilde{R}^{-1}\boldsymbol{\varphi} \Big]^{-1}. \tag{63}$$

The proof of Corollary 3 can be found in Appendix A.9. This result states that steady-state innovation policy depends on both the spillover network $\boldsymbol{\varphi}$ and the Planner's level of patience, as measured by their growth adjusted discount factor $\tilde{R}$. To better understand the intuition, note that the inverse matrix of Equation (63) follows the logic of a Leontief inverse. Today's innovations expand output according to income shares S and create spillovers for the following period. Iterating forward with Equation (62), these spillovers beget further innovations tomorrow, which generate their own spillovers in the following period, and so on. The sequence of ripple effects generates a geometric series that sums to the inverse matrix of Equation (63).[32]

[31] This is what happens in the steady-state under the parametric assumptions of Sections 2.5 and 5.2.

[32] Formally, we can write

$$\big[(1-\tilde{R}^{-1})\mathbf{I} - g\tilde{R}^{-1}\boldsymbol{\varphi}\big]^{-1} = \sum_{n\geq 0} \big(\frac{\mathbf{I}+g\boldsymbol{\varphi}}{\tilde{R}}\big)^n = \mathbf{I} + \big(\frac{\mathbf{I}+g\boldsymbol{\varphi}}{\tilde{R}}\big) + \big(\frac{\mathbf{I}+g\boldsymbol{\varphi}}{\tilde{R}}\big)^2 + \dots, \tag{64}$$

An informative case is when there are no cross-technology spillovers $\boldsymbol{\varphi} = \mathbf{0}$. Then, all innovation subsidies are set equal to $1/(\gamma-1)(1-\tilde{R}^{-1})$, which I call the *baseline innovation wedge*. This wedge reflects the difference between the private marginal benefit of innovation, $(\gamma - 1)$ times the contemporaneous income growth from innovation, and the social marginal benefit of innovation, a permanent increase in income from better technology, when there is no spillover network. In our setting, it is natural to think about innovation policy net of the baseline innovation wedge $\hat{\xi} \equiv (\gamma-1)(1-\tilde{R}^{-1})\tilde{\xi}$ because this term controls the *composition* of research effort. Without a spillover network, setting $\hat{\xi} = S$ in steady-state would be efficient because the lack of cross-technology spillovers implies innovation rents are proportional to the social value of innovation.

For the more general case with a spillover network, manipulating Equation (63), we have that $\hat{\xi}$ solves

$$(1-\tilde{R}^{-1})\big(\hat{\xi}' - S'\big) - g\tilde{R}^{-1}\hat{\xi}'\boldsymbol{\varphi} = 0. \tag{65}$$

This allows us to see the role of the Planner's level of patience. A Planner that is completely myopic, i.e. $\tilde{R}^{-1} \to 0$, will allow the composition of research effort to be set entirely by profit signals: $\hat{\xi} = S$. This maximizes the contemporaneous benefit of innovation by focusing exclusively on the immediate impact on output. A Planner that is perfectly patient, i.e. $\tilde{R}^{-1} \to 1$, will set innovation policy with an exclusive focus on the spillover benefit of innovation. They set innovation policy according to

$$\hat{\xi}'\boldsymbol{\varphi} = \hat{\xi}'\big(\tilde{\boldsymbol{\varphi}} - \mathbf{I}\big) = 0, \tag{66}$$

where $\tilde{\boldsymbol{\varphi}}$ is the gross spillover network discussed in Section 2.2. That is, $\hat{\xi}$ solves $\hat{\xi}'\tilde{\boldsymbol{\varphi}} = \hat{\xi}'$, so $\hat{\xi}$ is a measure of eigenvector centrality for the gross spillover network. In this case, the Planner prioritizes technologies that are central in the spillover network because these technologies are best able to produce spillovers over time.

In summary, technology stocks are used in the production of both physical goods and ideas, and Equation (65) states that steady-state innovation policy strikes a balance between these two benefits according to the Planner's degree of patience.

4.2 Second-Best: Innovation Subsidies with Incomplete Carbon Pricing

Having considered optimal policy with an unrestricted set of instruments, I will now consider a second-best case where the Planner cannot set the price on carbon. I view this as a realistic case given that pricing carbon has proved to be politically difficult.

Let $\{\hat{\tau}_t\}$ denote a sequence of externally imposed carbon prices. For example, these external carbon prices may simply be zero if carbon pricing is politically impossible. The loss of policy instruments implies that the Planner must now satisfy an incentive compatibility constraint. To keep focus on innovation, I will assume that intermediate subsidies continue to close the monopolist markup $\Upsilon = \gamma$.

so the Planner considers a full geometric series of spillover effects in setting steady-state innovation policy.

Definition 7 (Second-Best Planning Problem)**.** *The Planner solves*

$$\max_{\{c_t,\{A_{jt},s_{jt}\}\}} \sum_{t\geq 0} \frac{1}{(1+\rho)^t} u(c_t) \;\; s.t. \tag{67}$$
$$\mathcal{Y}_t = c_t + \sum_j r_j \Lambda_{jt}$$
$$A_{jt} = \gamma^{\chi_j s_{jt}^{\eta} \phi_{jt}} A_{jt-1}$$
$$\mathcal{S} = \sum_j s_{jt}$$
$$\{\Lambda_{jt}, \{\ell_{j\iota t}\}\} = \operatorname{argmax} \left[\mathcal{Y}_t - \sum_j r_j \Lambda_{jt} - \hat{\tau}_t \mathcal{E}_t\right] \;\; s.t. \;\; L = \sum_j \int_0^1 \ell_{j\iota t} d\iota.$$

The final constraint is an incentive compatibility constraint which reflects the private optimizing behavior of producers. The equilibrium of this economy can be represented as a simple revenue maximization problem, and the Planner must now choose an innovation allocation knowing that the choice of technology will affect production choices in equilibrium. The Planner still has a complete set of policy instruments on the innovation side of the economy, so they can back out the innovation subsidies that implement their desired allocation of scientists as an equilibrium. The following proposition characterizes those innovation subsidies.

Proposition 5. *When the Planner can no longer control the carbon price, they set innovation subsidies according to the recursion*

$$\xi_{jt}\Pi_{jt} = (1-\alpha)S_{jt}\mathcal{Y}_t - (\tau_t - \hat{\tau}_t)\mathcal{E}_t \frac{\partial \ln(\mathcal{E}_t)}{\partial \ln(A_{jt})} + \frac{1}{R_{t+1}}\left[\xi_{jt+1}\Pi_{jt+1} + \sum_i \xi_{it+1}\Pi_{it+1} g_{it+1} \varphi_{ijt+1}\right], \tag{68}$$

where τ_t follows the social cost of carbon formula of Equation (61).

The proof of Proposition 5 can be found in Appendix A.10. The first thing to note is that second-best innovation subsidies are remarkably similar to those of the first-best; the only difference is a new term that reflects the wedge between the external and efficient carbon price. The social cost of carbon is no longer properly internalized in production decisions, so to make up for this distortion, the Planner adjusts the direction of innovation in favor of clean technology.[33] This adjustment multiplies the wedge between the external and efficient carbon price, total emissions, and the elasticity of equilibrium emissions with respect to technology. For clean technology, this emissions elasticity will tend to be negative, implying an increase in clean innovation subsidies. Conversely, improvements in dirty technology will tend to increase emissions in equilibrium, so dirty innovation subsidies will be reduced.[34] As one might expect, this formula recovers the first-best when the external carbon price $\hat{\tau}_t$ happens to be set at the efficient level τ_t.

The similarity between first- and second-best innovation subsidies implies that the latter inherit the properties discussed in Section 4.1. All else equal, technologies are rewarded for their importance

[33] I will assume that $\tau_t \geq \hat{\tau}_t$ in my discussion of Proposition 5 as this is the empirically relevant case, but the result does not formally require this to be true.

[34] This opens the possibility that innovation subsidies are negative. In that case, the allocation of scientists will be a corner solution with zero scientists allocated to technologies with negative innovation subsidies. Formally, $\xi_{jt} \leq 0 \Rightarrow s_{jt} = 0$. Indeed, if all technologies lead to sufficiently large increases in equilibrium emissions, the Planner may want to stop growth altogether.

in production and their ability to create spillovers, as measured by Domar weights and eigenvector centrality respectively, with the weight between the two determined by the Planner's level of patience. Now that carbon pollution is improperly priced, the flow value of innovation includes the degree to which technological change exacerbates the economy's distortions. Thus, one can interpret the innovation subsidies of Equation (68) as adjusting for the possibility of "immiserizing growth", where innovation can reduce welfare in inefficient economies by further deepening distortions (Bhagwati, 1958, 1968). By this view, Proposition 5 provides a general way of thinking about innovation policy in the presence of distortions. The flow value of innovation that guides policy must include both Domar weights and the effect on distortions to prevent technological change from reducing welfare.

5 Calibration

In this section, I describe my calibration of the structural parameters of my model. My application is the transportation and electricity generation sectors in the United States. Throughout this section, I make the parametric assumptions for production and spillovers described in Section 2.5. I conclude the section with a model validation exercise that tests my model's ability to match the advances made by clean technology in both sectors from 2010 to 2021. I show that my model is able to match the evolution of technology in this period when it includes the spillover network but fails to do so when the spillover network is shut down.

A time period represents one year. Longer time periods, such as one representing five years, imply counterfactually rapid growth of clean transportation from 2010 to 2021. Finally, any data series referenced in this section that is published in nominal terms is adjusted using a US CPI index that takes 2012 as its base year. Thus, all dollar values referenced throughout the paper are in 2012 dollars.

5.1 Spillover Network

In this section, I describe how I calibrate one of the central objects in my model: the spillover network. I follow Liu and Ma (2021) in using the citation network of patents as a proxy for the spillover network.

For my sample of patents, I take granted US patents from PatentsView.[35] First, I need to assign patents to my five technology classes: clean transportation, dirty transportation, clean electricity generation, dirty electricity generation, and general technology. To do so, I make use of the International Patent Classification (IPC) and Cooperative Patent Classification (CPC) systems.[36] IPC codes pertaining to both clean and dirty transportation come from Aghion et al. (2016), who identify IPC codes for the transportation sector in their empirical study of directed innovation in the auto industry.[37] I also include categories from CPC subclass Y02T that pertain to transportation. For clean transportation, this includes electromobility (Y02T64-72), efficient charging and discharging systems (Y02T10/92), charging of electric

[35] I am abstracting from international patents in the construction of my spillover network, but I do not think this introduces bias for the following two reasons. First, US patents primarily cite other US patents. Indeed, 70% of US patent citations reference other US patents (Liu and Ma, 2021). Second, citation shares are scale independent, so for the exclusion of international patents to introduce bias, citations to international patents would have to systematically differ in their distribution across technology classes.

[36] The CPC system is a simple extension of the IPC system. While there is some discordance between the two systems, I manually verify that all of the IPC codes that I use map directly to the same CPC codes with the same meaning. For instance, B60L denotes electric vehicles in both the IPC and CPC system.

[37] For the full list of IPC codes pertaining to both clean and dirty transportation, see Table 1 of Aghion et al. (2016).

vehicles (Y02T90/10-16), hydrogen technology in transport (Y02T90/40), and hybrids (Y02T10/62). For dirty transportation, this includes ICE efficiency (Y02T10/12) and engine management (Y02T10/40).

For clean electricity generation, I take directly from the CPC subclass Y02E related to GHG reducing innovations in electricity generation. This includes renewables (Y02E10), nuclear (Y02E30), and energy storage (Y02E60/10-16). I also include integration of photovoltaics in buildings (Y02B10/10). IPC codes pertaining to dirty electricity generation come from Lanzi et al. (2011), who identify a list of IPC codes associated with fossil fuel technologies for electricity generation.[38] Finally, I define general technology patents as those patents that are not in any of the above categories. Table 1 gives a representative, but not exhaustive, list of my patent classification codes.

Table 1: Sampling of Patent Classification Codes

Transportation		Electricity Generation	
Description	IPC/CPC Codes	Description	IPC/CPC Codes
Clean		Clean	
Electric Vehicles	B60L	Renewables	Y02E10
Hybrids	Y02T10/62	Nuclear	Y02E30
Hydrogen Fuel Cells	H01M8	Energy Storage	Y02E60/10-16
Dirty		Dirty	
Internal Combustion Engines	F02B	Steam Engine Plants	F01K
Controlling Combustion Engines	F02D	Gas-Turbine Plants	F02C
Supplying Combustion Engines	F02M	Steam Generation	F22
Cylinders for Combustion Engines	F02F	Combustion Apparatus	F23

Notes: IPC codes for both clean and dirty transport patents come from Aghion et al. (2016), CPC codes for clean electric patents come directly from CPC subclass Y02E, and IPC codes for dirty electric patents come from Lanzi et al. (2011). General patents are those not classified as pertaining to either transport or electricity.

With patents placed into technology classes, I estimate the gross spillover network $\tilde{\boldsymbol{\varphi}}$ as the proportion of technology i's patent citations that reference patents in technology j. That is, I set gross spillovers according to

$$\tilde{\varphi}_{ij} = \frac{cites_{ij}}{\sum_k cites_{ik}}, \tag{69}$$

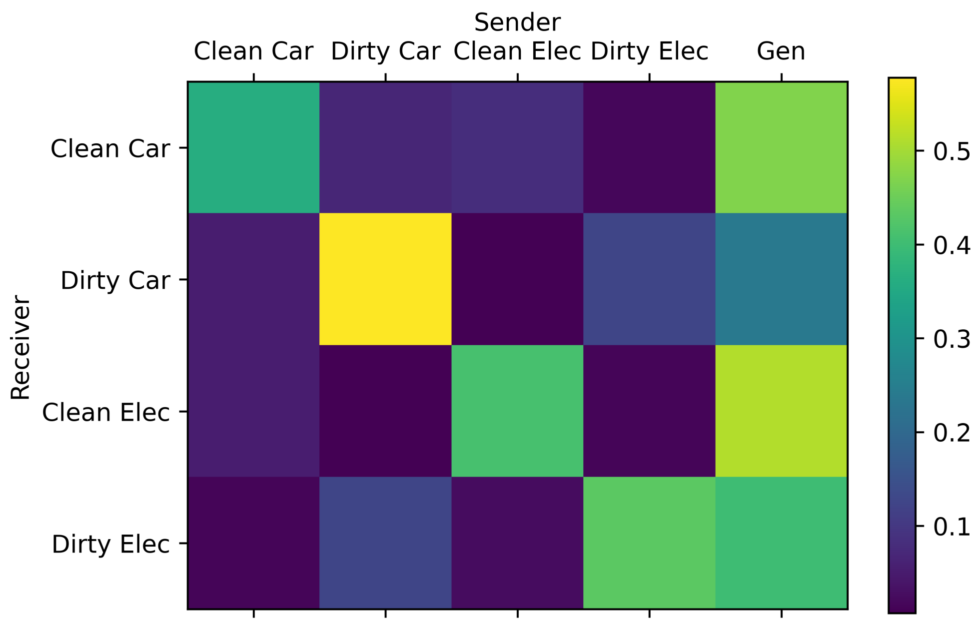
where $cites_{ij}$ is the number of citations from patents in technology class i that reference patents in technology class j. This approach leverages the widely accepted view that patent citations are a metric of knowledge spillovers (Jaffe et al., 1993; Hall et al., 2005). The (net) spillover network then follows mechanically from $\boldsymbol{\varphi} = \tilde{\boldsymbol{\varphi}} - \mathbf{I}$, and because Equation (69) implies that the rows of $\tilde{\boldsymbol{\varphi}}$ mechanically sum to one, this guarantees that the spillover function described in Equation (31) is homogeneous of degree zero.

Figure 2 contains a heat map that illustrates the gross spillover network $\tilde{\boldsymbol{\varphi}}$. It excludes the row pertaining to the general technology as almost all of the general technology's patent citations reference another general technology patent.[39] Figure 2 allows us to see several intuitive patterns in the spillover network. First, all of the climate-relevant technology classes extensively cite both themselves and the general technology. Second, the transportation sector has relatively high within-sector spillovers when compared with the electricity generation sector. As suggested by the story of Tesla's prototype, both

[38] For the full list of IPC codes pertaining to dirty electricity generation, see Table A1 of Lanzi et al. (2011).
[39] Specifically, 98.6% of general technology patent citations reference another general technology patent.

clean and dirty vehicles can learn from each other because they are fundamentally similar machines. This does not appear to be the case in the electricity generation sector where, for example, solar panels and coal-fired power plants generate electricity via entirely different means. Instead, clean electricity has an abnormally high citation share on the general technology, consistent with the solar panel inventors from Bell Labs building on knowledge from the ICT sector. Finally, clean and dirty technologies each form their own distinct spillover clusters. That is, dirty technologies of both sectors cite each other, and the same is true of clean technologies but to a lesser extent. Each of these features of the spillover network are relevant for convergence dynamics, as I will discuss in Section 6.1.

Figure 2: Heat Map of Gross Spillover Network $\tilde{\varphi}$



Notes: The y-axis denotes technologies receiving spillovers, while the x-axis denotes technologies sending spillovers. Granted US patents come from PatentsView.

5.2 Climate Module

To simulate optimal climate policy, I must add structure to my model specifying (i) how a sequence of emissions influences the Earth's climate system and (ii) how warming of the Earth's climate system impacts the economy's productive capacity. To this end, I follow Golosov et al. (2014) in my description of the carbon cycle and economic damage function.

Atmospheric carbon concentrations evolve according to

$$\mathcal{C}_t = \sum_{\hat{t}=1800}^{t} \left(\psi_p + (1-\psi_p)\psi_0\psi^{t-\hat{t}}\right)\mathcal{E}_{\hat{t}} + \bar{\mathcal{C}}, \tag{70}$$

where $\bar{\mathcal{C}}$ is the pre-industrial level of atmospheric carbon concentration. I select the year 1800 as the starting point of industrialization.[40] This specification of the carbon cycle provides a mapping from past

[40]One could debate the exact start date of industrialization, but given that carbon emissions were trivially small before

carbon emissions to the current level of atmospheric carbon concentrations. In particular, it states that fraction $(\psi_p + (1-\psi_p)\psi_0\psi^{t-\hat{t}})$ of carbon emitted at time $\hat{t} \leq t$ will remain in the atmosphere at time t. This remaining carbon has both a permanent and transitory component. The permanent component, fraction ψ_p, will remain in the atmosphere forever. The remainder, fraction $(1-\psi_p)$, is transitory. For the transitory component, a fraction $(1-\psi_0)$ exits the atmosphere within a period and is absorbed by the biosphere or surface oceans. Next, the remainder of the transitory component, fraction ψ_0, decays geometrically according to ψ. As argued in Archer (2005) and Golosov et al. (2014), this relatively simple specification of the Earth's climate system provides a good approximation of the complex relationship between carbon emissions and atmospheric carbon concentrations. In particular, it captures the slow process by which the deep oceans absorb carbon from the atmosphere.[41]

The pre-industrial level of atmospheric carbon concentration $\bar{\mathcal{C}}$ was 596.4 gigatons of carbons.[42] Throughout the paper, I list carbon quantities in terms of gigatons of carbon (GtC). For example, US greenhouse gas emissions in 2021 were equivalent to 1.7 GtC. I will, however, list carbon prices in terms of dollars per ton of CO_2 as this is the convention.[43] To map back and forth, one can note that a ton of CO_2 contains 12/44 tons of carbon.

To calibrate my climate module, I start by setting $\psi_p = 0.2$, in line with the 2007 IPCC report estimate that 20% of carbon emissions will remain in the atmosphere after thousands of years. For the remaining two parameters $\{\psi_0, \psi\}$, I perform a similar exercise as Acemoglu et al. (2016) and pick these values to match the relationship between carbon emissions and atmospheric carbon concentrations throughout the 1960-2020 period, selecting $\psi_0 = 0.377$ and $\psi = 0.994$ as the values that minimize the distance between the model's predictions and the data. A major advantage of Equation (70) is that it allows for the state of the climate to be written in terms of a two-dimensional recursion. The first dimension $\mathcal{C}_{1t}$ represents the permanent component of atmospheric carbon, and it follows

$$\mathcal{C}_{1t} = \psi_p \mathcal{E}_t + \mathcal{C}_{1t-1}. \tag{71}$$

The second dimension $\mathcal{C}_{2t}$ represents the transitory component of atmospheric carbon, and it follows

$$\mathcal{C}_{2t} = (1-\psi_p)\psi_0 \mathcal{E}_t + \psi \mathcal{C}_{2t-1}. \tag{72}$$

The sum of these two components is therefore atmospheric carbon concentration $\mathcal{C}_t = \mathcal{C}_{1t} + \mathcal{C}_{2t}$. Writing the state of the climate recursively allows for a simpler representation of the climate system as one does not need to keep track of the entire history of emissions. This requires the selection of an initial condition $\{\mathcal{C}_{1t_0}, \mathcal{C}_{2t_0}\}$. For any t_0, such as 1960 in the case of my calibration exercise, I set $\mathcal{C}_{1t_0} = \sum_{\hat{t}=1800}^{t_0} \psi_p \mathcal{E}_{\hat{t}} + \bar{\mathcal{C}}$

the year 1800, the choice is insignificant for present purposes.

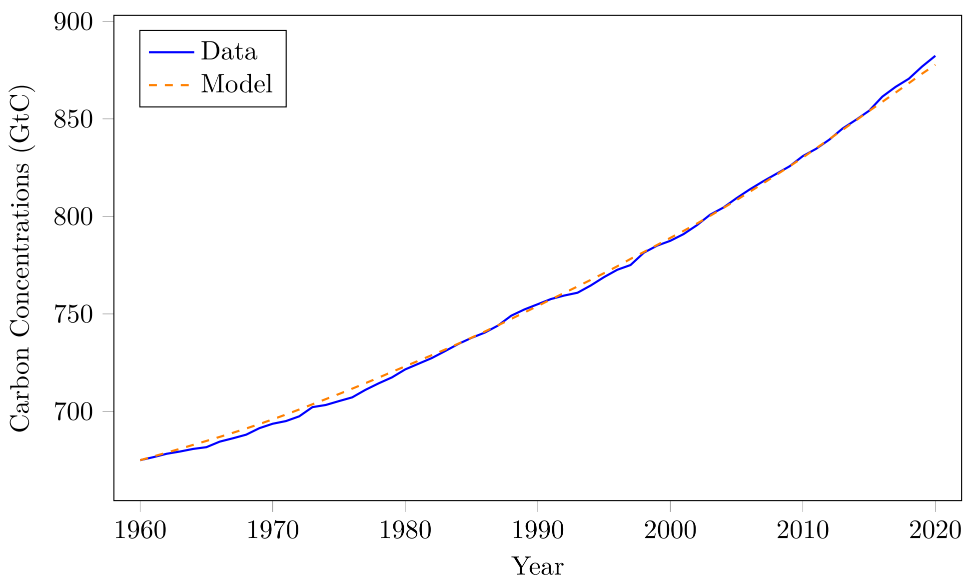
[41] This approach is an alternative to that of Nordhaus (2017), which models the Earth's climate system as consisting of three carbon reservoirs: the atmosphere, the upper oceans and biosphere, and the deep oceans. In that model, carbon concentrations of the three reservoirs evolve according to a system of linear difference equations. As discussed in Archer (2005), Archer et al. (2009), and Golosov et al. (2014), the linear, three reservoir specification implies a counterfactually rapid absorption of atmospheric carbon into the deep oceans. Instead, the approach taken here is better able to match the depreciation of atmospheric carbon concentrations, while also reducing the number of state variables needed to describe the climate. The implication of this difference in modeling choice is longer-lived, and therefore larger, impacts of anthropogenic carbon emissions on the climate.

[42] Atmospheric carbon concentrations are often stated in terms of parts per million (ppm) of CO_2. One ppm of CO_2 is equivalent to 2.13 GtC (O'Hara, 1990). This conversion was used to map the widely-accepted pre-industrial 280 ppm of CO_2 to GtC: $596.4 = 280 \times 2.13$.

[43] For instance, the White House lists its social cost of carbon in dollars per ton of CO_2.

and $\mathcal{C}_{2t_0} = \hat{\mathcal{C}}_{t_0} - \mathcal{C}_{1t_0}$, where $\hat{\mathcal{C}}_{t_0}$ is the observed level of atmospheric carbon concentrations. Data on carbon emissions dating back to 1800 comes from Our World in Data, and data on atmospheric carbon concentrations comes from the NOAA's Mauna Loa Observatory.[44] The results of my calibration exercise can be seen in Figure 3.

Figure 3: Climate Model Matches Historic Relationship of Emissions and Carbon Concentrations



Notes: Match of model predicted atmospheric carbon concentrations with data when $\psi_p = 0.2$, $\psi_0 = 0.377$, and $\psi = 0.994$. Carbon emissions come from Our World in Data, and atmospheric carbon concentrations come from the NOAA's Mauna Loa Observatory.

Finally, atmospheric carbon concentrations create damage to production via

$$\Omega_t = \exp\big(-\varrho(\mathcal{C}_t - \bar{\mathcal{C}})\big), \tag{73}$$

where $\varrho > 0$ is a scale parameter determining the sensitivity of output to the climate.[45] This specification implies that the semi-elasticity of final output with respect to atmospheric carbon concentrations is constant: $\frac{\partial \ln(\mathcal{Y}_t)}{\partial \mathcal{C}_t} = \varrho$. Assuming a constant semi-elasticity of damages with respect to atmospheric carbon concentrations does not allow for severe non-linearities in the climate system, e.g. tipping points, but this functional form for damages is consistent with the approach taken in much of the climate literature (Nordhaus, 1992, 2017; Stern, 2007; Krusell and Smith, 2022; Barrage and Nordhaus, 2023). In my choice of ϱ, I take the same value as Golosov et al. (2014) and set $\varrho = 5.3 \times 10^{-5}$. However, given the considerable uncertainties surrounding the damage that will result from climate change, I will also consider an alternative damages parameter that is four times as large, roughly in line with the

[44]Carbon emissions from land use change are only available going back to 1850, so I extrapolate the linear trend in those emissions from 1850 to 1950 back to 1800.

[45]One can also map from atmospheric carbon concentrations to temperature increases with the equation $\Delta T_t = \Gamma \ln(\mathcal{C}_t/\bar{\mathcal{C}})/\ln(2)$, where Γ represents the climate sensitivity parameter. That is, a doubling of atmospheric carbon concentrations over pre-industrial levels leads to warming of $\Gamma°$Celsius. The standard value for Γ is 3, but there remains considerable uncertainty over the value of Γ. The IPCC deems 1.5-4.5 to be the "likely" range of Γ, but this leaves a very real chance of much higher values (Wagner and Weitzman, 2016).

”catastrophic damages” scenario considered by Golosov et al. (2014).[46]

5.3 Structural Parameters & Initial Conditions

In this section, I describe my calibration strategy for the remaining structural parameters of my model $\{\sigma, \lambda, \{\nu_\theta, \omega_{\theta d}\}, L, \mathcal{S}, \gamma, \chi, \alpha, \{r_j\}, \eta, \vartheta, \rho\}$ as well as the initial condition for technology $\{A_{jt_0}\}$. In Appendix B.1, I provide details explaining how the equilibrium conditions of my model, and therefore the calibration moments discussed below, can be written in terms of just technology, policy instruments, and structural parameters. Table 2 provides a summary of my parameter choices.

Table 2: Parameter Choice Summary

Parameter	**Description**	**Value**	**Source**
σ	Clean/Dirty ES	1.86	Papageorgiou et al. (2014)
λ	Cross-Sector ES	0.1	Hassler et al. (2021)
ν_{car} ν_{elec}	CES Shares	0.029 0.025	See Text
γ	Innovation Step Size	1.07	Acemoglu et al. (2023)
α	Input Share	0.403	Barrage (2020)
$r_{\theta d}/r_{\theta c}$	Relative Dirty Input Price	2.25	BP (2022)
η	Research Elasticity	0.5	See Text
ϑ	Inverse Intertemporal ES	1	Standard
ρ	Rate of Pure Time Preference	0.01 0.15	Stern (2007) Nordhaus (2017)

For the within-sector elasticity of substitution between clean and dirty forms of production, I set $\sigma = 1.86$ in line with the average estimate from Papageorgiou et al. (2014).[47] The evidence for this estimate comes from the electricity generation sector, but I set it as the value for the transportation sector as well in part because it is of a similar magnitude to the estimate of Lanzi and Sue Wing (2011) who look at the energy sector more broadly. For the cross-sector elasticity of substitution, I set $\lambda = 0.1$ in line with the evidence from Hassler et al. (2021) who find that energy is a near perfect complement with capital/labor at annual frequencies. Across the literature, it is common to model energy as a complement to other factors (Van der Werf, 2008; Böhringer and Rutherford, 2009). For instance, Fried (2018) selects a near-zero elasticity of substitution between energy and other factors, making production almost Leontief.

For the sectoral CES shares, I normalize the sum of the shares to one and pick $\{\nu_{car}, \nu_{elec}\}$ so that the income shares of transportation and electricity generation in the laissez-faire steady-state match

[46] I abstract from the consideration of risk in optimal climate policy. There are substantial uncertainties regarding the climate system’s response to emissions, as well as the sensitivity of human civilization to changes in the Earth’s climate, so a large literature explicitly considers the role of risk and insurance in determining the social cost of carbon (Weitzman, 2009; Lemoine and Traeger, 2014; Gillingham et al., 2018; Cai and Lontzek, 2019). A thorough consideration of climate uncertainty is beyond the scope of this paper, but I will capture some elements of the insurance motive for climate policy by considering the possibility of catastrophic damages as an extension.

[47] See Table 4 of Papageorgiou et al. (2014).

the average values of those shares in the data from 2000 to 2020. Data on revenue for the electricity generation sector comes from the US Energy Information Agency (EIA).[48] Data on motor vehicle output and GDP come from the US Bureau of Economic Analysis (BEA). This approach yields $\nu_{car} = 0.029$ and $\nu_{elec} = 0.025$ to match average income shares of 2.9% and 2.3%, respectively.

I normalize the supply of scientists and workers to $\mathcal{S} = 1$ and $L = 100$, respectively. These are both WLOG normalizations, but they match the fraction of the labor force that works in R&D in the US (Jones and Vollrath, 2013). Following Acemoglu et al. (2023), I set the step size of innovation to $\gamma = 1.07$ to match the profit share of the petroleum and coal products, durable manufacturing, and wholesale trade sectors.[49] This is a similar value to that found in other models of step-ladder innovation (Acemoglu et al., 2016; Akcigit and Kerr, 2018). I set research productivity χ so that growth in the laissez-faire steady-state is 2% per year.

For the income share of inputs, I set $\alpha = 0.403$ in line with the estimate from Barrage (2020) of the labor share in electricity and resource production. For input costs, I normalize all of the clean input costs to one: $r_{\theta c} = 1$. I then set dirty input costs to match the relatively low efficiency of fossil fuel-based machinery at converting primary energy into useful energy. To compare the energy conversion efficiency of fossil fuels with non-fossil sources of energy, the BP Statistical Review of World Energy report assumes a thermal equivalent efficiency factor that rises linearly from 36% in 2000 to 45% in 2050 (BP, 2022). That is, in 2050 45 TWh of electricity produced from solar panels will require the burning of 100 TWh worth of coal, down from 125 TWh in 2000.[50] Therefore, I set the dirty input costs to $r_{\theta d} = 2.25 \approx 1/0.45$.

For my choice of the elasticity of innovation with respect to scientists η, I rely on a body of empirical evidence from the innovation literature. One body of studies considers the elasticity of patents with respect to R&D expenditures and generally finds a value of about 0.5 (Griliches, 1990; Hall and Ziedonis, 2001; Blundell et al., 2002). Another considers the elasticity of R&D expenditure with respect to the tax price of research and generally finds an elasticity of about unity (Hall, 1993; Hall and Van Reenen, 2000; Bloom et al., 2002; Wilson, 2009). As I discuss in Appendix B.1, both sets of findings correspond to a value of $\eta = 0.5$ in my model. Similar conclusions were reached by Akcigit and Kerr (2018), Acemoglu et al. (2018), and Bloom et al. (2021).

For the preference parameters, I set the inverse intertemporal elasticity of substitution ϑ equal to the standard value of unity, which implies a logarithmic period utility function. For the rate of pure time preference ρ, I consider both the Stern rate of 0.1% per year and the Nordhaus rate of 1.5% per year (Stern, 2007; Nordhaus, 2017). The former is the standard value for those who believe ethical considerations should drive the choice of discount rate, while the latter is the standard value for those who think the discount rate should match the rate of return on capital. There has been substantial debate in the climate literature over the appropriate choice of social discount rate (Stern, 2007; Nordhaus, 2007; Dasgupta, 2008; Barrage, 2018).[51] This debate is not the focus of this paper, but given the long-term

[48] Specifically, Forms EIA-826, EIA-861, and EIA-861M.

[49] Acemoglu et al. (2023) match the weighted average profit share for these sectors from 2004 to 2014 using the Quarterly Financial Report from the US Census Bureau.

[50] This efficiency factor is consistent with the US Department of Energy estimate that electric vehicles convert more than 77% of the electricity they pull from the grid into kinetic energy, while gasoline-powered vehicles can only convert 12-30% of the chemical energy in gasoline into kinetic energy. Given that gasoline-powered vehicles are a more mature technology, this superior conversion rate of electric vehicles likely reflects an intrinsic difference between the two technologies.

[51] It is worth noting that moral philosophers universally reject the idea that utility should be discounted simply because it occurs at a later date (Parfit, 1984; Ord, 2020). The only justification for discounting future utility that is widely accepted amongst philosophers is the risk of extinction, so it is acceptable to discount future utility in accordance with the probability that it may not happen. Indeed, this is the justification given by Stern (2007) for his choice of discount rate.

impact of both climate and innovation policy, it is important to understand how the choice of social discount rate shapes the optimal path of climate innovation policy.

For the initial condition of technology at time t_0, I first set within-sector relative technology $\{A_{\theta ct_0}/A_{\theta dt_0}\}$ to match the clean quantity share of each sector. I target clean quantity shares, as opposed to income shares, because of data availability, but as I show in Appendix B.1, there is a straightforward mapping between quantity shares and income shares. Data for quantity shares in the transportation sector comes from the US Department of Energy's Transportation Energy Data Book which lists the share of new light vehicles in the US that were hybrids, plug-in hybrids, and EVs from 1999 to 2021 (Davis and Boundy, 2022).[52] Data on the share of electricity generated from non-fossil sources comes from the EIA.[53] Next, I set cross-sector relative technology $\{A_{\theta dt_0}/A_{\Theta ct_0}\}$ to match the income share of each sector. In both cases, I include the clean input subsidies discussed below. Note that I am using data on sectoral income shares for my calibration of both the CES shares and the initial condition for technology. In the case of the CES shares, I am matching the average empirical income shares from 2000 to 2020 when technology is in steady-state, which guarantees that the income shares of these two small sectors will remain plausible as technology evolves in my simulations. In the case of the initial condition of technology, I am matching the empirical income shares at a given point in time to back out appropriate values for technology. Finally, I pin down the absolute level of technology by normalizing output to 100.

For my carbon intensity parameters $\{\omega_{\theta d}\}$, I match the emissions of transportation and electricity generation in 2021. Data on US GHG emissions by sector comes from the US Environmental Protection Agency (EPA). I find that the carbon intensity of electricity generation is about 65.2% larger than that of transportation. My optimal policy simulations track emissions from the US transportation and electricity generation sectors, but given that climate change is a global phenomenon, I need to select a sequence of emissions for the remainder of the world economy. To this end, I set outside emissions equal to the optimal emissions path of the 2010 RICE model, subtracting a little more than half of US emissions (Nordhaus, 2010).[54]

5.4 Model Validation

In this section, I test my model's ability to replicate the dynamics of clean technology from 2010 to 2021. During this period, clean technology in both transportation and electricity generation made substantial gains, leading to increases in the clean quantity share in both sectors.[55] As shown in Figure 4, the percentage of new light vehicles that were hybrid or electric rose from 2.4% to 9.8%, while the percentage of electricity generated from non-fossil sources rose from 30.1% to 39%. I show that my model can match these clean advances in both sectors, in spite of the fact that no model parameters was selected to target this change.[56] However, when I shut down the spillover network, my model predicts path dependence,

[52] In particular, see Table 6.02 of the Transportation Energy Data Book.

[53] Specifically, Form EIA-923.

[54] To be exact, I subtract 56.9% of US emissions as this is the average contribution of transportation and electricity generation to US emissions from 2000 to 2020.

[55] Solar photovoltaics and onshore wind are now cheaper than fossil fuels in terms of the private lifetime cost of new generation (Roser, 2020). These two forms of renewable electricity generation saw price declines of 89% and 70% from 2009 to 2019, respectively.

[56] The match of the model to the data over this period did inform my selection of the number of years in a time period, as I discuss at the start of Section 5.

implying that the 2010s would have seen dirty technology gain an increasing lead in both sectors.[57] Thus, the nascent clean energy transition of the 2010s is difficult to reconcile with endogenous growth models that assume path dependence.

To argue that technological path dependence is inconsistent with the 2010s advance in clean technology, it must be the case that US climate policy was too weak to push these sectors out of dirty basins of attraction. To account for US climate policy during this time, I allow for subsidies to both innovation and clean inputs. Further details on the selection of these subsidies are in Appendix B.2. For the clean input subsidies, the clean form of production in sector θ receives an input subsidy of $\bar{\xi}_{\theta c}$, so the effective input price becomes $(1-\bar{\xi}_{\theta c})r_{\theta c}$. For the transportation sector, I set $\bar{\xi}_{car,c}$ to match federal spending on the plug-in electric vehicle tax credit as a proportion of total spending in the transportation sector.[58] This yields a subsidy of $\bar{\xi}_{car,c} = 0.011$. For the electricity generation sector, I set $\bar{\xi}_{elec,c} = 0.3$ in line with the 30% investment tax credit (ITC) given to renewable electricity generation during this period.

In selecting innovation subsidies, I focus on relative innovation subsidies because the fixed supply of scientists implies that only relative innovation subsidy matter, as one can see in the research equilibrium condition (25). That is, subsidies that apply to innovation across the board, such as the R&D tax credit, will not affect the composition of scientists across technologies. To this end, I take data on public R&D spending by technology from the International Energy Agency (IEA). I then select innovation subsidies to match this spending as a proportion of total R&D spending, which I take from the BEA. This procedure involves simulating the equilibrium path of technology, so I conduct a separate calibration of innovation subsidies for the version of the model where I shut down the spillover network.

Using the procedure described in Section 5.3, I take 2010 as the start year to set my initial condition for technology. I then simulate the equilibrium path of technology with and without the spillover network and plot the implied dynamics of clean quantity shares in Figure 4.

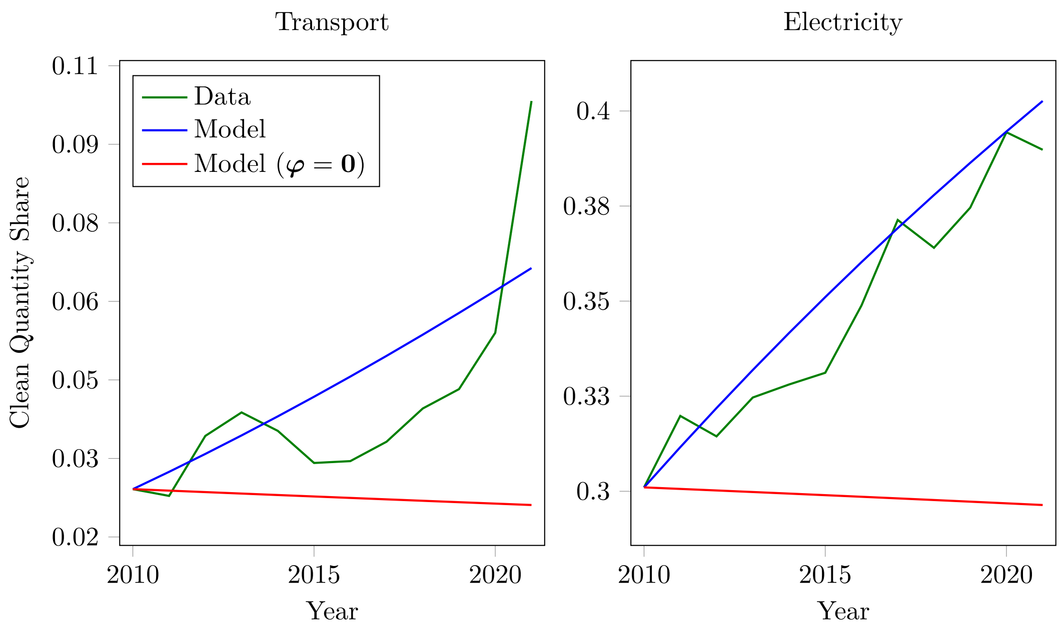
Comparing the simulations with the data, one can see that the model with spillovers broadly matches, as an untargeted moment, the empirical evolution of clean quantity shares in both sectors. This is especially true in the electricity generation sector. The simulated clean quantity shares start at the same value as the data in 2010 by construction, and by 2021, transportation and electricity generation reach clean quantity shares of 6.6% and 40.3% in the simulation, respectively. By comparison, these sectors had achieved clean quantity shares of 9.8% and 39% by 2021 in the data. When I shut down the spillover network, the model makes the counterfactual prediction that clean technology would have fallen even further behind during this period. Without spillovers, the clean quantity shares of transportation and electricity generation would have fallen to 2.1% and 29.6% by 2021. Therefore, a model without spillovers cannot match the qualitative direction of the data.

Using Propositions 2 and 3, we can better understand the role of spillovers in shaping the path of clean technology during this period. When the model includes the spillover network, the spectral radius of the transition matrix is equal to 0.993, indicating that technology is converging toward its steady-state. The model predicts that starting in 2010, transportation and electricity generation will converge halfway to their steady-states in 79 and 99 years, respectively. However, when I shutdown the spillover network, the spectral radius rises to 1.01, inducing path dependence in accordance with Corollary 1. In

[57] By shutting down the spillover network, I mean setting the spillover network to a matrix of zeros: $\boldsymbol{\varphi} = \mathbf{0}$. This is equivalent to setting the gross spillover matrix to the identity: $\tilde{\boldsymbol{\varphi}} = \mathbf{I}$.

[58] Data on federal spending on the plug-in electric vehicle tax credit during this period comes from US Congressional Research Service Report IF11017.

Figure 4: Model with Spillovers Matches 2010s Advance of Clean Technology



Notes: For transportation, the clean quantity share is the proportion of new light vehicles that are hybrid or electric, whereas for electricity generation, it is the proportion of electricity generated from non-fossil sources. Data for these two series come from Davis and Boundy (2022) and the EIA.

that case, the half-life of convergence of each sector becomes infinite as the economy diverges from its interior steady-state and moves deeper into the dirty basins of attraction. Thus, the model's ability to explain the nascent clean energy transition of the 2010s hinges on the inclusion of knowledge spillovers.

6 Simulation

In this section, I conduct three quantitative exercises. First, I examine the impact of introducing a carbon price and clean innovation subsidy. I do so for several levels of cross-technology spillovers to demonstrate their importance in determining the impact of the policy reform, both in the long-run and transition. To make use of the theoretical results in Section 3 at the level of the sector, I define sector-specific relative clean technology $\bar{B}_{\theta t} \equiv A_{\theta ct}/A_{\theta dt}$, which is log-linear in relative technology $\{\bar{A}_{jt}\}$.[59] To quantify convergence speeds, I will also describe half-lives of convergence at the level of the sector $t_\theta^{(1/2)}$.

Second, I simulate first-best climate innovation policy. This exercise highlights the distinction between the spillovers a technology *receives* and the spillovers it *sends*. Clean technologies are not particularly central in the spillover network, so they do not receive favorable innovation subsidies. Instead, clean technologies are beneficial for their ability to produce physical goods without pollution, but this attribute is already favored by the carbon price.

Finally, I simulate second-best innovation subsidies, where the price on carbon pollution is incomplete.

[59] To make use of Propositions 2 and 3 at the level of the sector, I can define

$$\bar{\mathcal{B}}_{\theta t} \equiv \ln(\bar{B}_{\theta t}) - \ln(\bar{B}_{\theta,ss}) = \bar{\mathcal{A}}_{\theta ct} - \bar{\mathcal{A}}_{\theta dt}. \tag{74}$$

Thus, there is a linear mapping from $\bar{\mathcal{A}}_t$ to $\bar{\mathcal{B}}_t$ which allows me to examine transition dynamics at the level of the sector.

I show that a small, growing carbon price can recover most of the welfare gains of the first-best. However, if pollution is unpriced, then the Planner must decarbonize with improved clean technology alone. This leads to slow emission reductions, due to a rebound effect of clean innovation, and slow economic growth, due to the loss of spillovers from dirty technology. Together, these result in massive welfare losses relative to the first-best.

6.1 Impact of Policy Reform

In this section, I examine the impact of introducing a policy reform. I consider a \$51 carbon price, the Biden Administration's current estimate of the social cost of carbon, as well as a uniform clean innovation subsidy equivalent to a 30% tax credit.[60] I pick the clean innovation subsidy to be quantitatively similar to the investment tax credits in the Inflation Reduction Act, but it should be noted that tax credits for production have different effects from tax credits for innovation. Consistent with my theoretical analysis, the policy reform's impact, in both the long-run and transition, depends critically on the level of cross-technology spillovers through its mediation of increasing returns to innovation.

I will assume the policy reform begins after research decisions are sunk in 2021, and for the sake of simplicity, I will make comparisons with laissez-faire. Table 3 summarizes the impact of the policy reform for three cases that differ in their level of cross-technology spillovers. First, I shut down the spillover network. Second, I take my calibrated spillover network. Third, I double the level of cross-technology spillovers in the economy. These three cases follow from a single parameterization which scales the spillover network by ζ. Variation in ζ controls the level of cross-technology spillovers by multiplying the off-diagonal components of the spillover network by factor ζ.[61] For example, the double spillovers case sets $\zeta = 2$, doubling the off-diagonal row sums of the spillover network. The third panel of Table 3 summarizes the difference between the three cases in terms of their degree of increasing returns to innovation, as measured by the spectral radius of the transition matrix.

The first panel of Table 3 considers the long-run impacts of the policy reform. The first two rows make use of Proposition 1 and show the first-order change in steady-state relative clean technology by sector. With calibrated spillovers, we have an increase of 116.6% for transportation and 120.3% for electricity generation.[62] Together, the carbon price and induced increase in clean technology increase within-sector clean income shares and reduce the economy's emissions intensity in steady-state.

The other two cases illustrate the role of cross-technology spillovers in determining the long-run impact of policy reforms. When the spillover network is shut down, the spectral radius of the transition matrix increases from 0.994 to 1.01, inducing path dependence. In that case, policy can only have a long-run impact on technology if it pushes the economy into one of its clean basins of attraction, but when it does so, the long-run impact is transformative. Table 3 shows the policy reform is strong enough to do so for electricity generation, but not for transportation.[63] This difference stems from the fact that clean technology starts out with a greater disadvantage, relative to its dirty rival, in transportation.

[60] This implies an innovation subsidy for clean transportation and electricity generation of $\xi_c = 1/(1-0.3) \approx 1.43$.

[61] One can also think in terms of the spillover matrix $\mathbf{\Phi}$, which measures the level of cross-technology spillovers in the economy. Multiplying the spillover network by ζ does the same for the spillover network, giving us $\zeta\mathbf{\Phi}$.

[62] Formally, these changes in relative technology are listed in log points.

[63] As discussed in Corollary 2, in the case of path dependence, policy reforms push the economy into clean basins of attraction by expanding those clean basins of attraction to include more initial conditions. Indeed, the policy reform of Table 3 *reduces* relative clean technology in the interior steady-state by 101.8 and 116.8 log points for transportation and electricity generation, respectively, thereby increasing the set of initial conditions that lead to clean growth. This expansion includes the initial condition for electricity generation but is too small to include that of transportation.

Table 3: Impact of Policy Reform

	No Spillovers	**Calibrated Spillovers**	**Double Spillovers**
Long-Run Impacts			
Relative Clean Technology by Sector			
$\%\Delta\bar{B}_{car}$	0%	+116.61%	+33.84%
$\%\Delta\bar{B}_{elec}$	$+\infty$%	+120.31%	+38.66%
Clean Income Shares by Sector			
ΔS_c^{car}	0 pp	+12.48 pp	+7.42 pp
ΔS_c^{elec}	+100 pp	+14.81 pp	+9.52 pp
Emissions Intensity			
$\%\Delta\bar{\omega}$	-74.36%	-74.26%	-57.23%
Transitional Impacts			
Half-Lives of Convergence by Sector			
$t_{car}^{(1/2)}$	–	121 years	20 years
$t_{elec}^{(1/2)}$	–	127 years	25 years
Carbon Emissions by Year			
$\%\Delta\mathcal{E}_{2035}$	-43.51%	-50.64%	-53.34%
$\%\Delta\mathcal{E}_{2060}$	-46.36%	-56.84%	-55.07%
Degree of Increasing Returns to Innovation			
Spectral Radius			
$\max\lvert\kappa_j\rvert$	1.01	0.994	0.977

Notes: Impact of introducing a carbon price at the Biden Administration's estimate of the SCC ($51) and clean innovation subsidy equivalent to a 30% tax credit ($\xi_c = 1.43$). Changes in relative technology are listed in log points. For path dependent economies, long-run impacts refer to corner, rather than interior, steady-states.

Conversely, when cross-technology spillovers are doubled, the spectral radius of the transition matrix decreases from 0.994 to 0.977, substantially reducing increasing returns to innovation. Now, the catchup growth generated by cross-technology spillovers reduces the long-run impact of policy because clean technologies cannot gain a large lead without enhancing research productivity in dirty technologies. Thus, the long-run change in relative technology for both sectors is reduced substantially, along with the change in clean income shares and emissions intensity.

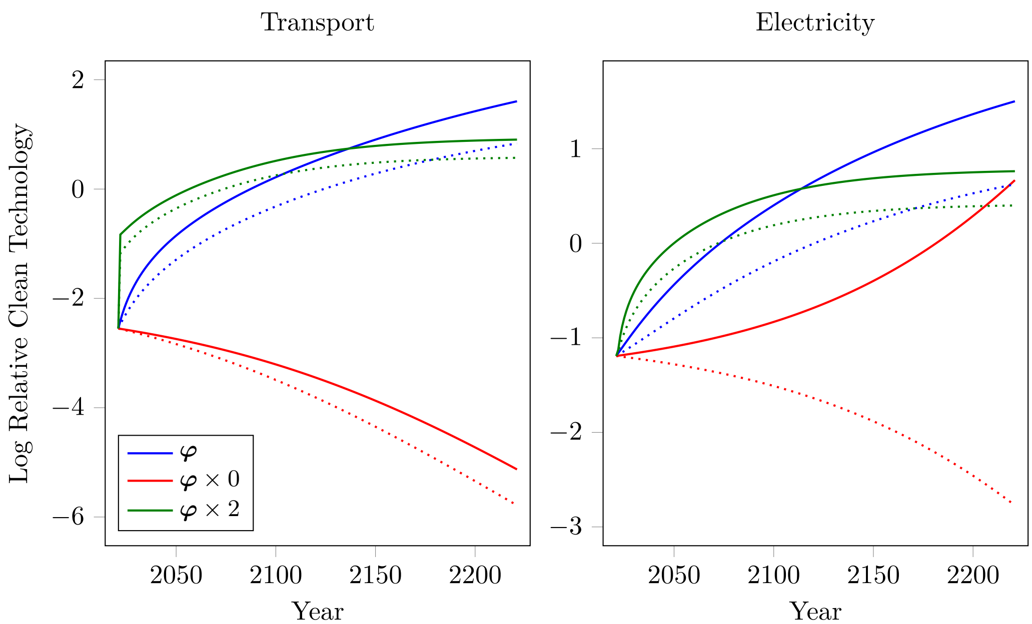
The second panel of Table 3 considers the transitional impacts of the policy reform. Using Proposition 2, I can characterize the speed of transition to the new steady-state.[64] For my calibrated spillover network, increasing returns to innovation are strong enough to create a slow transition but are not so strong that they generate path dependence. Following the policy reform, transportation and electricity generation converge halfway to their steady-states in 121 and 127 years, respectively. The policy reform's impact on

[64] Given that these results involve linearizing my model, one may be concerned that my results are not informative about such a large policy reform. To address this concern, I compare the linearized transition path to the full simulation in Figure C.3 and show that the two do not differ significantly.

both input prices and the path of technology reduces emissions by about half in 2035 and 2060, relative to the laissez-faire level in those years.[65]

As before, the other two cases illustrate how cross-technology spillovers shape the speed with which policy can influence the direction of innovation. When the spillover network is shut down, the economy doesn't converge to an interior steady-state, making the half-lives of convergence undefined. Moreover, the long-run emission reductions from pushing electricity generation into its clean basin of attraction are achieved only slowly because clean electricity cannot enjoy catchup growth. This is reflected in the weak emission reductions in the transition. Conversely, when cross-technology spillovers are doubled, the transition speeds up substantially, with comparatively quick half-lives of convergence of 20 years and 25 years. In the near term, the more rapid technological transition allows for greater emission reductions, but this lead is eventually lost as high cross-technology spillovers allow dirty technologies to stay relatively advanced.

Figure 5: Technology Path



Notes: Impact of introducing a carbon price at the Biden Administration's estimate of the SCC ($51) and clean innovation subsidy equivalent to a 30% tax credit ($\xi_c = 1.43$). Dotted lines indicate laissez-faire paths.

Figures 5, 6, and 7 provide further detail on the impact of the policy reform. The dotted lines indicate the paths that would have been taken in the absence of reform. With calibrated spillovers, the policy reform allows clean technologies to slowly gain a lead over their dirty counterparts as the economy transitions to a new, cleaner steady-state. This induced improvement in clean technology leads to greater clean incomes shares and reductions in total emissions over time, relative to laissez-faire.[66]

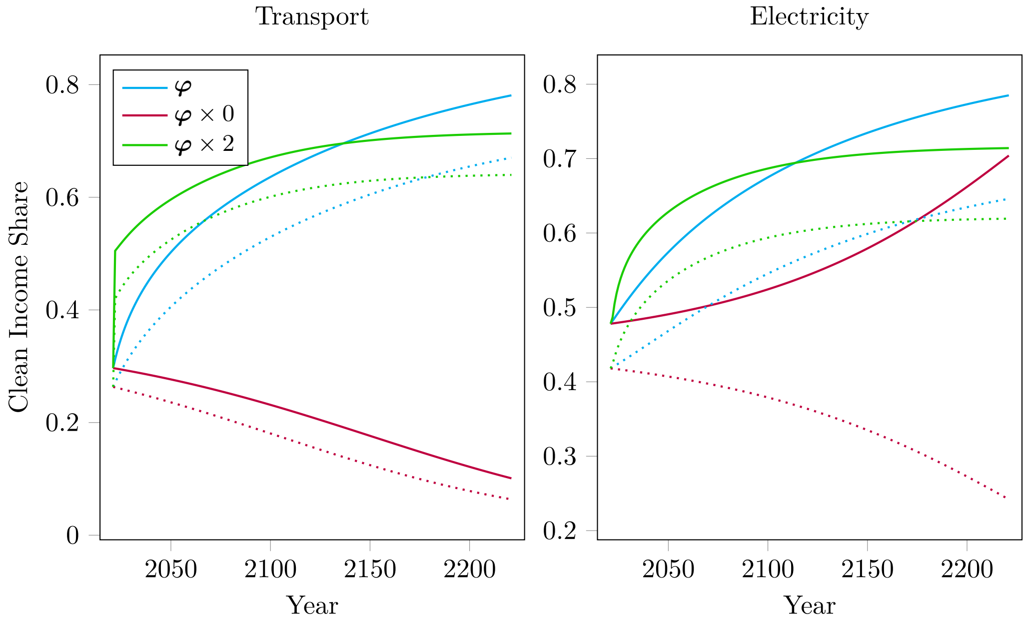
When the spillover network is shut down, the policy reform enables a slow shift toward clean technology for electricity generation while only delaying the advance of transportation into its dirty basin

[65] The emission reductions presented here abstract from differences in climate damages brought about by the different emission paths.

[66] Absolute emissions still rise steadily in the long-run because a fixed carbon price is eventually overcome by economic growth.

of attraction. The high increasing returns to innovation, to the point of path dependence, bifurcate the impact of policy. For large enough policy reforms, like in the case of electricity generation, the shift in the direction of innovation will be slow, but in the long-run, the impact will be massive. This contrasts sharply with the case where cross-technology spillovers are doubled. High spillovers allow for a rapid response to the policy reform, but the overall transition eventually peters out due to the small long-run change.

Figure 6: Clean Income Share Path

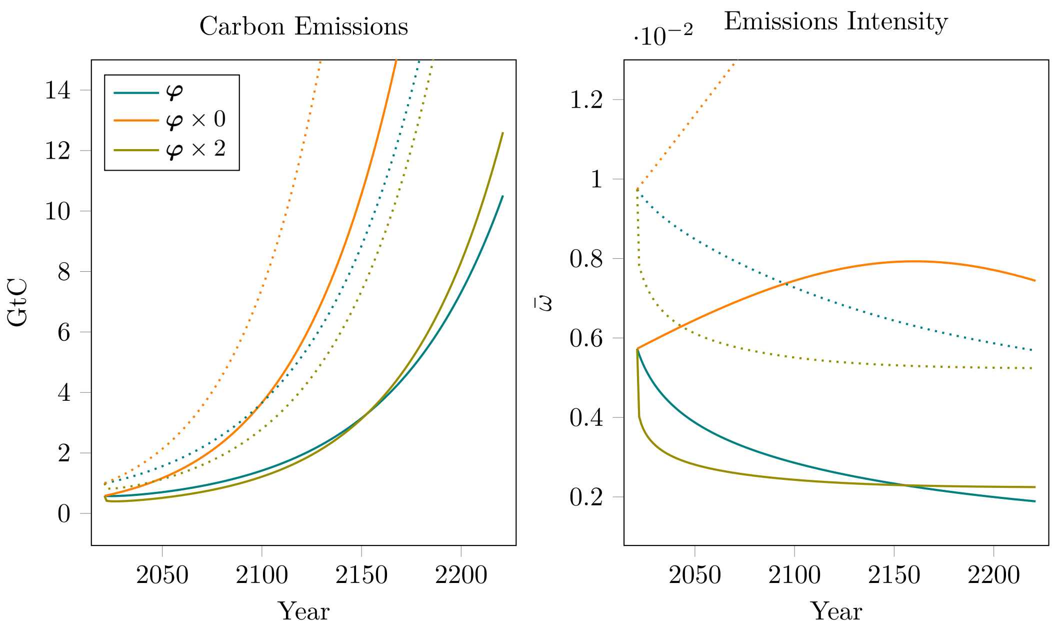


Notes: Impact of introducing a carbon price at the Biden Administration's estimate of the SCC (\$51) and clean innovation subsidy equivalent to a 30% tax credit ($\xi_c = 1.43$). Dotted lines indicate laissez-faire paths.

These differences showcase how both the long-run and transitional impacts of a policy reform are shaped by cross-technology spillovers. High cross-technology spillovers allow policy to shape the direction of innovation quickly by allowing less advanced technologies to achieve catchup growth. However, the same mechanism implies smaller long-run effects of policy as any advantage granted to some technologies will trickle down to others via spillovers. These effects are summarized by the degree of increasing returns to innovation, as measured by the spectral radius of the transition matrix. Low increasing returns allow for rapid transitions in response to policy reforms, but as argued in Corollary 2, the forces that prevent increasing returns also reduce the long-run impact of policy by preventing any technology from gaining a significant advantage.

This section has focused on cross-technology spillovers as a source of variation in increasing returns to innovation, but similar results hold for substitution patterns in production. In Figure C.4, I show how the steady-state impact of the policy reform varies in the level of cross-technology spillovers ζ and the elasticity of substitution between clean and dirty goods σ. As I have argued, the steady-state impact reduces as ζ increases. Similarly, the steady-state impact increases with the elasticity of substitution σ because higher substitutability allows clean goods to gain greater market share when favored by policy. In both cases, higher increasing returns to innovation increase the steady-state impact of policy while at the same time increasing the time needed for policy to take effect by slowing the transition. Figure C.5 illustrates how the degree of increasing returns to innovation varies in ζ and σ. As I have argued throughout this paper, cross-technology spillovers reduce increasing returns, with technology becoming path dependent if cross-technology spillovers are reduced to 66% of their calibrated value. Similarly, substitutability raises increasing returns, with technology becoming path dependent if the elasticity of substitution σ rises to 2.3.

Figure 7: Pollution Path



Notes: Impact of introducing a carbon price at the Biden Administration's estimate of the SCC ($51) and clean innovation subsidy equivalent to a 30% tax credit ($\xi_c = 1.43$). Dotted lines indicate laissez-faire paths.

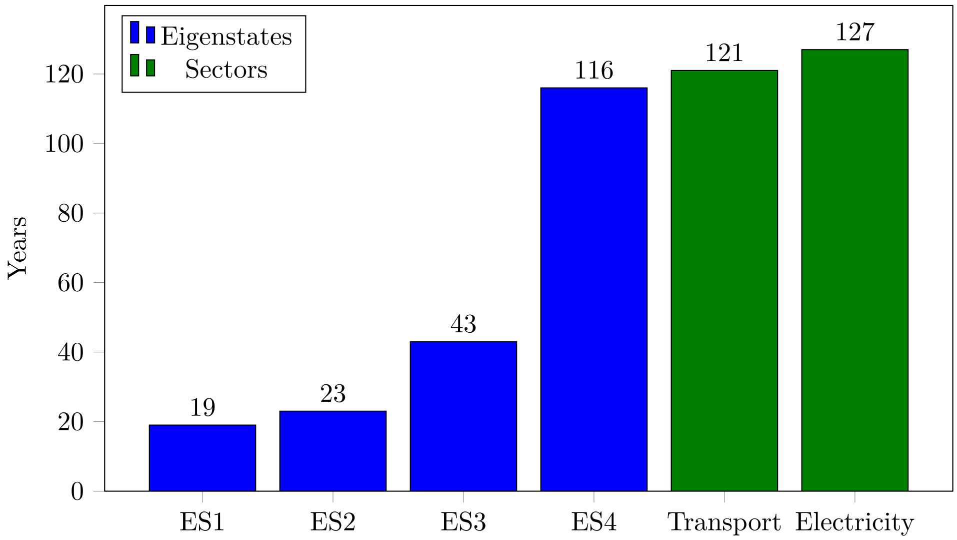
The EPA has recently proposed increasing the federal SCC to $190. I examine the impact of introducing a carbon price of $190 in Table C.1 as well as Figures C.6, C.7, and C.8. The results are qualitatively similar but quantitatively larger. In fact, a $190 carbon price is still not large enough to shift transportation into its clean basin of attraction when the spillover network is shut down. Figures C.9 and C.10 show the carbon prices and clean innovation subsidies necessary to shift each technology into its clean basin of attraction when the spillover network is shut down. With a clean innovation subsidy equivalent to a 30% tax credit ($\xi_c = 1.43$), the electricity generation sector would shift to clean innovation without a carbon price, while the transportation sector would require a carbon price of $245. With a carbon price at the Biden Administration's estimate of the SCC ($51), the electricity generation sector would shift to clean innovation with a clean innovation subsidy of $\xi_c = 1.04$, while the transportation sector would require an innovation subsidy of $\xi_c = 1.62$.[67]

Figure 8 shows the half-lives of convergence for each eigenstate and sector of the economy.[68] Proposition 3 allows me to unpack the substitution patterns and connections in the spillover network that drive each sector's convergence speed. Each eigenstate represents a region of state space with a distinct con-

[67] These innovation subsidies are equivalent to tax credits of 3.8% and 38.3%, respectively.

[68] I perform the same exercise for the case where cross-technology spillovers are doubled in Figure C.11.

Figure 8: Half-Lives of Convergence



Notes: Transition speeds following the introduction of a carbon price at the Biden Administration's estimate of the SCC (\$51) and clean innovation subsidy equivalent to a 30% tax credit ($\xi_c = 1.43$).

vergence speed, and Figure 8 shows that the slowest eigenstate drives the speed of convergence for both sectors. That is, technology's initial condition loads primarily on the eigenstate representing the region of state space where convergence is especially slow. This eigenstate represents initial conditions where dirty technology has an economy-wide advantage over clean technology; exactly the initial condition that policymakers must contend with.

What about this initial condition leads to slow convergence? First, the substitutability of clean and dirty goods makes switching the predominant technology in any sector slow as the small market size of the laggard will reduce innovation rents. However, this is not the whole story because the third eigenstate, which is substantially faster, represents initial conditions where dirty technology is advanced in one sector and clean technology is advanced in another. Moving from dirty to clean technology in both sectors concurrently is especially slow because, as argued in Section 5.1, clean and dirty technologies both form their own spillover clusters. Thus, catchup growth for clean technologies is reduced because clean technologies receive spillovers from their peers who are themselves trying to catch up. In spite of the forces that generate a slow transition, both sectors avoid path dependence. Overall, cross-technology knowledge spillovers are large enough to stimulate sufficient catchup growth for clean technologies. In particular, both clean technologies receive high spillovers from outside of the transportation and electricity generation sectors. Without those spillovers, the spectral radius would increase to 1.007, inducing path dependence.[69]

Another noteworthy feature of Figure 8 is that transportation converges a bit more quickly than electricity generation, with electricity generation taking about 5% longer to halve its log distance from steady-state. This difference is due to the higher within-sector spillovers in transportation. Another feature of the spillover network visible in Figure 2 is that clean transport receives spillovers from dirty

[69] To shut down spillovers from the general sector, I transfer all of the general spillover elasticities to the diagonal of the gross spillover network $\tilde{\boldsymbol{\varphi}}$. I take a similar approach to shutdown within-sector spillovers, transferring all of the within-sector spillover elasticities to the diagonal.

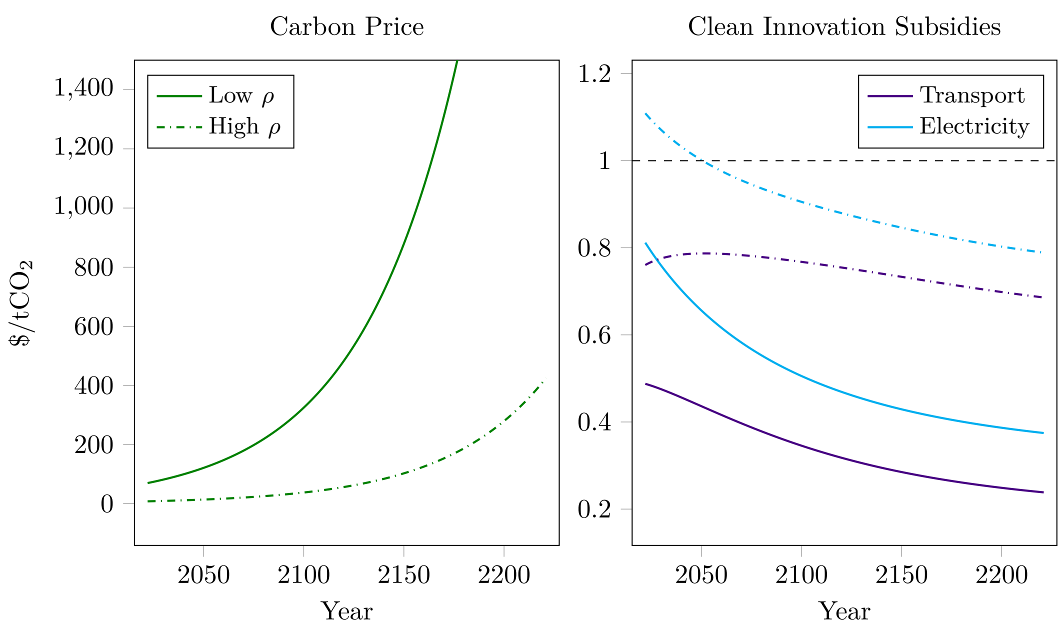
transport, while there are almost no spillovers from dirty electricity to clean electricity. This is consistent with the spillover anecdotes from my introduction, where electric vehicles receive spillovers from dirty vehicles, but solar panels receive spillovers from the ICT sector. To quantify the importance of within-sector spillovers, Figure C.12 shows half-lives of convergence when within-sector spillovers are shut down. Without these spillovers, the spectral radius rises to 0.998, and the half-lives for transportation and electricity generation rise to 434 years and 376 years, respectively. The loss of within-sector spillovers substantially slows the overall speed of convergence, but more importantly, the sectors switch their convergence speed rankings, signifying that transportation's more rapid transition stems from within-sector spillovers.

Finally, one may wonder whether the degree of increasing returns to innovation depends on the size of the policy reform. Figure C.13 shows that the eigenvalues of the transition matrix, and therefore the spectral radius, are not sensitive to the size of the policy reform. Thus, variation in transition speed across different policy reforms comes from differences in the loading of the initial condition on the eigenstates β, rather than the size of the eigenvalues $\{\kappa_j\}$ themselves.

6.2 First-Best Policy

In this section, I simulate the policy path that generates an optimal clean transition, starting in 2022.[70] As I have emphasized, optimal policy considers the spillovers a technology sends, rather than receives, because they create a benefit that innovators fail to internalize. Clean technologies create relatively few spillovers, so for this reason, clean innovation subsidies play a small role in the clean transition. Instead, carbon emissions are brought down primarily through pricing pollution.

Figure 9: First-Best Policy Path



Notes: Innovation subsidies are listed as a fraction of the baseline innovation wedge.

[70] I provide further details on my solution method in Appendix B.3.

Figure 9 shows the first-best policy path for two values of the discount rate. The left panel displays the optimal carbon price. As expected, a lower discount rate increases the carbon price by increasing the importance of future climate damages. The difference is substantial, with the less patient Planner taking more than a century to reach the more patient Planner's initial carbon price.

The right panel displays clean innovation subsidies as a fraction of the baseline innovation wedge. Strikingly, clean technologies do not receive favorable innovation subsidies. The only case where a clean technology receives an innovation subsidy above the baseline innovation wedge is clean electricity for the first few decades of policy when the discount rate is high. As shown in Proposition 4, the primary function of innovation subsidies is to correct the positive knowledge spillover externality. Clean technologies are valued for their ability to produce goods without pollution, not necessarily for their ability to produce spillovers. On the latter metric, they don't score particularly well. Clean transport's eigenvector centrality in the spillover network is barely more than a tenth of its steady-state income share (0.64 vs. 5.2%), while clean electricity's eigenvector centrality is slightly less than a quarter of its steady-state income share (0.91 vs. 3.7%).[71] This is why, in both cases, the innovation subsidies for clean electricity are uniformly higher than those for clean transport; the former is the superior clean technology for producing spillovers. Alternatively, dirty transport and electricity have eigenvector centralities of 1.14 and 0.83, respectively, relative to steady-state income shares of zero.

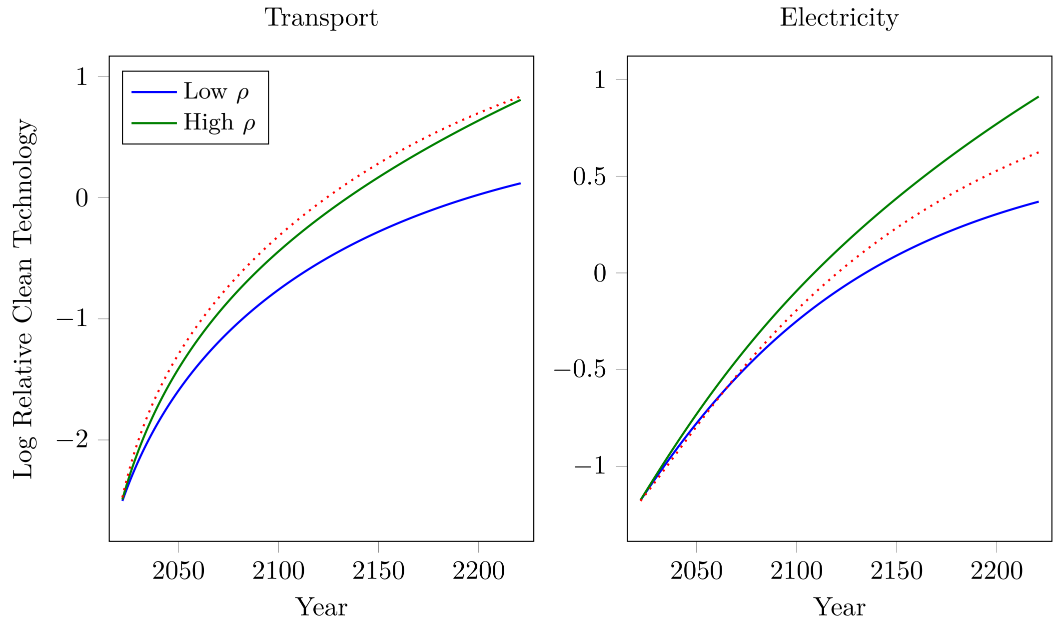
Moreover, a lower discount rate *reduces* clean innovation subsidies because, as shown in Corollary 3, a more patient Planner places more weight on spillover creation, rewarding technologies more for their ability to produce ideas via spillovers, rather than their ability to produce physical goods. The more patient Planner is more concerned about future climate damages, but this shows up in their higher carbon price. With the social cost of carbon properly priced, there is no reason for them to give additional support to clean technologies as such. For both values of the discount rate, clean innovation subsidies decline toward their steady-state value over time. This slope stems from the fact that clean income shares increase to their steady-state value over time. Thus, the private reward for clean innovation in early periods is low relative to the forward-looking social value, warranting a front-loaded path for clean innovation subsidies.

Figures 10, 11, and 12 provide further details on the first-best allocation. The dotted lines indicate laissez-faire paths. For the path of relative clean technology, optimal policy does not unambiguously favor clean technology. For transportation, the laissez-faire path actually has the highest relative clean technology, and for both transportation and electricity generation, the patient Planner ends up with the lowest relative clean technology. This again reflects the relatively low spillovers generated by clean technologies. Because clean technologies are actually disfavored by innovation subsidies, they do not gain as much of a lead over their dirty counterparts. This is especially true for transportation, where the clean technology is particularly bad at generating spillovers, and for the more patient Planner, who puts greater value on the generation of spillovers.

Nonetheless, the steadily rising price on carbon allows clean technology to slowly take over production in both sectors. For both levels of discounting, within-sector clean income shares rise to steady-state values of one, but this happens more quickly for the more patient Planner, who imposes a higher price on carbon. Clean technology starts at a greater disadvantage in the transportation sector, but catchup

[71] I have scaled eigenvector centrality to sum to 100, making it comparable to a percent. The steady-state income shares I have listed come from the low discounting allocation.

Figure 10: First-Best Technology Path



Notes: Dotted lines indicate laissez-faire paths.

growth via the spillover network allows clean transport to achieve a similar within-sector income share as clean electricity by 2200.

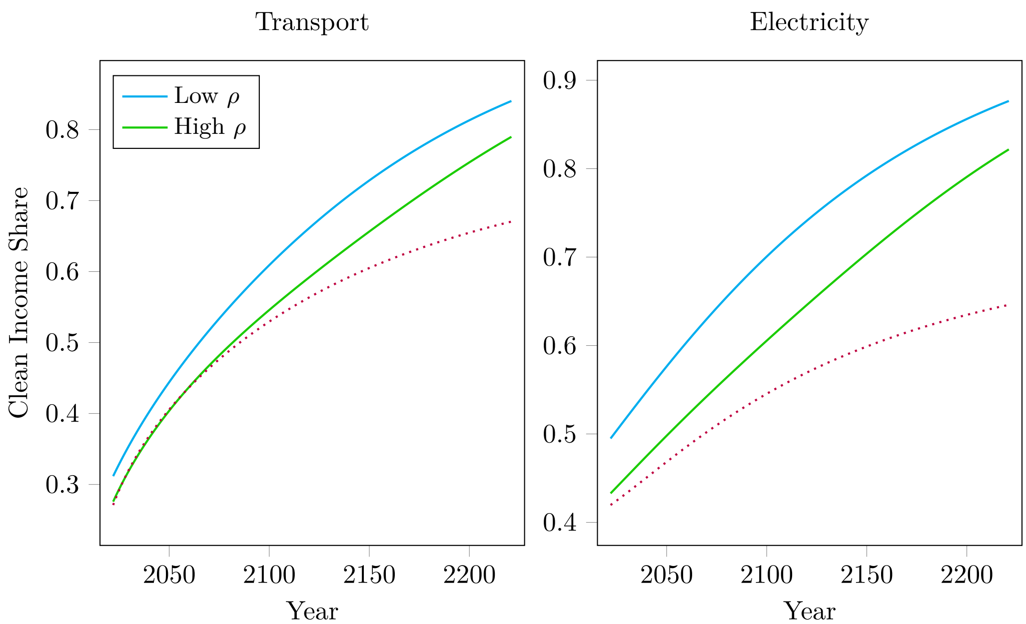
Finally, the carbon price substantially reduces carbon emissions, relative to laissez-faire. The more patient Planner immediately reduces carbon emissions by about half and steadily reduces them thereafter, whereas the less patient Planner allows carbon emissions to rise for a little more than a century. In both cases, the economy continues to grow, leading to steady reductions in emissions intensity. Together, these figures show that the Planner does not reduce emissions by abandoning dirty technologies. Instead, they push dirty technology out of production by raising the price on carbon, while at the same time continuing research in dirty technologies for the sake of spillovers.

In Figure C.14, I plot the temperature increases associated with first-best policy. Due to my focus on the transportation and electricity generation sectors in the US, most of the emissions that determine the global increase in temperature come from outside of my model. In spite of this, US climate policy still has a substantial influence on the level of warming beyond 2100. Warming peaks around 3°C or 3.2°C, depending on the discount rate, but in laissez-faire, warming continues unabated as the economy grows.

The damages we can expect from a warmer climate are highly uncertain, so I consider the possibility of substantially higher damages in Figure C.15. For that case, I quadruple the damage parameter ϱ, which substantially increases the social cost of carbon. Interestingly, despite the increased cost of climate change, clean innovation subsidies are reduced even further. The higher price on carbon leads to higher clean income shares, raising the private return to clean innovation. What does not change is the relatively low centrality of clean technologies in the spillover network. Thus, the social value of clean innovation increases by less than the private return, implying a reduction in innovation subsidies.

Finally, to further examine the role of spillovers in optimal policy, I shut down the spillover network in Figure C.16. In that case, carbon prices are largely unchanged, but clean technology receives a ”big

Figure 11: First-Best Clean Income Share Path



Notes: Dotted lines indicate laissez-faire paths.

push" with massive, yet temporary, innovation subsidies. Now innovation subsidies for clean transport and electricity start at more than triple and double the baseline innovation wedge, respectively, and slowly decrease down to the baseline level over the first century of policy. This is in sharp contrast to innovation policy in the previous case because, without a spillover network, clean technologies are beneficial in terms of both their ability to produce goods without pollution and their ability to produce spillovers. Before every technology could create spillovers on clean technologies, but now that only clean technologies can create clean spillovers, the usefulness of clean technology in production grants an exclusive value to their spillovers as well.

6.3 Second-Best Innovation Subsidies

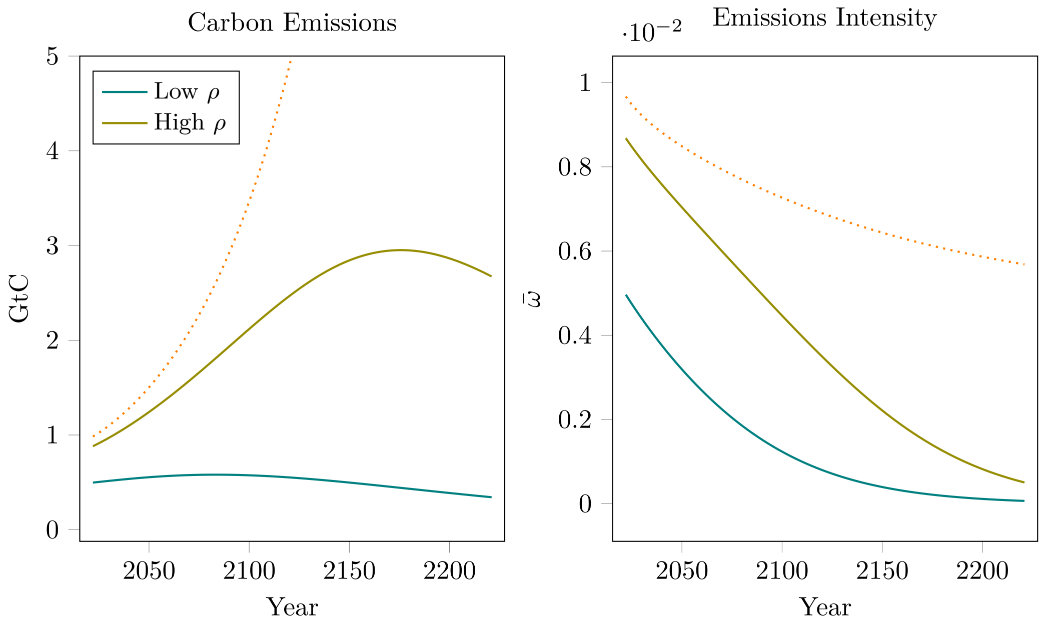
In this section, I simulate second-best innovation subsidies, where carbon prices are incomplete. I consider two cases: one where the external carbon price is zero and another where the external carbon price is ten percent of the social cost of carbon. I show that these cases differ substantially as even a small, but growing, price on carbon allows the Planner to achieve welfare similar to that of the first-best, but without a price on carbon, society suffers massive welfare losses relative to the first-best. As before, the policy simulation starts in 2022.[72] In the main text, I take my preferred specification of low discounting, but I consider high discounting as a robustness check.[73]

Figure 13 displays second-best innovation subsidies, as a fraction of the baseline innovation wedge,

[72] Appendix B.4 contains further details on my solution method.

[73] Another justification for low discounting of climate damages stems from the pricing of risk. Because climate damages are risky, limiting emissions is a form of insurance. Hence, abatement should carry a low, or even negative, risk premium, like that of a hedging asset (Wagner and Weitzman, 2016). I do not explicitly consider climate risk, so low discounting captures some of the insurance argument for limiting emissions.

Figure 12: First-Best Pollution Path



Notes: Dotted lines indicate laissez-faire paths.

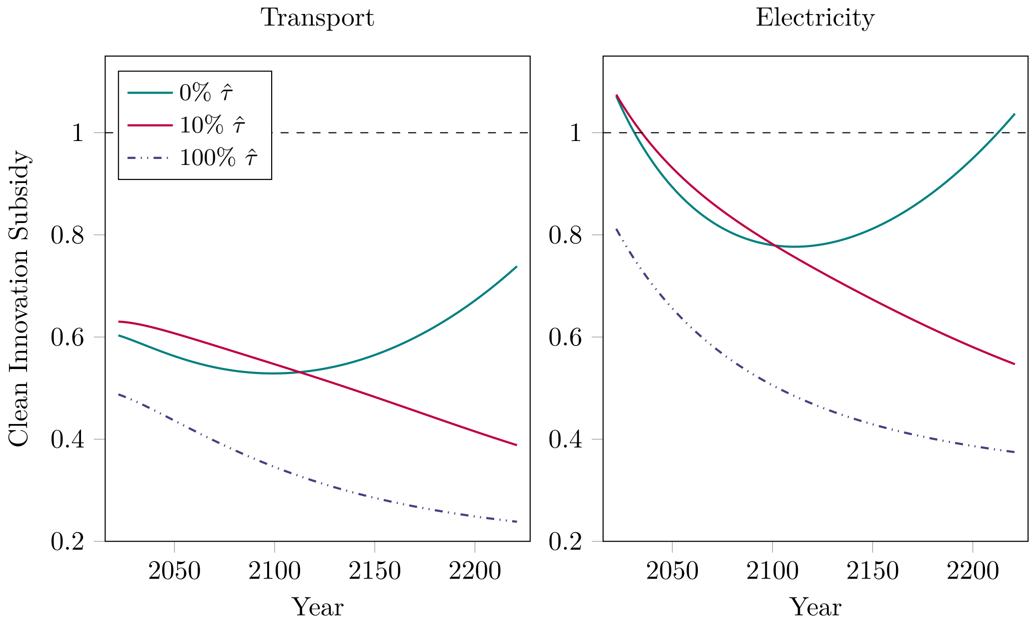
with the external carbon price $\hat{\tau}_t$ set as a proportion of the social cost of carbon. As shown in Proposition 5, innovation subsidies adjust to accommodate the distortion of incomplete carbon pricing. The dashed-dotted lines reference the first-best policy path. First, we can see that the case with $\hat{\tau}_t$ set to ten percent of the social cost of carbon is qualitatively similar to the first-best, with clean innovation subsidies shifted upwards by about 20-30% of the baseline innovation wedge. In that case, rising carbon prices eventually push dirty technology out of production, so the Planner can still develop dirty technologies to some degree for the sake of spillovers.

This contrasts sharply with the case where carbon pollution is entirely unpriced. In that case, clean innovation subsidies are similar to the ten percent case for the first century of policy, but beyond that, they begin to rise precipitously. Moreover, dirty innovation is immediately shut down. Without a carbon price, the only way for the Planner to push dirty technology out of production is to shut down dirty innovation altogether, forgoing spillovers in the process.

Figures 14, 15, and 16 provide further details on second-best policy. The dotted lines indicate laissez-faire paths. The difference between the two second-best cases emerges clearly in the path of relative clean technology. Again, a small, growing carbon price allows the Planner to continue dirty innovation for the sake of spillovers, so relative clean technology is only marginally higher than in the first-best. However, without a carbon price, the Planner must drive relative clean technology to infinity to push dirty production out of the economy. This opens a vast difference in relative clean technology for the zero case, relative to the others.

The substantial increase in relative clean technology in the zero case leads to similar clean income shares as in the first-best. Indeed, clean income shares in the zero case are larger than in the ten percent case. However, without a carbon price, emissions in the zero case are only slightly below their level in laissez-faire for the first few centuries. This is because increases in clean technology lead to expansions in

Figure 13: Second-Best Policy Path



Notes: Policy paths use the low discount rate. The external carbon price $\hat{\tau}$ is a proportion of the social cost of carbon, so 100% is the first-best. Innovation subsidies are listed as a fraction of the baseline innovation wedge.

output, which increase demand for dirty inputs in turn. Thus, reducing emissions via clean technological growth alone is ineffective because the substitution from clean to dirty is dampened by a rebound effect. Emissions in the ten percent case are still substantially larger than in the first-best, but a small carbon price goes a long way in preventing technological advance from translating into increases in emissions.

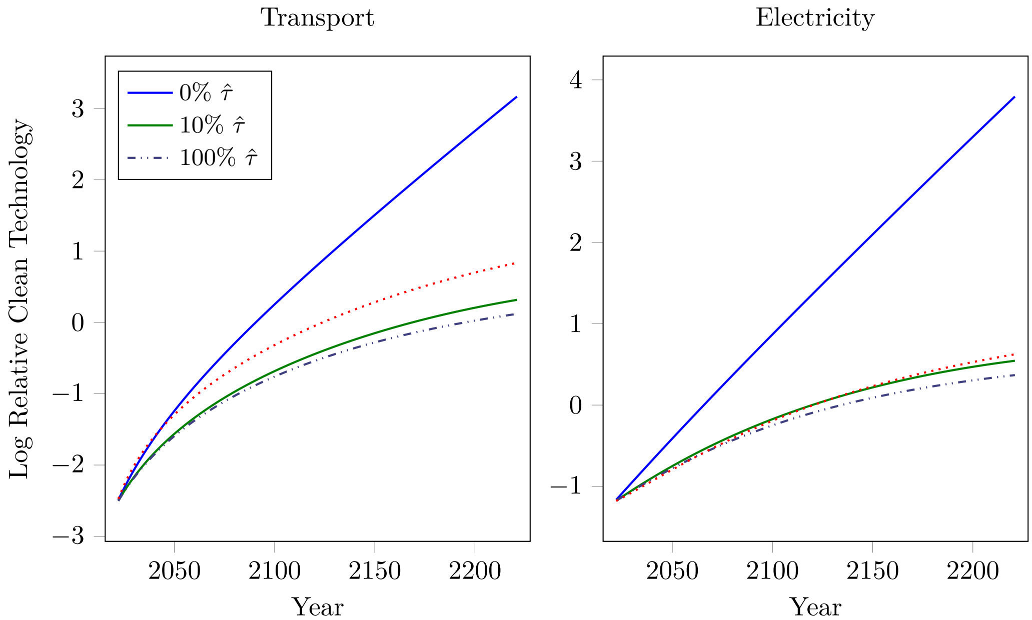
In Figure C.17, I show the path of temperature increases associated with each second-best policy path. As before, most global emissions come from outside of the model, but US climate policy has a substantial influence on the level of warming beyond 2100. As suggested by the path of emissions, the ten percent case leads to a fair amount of additional warming, with maximum warming of about 3.5°C, roughly half a degree higher than in the first-best. However, the zero case leads to disastrous warming of about 10°C and rising by 2500, underscoring the power of even a small carbon price.

These simulations illustrate the main tradeoff of second-best policy. Technology stocks are an input into the production of both physical goods and new ideas. As shown in the first-best, clean technologies should dominate production because they can produce goods without pollution, but in the production of ideas, dirty technologies are generally better due to their higher centrality in the spillover network. When the Planner can only control innovation subsidies, they must pick a path for technology that compromises between the two. A growing carbon price ensures that dirty technology is eventually pushed out of production, allowing dirty innovation to continue without increasing emissions. However, when pollution is unpriced, any dirty innovation will raise emissions, so the Planner must stop developing the dirty technology altogether.

In the end, decarbonizing by only improving clean technology leads to both slow emission reductions from a rebound effect and slow growth from the loss of spillovers.[74] For this reason, the welfare loss

[74] In the steady-state, the loss of dirty innovation leads growth to stop altogether. This stems from using a Cobb-Douglas

Figure 14: Second-Best Technology Path



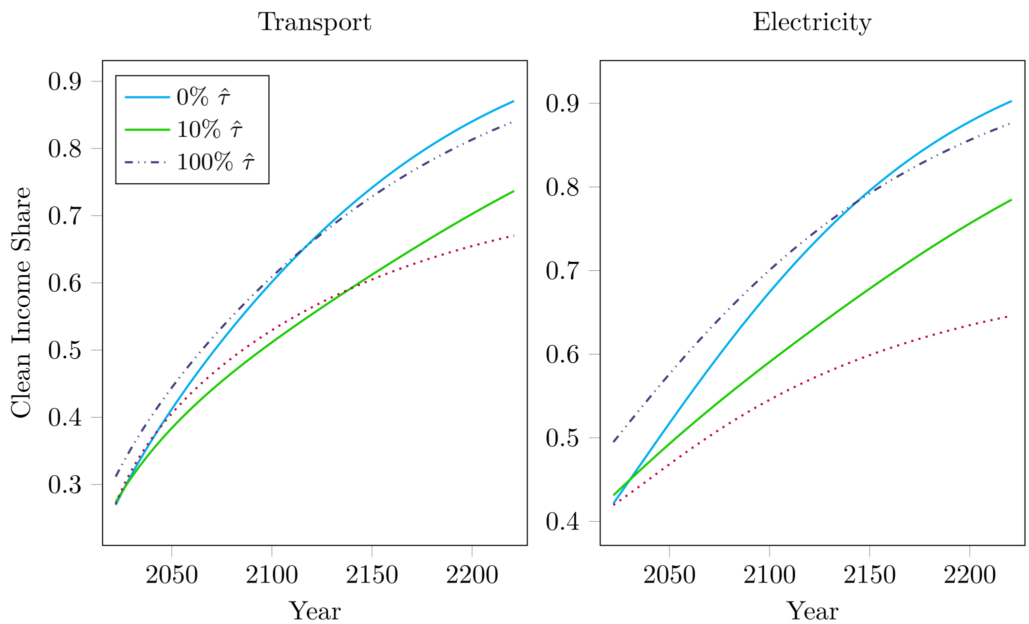
Notes: Policy paths use the low discount rate. The external carbon price $\hat{\tau}$ is a proportion of the social cost of carbon, so 100% is the first-best. Dotted lines indicate laissez-faire paths.

of moving from the first-best to the second-best increases sharply when carbon pollution is unpriced. Pricing carbon pollution at ten percent of the social cost of carbon leads to a consumption equivalent loss of 1.5%, implying that a small, growing carbon price can achieve most of the welfare of the first-best. However, if carbon pollution is unpriced, the consumption equivalent loss is almost 100%. Clearly, the clean technological transition requires both instruments so as to decouple the production of physical goods and new ideas.

Figure C.18 shows second-best innovation policy in the case of high discounting. The conclusions are broadly similar. Clean innovation subsidies are higher than in the first-best in both the zero and ten percent case, and those of the zero case grow substantially over time. However, the difference between the zero and ten percent case is much less stark for the less patient Planner. For instance, dirty innovation is now allowed to continue for about another 150 years in the zero case. This reduced difference stems from the fact that the higher emissions and slower growth of the zero case are both costs that materialize in the future. A less patient Planner puts less weight on these costs, with a consumption equivalent loss of 0.3% for the ten percent case and 0.9% for the zero case, relative to the first-best. Therefore, the cost of incomplete carbon pricing depends critically on the Planner's degree of patience.

spillover function in my simulations. Put differently, when one or more technologies are lost, the model switches from endogenous growth to semi-endogenous growth. If the spillover function had an elasticity of substitution above one, growth could continue without dirty innovation, though the abrupt stop in dirty innovation would still dampen growth.

Figure 15: Second-Best Clean Income Share Path



Notes: Policy paths use the low discount rate. The external carbon price $\hat{\tau}$ is a proportion of the social cost of carbon, so 100% is the first-best. Dotted lines indicate laissez-faire paths.

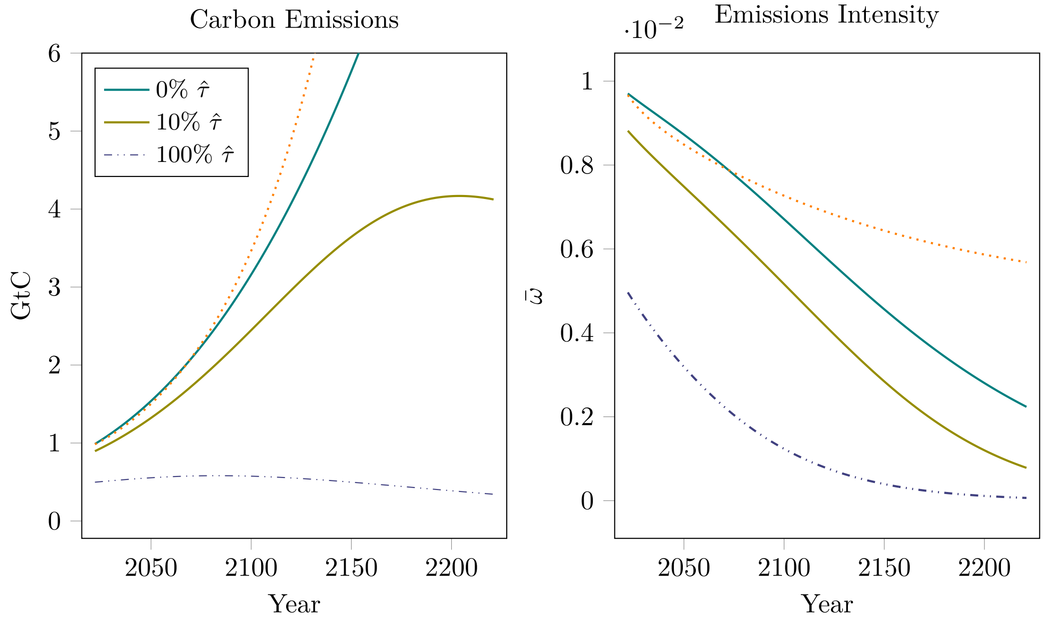
7 Conclusion

One of the main goals of climate policy is to redirect innovation from dirty to clean technology. Without such redirection, decarbonization will be prohibitively expensive. This paper has studied the role of policy in the transition to clean technology. I have argued that the two main forces that govern both the size and speed of technological transition, following a policy reform, are substitution patterns in productions and the network of knowledge spillovers. Moreover, I have shown that each of these forces can be summarized in terms of sufficient statistics matrices, providing a mapping from the model to the data. These two sufficient statistic matrices govern both the steady-state change and transition path by controlling the degree of increasing returns to innovation, which can be quantified in terms of a spectral radius. Interestingly, the long-run change in the direction of innovation is inversely related to the speed of transition because the mechanisms that speed up transitions also prevent technologies from gaining large leads over their peers in the long run.

For optimal policy, I have characterized the path of carbon prices and innovation subsidies that generate an optimal clean technological transition. In keeping with the Pigou principle, carbon prices correct the pollution externality, while innovation subsidies correct the knowledge spillover externality. That is, clean innovation should not be subsidized simply because clean technologies can create goods without pollution; it should be subsidized insofar as it produces spillovers. However, in the case where carbon pricing is incomplete, I have shown that innovation subsidies can be straightforwardly adjusted to account for the social cost of carbon.

To take my model to the data, I have performed four quantitative exercises, focusing on transportation and electricity generation in the US. First, I showed that my model can explain the 2010s advance in clean

Figure 16: Second-Best Pollution Path



Notes: Policy paths use the low discount rate. The external carbon price $\hat{\tau}$ is a proportion of the social cost of carbon, so 100% is the first-best. Dotted lines indicate laissez-faire paths.

technology when it includes the spillover network, but in the absence of cross-technology spillovers, my model counterfactually predicts that clean technology would have instead fallen further behind. Second, I showed that introducing a realistic carbon price and clean innovation subsidy has a large long-run effect on the direction of innovation but that this effect takes a long time to realize. Underpinning both the sizable long-run impact and sluggish transition is a spectral radius near one, indicating increasing returns to innovation that are large, but not so large as to induce path dependence. Third, I have shown that optimal climate innovation policy reduces emissions primarily through carbon prices, rather than clean innovation subsidies. Clean technologies do not receive favorable subsidies because they do not produce particularly large spillovers, as measured by their centrality in the spillover network. Fourth, I simulated second-best innovation subsidies and showed that if there is a small, growing carbon price, then the second-best can achieve most of the welfare of the first-best. However, if carbon pollution is unpriced, then decarbonization requires shutting down dirty innovation and relying exclusively on clean innovation, with disastrous consequences for welfare and the climate.

I have applied my model to the context of climate change because that is a salient case where policy must redirect innovation. However, my model is general and can be viewed as a framework for studying the impact of economic shocks – to policy, input prices, etc – on the direction of innovation. I view many of the insights of this paper as applying to a more general set of questions about the determination of the direction of innovation. I leave the study of those questions to future work.

References

Acemoglu, D. (1998). Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality. *The Quarterly Journal of Economics*, 113(4):1055–1089.

Acemoglu, D. (2002). Directed Technical Change. *The Review of Economic Studies*, 69(4):781–809.

Acemoglu, D. (2023). Distorted Innovation: Does the Market Get the Direction of Technology Right? *AEA Papers and Proceedings*, 113:1–28.

Acemoglu, D., Aghion, P., Barrage, L., and Hémous, D. (2023). Climate Change, Directed Innovation, and Energy Transition: The Long-Run Consequences of the Shale Gas Revolution. Working Paper 31657, National Bureau of Economic Research.

Acemoglu, D., Aghion, P., Bursztyn, L., and Hémous, D. (2012). The Environment and Directed Technical Change. *American Economic Review*, 102(1):131–66.

Acemoglu, D., Akcigit, U., Alp, H., Bloom, N., and Kerr, W. (2018). Innovation, Reallocation, and Growth. *American Economic Review*, 108(11):3450–91.

Acemoglu, D., Akcigit, U., Hanley, D., and Kerr, W. (2016). Transition to Clean Technology. *Journal of Political Economy*, 124(1):52–104.

Acemoglu, D. and Finkelstein, A. (2008). Input and Technology Choices in Regulated Industries: Evidence from the Health Care Sector. *Journal of Political Economy*, 116(5):837–880.

Acemoglu, D. and Linn, J. (2004). Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry. *The Quarterly Journal of Economics*, 119(3):1049–1090.

Acemoglu, D. and Restrepo, P. (2018). The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment. *American Economic Review*, 108(6):1488–1542.

Acemoglu, D. and Restrepo, P. (2022a). Demographics and Automation. *The Review of Economic Studies*, 89(1):1–44.

Acemoglu, D. and Restrepo, P. (2022b). Tasks, Automation, and the Rise in U.S. Wage Inequality. *Econometrica*, 90(5):1973–2016.

Aghion, P., Bénabou, R., Martin, R., and Roulet, A. (2023). Environmental Preferences and Technological Choices: Is Market Competition Clean or Dirty? *American Economic Review: Insights*, 5(1):1–20.

Aghion, P., Dechezleprêtre, A., Hémous, D., Martin, R., and Van Reenen, J. (2016). Carbon Taxes, Path Dependency, and Directed Technical Change: Evidence from the Auto Industry. *Journal of Political Economy*, 124(1):1–51.

Aghion, P., Hepburn, C., Teytelboym, A., and Zenghelis, D. (2019). Path Dependence, Innovation, and the Economics of Climate Change. In *Handbook on Green Growth*, pages 67–83. Edward Elgar Publishing.

Aghion, P. and Howitt, P. (1992). A Model of Growth Through Creative Destruction. *Econometrica*, 60(2):323–351.

Akcigit, U., Hanley, D., and Serrano-Velarde, N. (2020). Back to Basics: Basic Research Spillovers, Innovation Policy, and Growth. *The Review of Economic Studies*, 88(1):1–43.

Akcigit, U. and Kerr, W. R. (2018). Growth through Heterogeneous Innovations. *Journal of Political Economy*, 126(4):1374–1443.

Allen, T. and Donaldson, D. (2020). Persistence and Path Dependence in the Spatial Economy. Working Paper 28059, National Bureau of Economic Research.

Archer, D. (2005). Fate of Fossil Fuel CO2 in Geologic Time. *Journal of Geophysical Research: Oceans*, 110(C9).

Archer, D., Eby, M., Brovkin, V., Ridgwell, A., Cao, L., Mikolajewicz, U., Caldeira, K., Matsumoto, K., Munhoven, G., Montenegro, A., et al. (2009). Atmospheric Lifetime of Fossil Fuel Carbon Dioxide. *Annual Review of Earth and Planetary Sciences*, 37:117–134.

Arrow, K. J. (1962). Economic Welfare and the Allocation of Resources for Invention. In *The Rate and Direction of Inventive Activity*, pages 609–626. Princeton University Press.

Barrage, L. (2018). Be Careful What You Calibrate For: Social Discounting in General Equilibrium. *Journal of Public Economics*, 160:33–49.

Barrage, L. (2020). Optimal Dynamic Carbon Taxes in a Climate–Economy Model with Distortionary Fiscal Policy. *The Review of Economic Studies*, 87(1):1–39.

Barrage, L. and Nordhaus, W. D. (2023). Policies, Projections, and the Social Cost of Carbon: Results from the DICE-2023 Model. Working Paper 31112, National Bureau of Economic Research.

Bhagwati, J. (1958). Immiserizing Growth: A Geometrical Note. *The Review of Economic Studies*, 25(3):201–205.

Bhagwati, J. N. (1968). Distortions and Immiserizing Growth: A Generalization. *The Review of Economic Studies*, 35(4):481–485.

Blanchard, O. J. and Kahn, C. M. (1980). The Solution of Linear Difference Models under Rational Expectations. *Econometrica*, 48(5):1305–1311.

Bleakley, H. and Lin, J. (2012). Portage and Path Dependence. *The Quarterly Journal of Economics*, 127(2):587–644.

Bloom, N., Griffith, R., and Van Reenen, J. (2002). Do R&D Tax Credits Work? Evidence from a Panel of Countries 1979–1997. *Journal of Public Economics*, 85(1):1–31.

Bloom, N., Jones, C. I., Van Reenen, J., and Webb, M. (2020). Are Ideas Getting Harder to Find? *American Economic Review*, 110(4):1104–1144.

Bloom, N., Romer, P., Terry, S. J., and Van Reenen, J. (2021). Trapped Factors and China's Impact on Global Growth. *The Economic Journal*, 131(633):156–191.

Bloom, N., Schankerman, M., and Van Reenen, J. (2013). Identifying Technology Spillovers and Product Market Rivalry. *Econometrica*, 81(4):1347–1393.

Bloom, N., Van Reenen, J., and Williams, H. (2019). A Toolkit of Policies to Promote Innovation. *Journal of Economic Perspectives*, 33(3):163–84.

Blundell, R., Griffith, R., and Windmeijer, F. (2002). Individual Effects and Dynamics in Count Data Models. *Journal of Econometrics*, 108(1):113–131.

Böhringer, C. and Rutherford, T. F. (2009). Integrated Assessment of Energy Policies: Decomposing Top-Down and Bottom-Up. *Journal of Economic Dynamics and Control*, 33(9):1648–1661.

BP (2022). BP Statistical Review of World Energy, 71th Edition. Technical report.

Bryan, K. A. and Williams, H. L. (2021). Innovation: Market Failures and Public Policies. In *Handbook of Industrial Organization*, volume 5, pages 281–388. Elsevier.

Budish, E., Roin, B. N., and Williams, H. (2015). Do Firms Underinvest in Long-Term Research? Evidence from Cancer Clinical Trials. *American Economic Review*, 105(7):2044–85.

Cai, Y. and Lontzek, T. S. (2019). The Social Cost of Carbon with Economic and Climate Risks. *Journal of Political Economy*, 127(6):2684–2734.

Dasgupta, P. (2008). Discounting Climate Change. *Journal of Risk and Uncertainty*, 37:141–169.

Davis, S. and Boundy, R. G. (2022). Transportation Energy Data Book: Edition 40. Technical report, U.S. Department of Energy, Energy Vehicle Technologies Office, Oak Ridge National Laboratory.

Farhi, E. and Werning, I. (2007). Inequality and Social Discounting. *Journal of Political Economy*, 115(3):365–402.

Fried, S. (2018). Climate Policy and Innovation: A Quantitative Macroeconomic Analysis. *American Economic Journal: Macroeconomics*, 10(1):90–118.

Galor, O. (2007). *Discrete Dynamical Systems*. Springer.

Gans, J. S. (2012). Innovation and Climate Change Policy. *American Economic Journal: Economic Policy*, 4(4):125–145.

Gillingham, K., Nordhaus, W., Anthoff, D., Blanford, G., Bosetti, V., Christensen, P., McJeon, H., and Reilly, J. (2018). Modeling Uncertainty in Integrated Assessment of Climate Change: A Multimodel Comparison. *Journal of the Association of Environmental and Resource Economists*, 5(4):791–826.

Golosov, M., Hassler, J., Krusell, P., and Tsyvinski, A. (2014). Optimal Taxes on Fossil Fuel in General Equilibrium. *Econometrica*, 82(1):41–88.

Griliches, Z. (1990). Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, 28(4):1661–1707.

Grossman, G. M. and Helpman, E. (1991). Quality Ladders in the Theory of Growth. *The Review of Economic Studies*, 58(1):43–61.

Hall, B. and Van Reenen, J. (2000). How effective are fiscal incentives for r&d? a review of the evidence. *Research policy*, 29(4-5):449–469.

Hall, B. H. (1993). R&D Tax Policy during the 1980s: Success or Failure? *Tax Policy and the Economy*, 7:1–35.

Hall, B. H., Jaffe, A., and Trajtenberg, M. (2005). Market Value and Patent Citations. *The RAND Journal of Economics*, 36(1):16–38.

Hall, B. H. and Ziedonis, R. H. (2001). The Patent Paradox Revisited: An Empirical Study of Patenting in the U.S. Semiconductor Industry, 1979-1995. *The RAND Journal of Economics*, 32(1):101–128.

Hanlon, W. W. (2015). Necessity Is the Mother of Invention: Input Supplies and Directed Technical Change. *Econometrica*, 83(1):67–100.

Hassler, J., Krusell, P., and Olovsson, C. (2021). Directed Technical Change as a Response to Natural Resource Scarcity. *Journal of Political Economy*, 129(11):3039–3072.

Hicks, J. (1932). *The Theory of Wages*. London: Macmillan Company & Co.

Hulten, C. R. (1978). Growth Accounting with Intermediate Inputs. *The Review of Economic Studies*, 45(3):511–518.

Jaffe, A. B., Trajtenberg, M., and Henderson, R. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *The Quarterly Journal of Economics*, 108(3):577–598.

Jones, B. F. and Summers, L. H. (2021). A Calculation of the Social Returns to Innovation. In *Innovation and Public Policy*, pages 13–60. University of Chicago Press.

Jones, C. and Vollrath, D. (2013). *Introduction to Economic Growth*. W.W. Norton.

Jones, C. I. (1995). R&D-Based Models of Economic Growth. *Journal of Political Economy*, 103(4):759–784.

Jones, C. I. (2022). The Past and Future of Economic Growth: A Semi-Endogenous Perspective. *Annual Review of Economics*, 14:125–152.

Jones, C. I. and Williams, J. C. (1998). Measuring the Social Return to R&D. *The Quarterly Journal of Economics*, 113(4):1119–1135.

Kleinman, B., Liu, E., and Redding, S. J. (2023). Dynamic Spatial General Equilibrium. *Econometrica*, 91(2):385–424.

Krugman, P. (1991). Increasing Returns and Economic Geography. *Journal of Political Economy*, 99(3):483–499.

Krusell, P. and Smith, Anthony A, J. (2022). Climate Change Around the World. Working Paper 30338, National Bureau of Economic Research.

Känzig, D. R. (2023). The Unequal Economic Consequences of Carbon Pricing. Working Paper 31221, National Bureau of Economic Research.

Lanzi, E. and Sue Wing, I. (2011). Directed Technical Change in the Energy Sector: An Empirical Test of Induced Directed Innovation. *Working Paper*.

Lanzi, E., Verdolini, E., and Haščič, I. (2011). Efficiency-Improving Fossil Fuel Technologies for Electricity Generation: Data Selection and Trends. *Energy Policy*, 39(11):7000–7014.

Lemoine, D. and Traeger, C. (2014). Watch Your Step: Optimal Policy in a Tipping Climate. *American Economic Journal: Economic Policy*, 6(1):137–166.

Liu, E. and Ma, S. (2021). Innovation Networks and R&D Allocation. Working Paper 29607, National Bureau of Economic Research.

Moscona, J. and Sastry, K. (2022). Inappropriate Technology: Evidence from Global Agriculture. *Working Paper*.

Myers, K. R. and Lanahan, L. (2022). Estimating Spillovers from Publicly Funded R&D: Evidence from the US Department of Energy. *American Economic Review*, 112(7):2393–2423.

Newell, R. G., Jaffe, A. B., and Stavins, R. N. (1999). The Induced Innovation Hypothesis and Energy-Saving Technological Change. *The Quarterly Journal of Economics*, 114(3):941–975.

Nordhaus, W. D. (1992). An Optimal Transition Path for Controlling Greenhouse Gases. *Science*, 258(5086):1315–1319.

Nordhaus, W. D. (2002). Modeling Induced Innovation in Climate-Change Policy. In *Technological Change and the Environment*, pages 182–209. Routledge.

Nordhaus, W. D. (2007). A Review of the Stern Review on the Economics of Climate Change. *Journal of Economic Literature*, 45(3):686–702.

Nordhaus, W. D. (2010). Economic Aspects of Global Warming in a Post-Copenhagen Environment. *Proceedings of the National Academy of Sciences*, 107(26):11721–11726.

Nordhaus, W. D. (2017). Revisiting the Social Cost of Carbon. *Proceedings of the National Academy of Sciences*, 114(7):1518–1523.

O'Hara, F. M. (1990). *Glossary: Carbon Dioxide and Climate*. Oak Ridge National Laboratory.

Ord, T. (2020). *The Precipice: Existential Risk and the Future of Humanity*. Hachette Books.

Papageorgiou, C., Saam, M., and Schulte, P. (2014). Elasticity of Substitution between Clean and Dirty Energy Inputs - A Macroeconomic Perspective. *Working Paper*.

Parfit, D. (1984). *Reasons and Persons*. Oxford University Press.

Popp, D. (2002). Induced Innovation and Energy Prices. *American Economic Review*, 92(1):160–180.

Popp, D. (2004). ENTICE: Endogenous Technological Change in the DICE Model of Global Warming. *Journal of Environmental Economics and Management*, 48(1):742–768.

Romer, P. M. (1990). Endogenous Technological Change. *Journal of Political Economy*, 98(5, Part 2):S71–S102.

Roser, M. (2020). Why Did Renewables Become so Cheap so Fast? Our World in Data.

Stern, N. H. (2007). *The Economics of Climate Change: The Stern Review*. Cambridge University Press.

Uhlig, H. (1999). A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily. In *Computational Methods for the Study of Dynamic Economies*, pages 30–61. Oxford University Press.

Van der Werf, E. (2008). Production Functions for Climate Policy Modeling: An Empirical Analysis. *Energy Economics*, 30(6):2964–2979.

Wagner, G. and Weitzman, M. L. (2016). *Climate Shock: The Economic Consequences of a Hotter Planet*. Princeton University Press.

Weitzman, M. L. (2009). On Modeling and Interpreting the Economics of Catastrophic Climate Change. *The Review of Economics and Statistics*, 91(1):1–19.

Wilson, D. J. (2009). Beggar Thy Neighbor? The In-State, Out-of-State, and Aggregate Effects of R&D Tax Credits. *The Review of Economics and Statistics*, 91(2):431–436.

Appendix

A Theory Derivations & Proofs

A.1 Equilibrium Conditions

To derive relative prices of technology-specific goods, plug the intermediate limit price condition (20) into the intermediate demand condition (18) to obtain the equilibrium level of output for intermediates:

$$y_{j\iota t} = \frac{(1-\alpha)p_{jt}Y_{jt}\Upsilon a_{j\iota t}}{\gamma w_{\ell t}}. \tag{A.1}$$

Plugging this into the production function for technology-specific goods (2) gives us

$$Y_{jt} = \Lambda_{jt}\Big(\frac{(1-\alpha)p_{jt}\Upsilon A_{jt}}{\gamma w_{\ell t}}\Big)^{\frac{1-\alpha}{\alpha}}. \tag{A.2}$$

Plugging Equation (A.2) into the input demand condition (17) gives us

$$p_{jt} = \frac{\big(r_j + \omega_j \tau_t\big)^{\alpha}\big(\frac{1}{\alpha}\big)^{\alpha}\big(\frac{1}{1-\alpha}\big)^{1-\alpha}\big(\gamma w_{\ell t}/\Upsilon\big)^{1-\alpha}}{A_{jt}^{1-\alpha}}, \tag{A.3}$$

which implies that relative prices of technology-specific goods follow

$$\frac{p_{jt}}{p_{Jt}} = \Big(\frac{r_j + \omega_j\tau_t}{r_J + \omega_J\tau_t}\Big)^{\alpha}\Big(\frac{A_{jt}}{A_{Jt}}\Big)^{\alpha-1}, \tag{A.4}$$

which gives us Equation (22) from the main text.

Turning to the innovation side of the economy, the research firm solves the problem described in Equation (23). This yields the condition

$$\xi_{jt}\chi_j\eta s_{jt}^{\eta-1}\phi_{jt}\Pi_{jt} = w_{st}, \tag{A.5}$$

which gives us the research condition (25) in the main text.

To derive demand for technology-specific goods under the nested-CES production structure described in Section 2.5, first consider the final producer's demand for sector-level output. The final producer solves

$$\max_{\{E_{\theta t}\}} \mathcal{Y}_t - \sum_{\theta} p_{\theta t}E_{\theta t}, \tag{A.6}$$

where $p_{\theta t}$ is the price for sector θ. This yields the condition

$$\Omega_t^{\frac{\lambda-1}{\lambda}}\nu_\theta^{\frac{1}{\lambda}}\Big(\frac{\mathcal{Y}_t}{E_{\theta t}}\Big)^{\frac{1}{\lambda}} = p_{\theta t}. \tag{A.7}$$

Next, producers at the sector level solve

$$\max_{Y_{\theta dt}, Y_{\theta ct}} p_{\theta t}E_{\theta t} - p_{\theta ct}Y_{\theta ct} - p_{\theta dt}Y_{\theta dt} \tag{A.8}$$

where $p_{\theta ct}$ and $p_{\theta dt}$ are the price of clean and dirty production in sector θ, respectively. This yields the condition

$$p_{\theta t}\Big(\frac{E_{\theta t}}{Y_{\theta et}}\Big)^{\frac{1}{\sigma}} = p_{\theta et}, \tag{A.9}$$

where $e \in \{c, d\}$.

A.2 Task-Based Microfoundation for CES Shares

Following Acemoglu and Restrepo (2022b), let final output be a CES aggregator over a unit interval of tasks

$$\mathcal{Y}_t = \Omega_t\Big(\int_0^1 q_{xt}^{\frac{\lambda-1}{\lambda}}\, dx\Big)^{\frac{\lambda}{\lambda-1}}, \tag{A.10}$$

where q_{xt} is task-level output.

Task-level output follows

$$q_{xt} = \sum_{\theta} \upsilon_{\theta x} e_{\theta xt}, \tag{A.11}$$

where $e_{\theta xt}$ is the demand for sector θ goods in task x, and $\upsilon_{\theta x}$ is the productivity of sector θ goods in task x. Market clearing requires total demand for sector θ goods across tasks to equal supply:

$$\int_0^1 e_{\theta xt} dx = E_{\theta t}. \tag{A.12}$$

Task-level output is perfect substitutes, so the sector with the lowest marginal cost is used to produce a given task. Thus, task-level prices follow

$$p_{xt} = \min_{\theta} \frac{p_{\theta t}}{\upsilon_{\theta x}}, \tag{A.13}$$

and task-level demand follows

$$\Omega_t^{\frac{\lambda-1}{\lambda}}\Big(\frac{\mathcal{Y}_t}{q_{xt}}\Big)^{\frac{1}{\lambda}} = p_{xt}. \tag{A.14}$$

Suppose that the unit interval of tasks is partitioned into subsets where only one sector can produce. That is, each sector has a set of tasks

$$\begin{aligned} \mathcal{X}_\theta &= \big\{x \in [0,1] \big| \upsilon_{\theta x} = 1\big\} \\ \forall x \notin \mathcal{X}_\theta: &\quad \upsilon_{\theta x} = 0 \\ \cup_\theta \mathcal{X}_\theta &= [0,1] \\ \forall \theta, \theta': &\quad \mathcal{X}_\theta \cap \mathcal{X}_{\theta'} = \emptyset. \end{aligned} \tag{A.15}$$

Plugging task-level demand (A.14) into the market clearing condition (A.12) allows us to derive equilibrium sector prices

$$p_{\theta t} = |\mathcal{X}_\theta|^{\frac{1}{\lambda}} \Omega_t^{\frac{\lambda-1}{\lambda}}\Big(\frac{\mathcal{Y}_t}{E_{\theta t}}\Big)^{\frac{1}{\lambda}}, \tag{A.16}$$

where $|\mathcal{X}_\theta| \equiv \int_{x\in\mathcal{X}_\theta} dx$ is the mass of tasks produced by sector θ. Plugging this into the ideal price index

$$1 = \Omega_t^{-1}\Big(\int_0^1 p_{xt}^{1-\lambda} dx\Big)^{\frac{1}{1-\lambda}} \tag{A.17}$$

gives us an aggregate representation of final output:

$$\mathcal{Y}_t = \Omega_t \Big(\sum_\theta |\mathcal{X}_\theta|^{\frac{1}{\lambda}} E_{\theta t}^{\frac{\lambda-1}{\lambda}} \Big)^{\frac{\lambda}{\lambda-1}}. \tag{A.18}$$

Therefore, we can interpret the CES shares ν_θ of Equation (1) as the mass of tasks $|\mathcal{X}_\theta|$ produced by sector θ, as claimed in the main text.

A.3 Substitution Matrix for Nested-CES Production

I will omit time subscripts for the sake of parsimony. To derive the substitution matrix $\mathbf{\Sigma}$, consider the function $G : \mathbb{R}^{2(J-1)} \to \mathbb{R}^{J-1}$, which follows

$$G_i\Big(\big\{ \ln\big(Y_j/Y_J\big), \ln\big(p_J/p_j\big) \big\} \Big) = \ln\big(p_J/p_i\big) - \Big[\frac{1}{\lambda} \ln\big(\nu_{\theta(J)}/\nu_{\theta(i)}\big) + \big(\frac{1}{\sigma} - \frac{1}{\lambda}\big) \ln\big(E_{\theta(J)}/E_{\theta(i)}\big) - \frac{1}{\sigma} \ln\big(Y_J/Y_i\big) \Big]. \tag{A.19}$$

That is, G is a function valued at the zero vector when the equilibrium price conditions (A.7) and (A.9) are satisfied. Note that $\theta(i)$ is the sector associated with technology i. To derive demand responses, we can apply the Implicit Function Theorem to G. This gives us

$$\frac{\partial G_i}{\partial \ln\big(Y_j/Y_J\big)} = -\Big[\frac{1}{\sigma} \mathbb{1}(i=j) + \big(\frac{1}{\lambda} - \frac{1}{\sigma}\big) \varepsilon^E_{\theta(i)j} \mathbb{1}(\theta(i) = \theta(j)) \Big] \tag{A.20}$$

$$\frac{\partial G_i}{\partial \ln\big(p_J/p_j\big)} = \mathbb{1}(i=j), \tag{A.21}$$

where $\varepsilon^E_{\theta(i)j}$ is the elasticity of sector $\theta(i)$ with respect to good j. Equation (A.20) relies on the fact that $E_{\theta(J)} = Y_J$ because the final sector Θ only has a clean form of production.

Applying the Implicit Function Theorem gives us

$$\mathbf{\Sigma} = -\mathbf{D_y G}^{-1}, \tag{A.22}$$

where $\mathbf{D_y G}$ is the Jacobian of G with respect to log relative quantities. We can ignore the Jacobian of G with respect to log relative prices because it is just the identity matrix $\mathbf{I}$. Using the fact that the Jacobian of G with respect to log relative quantities is block diagonal, we can derive

$$\mathbf{\Sigma} = \begin{pmatrix} \tilde{\mathbf{\Sigma}}_1 & 0 & \dots & 0 \\ 0 & \tilde{\mathbf{\Sigma}}_2 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \tilde{\mathbf{\Sigma}}_{\Theta-1} \end{pmatrix}, \tag{A.23}$$

where

$$\tilde{\mathbf{\Sigma}}_\theta = \begin{pmatrix} \lambda + (\sigma - \lambda)\varepsilon^E_{\theta d} & (\lambda - \sigma)\varepsilon^E_{\theta d} \\ (\lambda - \sigma)\varepsilon^E_{\theta c} & \lambda + (\sigma - \lambda)\varepsilon^E_{\theta c} \end{pmatrix}, \tag{A.24}$$

where, again, $\varepsilon^E_{\theta d}$ is the elasticity of sector θ with respect to its dirty form of production, and $\varepsilon^E_{\theta c}$ is defined analogously. We can also interpret these elasticities as income shares. Note that in the special case where $\sigma = \lambda$, we have $\mathbf{\Sigma} = \sigma \mathbf{I}$ as $\sigma = \lambda$ would imply a standard CES.

A.4 Proof of Proposition 1

I will omit time subscripts from variables that are time invariant in steady-state. Any balanced growth steady-state requires equal levels of innovation across technologies: $z_i = z_J$. From Equation (10), this implies

$$\ln\big(s_i/s_J\big) = -\frac{1}{\eta}\ln\big(\chi_i/\chi_J\big) - \frac{1}{\eta}\ln\big(\phi_i/\phi_J\big). \tag{A.25}$$

Plugging this into the research condition (25) and using the no-arbitrage condition (24) and relative good prices (22), we have

$$(1-\alpha)\ln(\bar{A}_i) = \ln(\Xi_i) + \frac{1}{\eta}\ln\big(\chi_i/\chi_J\big) + \frac{1}{\eta}\ln\big(\phi_i/\phi_J\big) + \alpha\ln(\mathcal{R}_i) + \ln\big(Y_{it}/Y_{Jt}\big), \tag{A.26}$$

where $\Xi_i \equiv \xi_i/\xi_J$ denotes relative innovation subsidies, and $\mathcal{R}_i \equiv (r_i + \omega_i\tau)/(r_J + \omega_J\tau)$ denotes relative input costs inclusive of the carbon price. Note that assuming that production is constant returns implies we can write relative quantities demanded just in terms of relative prices, and assuming spillovers are homogeneous of degree zero implies we can write spillovers just in terms of relative technology. Thus, relative technology in any balanced growth steady-state must satisfy Equation (A.26).

To derive the Jacobian of log steady-state relative technology with respect to log relative innovation subsidies $\mathbf{D}_{\boldsymbol{\Xi}}\bar{\mathbf{A}}_{\mathbf{ss}}$, we can differentiate

$$(1-\alpha)\frac{\partial\ln(\bar{A}_i)}{\partial\ln(\Xi_j)} = \mathbb{1}(i=j) - \frac{1}{\eta}\sum_q \Phi_{iq}\frac{\partial\ln(\bar{A}_q)}{\partial\ln(\Xi_j)} + (1-\alpha)\sum_q \Sigma_{iq}\frac{\partial\ln(\bar{A}_q)}{\partial\ln(\Xi_j)}, \tag{A.27}$$

which in matrix notation gives us

$$\begin{aligned}(1-\alpha)\mathbf{D}_{\boldsymbol{\Xi}}\bar{\mathbf{A}}_{\mathbf{ss}} &= \mathbf{I} - \frac{1}{\eta}\boldsymbol{\Phi}\mathbf{D}_{\boldsymbol{\Xi}}\bar{\mathbf{A}}_{\mathbf{ss}} + (1-\alpha)\boldsymbol{\Sigma}\mathbf{D}_{\boldsymbol{\Xi}}\bar{\mathbf{A}}_{\mathbf{ss}} \\ \Rightarrow \mathbf{D}_{\boldsymbol{\Xi}}\bar{\mathbf{A}}_{\mathbf{ss}} &= \eta\big[\boldsymbol{\Phi} - \eta(1-\alpha)\big(\boldsymbol{\Sigma}-\mathbf{I}\big)\big]^{-1}.\end{aligned} \tag{A.28}$$

To derive the Jacobian of log steady-state relative technology with respect to log relative input prices $\mathbf{D}_{\mathbf{R}}\bar{\mathbf{A}}_{\mathbf{ss}}$, we can differentiate

$$(1-\alpha)\frac{\partial\ln(\bar{A}_i)}{\partial\ln(\mathcal{R}_j)} = \alpha\mathbb{1}(i=j) - \frac{1}{\eta}\sum_q \Phi_{iq}\frac{\partial\ln(\bar{A}_q)}{\partial\ln(\mathcal{R}_j)} + (1-\alpha)\sum_q \Sigma_{iq}\frac{\partial\ln(\bar{A}_q)}{\partial\ln(\mathcal{R}_j)} - \alpha\Sigma_{ij}, \tag{A.29}$$

which in matrix notation gives us

$$\begin{aligned}(1-\alpha)\mathbf{D}_{\mathbf{R}}\bar{\mathbf{A}}_{\mathbf{ss}} &= \alpha\mathbf{I} - \frac{1}{\eta}\boldsymbol{\Phi}\mathbf{D}_{\mathbf{R}}\bar{\mathbf{A}}_{\mathbf{ss}} + (1-\alpha)\boldsymbol{\Sigma}\mathbf{D}_{\mathbf{R}}\bar{\mathbf{A}}_{\mathbf{ss}} - \alpha\boldsymbol{\Sigma} \\ \Rightarrow \mathbf{D}_{\mathbf{R}}\bar{\mathbf{A}}_{\mathbf{ss}} &= -\eta\alpha\big[\boldsymbol{\Phi} - \eta(1-\alpha)\big(\boldsymbol{\Sigma}-\mathbf{I}\big)\big]^{-1}\big(\boldsymbol{\Sigma}-\mathbf{I}\big).\end{aligned} \tag{A.30}$$

Combining Equations (A.28) and (A.30), we have

$$d\ln(\bar{A}_{ss}) = \eta\big[\boldsymbol{\Phi} - \eta(1-\alpha)\big(\boldsymbol{\Sigma}-\mathbf{I}\big)\big]^{-1}\big[d\ln(\Xi) - \alpha\big(\boldsymbol{\Sigma}-\mathbf{I}\big)d\ln(\mathcal{R})\big], \tag{A.31}$$

which gives us Equations (38) and (39) from Proposition 1.

$\square$

A.5 Steady-State under Parametric Assumptions of Section 2.5

Under the parametric assumption of Section 2.5, we can further characterize the steady-state for relative technology. Plugging the equilibrium price conditions (A.7) and (A.9) into Equation (A.26), we have

$$\sum_{j<J} \Phi_{ij} \ln(\bar{A}_j) = \eta \ln(\Xi_i) + \eta(1-\alpha)(\sigma-1)\ln(\bar{A}_i) + \eta(\sigma-\lambda)\ln\big(p_{\theta(i)t}/p_{\theta(J)t}\big) - \eta\alpha(\sigma-1)\ln(\mathcal{R}_i). \tag{A.32}$$

Denote relative sector prices by $\mathcal{P}_i \equiv p_{\theta(i)t}/p_{\theta(J)t}$. Writing Equation (A.32) in terms of matrix notation, we have

$$\ln(\bar{A}_{ss}) = \eta\big[\mathbf{\Phi} - \eta(1-\alpha)(\sigma-1)\mathbf{I}\big]^{-1}\big[\ln(\Xi) + (\sigma-\lambda)\ln(\mathcal{P}) - \alpha(\sigma-1)\ln(\mathcal{R})\big]. \tag{A.33}$$

Using the sector-level ideal price index, we have that relative sector prices follow

$$\mathcal{P}_i = \frac{p_{\theta(i)t}}{p_{\theta(J)t}} = \left(\frac{p_{\theta ct}^{1-\sigma} + p_{\theta dt}^{1-\sigma}}{p_{Jt}^{1-\sigma}}\right)^{\frac{1}{1-\sigma}} = \left(\mathcal{R}_{\theta c}^{\alpha(1-\sigma)}\bar{A}_{\theta c}^{(1-\alpha)(\sigma-1)} + \mathcal{R}_{\theta d}^{\alpha(1-\sigma)}\bar{A}_{\theta d}^{(1-\alpha)(\sigma-1)}\right)^{\frac{1}{1-\sigma}}, \tag{A.34}$$

where we have used the fact that the final sector Θ only has a clean form of production. Steady-state relative technology is then the solution to a fixed-point problem defined by Equation (A.33).

Furthermore, Equation (A.25) implies that

$$\ln(\bar{s}_{ss}) = \ln(\mathcal{V}) + \frac{1}{\eta}\mathbf{\Phi}\ln(\bar{A}_{ss}), \tag{A.35}$$

where $\bar{s}_j \equiv s_j/s_J$ are relative scientists, and $\mathcal{V}_j \equiv \nu_{\theta(j)}/\nu_{\theta(J)}$ are relative CES shares. Thus, we have

$$\ln(\bar{s}_{ss}) = \ln(\mathcal{V}) + \mathbf{\Phi}\big[\mathbf{\Phi} - \eta(1-\alpha)(\sigma-1)\mathbf{I}\big]^{-1}\big[\ln(\Xi) + (\sigma-\lambda)\ln(\mathcal{P}) - \alpha(\sigma-1)\ln(\mathcal{R})\big]. \tag{A.36}$$

We can also characterize corner steady-states in the case where the spillover network is shut down. Then, each sector converges to a single technology, either clean or dirty. First, the technologies that lose out in their sector will converge to zero in relative terms. For the remaining technologies, we can use the same strategy as above. We have

$$\ln(\bar{A}_{ss}) = \frac{1}{(1-\alpha)(1-\lambda)}\ln(\Xi) + \frac{\alpha}{1-\alpha}\ln(\mathcal{R}), \tag{A.37}$$

but only for technologies that persist in the long-run, one for each sector.

A.6 Proof of Proposition 2

I will describe the equilibrium evolution of technology as a dynamic process with log relative technology $\{\ln(\bar{A}_{jt})\}$ as the state variable and scientists $\{s_{jt}\}$ as the control variable. Thus, we can write the dynamic process in vector notation as

$$\ln(\bar{A}_t) = \mathcal{H}\Big(s_t\big(\ln(\bar{A}_{t-1})\big), \ln(\bar{A}_{t-1})\Big), \tag{A.38}$$

where $s_t\big(\ln(\bar{A}_{t-1})\big)$ is the equilibrium mapping of the control to the previous period's state. By definition, the steady-state $\ln(\bar{A}_{ss})$ is a fixed point of $\mathcal{H}$, so taking a first-order approximation around the steady-state, we have

$$\ln(\bar{A}_t) - \ln(\bar{A}_{ss}) \approx \boldsymbol{\mathcal{J}}\big(\ln(\bar{A}_{t-1}) - \ln(\bar{A}_{ss})\big), \tag{A.39}$$

where $\boldsymbol{\mathcal{J}}$ is the Jacobian of $\mathcal{H}$, evaluated at the steady-state. Thus, because $\mathcal{A}_t$ is defined as the log deviation of relative technology from steady-state, we have Equation (41) of Proposition 2.

To unpack the Jacobian $\boldsymbol{\mathcal{J}}$, we can describe the dynamic process $\mathcal{H}$ using the law of motion for technology (8), the research condition (25), the no-arbitrage condition (24), and the market clearing condition for scientists (12). I will omit time subscripts on policy variables. This yields

$$\begin{aligned}\ln(\bar{A}_{it}) &= \ln(\gamma)\big[\chi_i s_{it}^{\eta}\phi_{it} - \chi_J s_{Jt}^{\eta}\phi_{Jt}\big] + \ln(\bar{A}_{it-1}) \qquad (A.40)\\ (1-\eta)\big[\ln(s_{it}) - \ln(s_{Jt})\big] &= \ln\big(\chi_i/\chi_J\big) + \ln\big(\xi_i/\xi_J\big) + \ln\big(\phi_{it}/\phi_{Jt}\big) + \ln\big(p_{it}/p_{Jt}\big) + \ln\big(Y_{it}/Y_{Jt}\big)\\ \sum_i s_{it} &= \mathcal{S}.\end{aligned}$$

Note that assuming that production is constant returns implies we can write relative quantities demanded just in terms of relative prices, and assuming spillovers are homogeneous of degree zero implies we can write spillovers just in terms of relative technology.

Denote the elasticity of variable x_{it} with respect to $\bar{A}_{jt-1}$ by $\varepsilon_{ij}^{x} \equiv \frac{\partial \ln(x_{it})}{\partial \ln(\bar{A}_{jt-1})}$, which implies $\varepsilon_{ij}^{\bar{A}} = \mathcal{J}_{ij}$. Differentiating the dynamical system and evaluating at the steady state, we have

$$\begin{aligned}\varepsilon_{ij}^{\bar{A}} &= g\big[\eta(\varepsilon_{ij}^{s} - \varepsilon_{Jj}^{s}) - \Phi_{ij}\big] + \mathbb{1}(i=j) \qquad (A.41)\\ (1-\eta)(\varepsilon_{ij}^{s} - \varepsilon_{Jj}^{s}) &= -\Phi_{ij} - (1-\alpha)\epsilon_{ij}^{\bar{A}} + \sum_q \Sigma_{iq}(1-\alpha)\epsilon_{qj}^{\bar{A}}\\ \Rightarrow \big[(1-\eta) + g\eta(1-\alpha)\big]\epsilon_{ij}^{\bar{A}} &= (1-\eta)\mathbb{1}(i=j) - g\Phi_{ij} + g\eta(1-\alpha)\sum_q \Sigma_{iq}\epsilon_{qj}^{\bar{A}},\end{aligned}$$

where the derivative of relative good prices with respect to relative technology comes from Equation (A.4). Transforming into matrix notation, we have Equation (42) from Proposition 2:

$$\begin{aligned}\big[(1-\eta)\mathbf{I} - g\eta(1-\alpha)\big(\boldsymbol{\Sigma} - \mathbf{I}\big)\big]\boldsymbol{\mathcal{J}} &= (1-\eta)\mathbf{I} - g\boldsymbol{\Phi} \qquad (A.42)\\ \Rightarrow \boldsymbol{\mathcal{J}} &= \big[(1-\eta)\mathbf{I} - g\eta(1-\alpha)\big(\boldsymbol{\Sigma} - \mathbf{I}\big)\big]^{-1}\big[(1-\eta)\mathbf{I} - g\boldsymbol{\Phi}\big].\end{aligned}$$

□

A.7 Proof of Corollary 2

I will assume that the transition matrix $\boldsymbol{\mathcal{J}}$ has $J-1$ distinct real eigenvalues. We have

$$\begin{aligned} \boldsymbol{\mathcal{J}}\boldsymbol{Q} &= \boldsymbol{Q}\boldsymbol{D}(\kappa) \qquad \text{(A.43)} \\ \Rightarrow \big[(1-\eta)\mathbf{I} - g\boldsymbol{\Phi}\big]\boldsymbol{Q} &= \big[(1-\eta)\mathbf{I} - g\eta(1-\alpha)\big(\boldsymbol{\Sigma}-\mathbf{I}\big)\big]\boldsymbol{Q}\boldsymbol{D}(\kappa) \\ \Rightarrow \big[(1-\eta)\mathbf{I} - g\eta(1-\alpha)\big(\boldsymbol{\Sigma}-\mathbf{I}\big)\big]\boldsymbol{Q}\boldsymbol{D}(1-\kappa) &= g\big[\boldsymbol{\Phi} - \eta(1-\alpha)\big(\boldsymbol{\Sigma}-\mathbf{I}\big)\big]\boldsymbol{Q} \\ \Rightarrow \boldsymbol{\mathcal{M}} &= g\boldsymbol{Q}\boldsymbol{D}(1-\kappa)^{-1}\boldsymbol{Q}^{-1}\big[(1-\eta)\mathbf{I} - g\eta(1-\alpha)\big(\boldsymbol{\Sigma}-\mathbf{I}\big)\big]^{-1}, \end{aligned}$$

which gives us Equation (50) of Corollary 2.

□

A.8 Proof of Proposition 4

As described in the Definition 6, the Planner solves

$$\max_{\{c_t,\{\Lambda_{jt},\{\ell_{j\iota t}\},A_{jt},s_{jt}\}\}} \sum_{t\geq 0} \frac{1}{(1+\rho)^t} u(c_t) \;\; s.t. \qquad \text{(A.44)}$$
$$\mathcal{Y}_t = c_t + \sum_j r_j \Lambda_{jt} : \quad \varkappa_t$$
$$L = \sum_j \int_0^1 \ell_{j\iota t} d\iota : \quad \varpi_{\ell t}$$
$$\ln(A_{jt}) = \ln(\gamma)\chi_j s_{jt}^{\eta}\phi_{jt} + \ln(A_{jt-1}) : \quad \epsilon_{jt}$$
$$\mathcal{S} = \sum_j s_{jt} : \quad \varpi_{st},$$

where next to each constraint is the corresponding multiplier. Note that I have transformed the law of motion for technology to be in terms of logs.

Taking the FOC with respect to c_t, we have

$$\frac{u'_t}{(1+\rho)^t} = \varkappa_t, \qquad \text{(A.45)}$$

which I will use to solve out $\varkappa_t$ in the remaining FOCs. Taking the FOC with respect to Λ_{jt}, we have

$$\frac{\partial \mathcal{Y}_t}{\partial \Lambda_{jt}} = r_j - \omega_j \sum_{\hat{t}\geq t} \frac{\mathcal{Y}_{\hat{t}}}{(1+\rho)^{\hat{t}-t}} \frac{u'_{\hat{t}}}{u'_t} \frac{\partial \ln(\Omega_{\hat{t}})}{\partial \mathcal{E}_t}. \qquad \text{(A.46)}$$

Taking the FOC with respect to $\ell_{j\iota t}$, we have

$$\frac{\partial \mathcal{Y}_t}{\partial \ell_{j\iota t}} = \frac{\varpi_{\ell t}}{u'_t/(1+\rho)^t}. \qquad \text{(A.47)}$$

Taking the FOC with respect to A_{jt}, we have

$$\frac{\epsilon_{jt}}{A_{jt}} = \frac{u'_t}{(1+\rho)^t}\frac{\partial \mathcal{Y}_t}{\partial A_{jt}} + \frac{\epsilon_{jt+1}}{A_{jt}} + \sum_i \epsilon_{it+1} \ln(\gamma)\chi_i s_{it+1}^{\eta} \frac{\partial \phi_{it+1}}{\partial A_{jt}}. \qquad \text{(A.48)}$$

Taking the FOC with respect to s_{jt}, we have

$$\varpi_{st} = \epsilon_{jt} \ln(\gamma) \chi_j \eta s_{jt}^{\eta-1} \phi_{jt}. \tag{A.49}$$

We can now determine the prices and policy instruments that support this allocation as a competitive equilibrium. First, to close the markup of the intermediate producer, the Planner sets the intermediate subsidy equal to the markup: $\Upsilon = \gamma$. This guarantees that intermediates are produced at marginal cost. Next, the Planner sets the wage for workers to reflect the shadow price of labor: $w_{\ell t} = \frac{\varpi_{\ell t}}{u_t'/(1+\rho)^t}$. Then, if the carbon price is set properly, prices $\{p_{jt}, \{p_{jit}\}\}$ set equal to marginal products/marginal costs will be efficient. Note that, with these prices and policies, we now have average intermediate producer profit $\Pi_{jt} = (\gamma - 1)(1-\alpha) p_{jt} Y_{jt}$.

Next, define $R_{t+1} \equiv (1+\rho) u_t'/u_{t+1}'$ as the Planner's intertemporal marginal rate of substitution. To align the input demand condition (17) with the Planner's input FOC (A.46), the Planner sets the carbon price according to

$$\tau_t = -\sum_{\hat{t} \geq t} \prod_{\hat{s}=1}^{\hat{t}-t} \frac{1}{R_{t+\hat{s}}} \mathcal{Y}_{\hat{t}} \frac{\partial \ln(\Omega_{\hat{t}})}{\partial \mathcal{E}_t}, \tag{A.50}$$

which is the social cost of carbon. This Pigouvian correction gives us Equation (61) of Proposition 4.

For innovation subsidies, we can compare the optimality condition of the research problem (A.5) with the Planner's scientist FOC (A.49). We have

$$\begin{aligned} \chi_j \eta s_{jt}^{\eta-1} \phi_{jt} \xi_{jt} \Pi_{jt} &= w_{st} \\ \chi_j \eta s_{jt}^{\eta-1} \phi_{jt} \epsilon_{jt} &= \frac{\varpi_{st}}{\ln(\gamma)}. \end{aligned} \tag{A.51}$$

The Planner can set the wage for scientists according to

$$w_{st} = \frac{\varpi_{st} \;\ln(\gamma)}{u_t'/(1+\rho)^t}. \tag{A.52}$$

Plugging this into the two research conditions (A.51), we can see that the incentives for innovation will be efficient if the Planner sets innovation subsidies according to

$$\xi_{jt} \Pi_{jt} = \frac{\epsilon_{jt}}{u_t'/(1+\rho)^t}. \tag{A.53}$$

Note that any positive multiple of $\{w_{st}, \{\xi_{jt}\}\}$ would achieve the same outcome, but I am focusing on what I view as a natural normalization for these prices and policies. Plugging this into the Planner's technology FOC (A.48), we get the recursion

$$\xi_{jt} \Pi_{jt} = (1-\alpha) S_{jt} \mathcal{Y}_t + \frac{1}{R_{t+1}} \left[\xi_{jt+1} \Pi_{jt+1} + \sum_i \xi_{it+1} \Pi_{it+1} g_{it+1} \varphi_{ijt+1} \right], \tag{A.54}$$

which gives us Equation (62) of Proposition 4. Finally, the Planner sets the lump-sum tax D_t to balance the government's budget (27) in each period.

□

A.9 Proof of Corollary 3

Manipulating Equation (62), we have

$$\tilde{\xi}_{jt} = \frac{S_{jt}}{\gamma - 1} + \frac{\mathcal{Y}_{t+1}/\mathcal{Y}_t}{R_{t+1}}\big[\tilde{\xi}_{jt+1} + \sum_i \tilde{\xi}_{it+1} g_{it+1}\varphi_{ijt+1}\big]. \tag{A.55}$$

I have assumed the economy is in steady-state, so income shares for technologies are constant and the growth rate of output is equal to a constant g_y. For example, if carbon pollution stops and the damage function settles to a constant, then the growth rate of output g_y will be equal to the growth rate of technology g.[75] More generally, we would have $g_y \leq g$ as damages could worsen with continued carbon pollution.

Define $\tilde{R} \equiv R/(1+g_y)$ as the Planner's growth-adjusted intertemporal marginal rate of substitution, evaluated at the steady-state. Then, evaluating Equation (A.55) in matrix notation at the steady-state, we have

$$\tilde{\xi}' = \frac{1}{\gamma - 1} S' \Big[(1 - \tilde{R}^{-1})\mathbf{I} - g\tilde{R}^{-1}\boldsymbol{\varphi}\Big]^{-1}, \tag{A.56}$$

which gives us Equation (63) of Corollary 3. If all income shares remain positive in steady-state, this is equivalent to

$$\xi' = \frac{1}{\gamma - 1}\vec{1}'\Big[(1 - \tilde{R}^{-1})\mathbf{I} - g\tilde{R}^{-1}\boldsymbol{D}(S)\boldsymbol{\varphi}\boldsymbol{D}(S)^{-1}\Big]^{-1}. \tag{A.57}$$

□

A.10 Proof of Proposition 5

As described in the Definition 7, the Planner solves

$$\max_{\{c_t, \{\Lambda_{jt}, \{\ell_{j\iota t}\}, A_{jt}, s_{jt}\}\}} \sum_{t \geq 0} \frac{1}{(1+\rho)^t} u(c_t) \;\; s.t. \tag{A.58}$$

$$\mathcal{Y}_t = c_t + \sum_j r_j \Lambda_{jt} : \;\; \varkappa_t$$

$$\ln(A_{jt}) = \ln(\gamma)\chi_j s_{jt}^{\eta}\phi_{jt} + \ln(A_{jt-1}) : \;\; \epsilon_{jt}$$

$$\mathcal{S} = \sum_j s_{jt} : \;\; \varpi_{st}$$

$$\{\Lambda_{jt}, \{\ell_{j\iota t}\}\} = \operatorname{argmax}\big[\mathcal{Y}_t - \sum_j r_j \Lambda_{jt} - \hat{\tau}_t \mathcal{E}_t\big] \;\; s.t. \;\; L = \sum_j \int_0^1 \ell_{j\iota t} d\iota,$$

where next to each constraint is the corresponding multiplier. As before, I have transformed the law of motion for technology to be in terms of logs. Now the Planner must consider the incentive compatibility constraint for factors of production. Note that I am assuming the intermediate subsidy is set according to $\Upsilon = \gamma$.

Taking the FOC with respect to c_t, we have

$$\frac{u_t'}{(1+\rho)^t} = \varkappa_t, \tag{A.59}$$

[75] This is the steady-state under the parametric assumptions of Sections 2.5 and 5.2.

which I will use to solve out $\varkappa_t$ in the remaining FOCs. Taking the FOC with respect to s_{jt}, we have

$$\varpi_{st} = \epsilon_{jt} \ln(\gamma) \chi_j \eta s_{jt}^{\eta-1} \phi_{jt}. \tag{A.60}$$

Taking the FOC with respect to A_{jt}, we have

$$\frac{\epsilon_{jt}}{A_{jt}} = \frac{u'_t}{(1+\rho)^t} \Big[\frac{\partial \mathcal{Y}_t}{\partial A_{jt}} + \sum_i \int_0^1 \frac{\partial \mathcal{Y}_t}{\partial \ell_{i\iota t}} \frac{\partial \ell_{i\iota t}}{\partial A_{jt}} d\iota + \sum_i \big(\frac{\partial \mathcal{Y}_t}{\partial \Lambda_{it}} - r_i \big) \frac{\partial \Lambda_{it}}{\partial A_{jt}} - \tau_t \frac{\partial \mathcal{E}_t}{\partial A_{jt}} \Big] \tag{A.61}$$
$$+ \frac{\epsilon_{jt+1}}{A_{jt}} + \sum_i \epsilon_{it+1} \ln(\gamma) \chi_i s_{it+1}^{\eta} \frac{\partial \phi_{it+1}}{\partial A_{jt}},$$

where τ_t is the same social cost of carbon from Equation (61). Now consider the producer optimality conditions

$$\frac{\partial \mathcal{Y}_t}{\partial \Lambda_{jt}} = r_j + \omega_j \hat{\tau}_t \tag{A.62}$$
$$\frac{\partial \mathcal{Y}_t}{\partial \ell_{j\iota t}} = w_{\ell t}.$$

Plugging these conditions into the Planner's optimality condition for technology (A.61), and noting that $\sum_j \int_0^1 \frac{\partial \ell_{j\iota t}}{\partial x} d\iota = 0$ for any x due to the fixed supply of labor, we have

$$\frac{\epsilon_{jt}}{A_{jt}} = \frac{u'_t}{(1+\rho)^t} \Big[\frac{\partial \mathcal{Y}_t}{\partial A_{jt}} - (\tau_t - \hat{\tau}_t) \frac{\partial \mathcal{E}_t}{\partial A_{jt}} \Big] + \frac{\epsilon_{jt+1}}{A_{jt}} + \sum_i \epsilon_{it+1} \ln(\gamma) \chi_i s_{it+1}^{\eta} \frac{\partial \phi_{it+1}}{\partial A_{jt}}. \tag{A.63}$$

As before, the Planner can create efficient incentive for innovation by satisfying Equations (A.52) and (A.53). Thus, innovation subsidies follow the recursion

$$\xi_{jt} \Pi_{jt} = (1-\alpha) S_{jt} \mathcal{Y}_t - (\tau_t - \hat{\tau}_t) \mathcal{E}_t \frac{\partial \ln(\mathcal{E}_t)}{\partial \ln(A_{jt})} + \frac{1}{R_{t+1}} \big[\xi_{jt+1} \Pi_{jt+1} + \sum_i \xi_{it+1} \Pi_{it+1} g_{it+1} \varphi_{ijt+1} \big], \tag{A.64}$$

which gives us Equation (68) of Proposition 5. By construction, prices on the production side of the economy are set in equilibrium. Finally, the Planner sets the lump-sum tax D_t to balance the government's budget (27) in each period.

□

B Details on Calibration & Simulation

B.1 Details on Numerical Representations

This appendix contains details on how I can represent equilibrium outcomes of my model in terms of fundamentals: technology, policy, and structural parameters. I make the parametric assumptions of Section 2.5. First, I will define what I call pseudo prices. These follow

$$\tilde{p}_{jt} \equiv \big(r_j + \omega_j \tau_t \big)^{\alpha} / A_{jt}^{1-\alpha}. \tag{B.1}$$

From Equation (A.3), we have that pseudo prices are proportional to prices and satisfy

$$p_{jt} = \left(\frac{1}{\alpha}\right)^{\alpha}\left(\frac{1}{1-\alpha}\right)^{1-\alpha}\left(\frac{\gamma w_{\ell t}}{\Upsilon}\right)^{1-\alpha}\tilde{p}_{jt}. \tag{B.2}$$

We can define similar pseudo prices at higher levels using ideal price indices

$$\tilde{p}_{\theta t} \equiv \left(\tilde{p}_{\theta ct}^{1-\sigma} + \tilde{p}_{\theta dt}^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \tag{B.3}$$

$$\tilde{P}_t \equiv \Omega_t^{-1}\Big(\sum_{\theta}\nu_\theta \tilde{p}_{\theta t}^{1-\lambda}\Big)^{\frac{1}{1-\lambda}}, \tag{B.4}$$

which must also be proportional to their corresponding price.

Next, combining the intermediate production function (4), equilibrium intermediate output (A.1), and labor supply (6) gives us

$$\frac{\gamma w_{\ell t}}{\Upsilon} = (1-\alpha)\frac{\mathcal{Y}_t}{L}. \tag{B.5}$$

Combining the final pseudo price (B.4) with the fact that output is the numeraire, we have for output

$$\mathcal{Y}_t = \alpha^{\frac{\alpha}{1-\alpha}} L \tilde{P}_t^{\frac{-1}{1-\alpha}}. \tag{B.6}$$

Next, using the demand conditions for sectors (A.7) and goods within sectors (A.9), we can derive income shares for sectors and goods within sectors. We have

$$S_{\theta t} \equiv \frac{p_{\theta t}E_{\theta t}}{\mathcal{Y}_t} = \frac{\nu_\theta \tilde{p}_{\theta t}^{1-\lambda}}{\Omega_t^{1-\lambda}\tilde{P}_t^{1-\lambda}} \tag{B.7}$$

$$S_{et}^{\theta} \equiv \frac{p_{\theta et}Y_{\theta et}}{p_{\theta t}E_{\theta t}} = \frac{\tilde{p}_{\theta et}^{1-\sigma}}{\tilde{p}_{\theta t}^{1-\sigma}}, \tag{B.8}$$

where $e \in \{c, d\}$. The income share for technology $j = \theta e$ is then $S_{jt} = S_{\theta t} \cdot S_{et}^{\theta}$. The input demand condition (17) implies

$$\Lambda_{jt} = \frac{\alpha S_{jt}\mathcal{Y}_t}{r_j + \omega_j \tau_t}, \tag{B.9}$$

which gives us total emissions via $\mathcal{E}_t = \sum_j \omega_j \Lambda_{jt}$ and consumption via $c_t = \mathcal{Y}_t - \sum_j r_j \Lambda_{jt}$.

I define clean quantity shares by sector as $q_{ct}^{\theta} \equiv Y_{\theta ct}/(Y_{\theta ct} + Y_{\theta dt})$. To map between clean income shares and clean quantity shares, note that the demand condition within sectors (A.9) implies that

$$S_{ct}^{\theta} = \frac{p_{\theta ct}Y_{\theta ct}}{p_{\theta ct}Y_{\theta ct} + p_{\theta dt}Y_{\theta dt}} = \frac{1}{1 + \left((1-q_{ct}^{\theta})/q_{ct}^{\theta}\right)^{\frac{\sigma-1}{\sigma}}}. \tag{B.10}$$

Finally, to map my model of innovation to the empirical evidence that informs my choice of the elasticity of innovation with respect to scientists η, first define technology specific R&D expenditures as $R\&D_{jt} = w_{st}s_{jt}$. To consider the elasticity of patents with respect to R&D expenditure, plug this definition into the innovation production function (10) and note

$$\begin{aligned} z_{jt} &= \chi_j \left(\frac{R\&D_{jt}}{w_{st}} \right)^{\eta} \phi_{jt} \\ \Rightarrow \frac{\partial \ln (z_{jt})}{\partial \ln (R\&D_{jt})} &= \eta, \end{aligned} \tag{B.11}$$

where I am taking the mass of innovations z_{jt} as the object in my model analogous to patents. Next, to consider the elasticity of R&D expenditure with respect to the price of research, note that the optimality condition of the research problem (A.5) implies

$$\begin{aligned} \eta \chi_j \left(\frac{R\&D_{jt}}{w_{st}} \right)^{\eta} \phi_{jt} &= R\&D_{jt} \\ \Rightarrow \frac{\partial \ln (R\&D_{jt})}{\partial \ln (w_{st})} &= -\frac{\eta}{1-\eta}, \end{aligned} \tag{B.12}$$

which gives a demand elasticity of unity when $\eta = 0.5$.

B.2 Details on Selection of US Policy in the 2010s

This appendix provides further details on my procedure for selecting subsidies that describe US climate policy throughout the 2010s. These subsidies are an input into the simulations of Section 5.4. To determine the clean input subsidy for the transportation sector $\bar{\xi}_{car,c}$, consider public spending on the subsidy as a proportion of total spending in the transportation sector. We have

$$\frac{\bar{\xi}_{car,c} r_{car,c} \Lambda_{car,ct}}{p_{car,t} E_{car,t}} = \frac{\bar{\xi}_{car,c}}{1-\bar{\xi}_{car,c}} \frac{\alpha p_{car,ct} Y_{car,ct}}{p_{car,t} E_{car,t}} = \frac{\bar{\xi}_{car,c}}{1-\bar{\xi}_{car,c}} \alpha S_{ct}^{car}, \tag{B.13}$$

where the first equality comes from the input demand condition (17).

Let $\hat{cred}_{car,ct}$ denote data on federal spending on the plug-in electric vehicle tax credit as a proportion of total spending in the transportation sector. The numerator comes from US Congressional Research Service Report IF11017, which contains federal outlays on the tax credit for the years 2011 to 2018. This report was published in 2019 and contains projections out to 2022, but I only include spending that occurred before the report was published. The denominator comes from the series on motor vehicle output discussed in Section 5.3. Next, I take estimates of clean income shares in the transportation sector $\hat{S}_{ct}^{car}$ from my data on clean quantity shares using the mapping from Equation (B.10). Plugging these data series into Equation (B.13), we have

$$\frac{\bar{\xi}_{car,c}}{1-\bar{\xi}_{car,c}} \sum_{t=2011}^{2018} \alpha \hat{S}_{ct}^{car} = \sum_{t=2011}^{2018} \hat{cred}_{car,ct}, \tag{B.14}$$

which allows me to back out the value $\bar{\xi}_{car,c} = 0.011$.

To determine innovation subsidies $\{\xi_j\}$, consider public spending on each subsidy as a proportion of total spending on R&D. Using the optimality condition of the research problem (A.5) and scientist supply (12), we have

$$\frac{(\xi_j - 1) z_{jt} \Pi_{jt}}{w_{st} \sum_j s_{jt}} = \frac{\xi_j - 1}{\xi_j} \frac{s_{jt}}{\eta \mathcal{S}}. \tag{B.15}$$

As discussed in the main text, I normalize the innovation subsidy on the general technology to one as only relative innovation subsidies influence the composition of scientists across technologies.

Let $\hat{pubrd}_{jt}$ denote data on public spending on R&D in technology j as a proportion of total spending on R&D. The numerator comes from IEA data on public spending on R&D by technology. This series goes until 2015. Table B.1 contains information on the assignment of spending types to the technologies in my model. The denominator comes from the BEA. I then set innovation subsidies to match this series according to

$$\frac{\xi_j - 1}{\xi_j} \sum_{t=2010}^{2015} \frac{s_{jt}}{\eta \mathcal{S}} = \sum_{t=2010}^{2015} \hat{pubrd}_{jt}. \tag{B.16}$$

Unlike clean input subsidies, which can be set directly from the data, finding the innovation subsidies that match the data requires simulating the model to solve Equation (B.16), so I calibrate separate innovation subsidies for the versions of the model with and without spillovers. In the model with spillovers this yields $(\xi_{car,c}, \xi_{car,d}, \xi_{elec,c}, \xi_{elec,d}) = (1.005, 1.015, 1.057, 1.013)$, and in the model without spillovers this yields $(\xi_{car,c}, \xi_{car,d}, \xi_{elec,c}, \xi_{elec,d}) = (1.212, 1.005, 1.293, 1.016)$. The clean innovation subsidies are higher in the model without spillovers because clean technologies receive less R&D overall, so the same public R&D spending requires higher subsidy rates.

Table B.1: Assignment of Public R&D Spending to Technologies

Transportation		Electricity Generation	
Description	Codes	Description	Codes
Clean		Clean	
Vehicle Batteries/Storage Technologies	1311	Renewable Energy Sources	3
Advanced EV/HEV/FCV Systems	1312	–Excluding Biofuels	34
Electric Vehicle Infrastructure	1314	Nuclear	4
Dirty		Dirty	
Advanced Combustion Engines	1313	Coal	22
Oil & Gas (1/2)	21	Oil & Gas (1/2)	21

Notes: Clean electricity generation is assigned all of the spending in the renewables category except biofuels. I assign spending in the oil and gas category equally between dirty transportation and electricity generation because these fuels are used in both sectors. Data on public R&D spending by technology comes from the IEA.

B.3 First-Best Simulation Solution Method

To simulate the first-best policy path, I solve a root-finding problem. I make the parametric assumptions of Sections 2.5 and 5.2. Using the results described in Appendix B.1, my economy can be reduced to the sequence $\{\tau_{1t}, \tau_{2t}, \mathcal{C}_{1t}, \mathcal{C}_{2t}, \{\tilde{\xi}_{jt}, A_{jt}\}\}$.

For the carbon price, combining atmospheric carbon concentrations (70), the damage function (73), and the social cost of carbon (61) gives us

$$\tau_t = \varrho \sum_{\hat{t} \geq t} \prod_{\hat{s}=1}^{\hat{t}-t} \frac{1}{R_{t+\hat{s}}} \mathcal{Y}_{\hat{t}} \big(\psi_p + (1-\psi_p)\psi_0 \psi^{\hat{t}-t}\big). \tag{B.17}$$

This allows us to write the carbon price τ_t as a two-dimensional recursion $\{\tau_{1t}, \tau_{2t}\}$, which provides separate Pigouvian corrections for the permanent and transitory component of carbon pollution. These follow

$$\tau_{1t} = \varrho\psi_p\mathcal{Y}_t + \frac{1}{R_{t+1}}\tau_{1t+1} \tag{B.18}$$

$$\tau_{2t} = \varrho(1-\psi_p)\psi_0\mathcal{Y}_t + \frac{1}{R_{t+1}}\psi\tau_{2t+1}. \tag{B.19}$$

The sum of these two components gives us the carbon price $\tau_t = \tau_{1t} + \tau_{2t}$. The recursive formulation for $\{\mathcal{C}_{1t}, \mathcal{C}_{2t}\}$ is given by Equations (71) and (72), and these together sum to atmospheric carbon concentrations $\mathcal{C}_t = \mathcal{C}_{1t} + \mathcal{C}_{2t}$.

Innovation subsidies times income shares $\{\tilde{\xi}_{jt}\}$ follow the recursive formula described in Equation (A.55), and technology $\{A_{jt}\}$ follows the law of motion (8). Note that the research condition (25) and scientist supply (12) allow us to write $\{s_{jt}\}$ as a function of $\{\tilde{\xi}_{jt}\}$ and $\{A_{jt-1}\}$. That is, we have

$$(1-\eta)\ln(\bar{s}_t) = \ln(\tilde{\Xi}_t) - \eta\ln(\mathcal{V}) - \mathbf{\Phi}\ln(\bar{A}_{t-1}) \tag{B.20}$$

where $\bar{s}_{jt} \equiv s_{jt}/s_{Jt}$ are relative scientists, $\tilde{\Xi}_t \equiv \tilde{\xi}_{jt}/\tilde{\xi}_{Jt}$ are relative innovation subsidies times income shares, and $\mathcal{V}_j \equiv \nu_{\theta(j)}/\nu_{\theta(J)}$ are relative CES shares. Scientist supply (12) then pins down the scale of the scientist allocation. Furthermore, as I describe in Appendix B.1, all of the other endogenous outcomes of the economy, such as output $\mathcal{Y}_t$, consumption c_t, income shares S_{jt}, and emissions $\mathcal{E}_t$, can be written as functions of τ_t and $\{A_{jt}\}$.

In summary, the root system follows

$$\begin{aligned}
\tau_{1t} &= \varrho\psi_p\mathcal{Y}_t + \frac{1}{R_{t+1}}\tau_{1t+1} \\
\tau_{2t} &= \varrho(1-\psi_p)\psi_0\mathcal{Y}_t + \frac{1}{R_{t+1}}\psi\tau_{2t+1} \\
\mathcal{C}_{1t} &= \psi_p\mathcal{E}_t + \mathcal{C}_{1t-1} \\
\mathcal{C}_{2t} &= (1-\psi_p)\psi_0\mathcal{E}_t + \psi\mathcal{C}_{2t-1} \\
\tilde{\xi}_{jt} &= \frac{S_{jt}}{\gamma-1} + \frac{\mathcal{Y}_{t+1}/\mathcal{Y}_t}{R_{t+1}}\big[\tilde{\xi}_{jt+1} + \sum_i \tilde{\xi}_{it+1}g_{it+1}\varphi_{ijt+1}\big] \\
\ln(A_{jt}) &= \ln(\gamma)\chi\Big(\frac{s_{jt}}{\nu_{\theta(j)}}\Big)^{\eta}\phi_{jt} + \ln(A_{jt-1}).
\end{aligned} \tag{B.21}$$

The state variables $\{\mathcal{C}_{1t}, \mathcal{C}_{2t}, \{A_{jt}\}\}$ are backward-looking, so I specify initial conditions using the strategies described in Sections 5.2 and 5.3. The policy variables $\{\tau_{1t}, \tau_{2t}, \{\tilde{\xi}_{jt}\}\}$ are forward-looking, so I specify a terminal condition by assuming the economy is in steady-state beyond my final period T. My simulation extends for 500 periods; long enough for the economy to be near its steady-state. Denote carbon prices relative to output by $\tilde{\tau}_t \equiv \tau_t/\mathcal{Y}_t$. These achieve a steady-state which follows

$$\tilde{\tau}_1 = \frac{\varrho\psi_p}{1-\tilde{R}^{-1}} \tag{B.22}$$

$$\tilde{\tau}_2 = \frac{\varrho(1-\psi_p)\psi_0}{1-\psi\tilde{R}^{-1}}, \tag{B.23}$$

where, as before, $\tilde{R} \equiv R/(1+g)$ is the Planner's growth-adjusted intertemporal marginal rate of substitution, evaluated at the steady-state. From this, we can see that carbon prices asymptote to infinity as the economy continues to grow, so emissions and the income shares of dirty technologies must go to zero. Hence, the growth rate of output g_y goes to the growth rate of technology g because climate damages asymptote to a constant as the transitory component of carbon concentrations decays to zero. Therefore, I assume that both consumption and output grow at rate g in the periods beyond my simulation. Terminal carbon prices then follow $\tau_{1T+1} = (1+g)\mathcal{Y}_T\tilde{\tau}_1$ and $\tau_{2T+1} = (1+g)\mathcal{Y}_T\tilde{\tau}_2$.

By Corollary 3, the steady-state for $\{\tilde{\xi}_j\}$ follows

$$\tilde{\xi}' = \frac{1}{\gamma-1}S'\Big[(1-\tilde{R}^{-1})\mathbf{I} - g\tilde{R}^{-1}\boldsymbol{\varphi}\Big]^{-1}. \tag{B.24}$$

To select the steady-state income shares S and growth rate g, we can derive steady-state relative technology by combining Equations (A.35) and (B.20) to achieve

$$\ln(\bar{A}_{ss}) = \eta\boldsymbol{\Phi}^{-1}\big[\ln(\tilde{\Xi}) - \ln(\mathcal{V})\big], \tag{B.25}$$

which, combined with the requirement that dirty income shares are zero, defines a fixed-point problem to solve for steady-state $\{\tilde{\xi}_j\}$. Note that the growth rate g comes from Equation (9).

Finding the steady-state when there are no cross-technology spillovers $\boldsymbol{\varphi} = \mathbf{0}$ is special case because the inverse of $\boldsymbol{\Phi}$ does not exist. In that case, we have $\tilde{\xi}_j = S_j/(\gamma-1)(1-\tilde{R}^{-1})$, which implies $\tilde{\xi}_j = 0$ for dirty technologies because they have zero income share. Thus, scientists for dirty technologies must also be zero. The steady-state scientist condition (A.35) and scientist supply (12) can be satisfied by setting $s_j = \nu_{\theta(j)}$ for all clean technologies. Plugging this into Equation (B.20) gives us that income shares for clean technologies also follow $S_j = \nu_{\theta(j)}$. From Equation (B.7), this implies that $\bar{A}_j = 1$ for clean technologies and $\bar{A}_j = 0$ for dirty technologies.

B.4 Second-Best Simulation Solution Method

To simulate the second-best policy path, I use a similar strategy as in the first-best, reducing the economy to the sequence $\{\tau_{1t}, \tau_{2t}, \mathcal{C}_{1t}, \mathcal{C}_{2t}, \{\tilde{\xi}_{jt}, A_{jt}\}\}$.

Carbon concentrations $\{\mathcal{C}_{1t}, \mathcal{C}_{2t}\}$ and technology $\{A_{jt}\}$ follow the same laws of motion as before. The two components of the social cost of carbon $\{\tau_{1t}, \tau_{2t}\}$ follow the same recursion as before, except now it is $\hat{\tau}_t$ that influences production decisions.

Manipulating Equation (68), we have that second-best innovation subsidies times income shares $\{\tilde{\xi}_{jt}\}$ follow the recursive formula

$$\tilde{\xi}_{jt} = \frac{S_{jt}}{\gamma-1} - \frac{\tau_t - \hat{\tau}_t}{(\gamma-1)(1-\alpha)}\frac{\mathcal{E}_t}{\mathcal{Y}_t}\frac{\partial\ln(\mathcal{E}_t)}{\partial\ln(A_{jt})} + \frac{\mathcal{Y}_{t+1}/\mathcal{Y}_t}{R_{t+1}}\big[\tilde{\xi}_{jt+1} + \sum_i \tilde{\xi}_{it+1}g_{it+1}\varphi_{ijt+1}\big]. \tag{B.26}$$

As before, we can write $\{s_{jt}\}$ as a function of $\{\tilde{\xi}_{jt}\}$ and $\{A_{jt-1}\}$ using Equation (B.20), but now that innovation subsidies can be negative, we will set $s_{jt} = 0$ whenever $\tilde{\xi}_{jt} \leq 0$. For the effect of innovation on equilibrium emissions, we can use the results of Appendix B.1 to derive

$$\frac{\mathcal{E}_t}{\mathcal{Y}_t}\frac{\partial \ln(\mathcal{E}_t)}{\partial \ln(A_{jt})} = \alpha \sum_{\theta} \frac{\omega_{\theta d} S_{\theta dt}}{r_{\theta d} + \omega_{\theta d}\hat{\tau}_t}\big[(1+(1-\alpha)(1-\lambda))S_{jt} \\ + \mathbb{1}(\theta = \theta(j))(1-\alpha)(\lambda-\sigma)S^{\theta}_{jt} \\ + \mathbb{1}(\theta d = j)(1-\alpha)(\sigma-1)\big]. \tag{B.27}$$

In summary, the root system follows

$$\begin{aligned}
\tau_{1t} &= \varrho\psi_p\mathcal{Y}_t + \frac{1}{R_{t+1}}\tau_{1t+1} \\
\tau_{2t} &= \varrho(1-\psi_p)\psi_0\mathcal{Y}_t + \frac{1}{R_{t+1}}\psi\tau_{2t+1} \\
\mathcal{C}_{1t} &= \psi_p\mathcal{E}_t + \mathcal{C}_{1t-1} \\
\mathcal{C}_{2t} &= (1-\psi_p)\psi_0\mathcal{E}_t + \psi\mathcal{C}_{2t-1} \\
\tilde{\xi}_{jt} &= \frac{S_{jt}}{\gamma-1} - \frac{\tau_t - \hat{\tau}_t}{(\gamma-1)(1-\alpha)}\frac{\mathcal{E}_t}{\mathcal{Y}_t}\frac{\partial \ln(\mathcal{E}_t)}{\partial \ln(A_{jt})} + \frac{\mathcal{Y}_{t+1}/\mathcal{Y}_t}{R_{t+1}}\big[\tilde{\xi}_{jt+1} + \sum_i \tilde{\xi}_{it+1} g_{it+1}\varphi_{ijt+1}\big] \\
\ln(A_{jt}) &= \ln(\gamma)\chi\Big(\frac{s_{jt}}{\nu_{\theta(j)}}\Big)^{\eta}\phi_{jt} + \ln(A_{jt-1}),
\end{aligned} \tag{B.28}$$

where endogenous outcomes, such as emissions $\mathcal{E}_t$, are determined by $\hat{\tau}_t$, rather than the social cost of carbon. The state variables $\{\mathcal{C}_{1t}, \mathcal{C}_{2t}, \{A_{jt}\}\}$ are backward-looking, so I specify initial conditions as before. The policy variables $\{\tau_{1t}, \tau_{2t}, \{\tilde{\xi}_{jt}\}\}$ are forward-looking, so I again assume the economy is in steady-state beyond my final period T. The steady-states for the two components of the social cost of carbon relative to output again follow Equations (B.22) and (B.23).

Denote by $\mathcal{T}_{jt} \equiv \frac{\tau_t - \hat{\tau}_t}{1-\alpha}\frac{\mathcal{E}_t}{\mathcal{Y}_t}\frac{\partial \ln(\mathcal{E}_t)}{\partial \ln(A_{jt})}$ the component of the innovation subsidy that reflects the pollution distortion, excluding $(\gamma-1)$. Similar to Corollary 3, the steady-state for $\{\tilde{\xi}_j\}$ follows

$$\tilde{\xi}' = \frac{1}{\gamma-1}\big(S' - \mathcal{T}'\big)\Big[(1-\tilde{R}^{-1})\mathbf{I} - g\tilde{R}^{-1}\boldsymbol{\varphi}\Big]^{-1}. \tag{B.29}$$

The question, then, is the steady-state behavior of $\mathcal{T}$. I will assume that $\hat{\tau}_t$ either converges to a weakly positive constant or grows to infinity. If it grows to infinity, I will assume it remains below the social cost of carbon, so the asymptote of $\hat{\tau}_t/\mathcal{Y}_t$ is below $\tilde{\tau}_1 + \tilde{\tau}_2$. We have

$$\mathcal{T}_j = (\tilde{\tau}_1 + \tilde{\tau}_2 - \hat{\tau}_t/\mathcal{Y}_t)\frac{\alpha}{1-\alpha}\mathcal{Y}_t \sum_{\theta} \frac{\omega_{\theta d} S_{\theta d}}{r_{\theta d} + \omega_{\theta d}\hat{\tau}_t}\big[(1+(1-\alpha)(1-\lambda))S_j \\ + \mathbb{1}(\theta = \theta(j))(1-\alpha)(\lambda-\sigma)S^{\theta}_j \\ + \mathbb{1}(\theta d = j)(1-\alpha)(\sigma-1)\big]. \tag{B.30}$$

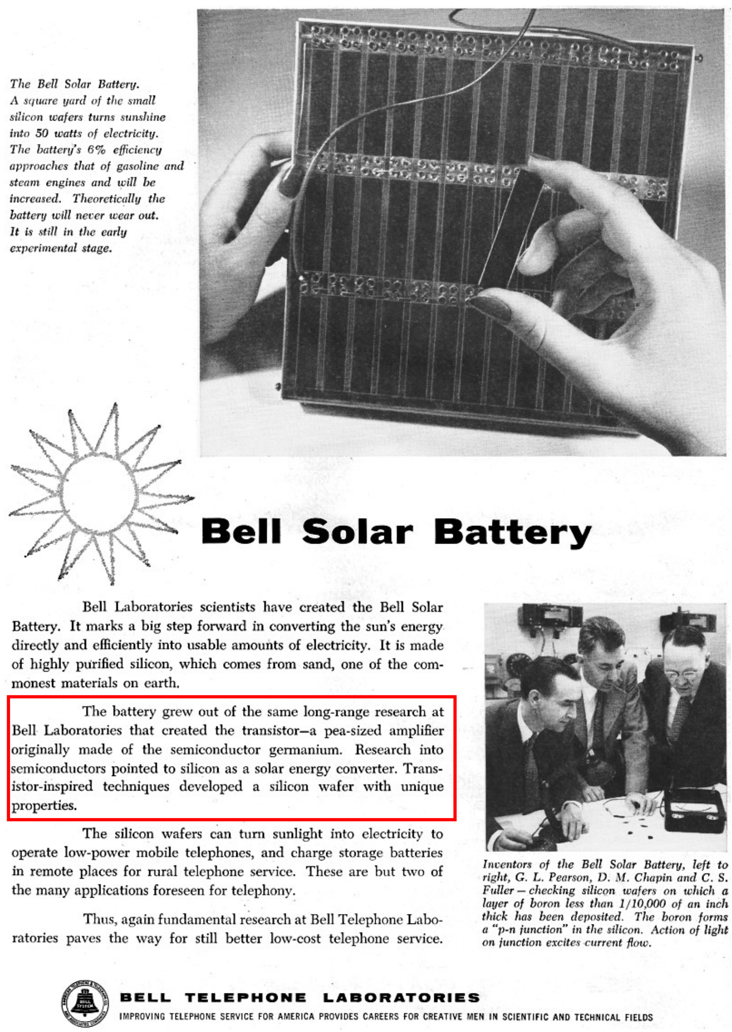
Thus, there are two cases. If $\hat{\tau}_t$ grows over time, then dirty income shares must be zero in steady-state and we will have $\mathcal{T} = \vec{0}$. If $\hat{\tau}_t$ converges to a weakly positive constant, then $\mathcal{T}$ will diverge to infinity unless steady-state dirty income shares are zero. Without a growing carbon price, this requires steady-state relative dirty technologies to be zero as well. In either case, $\mathcal{T} = \vec{0}$ and steady-state $\{\tilde{\xi}_j\}$ is the solution to the fixed-point problem defined by Equation (B.29), but in the latter case, we have the added restriction that steady-state relative dirty technologies must be zero.

C Additional Tables & Figures

Figure C.1: Tesla Mule 1

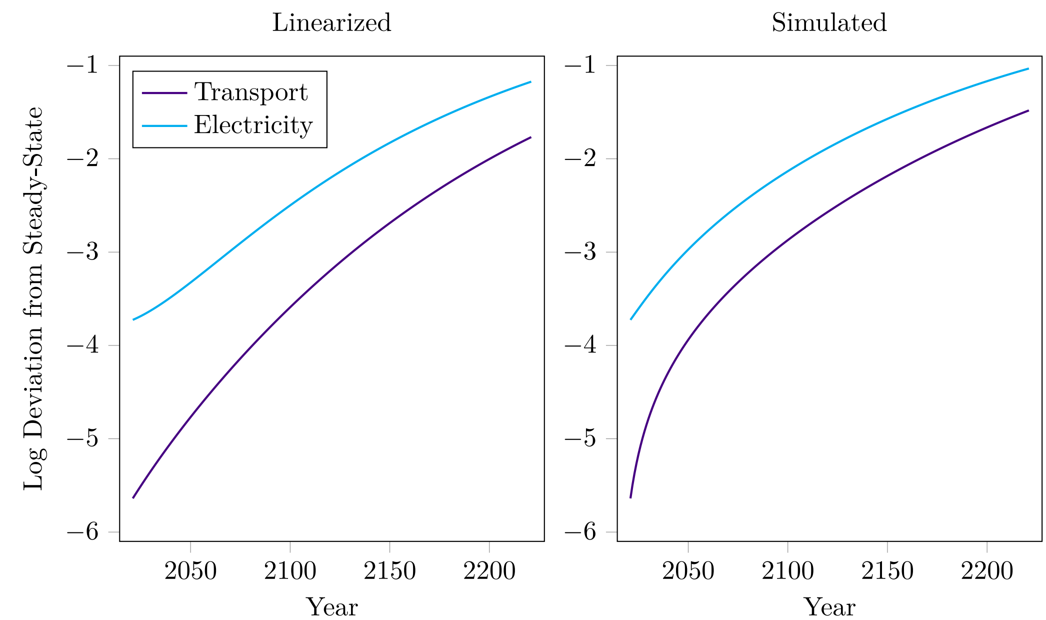
Notes: Tesla engineers modifying a gas-powered Lotus Elise to make a prototype electric vehicle.

Figure C.2: Bell Labs Solar Cell



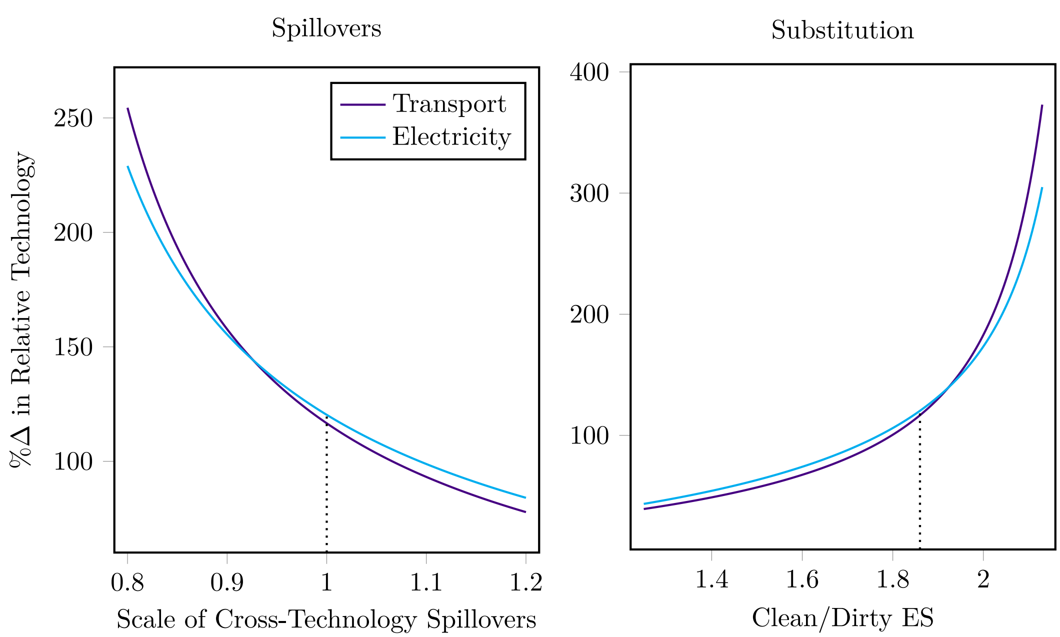
Notes: Bell Telephone Laboratories July 1954 Ad. Red box added to emphasize the description of knowledge spillovers.

Figure C.3: Linearized Transition Path Accurately Approximates Full Simulation



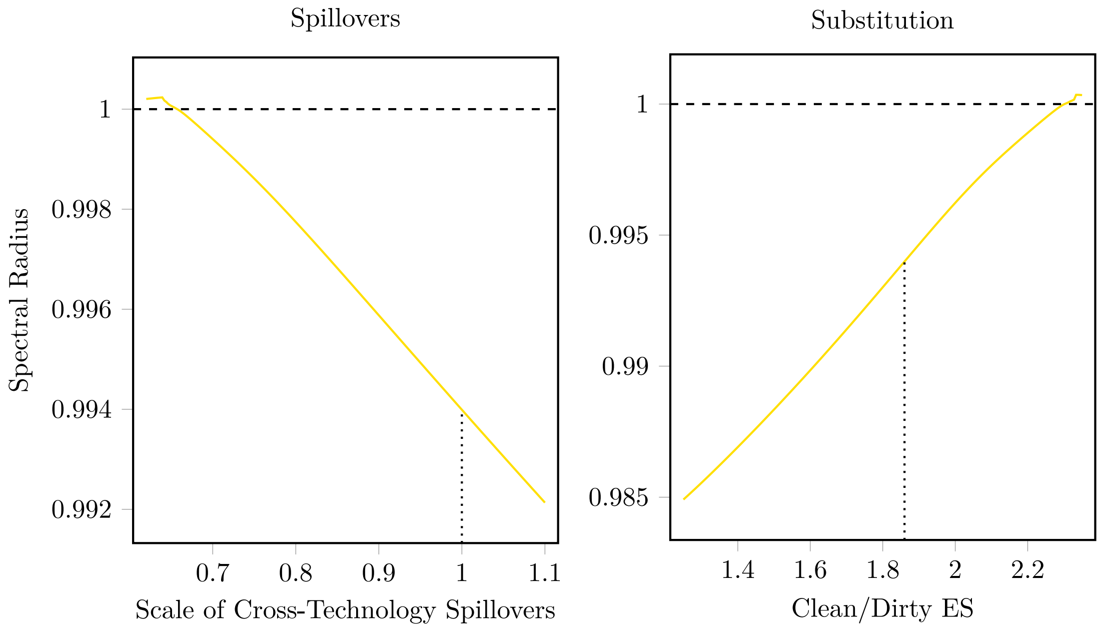
Notes: Technology's transition path following the policy reform of Section 6.1. The left panel displays the linearized path using Proposition 2, while the right panel displays the fully simulated path.

Figure C.4: Determinants of Steady-State Policy Impact



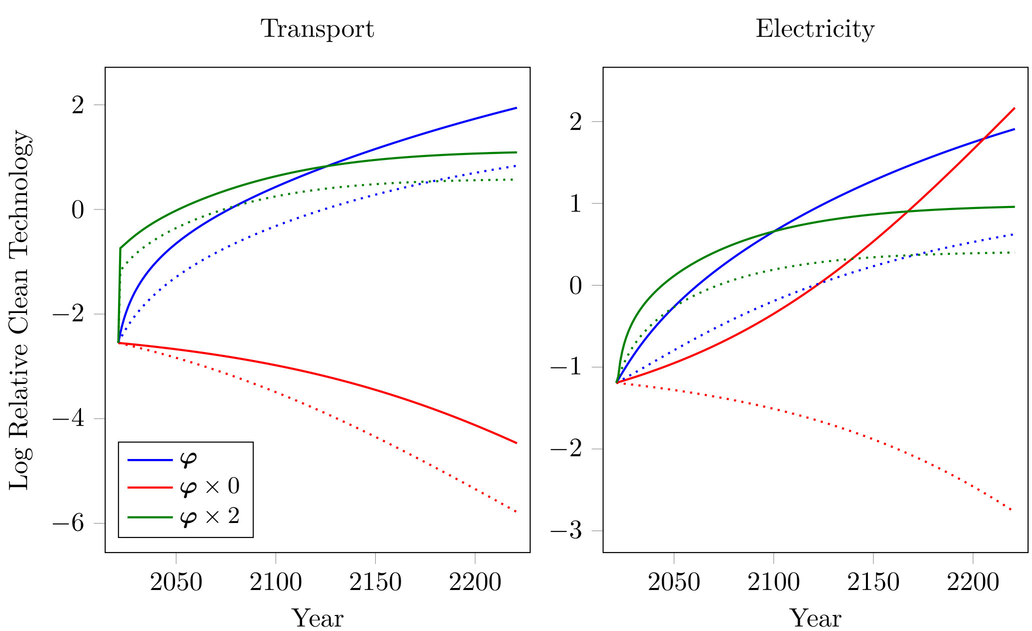
Notes: Impact of the policy reform of Section 6.1 as a function of the level of cross-technology knowledge spillovers and the elasticity of substitution between clean and dirty goods. Vertical dotted lines represent the benchmark calibration.

Figure C.5: Determinants of Increasing Returns to Innovation



Notes: Spectral radius of the transition matrix as a function of the level of cross-technology knowledge spillovers and the elasticity of substitution between clean and dirty goods. The spectral radius passes one when spillovers reduce to 66% of their calibrated level or the elasticity of substitution increases to 2.3. Vertical dotted lines represent the benchmark calibration.

Figure C.6: Technology Path ($190 Carbon Price)



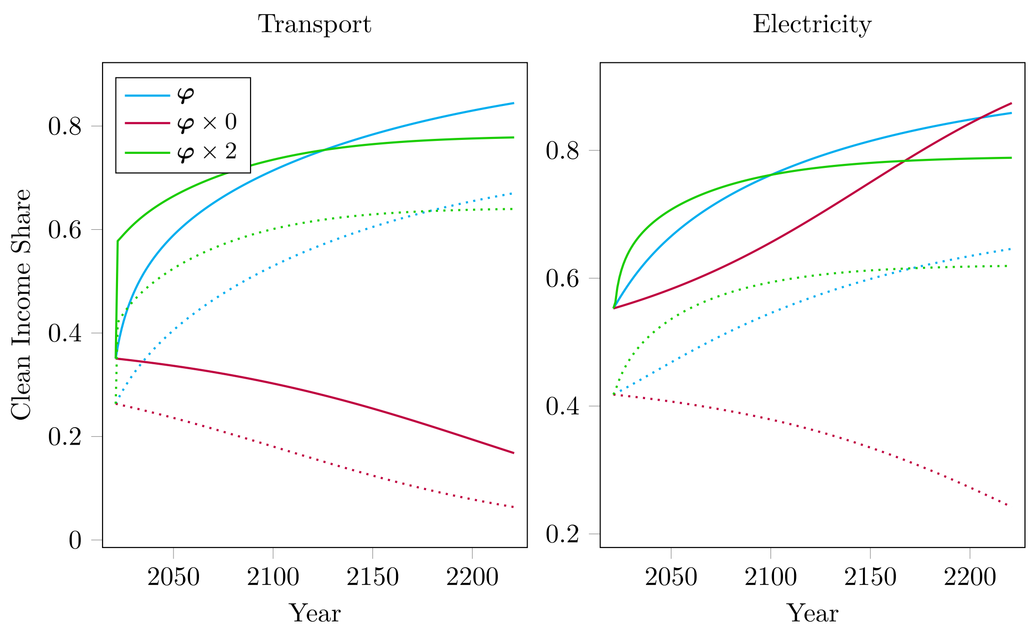
Notes: Impact of introducing a carbon price at the EPA's proposed SCC ($190) and clean innovation subsidy equivalent to a 30% tax credit ($\xi_c = 1.43$). Dotted lines indicate laissez-faire paths.

Table C.1: Impact of Policy Reform ($190 Carbon Price)

	No Spillovers	Calibrated Spillovers	Double Spillovers
Long-Run Impacts			
Relative Clean Technology by Sector			
$\%\Delta\bar{B}_{car}$	0%	+218.81%	+56.98%
$\%\Delta\bar{B}_{elec}$	$+\infty$%	+216.11%	+63.9%
Clean Income Shares by Sector			
ΔS_c^{car}	0 pp	+17.93 pp	+13.94 pp
ΔS_c^{elec}	+100 pp	+21.25 pp	+16.98 pp
Emissions Intensity			
$\%\Delta\bar{\omega}$	-87.42%	-93.49%	-84.39%
Transitional Impacts			
Half-Lives of Convergence by Sector			
$t_{car}^{(1/2)}$	–	149 years	21 years
$t_{elec}^{(1/2)}$	–	145 years	25 years
Carbon Emissions by Year			
$\%\Delta\mathcal{E}_{2035}$	-73.79%	-79.76%	-81.43%
$\%\Delta\mathcal{E}_{2060}$	-76.22%	-83.75%	-82.68%
Degree of Increasing Returns to Innovation			
Spectral Radius			
$\max\lvert\kappa_j\rvert$	1.01	0.994	0.978

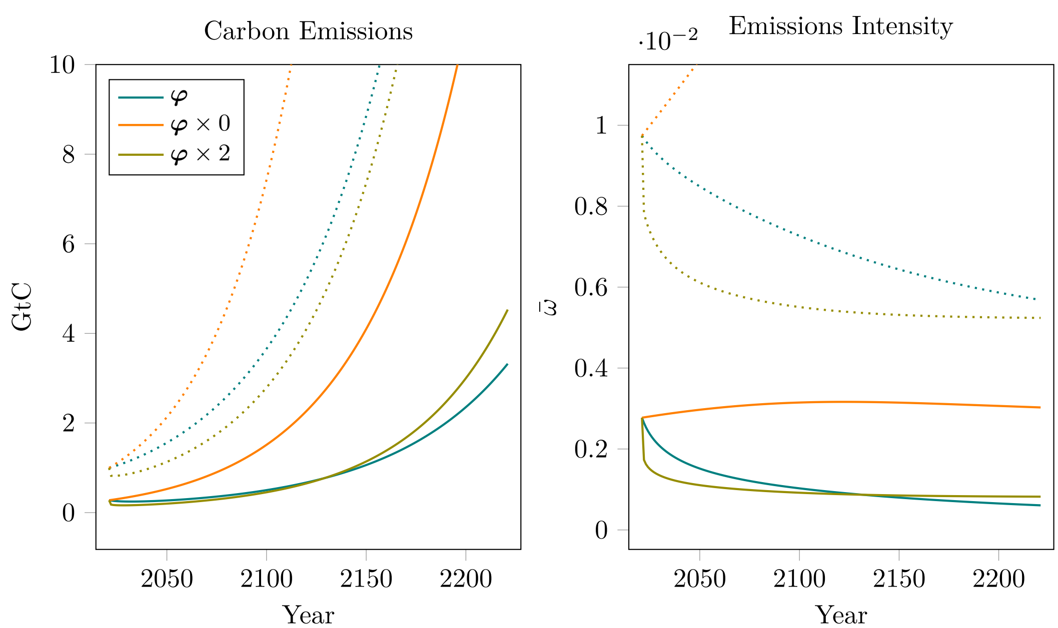
Notes: Impact of introducing a carbon price at the EPA's proposed SCC ($190) and clean innovation subsidy equivalent to a 30% tax credit ($\xi_c = 1.43$). Changes in relative technology are listed in log points. For path dependent economies, long-run impacts refer to corner, rather than interior, steady-states.

Figure C.7: Clean Income Share Path ($190 Carbon Price)



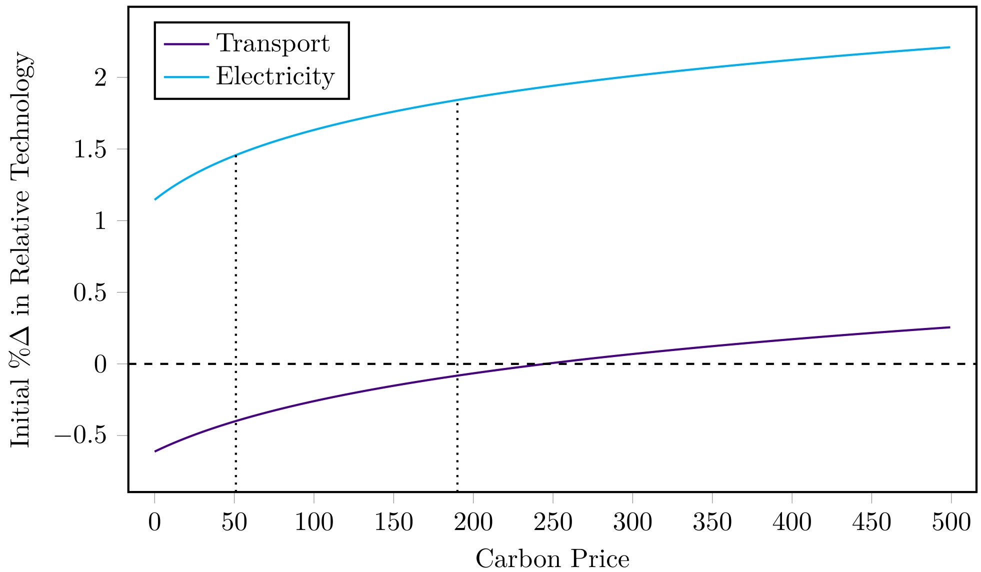
Notes: Impact of introducing a carbon price at the EPA's proposed SCC ($190) and clean innovation subsidy equivalent to a 30% tax credit ($\xi_c = 1.43$). Dotted lines indicate laissez-faire paths.

Figure C.8: Pollution Path ($190 Carbon Price)



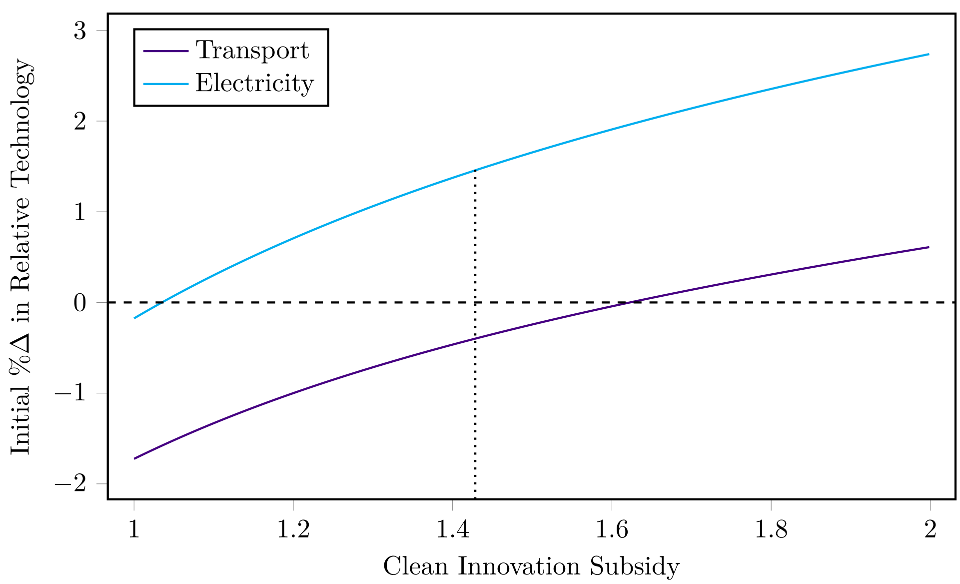
Notes: Impact of introducing a carbon price at the EPA's proposed SCC ($190) and clean innovation subsidy equivalent to a 30% tax credit ($\xi_c = 1.43$). Dotted lines indicate laissez-faire paths.

Figure C.9: Carbon Price Necessary for Clean Growth (No Spillovers)



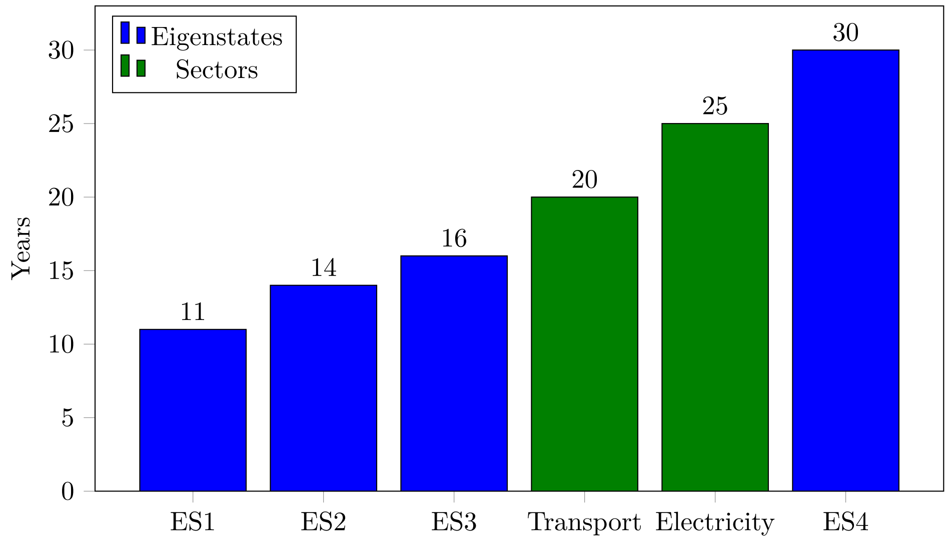
Notes: Direction of innovation as a function of the carbon price when the spillover network has been shut down. Values above zero indicate a sector is in its clean basin of attraction. Carbon prices are in addition to a clean innovation subsidy equivalent to a 30% tax credit ($\xi_c = 1.43$). Vertical dotted lines reference the Biden Administration's SCC ($51) and the EPA's proposed SCC ($190). Transportation switches to clean growth with a carbon price of about $245, while electricity generation switches without a carbon price.

Figure C.10: Clean Innovation Subsidy Necessary for Clean Growth (No Spillovers)



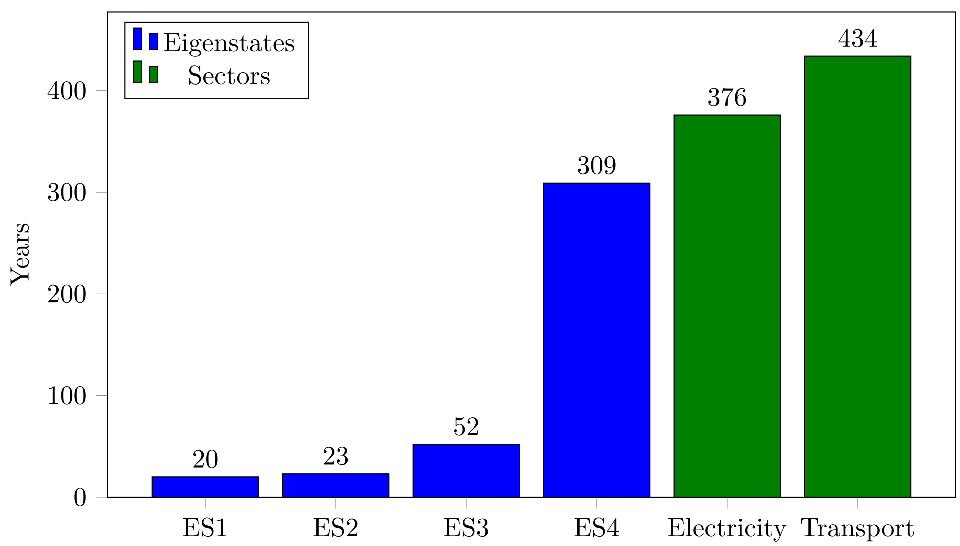
Notes: Direction of innovation as a function of the clean innovation subsidy when the spillover network has been shut down. Values above zero indicate a sector is in its clean basin of attraction. Clean innovation subsidies are in addition to a carbon price at the Biden Administration's SCC ($51). Vertical dotted line references clean innovation subsidy equivalent to a 30% tax credit ($\xi_c = 1.43$). Transportation switches to clean growth with a subsidy of 1.62, while electricity generation switches with a subsidy of 1.04.

Figure C.11: Half-Lives of Convergence (Double Spillovers)



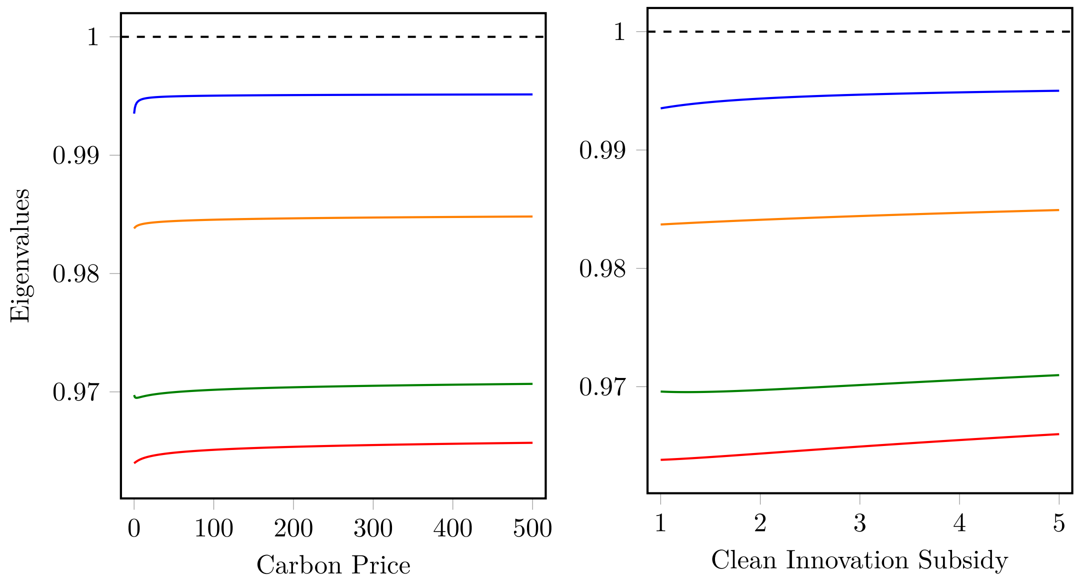
Notes: Transition speeds following the policy reform of Section 6.1. Cross-technology spillovers have been doubled.

Figure C.12: Half-Lives of Convergence (No Within-Sector Spillovers)



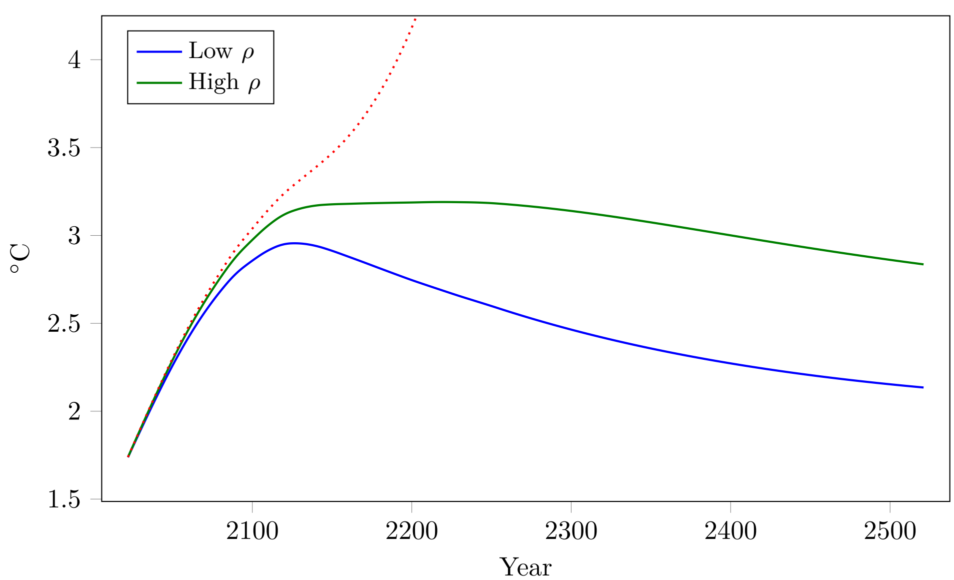
Notes: Transition speeds following the policy reform of Section 6.1. Within-sector spillovers are shut down by transferring them to the diagonal of the gross spillover network.

Figure C.13: Degree of Increasing Returns Doesn't Depend on Policy Reform



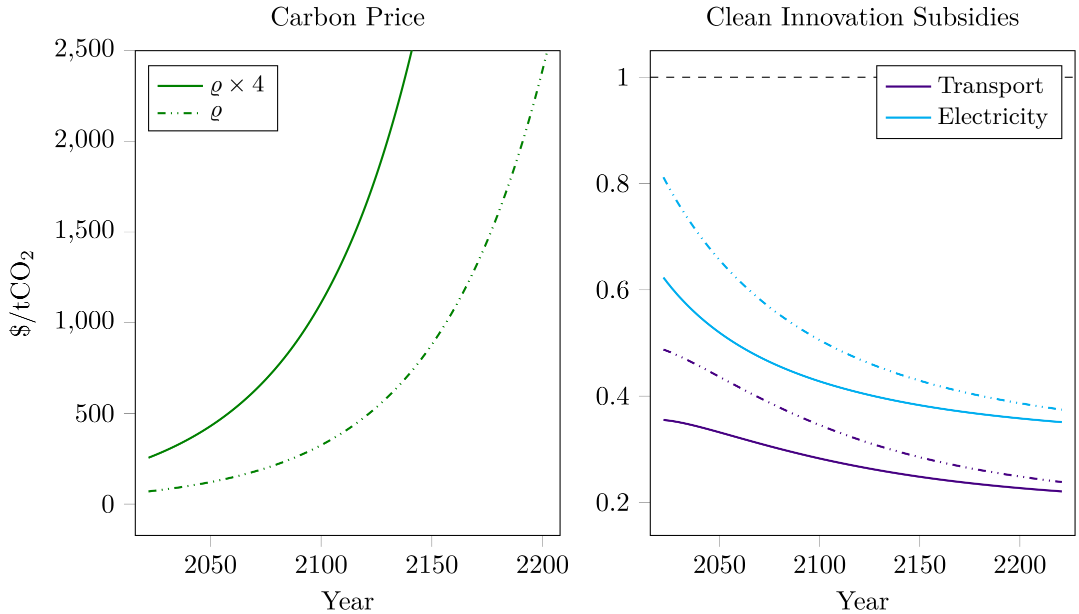
Notes: Eigenvalues of the transition matrix as a function of carbon prices and clean innovation subsidies. In each case, the other policy instrument is kept at its value from the policy reform of Section 6.1.

Figure C.14: First-Best Temperature Path



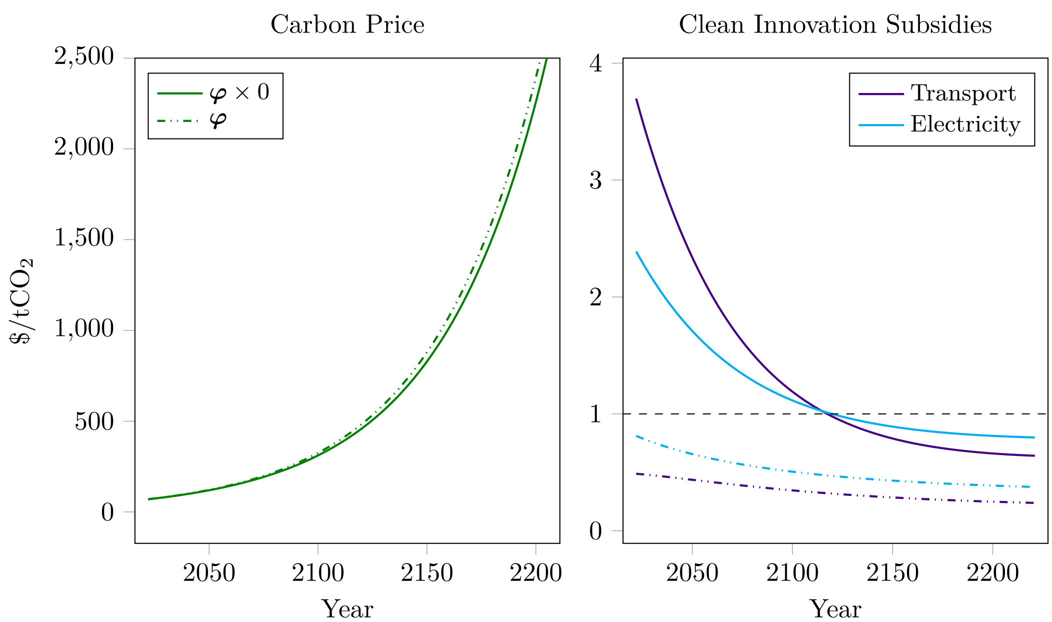
Notes: Temperature increases follow from the equation $\Delta T_t = \Gamma \ln(\mathcal{C}_t/\bar{\mathcal{C}})/\ln(2)$, with $\Gamma = 3$. See Footnote 45 for further discussion. Outside emissions come from the 2010 RICE model. The dotted line indicates the laissez-faire path.

Figure C.15: First-Best Policy Path (High Damages)



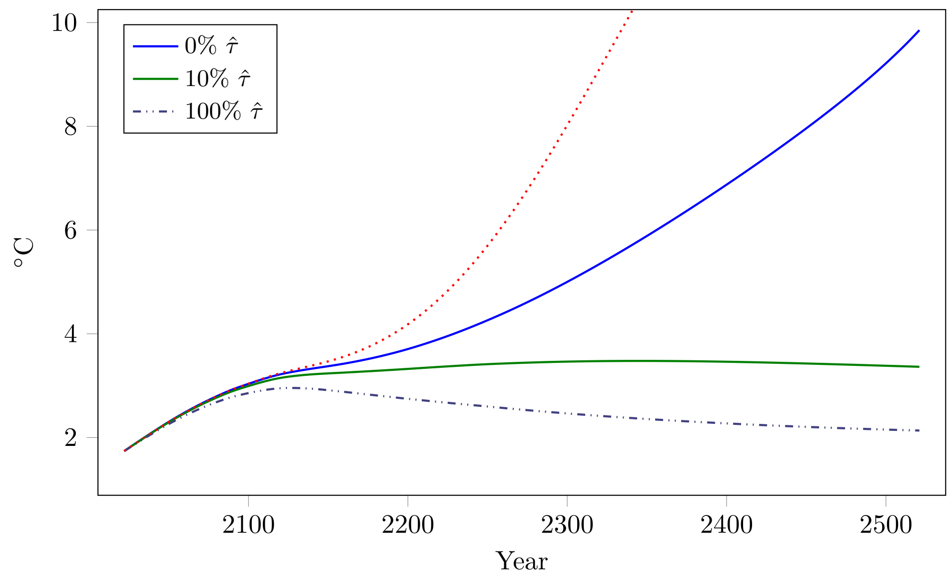
Notes: Optimal policy when climate damages are quadrupled. Dashed-dotted lines represent the benchmark calibration. Both cases use the low discount rate. Innovation subsidies are listed as a fraction of the baseline innovation wedge.

Figure C.16: First-Best Policy Path (No Spillovers)



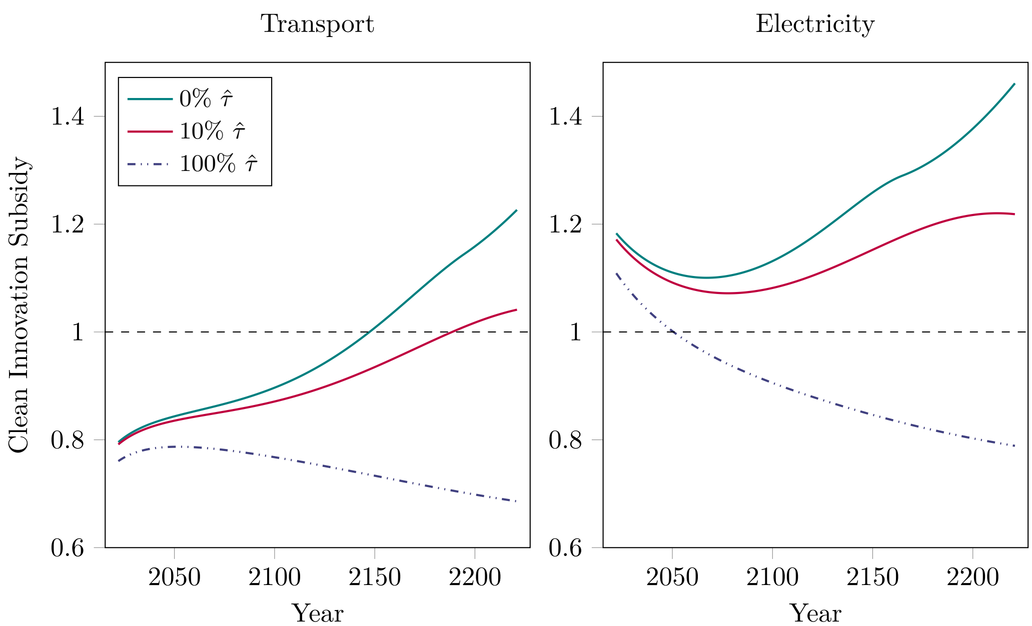
Notes: Optimal policy when the spillover network is shut down. Dashed-dotted lines represent the benchmark calibration. Both cases use the low discount rate. Innovation subsidies are listed as a fraction of the baseline innovation wedge.

Figure C.17: Second-Best Temperature Path



Notes: Temperature increases follow from the equation $\Delta T_t = \Gamma \ln(\mathcal{C}_t/\bar{\mathcal{C}})/\ln(2)$, with $\Gamma = 3$. See Footnote 45 for further discussion. Outside emissions come from the 2010 RICE model. Policy paths use the low discount rate. The external carbon price $\hat{\tau}$ is a proportion of the social cost of carbon, so 100% is the first-best. The dotted line indicates the laissez-faire path.

Figure C.18: Second-Best Policy Path (High Discounting)



Notes: Optimal second-best policy with the high discount rate. The external carbon price $\hat{\tau}$ is a proportion of the social cost of carbon, so 100% is the first-best. Innovation subsidies are listed as a fraction of the baseline innovation wedge.