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

Eric Donald†
University of Pittsburgh

February 19, 2025
Click here for latest version

Abstract

This paper argues that cross-technology knowledge spillovers are critical for understanding policy's role in the transition to clean technology. I develop an endogenous growth model with clean and dirty technologies and a network of cross-technology spillovers. The resulting formulas for the size and speed of technological transition, following a policy reform, show that greater spillovers induce faster transitions but reduce the long-run impact of policy. Such spillovers also prevent lock-in of dirty technology. Using patent citation data, I apply my model to US transportation and electricity generation and find that spillovers are mid-sized: they prevent lock-in but imply slow transitions with high long-run impact of policy. I then examine how cross-technology spillovers affect optimal policy, deriving innovation subsidy formulas under arbitrary carbon prices. Quantitatively, a "big push" of clean innovation subsidies is no longer optimal because the spillover network allows us to build toward clean technology in the future by innovating in other technologies today.

*I am deeply grateful to my advisors Pascual Restrepo, David Lagakos, and Yuhei Miyauchi for their guidance and support. I thank John Firth, Masao Fukui, Stefania Garetto, Adam Guren, Chad Jones, Bob King, Kevin Kuruc, Bernardo Ribeiro, John Sturm, and Stephen Terry as well as seminar participants at Arizona State, Bank of Canada, Boston University, Global Priorities Institute, Green Line Macro Meeting, Harvard Business School, SEA Annual Meeting, SED Annual Meeting, SED Winter Meeting, University of Georgia, University of Kentucky, and University of Pittsburgh for their helpful comments. I gratefully acknowledge funding as an Oxford Global Priorities Fellow.

†Email: eric.donald@pitt.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. Indeed, the European Commission states that one of the main objectives of its emissions trading scheme is to "promote investment in innovative, low-carbon technologies". The United States, where carbon pricing is absent, is on track to spend roughly $360 billion on subsidies for clean technology via the Inflation Reduction Act, much of which is explicitly geared toward encouraging clean innovation.

This paper argues that *cross-technology knowledge spillovers* are crucial for understanding the role of policy in transitioning the economy to cleaner production technologies. Such spillovers are typically excluded from existing models, but they underlied the invention of many clean technologies that are now at the vanguard of the energy transition. For instance, the first Tesla prototype – the Mule 1 – was a combustion engine car that the engineers at Tesla reconfigured by ripping out the engine and stuffing the engine compartment full of 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]

To study how cross-technology knowledge spillovers shape climate policy, I develop a general endogenous growth model with clean and dirty technologies whose key feature is a network of cross-technology knowledge spillovers. For policy instruments, I consider a price on carbon and technology-specific innovation subsidies. I show that cross-technology spillovers matter for both describing the impact of a policy reform and prescribing optimal clean innovation policy. Greater cross-technology spillovers allow for a more rapid transition to clean technology because they allow less advanced technologies to achieve catchup growth. However, by generating catchup growth, cross-technology spillovers also reduce the long-run impact of policy as they prevent any one technology from gaining too large a lead. Thus, depending on the level of cross-technology spillovers, a policy reform that favors clean technologies may generate rapid change in the short run or substantial redirection of innovation in the long run, but not both.

My main result is to provide formulas for the size and speed of technology's transition, following a policy reform, that depend on two matrices: one summarizing the network of knowledge spillovers and another summarizing substitution patterns in production. These formulas show that cross-technology spillovers provide a countervailing force to the substitutability of clean and dirty goods: the mechanism studied by Acemoglu et al. (2012). Indeed, higher substitutability of clean and dirty goods has the opposite effect of cross-technology spillovers – slowing down the transition and increasing the long-run impact of policy – by increasing the market size of more advanced technologies. My formulas summarize these forces in terms of sufficient statistic matrices, 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.

These formulas allow me to aggregate the two forces – spillovers and substitutability – into a single measure of *increasing returns to innovation*, which quantifies the degree to which more advanced technologies offer greater incentives for innovation. Formally, this measure is in terms of the spectral radius of a matrix that describes technology's transition path. As this measure increases, the transition slows

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

down while the long-run impact of policy grows, and if this measure exceeds one, technology becomes path dependent, locking in an inefficient dirty equilibrium. Thus, by reducing increasing returns to innovation, cross-technology spillovers also help to prevent technological lock-in, whereas the reverse is true for the substitutability between clean and dirty goods.

Next, I examine how cross-technology spillovers shape optimal policy, in particular optimal clean innovation subsidies. I derive a recursive formula for optimal innovation subsidies that holds for arbitrary, potentially suboptimal, carbon prices. I view this as a useful property because placing a proper price on carbon has been one of the main political challenges of climate policy. This formula has two components. The first term reflects the social value of knowledge spillovers sent through the spillover network, which can be measured in terms of a technology's centrality in the spillover network. Second, reflecting the pollution externality, is a term comprised of the wedge between the social cost of carbon and the price of carbon multiplied by the elasticity of emissions with respect to innovation. As a result, clean technologies that reduce equilibrium emissions are further subsidized, whereas dirty technologies that increase equilibrium emissions are punished. In the first-best, when there is an efficient price on carbon, the pollution externality term disappears, so innovation subsidies exclusively target spillover creation. In the second-best case where the carbon price is suboptimal, innovation subsidies are a compromise between the two externalities.

I then consider 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. Using the citation network of granted US patents as a proxy for the spillover network, I find that cross-technology spillovers are large enough to prevent lock-in of dirty technology in both sectors, but that the substitutability of clean and dirty goods still creates slow transitions with large long-run impacts of policy reforms. That is, my measure of increasing returns to innovation is near one from below. I validate my model by examining the progress made in clean transportation and electricity generation during the 2010s, showing that my model can match this episode as an untargeted moment. However, without cross-technology spillovers, my model makes the counterfactual prediction that dirty technologies would have become *more* entrenched in both sectors, not less.

I apply my model to the data with several quantitative exercises. First, I examine the impact of introducing a $51 carbon price (the Biden Administration's estimate of the social cost of carbon) and a uniform clean innovation subsidy equivalent to a 30% tax credit (consistent with the subsidies in the Inflation Reduction Act). I find a large long-run impact of this policy reform, with the steady-state prevalence of clean technology increasing by 117.3% and 120.8% for transportation and electricity generation, respectively. However, the transition to this new steady-state is slow, with half-lives of convergence of 119 and 131 years. In the absence of cross-technology spillovers, the economy would start out locked in a dirty equilibrium, and the policy reform would only be enough to induce a slow switch to clean innovation in electricity generation.

Second, I simulate the optimal path of innovation subsidies for a variety of potential restrictions on the carbon price. Conventional wisdom argues for a "big push" of large, temporary clean innovation subsidies to jolt the economy away from dirty technology (Acemoglu et al., 2012, 2016). Indeed, in the absence of cross-technology spillovers, this is what my model recommends: clean innovation subsidies for transport and electricity generation that start out at 365.9% and 217.5% of the common innovation

wedge and reduce down to baseline over the first century of policy. This prescription is motivated by the problem of technological lock-in, but according to my calibration, this is not the empirically relevant case for the US economy. Instead, I find that lower, more stable subsidies for clean technology are optimal, and this holds quantitatively even when the carbon price is set below the social cost of carbon.

In the first-best, when carbon pollution is priced at its social marginal cost, clean transportation and electricity generation receive average innovation subsidies of only 39.4% and 62.4% of the common innovation wedge over the first century of policy. This is because, when spillovers occur across technologies, we can build toward clean production in the future by innovating in a broader range of technologies as they will produce future spillovers onto clean technologies through the network. In the absence of a spillover network, society can only build towards a clean production system in the future with clean innovation today. Thus, a spillover network prevents the need to immediately shift innovation incentives toward clean technology. The undesirability of a big push is also robust to the choice of social discount rate and elasticity of substitution for the spillover functions.

Even in the second-best, where there is an additional justification for clean innovation subsidies, a big push is not warranted with an empirically disciplined spillover network. I consider two cases for restrictions on the carbon price: one where the carbon price is ten percent of the social cost of carbon and another where it is zero. In the case with a small carbon price, clean innovation subsidies shift upward by about 20-30% of the common innovation wedge but otherwise retain the same pattern as in the first-best. When a price on carbon is lost altogether, I find that clean innovation subsidies should start with a similar upward shift as in the small carbon price case and then rise rapidly after the first century of policy; an ongoing nudge rather than a big push.

Inefficient carbon prices only result in a small increase in optimal clean innovation subsidies because clean innovation alone is ineffective at reducing emissions. That is, better clean technology reduces emissions intensity through substitution away from dirty goods, but it also generates a rebound effect by increasing the economy's productive capacity. Relatedly, my model allows me to quantify the welfare losses from restrictions on the carbon price. A small, growing carbon price eventually pushes dirty technology out of production, allowing some dirty innovation to continue for the sake of spillover creation. As a result, the welfare loss relative to the first-best is small. 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 the 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. Hence, the welfare cost of restrictions on the carbon price is highly nonlinear.

1.1 Related Literature

This paper contributes to several strands of literature. First is the extensive literature on endogenous innovation and the climate. Typical macro models of climate innovation exclude the possibility of cross-technology knowledge spillovers, assuming that spillovers take place exclusively within technologies (Gans, 2012; Acemoglu et al., 2012, 2016, 2023; Hassler et al., 2021). As argued in the seminal work of Acemoglu et al. (2012), the substitutability of clean and dirty goods alone generates destabilizing increasing returns to innovation that must be overcome with a big push of clean innovation subsidies. Fried (2018) allows for cross-technology spillovers in a quantitative framework but does not analytically

characterize their relevance for describing the impact of a policy reform or prescribing optimal clean innovation policy. Aghion et al. (2024) studies the decision to adopt clean technologies, focusing on the strategic complementarity of adoption decisions in the context of a supply chain.

My paper also relates to the topic of technological path dependence. 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 an 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). Examples of theoretical contributions from the skill-biased technological change literature include Acemoglu (1998, 2002) and Acemoglu and Restrepo (2018, 2022a). This literature has long argued that the degree of substitutability across goods is a fundamental determinant of the impact of input prices on the direction of innovation. My paper argues that cross-technology knowledge spillovers deserve a similar status, while also providing guidance for the measurement and aggregation of both forces.

My paper also relates to the vast literature on integrated assessment models (IAMs) that seeks to quantify the social harm of carbon pollution. Examples from this literature include Stern (2007), Golosov et al. (2014), and Nordhaus (2017).[2] 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 incomplete. 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 paper is closely related to Liu and Ma (2021).[3] They argue that innovation policy should consider the value of technology in producing both goods and future innovations, with the weight between the two determined by the Planner's level of patience. My innovation subsidies recover their formulas as a special case whenever carbon prices are efficient and the economy is in steady-state. 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). Their paper considers 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 I use 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 Section 4, I derive optimal innovation subsidies in the face of arbitrary carbon prices. In Section 5, I describe the calibration of the spillover network and other structural parameters of my model. In Section 6, I perform quantitative exercises that simulate the impact of a policy reform as well as optimal

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

[3]In general, knowledge spillovers have long been a central consideration of innovation policy (Arrow, 1962; Romer, 1990; Bloom et al., 2019; Akcigit et al., 2021; 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).

climate policy. Section 7 concludes.

2 Model

This section lays out the endogenous growth model used throughout the paper. Time is discrete and indexed by t. There are J-many technologies in the economy, each of which is classified as either clean and dirty. The innovation process 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 dirty technologies (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.

2.1 Production

Each technology in this economy produces a distinct good, which 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

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

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. 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.

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:

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

The economy has a fixed endowment of labor that is supplied inelastically, so the allocation of 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.[4]

Finally, the resource constraint requires that

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

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 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

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

[4] 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).

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 $\varphi_{ijt}|_{i\neq j} \geq 0$, which guarantees that technologies receive higher spillovers when they are relatively less advanced. By assuming that spillover functions are homogeneous of degree zero, I ensure balanced growth in the presence of a fixed supply of scientists.[5] It will occasionally be helpful to refer to *gross spillover network* $\tilde{\boldsymbol{\varphi}}_t$, 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 sum to one, rather than zero.

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. Note that there would still be spillovers *within* technologies as Equation (8) states that innovation builds multiplicatively on a technology's existing knowledge 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). In a setting where scientists have an outside option such as

[5] Allowing for $\sum_j \varphi_{ijt} < 0$ would be consistent with the fishing out feature of semi-endogenous growth models (Jones, 2022). The ongoing work of Donald et al. (2025) shows that all the theoretical results presented here carry over with minor adjustment to the semi-endogenous case; balanced growth just requires an exponential increase in the supply of scientists. In the present setting, 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. However, if all technologies advance concurrently, then reductions in research productivity can be avoided.

production work or leisure, the equilibrium supply of scientists is likely to be below the optimum. This has been the traditional justification of innovation subsidies (Romer, 1990). Instead, when the supply of scientists is fixed and there are multiple technologies, the concern is that the equilibrium composition of scientific effort may not be optimal.

2.3 Preferences & Policy Instruments

The Planner seeks to maximize the utility of a representative household with two main policy instruments: a carbon price τ_t and technology-specific innovation subsidies 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.[6]

As discussed below, producers of intermediates have market power, so I will allow the Planner to subsidize away 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 Υ are constant, which would be the case in both laissez-faire and the first-best. 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 in a given technology grants exclusive ownership of the most advanced form of production for a random 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

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

[6] A distinction can be drawn between the household's (private) discount rate and the Planner's (social) discount rate, but this difference only matters when the household is making an intertemporal decision that the Planner can influence with policy (e.g. Farhi and Werning (2007)). None of the equilibrium outcomes of this economy depend on the household's discount rate, so I can ignore the distinction and interpret ρ as the social discount rate.

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

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

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

One can see from Equation (17) that carbon pricing raises the cost of dirty inputs.

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

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

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

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

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 for innovating in technology j scales 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 boosting the research productivity of relatively less advanced technologies. Third, technologies with larger markets have larger innovation rents, and therefore, stronger incentives to innovate. It is through this market size channel that substitution patterns in production will be relevant because substitution patterns determine how price changes, from changes in carbon prices or technology, influence a technology's market size.

Finally, the household owns all 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 Appendix A.1, I provide a sufficient condition for the existence and uniqueness of the economy's equilibrium path, given some initial condition.

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 functional forms. 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 share parameters $\{\nu_\theta\}$.

Output at the sector level follows

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

where $Y_{\theta c}$ and $Y_{\theta d}$ are the clean and dirty forms of production for sector θ. 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 only has a clean form of production. My application considers transportation and electricity generation, 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.

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

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

The task-based framework shows that one can interpret sectoral CES shares ν_θ as the share of tasks in the economy performed by sector θ (Acemoglu and Restrepo, 2022b). Thus, the relevant notion of effective scientific effort is the number of scientists per economic task associated with a technology.

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, so the Cobb-Douglas exponents define the gross spillover network $\tilde{\varphi}$. The denominator represents the recipient's own knowledge stock. 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.

3 Policy's Impact on the Direction of Innovation

This section characterizes how the direction of innovation responds to a policy reform. Specifically, I will imagine there is a once-and-for-all policy reform $\{d\tau, \{d\xi_j\}\}$.[7] 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*.[8]

Both the amplification and transition matrix are composed of the same two sufficient statistic matrices: one describing the network of knowledge spillovers and another describing substitution patterns in production. These two matrices show how two forces, spillovers and substitutability, determine both the change in technology's steady-state and the speed of convergence to the 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 its spectral radius. 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.

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) the network of cross-technology knowledge spillovers and (ii) substitution patterns in production.

First, consider the spillover matrix:

Definition 3 (Spillover Matrix $\mathbf{\Phi}$)**.** *The spillover matrix is a $J-1 \times J-1$ matrix with elements*

$$\Phi_{ijt} \equiv \frac{\partial \ln\left(\phi_{Jt}/\phi_{it}\right)}{\partial \ln\left(A_{jt-1}/A_{Jt-1}\right)}. \tag{33}$$

[7] 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.

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

The spillover matrix $\mathbf{\Phi_t}$ describes how relative spillovers change in response to changes in relative technology, where the base technology is J.[9] The diagonal of $\mathbf{\Phi}_t$ provides a measure of the scale of cross-technology spillovers in an economy. The fact that spillover functions are homogeneous implies that

$$\Phi_{iit} = \sum_{j \neq i} \left(\varphi_{ijt} - \varphi_{Jjt} \right), \tag{34}$$

so the diagonal elements of $\mathbf{\Phi}_t$ will be large when technologies receive high cross-technology spillovers, relative to the spillovers received by the base technology. The diagonal is thus informative of the extent to which cross-technology spillovers will stimulate catch-up growth. The off-diagonal elements reflect the degree to which other technologies provide greater spillovers to the base technology than technology i. These terms offset the diagonal insofar as the base technology tends to receive greater spillovers.[10] Therefore, the spillover matrix $\mathbf{\Phi}_t$ provides a general description of an economy's spillover structure for a wide class of spillover functions.

Second, consider the substitution matrix:

Definition 4 (Substitution Matrix $\mathbf{\Sigma}$)**.** *The substitution matrix is a* $J-1 \times J-1$ *matrix with elements*

$$\Sigma_{ijt} \equiv \frac{\partial \ln \left(Y_{it} / Y_{Jt} \right)}{\partial \ln \left(p_{Jt} / p_{jt} \right)}. \tag{35}$$

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, so we have the usual interpretation that values above one describe substitutes while values below one describe complements. As an illustration, if the output aggregator were a standard CES with elasticity of substitution σ, we would have $\mathbf{\Sigma}_t = \sigma \mathbf{I}$. In the general case, the off-diagonal elements allow for more complicated substitution patterns where the relative demand for one good depends on the relative price of another. Therefore, the demand responses described by the substitution matrix $\mathbf{\Sigma}_t$ provide a general characterization of substitution patterns in production.[11]

3.2 Steady-State Impact of Policy

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

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

[9] One can map from the spillover network to the spillover matrix using the equation $\Phi_{ijt} = \varphi_{Jjt} - \varphi_{ijt}$, which implies that any estimate of the spillover network also provides an estimate of the spillover matrix.

[10] A symmetric interpretation of the spillover matrix notes that $d \ln \left(\phi_{it} / \phi_{Jt} \right) = - \sum_{j<J} \Phi_{ijt} d \ln \left(A_{jt-1} / A_{Jt-1} \right)$. Thus, the diagonal of $-\mathbf{\Phi_t}$ reflects the degree to which technologies harm their own relative research productivity by becoming relatively more advanced, while the off-diagonal reflects the offsetting impact of cross-technology spillovers. The interpretation of $\mathbf{\Phi}_t$ simply flips this logic and notes that, when spillover functions are homogeneous of degree zero, cross-technology spillovers must exactly balance out one's own fishing out effect.

[11] I show in Appendix A.2 that with nested-CES production, $\mathbf{\Sigma}_t$ is a block-diagonal matrix that depends on the elasticities of substitution of both nests. I use this closed-form solution to estimate $\mathbf{\Sigma}_t$ in Sections 5 and 6.

I will describe the steady-state 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

$$\bar{A}_{jt} \equiv \frac{A_{jt}}{A_{Jt}}, \tag{36}$$

where the base technology is J. In Appendix A.3, I provide conditions for the existence and uniqueness of the economy's steady-state, and in Appendix A.7, I derive a closed-form solution for the steady-state growth rate g.

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*

$$d\ln(\bar{A}_{ss}) = \eta \boldsymbol{\mathcal{M}} \big[d\ln(\Xi) - \alpha \big(\boldsymbol{\Sigma} - \boldsymbol{I} \big) d\ln(\mathcal{R}) \big], \tag{37}$$

where $d\ln(\Xi_j) \equiv d\ln\big(\xi_j/\xi_J\big)$ and $d\ln(\mathcal{R}_j) \equiv d\ln\big((r_j + \omega_j\tau)/(r_J + \omega_J\tau)\big)$ are the changes in 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

$$\boldsymbol{\mathcal{M}} = \big[\boldsymbol{\Phi} - \eta(1-\alpha)\big(\boldsymbol{\Sigma} - \boldsymbol{I}\big) \big]^{-1}. \tag{38}$$

The proof of Proposition 1 can be found in Appendix A.3. Proposition 1 defines the amplification matrix $\boldsymbol{\mathcal{M}}$, which describes the role of spillovers and substitutability 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.[12] Thus, increases in relative subsidies raise the prevalence of favored technologies in steady-state, while higher carbon prices reduce the prevalence of dirty technologies in steady-state, insofar as goods are substitutes.

The amplification matrix $\boldsymbol{\mathcal{M}}$ describes the full general equilibrium impact of these policy reforms, and it depends on the two forces of spillovers and substitutability. Greater cross-technology spillovers reduce the long-run impact of policy by increasing the research productivity of less advanced technologies, thereby preventing technologies favored by policy from gaining too large of a lead. Conversely, greater substitutability increases the long-run impact of policy by increasing the market size of more advanced technologies, raising the incentive to further innovate in technologies favored by policy. The opposite is true when goods are complements because complementarity lowers the market size of more advanced technologies.

The amplification matrix reflects these two forces via an inverse matrix because, following the logic of a Leontief inverse, policy reforms will generate a full sequence of ripple effects that sum to the amplification matrix.[13] Intuitively, the spillover matrix decreases the "size" of the amplification matrix

[12] Corollaries 1 and 2 will provide an economic interpretation for the case when the amplification matrix does flip the sign of policy impacts. When technology is path dependent, the steady-state defines the boundary of the different basins of attraction. Thus, policy reforms that favor clean technology will shrink the dirty basin of attraction, expanding the set of initial conditions that lead to clean growth.

[13] Formally, we have that

$$\boldsymbol{\mathcal{M}} \propto \mathbf{I} + \big(\boldsymbol{\Sigma} - \frac{1}{\eta(1-\alpha)} \boldsymbol{\Phi} \big) + \big(\boldsymbol{\Sigma} - \frac{1}{\eta(1-\alpha)} \boldsymbol{\Phi} \big)^2 + \ldots,$$

as it enters with a positive in the "denominator", while the substitution matrix increases the "size" of the amplification matrix as it enters with a negative in the "denominator".

3.3 Technology's Transition Path

Having examined how policy affects the 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. This matrix allows me to describe the transition path of technology in terms of the log deviation of relative technology from steady-state

$$\bar{\mathcal{A}}_t \equiv \ln\left(\bar{A}_t\right) - \ln\left(\bar{A}_{ss}\right). \tag{39}$$

The following proposition provides a first-order characterization of technology's transition.

Proposition 2. *Given $\bar{\mathcal{A}}_0$, the transition path of technology follows, to a first-order, the linear process*

$$\bar{\mathcal{A}}_t \approx \boldsymbol{\mathcal{J}}\bar{\mathcal{A}}_{t-1}, \tag{40}$$

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

$$\boldsymbol{\mathcal{J}} = \left[(1-\eta)\boldsymbol{I} - g\eta(1-\alpha)\left(\boldsymbol{\Sigma} - \boldsymbol{I}\right)\right]^{-1}\left[(1-\eta)\boldsymbol{I} - g\boldsymbol{\Phi}\right]. \tag{41}$$

The proof of Proposition 2 can be found in Appendix A.5. The main object of interest in Proposition 2 is the transition matrix $\boldsymbol{\mathcal{J}}$. We can see that the speed of transition depends on how quickly the transition matrix $\boldsymbol{\mathcal{J}}$ shrinks technology's initial log deviation from steady-state $\bar{\mathcal{A}}_0$. 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. In the following section, I will describe how to quantify the speed of convergence via an eigendecomposition.

The transition matrix $\boldsymbol{\mathcal{J}}$ depends on the same two sufficient statistic matrices as before, except now their influence is reversed. Now, cross-technology spillovers speed up the transition because increasing the research productivity of less advanced technologies allows for catchup growth. Conversely, greater substitutability slows down the transition because it allows more advanced technologies to maintain a larger market size for longer. This makes it makes it more difficult to dislodge incumbents. Again, the opposite will be true for complements.

The rightmost matrix of Equation (41) – the one that contains the spillover matrix – describes the initial impact of perturbing lagged technology because spillovers depend on the inherited state of technology. However, the substitution matrix shows up through an inverse matrix because market size depends on the realized state of technology. So we again have the logic of a Leontief inverse as the initial impact of perturbing lagged technology generates a full sequence of market size ripple effects that culminates in an inverse matrix.[14] Intuitively, the spillover matrix decreases the "size" of the transition

reflecting the full sequence of ripple effects from the policy reform.

[14] Formally, we have that

$$\left[(1-\eta)\mathbf{I} - g\eta(1-\alpha)\left(\boldsymbol{\Sigma} - \mathbf{I}\right)\right]^{-1} \propto \mathbf{I} + \left(\frac{g\eta(1-\alpha)}{1-\eta}(\boldsymbol{\Sigma} - \mathbf{I})\right) + \left(\frac{g\eta(1-\alpha)}{1-\eta}(\boldsymbol{\Sigma} - \mathbf{I})\right)^2 + \ldots,$$

matrix as it enters with a negative in the "numerator", while the substitution matrix increases the "size" of the transition matrix as it enters with a negative in the "denominator".

3.4 Degree of Increasing Returns to Innovation

In this section, I eigendecompose the transition matrix and show that the spectral radius – its largest eigenvalue – gives us a summary statistic of the degree of increasing returns to innovation. This measure of increasing returns allows for scalar interpretation of both the amplification and transition matrix, with the implication that the long-run effect of policy on innovation and the speed with which policy can redirect innovation are inversely related.

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. First, consider the eigendecomposition of the transition matrix

$$\boldsymbol{\mathcal{J}} = \boldsymbol{Q}\boldsymbol{D}(\kappa)\boldsymbol{Q}^{-1}, \tag{42}$$

where $\boldsymbol{D}(\kappa)$ is a diagonal matrix whose diagonal elements are the eigenvalues of the transition matrix, and $\boldsymbol{Q}$ is a 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). Next, iterating forward from technology's initial condition, we have

$$\bar{\mathcal{A}}_t \approx \boldsymbol{Q}\boldsymbol{D}(\kappa)^t\boldsymbol{Q}^{-1}\bar{\mathcal{A}}_0. \tag{43}$$

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, while eigenvalues slightly below one will gradually shrink.

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

$$\beta = \boldsymbol{Q}^{-1}\bar{\mathcal{A}}_0. \tag{44}$$

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.[15] 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.[16] 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

which describes the full impact of market size effects from perturbing lagged technology.

[15] 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.

[16] 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.

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*

$$\bar{\mathcal{A}}_t \approx \sum_j \kappa_j^t \beta_j Q_j, \tag{45}$$

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 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. We can interpret the eigenvalues as a measure of increasing returns to innovation because they tell us the degree to which more advanced technologies offer greater incentives for innovation. If more advanced, incumbent technologies are difficult to dislodge, this will show up as large eigenvalues. Moreover, to aggregate the two forces of spillovers and substitutability into a single summary statistic of increasing returns to innovation, we can use the spectral radius, or the largest eigenvalue, as this will tend to have an outsized influence on the speed of convergence.

Measuring increasing returns with the spectral radius allows me to derive a condition for technological path dependence, where 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 approach one, and once an eigenvalue 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 spectral radius of the transition matrix $\max|\kappa_j|$ exceeds one.*

If the initial state of technology loads on an eigenstate with an eigenvalue exceeding one, then technology will diverge in the direction of this unstable eigenstate. Consequently, the 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 and transition matrix.

Corollary 2. *The amplification matrix follows*

$$\boldsymbol{\mathcal{M}} = g\boldsymbol{Q}\boldsymbol{D}(1-\kappa)^{-1}\boldsymbol{Q}^{-1}\big[(1-\eta)\boldsymbol{I} - g\eta(1-\alpha)\big(\boldsymbol{\Sigma}-\boldsymbol{I}\big)\big]^{-1}. \tag{46}$$

The proof of Corollary 2 can be found in Appendix A.6. This result establishes that increasing returns to innovation increase the long-run impact of policy reforms through the term $\boldsymbol{D}(1-\kappa)^{-1}$, which implies that the steady-state 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 will be large, but the transition will be slow. Conversely, when increasing returns are low ($\kappa_j \to 0$), the steady-state impact 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, but the same mechanism leads to sluggish transitions by preventing less advanced technologies from catching up. This is why the two forces of spillovers and substitutability have opposite effects in the long run and transition.

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

$$t_{ES_j}^{(1/2)} = \left\lceil \frac{\ln(1/2)}{\ln(\kappa_j)} \right\rceil, \tag{47}$$

where $\lceil . \rceil$ is the ceiling function. Thus, if the initial state of technology is proportional to the jth eigenstate, technology's half-life will also 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.

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 $\bar{A}_t = A_{ct}/A_{dt}$. I will assume a mild regularity condition to ease exposition.[17]

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

$$\Sigma = \sigma \tag{48}$$

$$\Phi = \varphi_{cd} + \varphi_{dc}. \tag{49}$$

That is, Σ is the elasticity of substitution between clean and dirty production, and Φ is the sum of cross-technology knowledge spillovers in the economy. Consistent with Proposition 1, the steady-state for relative technology follows

$$\ln(\bar{A}_{ss}) = \eta \frac{\ln(\Xi) - \alpha(\Sigma - 1)\ln(\mathcal{R})}{\Phi - \eta(1-\alpha)(\Sigma - 1)}. \tag{50}$$

From this, we can see that the amplification matrix follows

$$\mathcal{M} = \frac{1}{\Phi - \eta(1-\alpha)(\Sigma - 1)}. \tag{51}$$

[17] 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 $g = 0.02$, shrinking the right-hand side of both inequalities.

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

$$\mathcal{J} = \frac{(1-\eta) - g\Phi}{(1-\eta) - g\eta(1-\alpha)(\Sigma-1)}. \tag{52}$$

The speed of convergence is increasing in the level of cross-technology spillovers Φ and decreasing in the elasticity of substitution Σ, as argued in Section 3.3. In particular, technology's half-life follows

$$t^{(1/2)} = \left\lceil \frac{\ln(1/2)}{\ln(\mathcal{J})} \right\rceil. \tag{53}$$

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

$$\Phi < \eta(1-\alpha)(\Sigma-1). \tag{54}$$

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. 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. This process will continue until, in the limit, all production and innovation is devoted to the more advanced technology. This is the path dependence scenario considered by Acemoglu et al. (2012).

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

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

Thus, increasing returns to innovation, as measured by $\mathcal{J}$, slow the transition while at the same time increasing the long-run impact of policy.

4 Optimal Policy

We have seen the response of the direction of innovation to an arbitrary policy reform, so I will now characterize optimal climate innovation policy. I consider a realistic second-best problem where the Planner must choose innovation subsidies when there is some externally imposed carbon price. In this setting, optimal innovation subsidies follow a recursive formula that both corrects the knowledge spillover externality as well as indirectly reduces emissions by rewarding clean technologies. That is, the Planner must correct both externalities with a single instrument. This formula nests the first-best because if carbon is priced at its social marginal cost, innovation subsidies focus exclusively on knowledge spillovers, in keeping with the Pigou Principle.

4.1 Planner's Problem

I start by defining the Planner's problem when they cannot set the price on carbon. I view this as a realistic restriction because pricing carbon has proved to be politically difficult.

Let $\{\tau_t\}$ 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 Planner still has a complete set of policy instruments on the innovation side of the economy, but the inability to specify the carbon price implies the Planner must satisfy an incentive compatibility constraint. To keep focus on innovation, I will assume that intermediate subsidies are set to close the monopolist markup $\Upsilon = \gamma$. The Planner then solves a primal problem where they pick the allocation of scientists directly and back out the innovation subsidies that implement this allocation as an equilibrium.

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

$$\begin{aligned}
\max_{\{c_t,\{A_{jt},s_{jt}\}\}} \sum_{t\geq 0} \frac{1}{(1+\rho)^t} u(c_t) \;\; s.t. \qquad (56)\\
\mathcal{Y}_t &= c_t + \sum_j r_j \Lambda_{jt} \\
A_{jt} &= \gamma^{z_{jt}} A_{jt-1} \\
\mathcal{S} &= \sum_j s_{jt} \\
\{\Lambda_{jt}, \{\ell_{j\iota t}\}\} = \text{argmax}\left[\mathcal{Y}_t - \sum_j r_j \Lambda_{jt} - \tau_t \mathcal{E}_t\right] \;\; s.t. \;\; L = \sum_j \int_0^1 \ell_{j\iota t} d\iota.
\end{aligned}$$

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 choose an innovation allocation knowing that the choice of technology will affect production choices in equilibrium. It will be helpful to define the Planner's intertemporal marginal rate of substitution $R_{t+1} \equiv (1+\rho)u'_t/u'_{t+1}$, which is the Planner's discount factor over consumption goods. The following proposition characterizes the innovation subsidies ξ_{jt} that support the Planner's allocation as an equilibrium.

Proposition 4. *Optimal innovation subsidies follow the recursion*

$$\xi_{jt}\Pi_{jt} = (1-\alpha)S_{jt}\mathcal{Y}_t - \sum_{\hat{t}\geq t}\prod_{\hat{s}=1}^{\hat{t}-t}\frac{1}{R_{t+\hat{s}}}(SCC_{\hat{t}} - \tau_{\hat{t}})\mathcal{E}_{\hat{t}}\frac{\partial \ln(\mathcal{E}_{\hat{t}})}{\partial \ln(A_{jt})} + \frac{1}{R_{t+1}}\left[\xi_{jt+1}\Pi_{jt+1} + \sum_i \xi_{it+1}\Pi_{it+1}g_{it+1}\varphi_{ijt+1}\right], \tag{57}$$

where the social cost of carbon follows

$$SCC_t = -\sum_{\hat{t}\geq t}\prod_{\hat{s}=1}^{\hat{t}-t}\frac{1}{R_{t+\hat{s}}}\mathcal{Y}_{\hat{t}}\frac{\partial \ln(\Omega_{\hat{t}})}{\partial \mathcal{E}_t}. \tag{58}$$

Furthermore, if $\tau_t = SCC_t$, these innovation subsidies achieve the first-best.

The proof of Proposition 4 can be found in Appendix A.8. This result tells us that second-best

innovation subsidies balance correcting spillover externalities with indirectly correcting the pollution externality by rewarding technologies that reduce carbon emissions in equilibrium. Innovation subsidies set according to the recursive formula of Equation (57) guarantee that the private reward for innovation $\xi_{jt}\Pi_{jt}$ equals the social value of innovation, which is composed of three 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 (Domar, 1961; Hulten, 1978). Second, when there is a wedge between the price of carbon and its true social cost, producers demand too much of the dirty inputs. To make up for this distortion, the Planner adjusts the direction of innovation in favor of clean technology. This adjustment multiplies the carbon price wedge, total emissions, and the elasticity of equilibrium emissions with respect to technology, summed over all future periods.

For clean technology, the contemporaneous emissions elasticity will tend to be negative, implying an increase in clean innovation subsidies. Conversely, improvements in dirty technology will tend to contemporaneously increase emissions in equilibrium, so dirty innovation subsidies will be reduced.[18] Equation (57) also includes changes in emissions in future periods because of dynamic effects of climate damages: polluting less (more) today implies an increase (decrease) in future emissions as damages will be lower (higher) in the future. Quantitatively, I find that these dynamic effects are unimportant as the contemporaneous shifts in demand from technological improvement are what drive the carbon price wedge term, as one would expect.

Third, the term in brackets represents the spillover benefits generated 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, represented by the $\sum_i \xi_{it+1}\Pi_{it+1}g_{it+1}\varphi_{ijt+1}$ term. Therefore, Proposition 4 provides a general formula for climate innovation policy as it relies on minimal assumptions about the structure of production and spillovers and does not require efficient carbon prices.

A visible implication of Equation (57) is that efficient carbon prices allow the Planner to separately correct the economy's two externalities and achieve the first-best.[19] This Pigou 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 (58), innovation rents will be properly adjusted to reward clean technology. Thus, first-best innovation subsidies do not need to provide additional support to clean innovation as such.

4.2 Steady-State Policy

To provide further intuition for the determinants of optimal innovation subsidies, I will now consider steady-state innovation policy. First, 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. In describing the steady-state, I will normalize by the *common innovation wedge* $1/(\gamma-1)(1-\tilde{R}^{-1})$: the

[18] 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.

[19] I describe the first-best planning problem in detail in Appendix A.10.

steady-state difference between the private and social benefit of innovation that is common across all technologies. This wedge stems from the fact that innovators only enjoy a fraction of the marginal surplus from innovation for a single period, whereas society enjoys the full marginal surplus in perpetuity. In a model with an endogenous supply of scientists (e.g. where workers could choose between being a scientist or a production worker), correcting the common innovation wedge optimizes the total supply of scientists.

Therefore, it will be natural to describe the steady-state in terms of innovation subsidies multiplied by income shares and normalized by the common innovation wedge, $\hat{\xi}_{jt} \equiv (\gamma - 1)(1 - \tilde{R}^{-1})\xi_{jt}S_{jt}$, as this isolates the portion of innovation incentives that controls the *composition* of research effort. Finally, denote the carbon price wedge term by $\mathcal{T}_{jt} \equiv \sum_{\hat{t} \geq t} \prod_{\hat{s}=1}^{\hat{t}-t} \frac{1}{\tilde{R}_{t+\hat{s}}} \frac{SCC_{\hat{t}} - \tau_{\hat{t}}}{1-\alpha} \bar{\omega}_{\hat{t}} \frac{\partial \ln(\mathcal{E}_{\hat{t}})}{\partial \ln(A_{jt})}$. The following corollary characterizes steady-state innovation policy.

Corollary 3. *In steady-state, innovation incentives satisfy*

$$(1 - \tilde{R}^{-1})\big(\hat{\xi}' - [S' - \mathcal{T}']\big) - g\tilde{R}^{-1}\hat{\xi}'\boldsymbol{\varphi} = \vec{0}'. \tag{59}$$

The proof of Corollary 3 can be found in Appendix A.9. This result mirrors those of Liu and Ma (2021), implying that my innovation subsidy formula recovers theirs as a special case whenever (i) pollution is priced efficiently $\mathcal{T} = \vec{0}$ and (ii) the economy is in steady-state. Equation (59) states that steady-state innovation incentives solve a convex combination of two root-solving problems. The weight between the two, and thus the focus of innovation policy, is determined by the Planner's level of patience, as measured by their growth adjusted discount factor $\tilde{R}^{-1}$.[20]

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 the flow value of innovation: $\hat{\xi} = S - \mathcal{T}$. This maximizes the contemporaneous benefit of innovation by focusing exclusively on the immediate impact on output. If pollution is priced efficiently $\mathcal{T} = \vec{0}$, this also implies that the composition of research effort can be driven by profit signals in the steady-state. With a proper price on pollution, the absence of cross-technology spillovers $\boldsymbol{\varphi} = \mathbf{0}$ would also imply that the laissez-faire composition of research effort is efficient in steady-state because innovation rents will be proportional to the social value of innovation.

A Planner that is perfectly patient, i.e. $\tilde{R}^{-1} \to 1$, will set innovation incentives with an exclusive focus on the spillover benefits of innovation.[21] They set innovation incentives according to

$$\hat{\xi}'\boldsymbol{\varphi} = \hat{\xi}'\big(\tilde{\boldsymbol{\varphi}} - \mathbf{I}\big) = \vec{0}', \tag{60}$$

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 $\hat{\xi}$ represent technologies' eigenvector centrality in 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. This selects the composition of research effort which maximizes the growth rate.

In summary, ideas are used to produce both physical goods and new ideas, and steady-state innovation policy should balance the two according to the Planner's degree of patience.

[20] The intuition of steady-state policy can also be seen in terms of a Leontief inverse, which I describe in Appendix A.9.

[21] I am assuming here that $\lim_{\tilde{R}^{-1} \to 1}(1 - \tilde{R}^{-1})\mathcal{T} = \vec{0}$.

5 Calibration

In this section, I describe my calibration of the structural parameters of my model. My application is transportation and electricity generation in the US. Throughout this section, I make the parametric assumptions 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 can 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 the US CPI with 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 the spillover network. I follow Liu and Ma (2021) in using the citation network of patents as a proxy for the spillover network. They provide evidence that a spillover network, constructed in this way, is predictive of downstream patenting activity.

For my sample of patents, I take the universe of granted US patents from PatentsView.[22] First, I need to assign patents to my five technology classes: clean/dirty transportation, clean/dirty electricity generation, and general technology. To do so, I make use of the International Patent Classification (IPC) and Cooperative Patent Classification (CPC) systems.[23] 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.[24] 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). 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.[25] Finally, I define general technology patents as those patents that are not in any of the above categories. Table C.1 gives a representative, but not exhaustive, list of my patent classification codes.

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

$$\tilde{\varphi}_{ij} = \frac{cites_{ij}}{\sum_k cites_{ik}}, \tag{61}$$

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

[22] 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.

[23] 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 to the same CPC codes with the same meaning.

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

[25] For the full list of IPC codes pertaining to dirty electricity generation, see Table A1 of Lanzi et al. (2011).

of knowledge spillovers (Jaffe et al., 1993; Hall et al., 2005). The (net) spillover network then follows mechanically from $\boldsymbol{\varphi} = \tilde{\boldsymbol{\varphi}} - \mathbf{I}$.

One potential concern with this approach is that the spillover network may change over time. This concern is especially relevant for young, clean technologies as forward-looking policy would need to know the spillovers these technologies will create in the future. To address this concern, I compute the spillover network for each of the five year bins between 1970 and 2015 and find there is little change over this period. In particular, I find no trend in the spillover creation or network centrality of clean technologies. Hence, the evidence gives us no reason to expect a substantial change in the spillover network going forward. Further details can be found in Appendix B.1.

Figure 1 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.[26] Figure 1 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 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.

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 damage function.

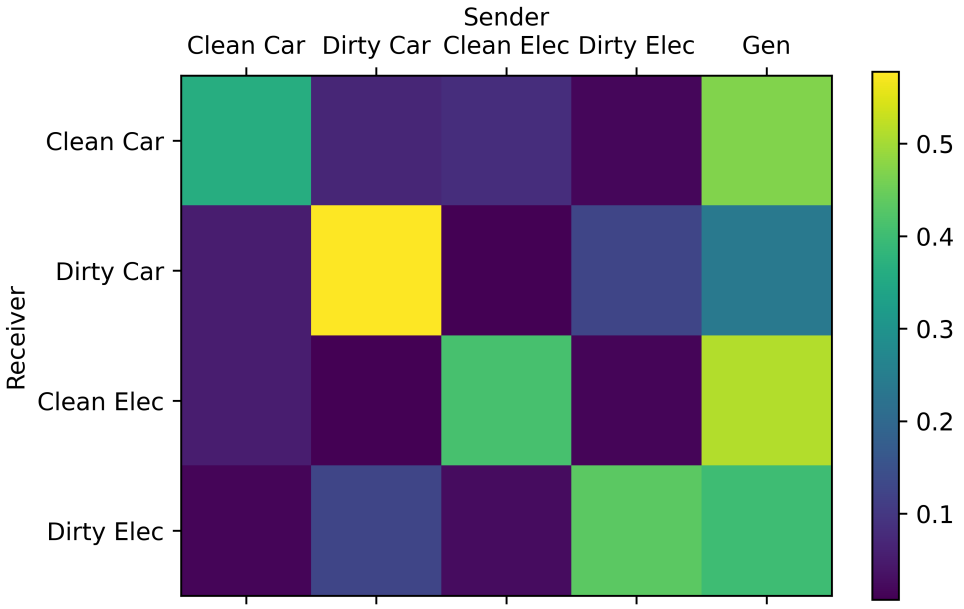
Atmospheric carbon concentrations evolve according to

$$\mathcal{C}_t = \sum_{\hat{t}=1800}^{t} \left(\psi_p + (1-\psi_p)\psi_0\psi^{t-\hat{t}}\right)\mathcal{E}_{\hat{t}} + \bar{\mathcal{C}}, \tag{62}$$

where $\bar{\mathcal{C}}$ is the pre-industrial level of atmospheric carbon concentration. I select the year 1800 as the starting point of industrialization. This specification of the carbon cycle provides a mapping from past 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 $\hat{t}$ 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. For the transitory component, fraction $(1-\psi_0)$ exits the

[26] Specifically, 98.6% of general technology patent citations reference another general technology patent.

Figure 1: 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.

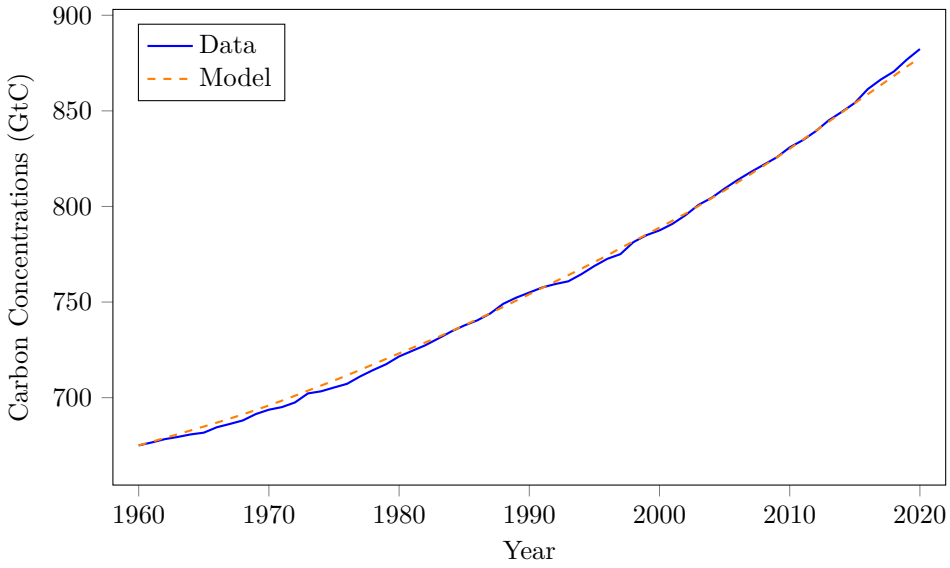
atmosphere within a period and is absorbed by the biosphere or surface oceans. The remainder decays geometrically according to ψ. As argued in Archer (2005) and Golosov et al. (2014), this relatively simple specification of the climate system provides a good approximation of the complex relationship between carbon emissions and atmospheric carbon concentrations.

The pre-industrial level of atmospheric carbon concentration $\bar{\mathcal{C}}$ was 596.4 gigatons of carbons.[27] 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. 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. I describe in Appendix B.2 how the state of the climate can be written as a two-dimensional recursion and how I use this recursion to select initial conditions for the climate. 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. The results of my calibration exercise can be seen in Figure 2.

[27] Atmospheric carbon concentrations are often stated in terms of parts per million (ppm) of CO_2, which is equivalent to 2.13 GtC. This conversion was used to map the widely-accepted pre-industrial 280 ppm of CO_2 to GtC.

Figure 2: 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

$$\Omega_t = \exp\big(-\varrho(\mathcal{C}_t - \bar{\mathcal{C}})\big), \tag{63}$$

where $\varrho > 0$ is a scale parameter stipulating the semi-elasticity of final output with respect to atmospheric carbon concentrations.[28] Assuming a constant semi-elasticity of damages with respect to atmospheric carbon concentrations does not allow for severe non-linearities in the climate system such as tipping points, but this functional form for damages is consistent with the approach taken in much of the climate literature (Nordhaus, 2017; Stern, 2007; Barrage and Nordhaus, 2024). 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 scale of climate damages, I will also consider an alternative damages parameter that is four times as large, in line with the "catastrophic damages" scenario considered by Golosov et al. (2014).[29]

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 Ap-

[28] 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. The standard value for Γ is 3, but there remains considerable uncertainty over the value of Γ (Wagner and Weitzman, 2016).

[29] I abstract from the consideration of risk in optimal climate policy. There are substantial uncertainties regarding the severity of climate change, so a large literature considers the role of risk 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 insuring against climate risk by considering the possibility of catastrophic damages.

pendix B.2, I provide details explaining how the equilibrium conditions of my model, and therefore the calibration moments discussed below, can be reduced to technology, policy instruments, and structural parameters. Table 1 provides a summary of my parameter choices.

Table 1: Parameter Choice Summary

Parameter	Description	Value	Source
σ	Clean/Dirty ES	1.86	Papageorgiou et al. (2017)
λ	Cross-Sector ES	0.1	Hassler et al. (2021)
ν_{car} ν_{elec}	CES Shares	0.028 0.023	Income Shares
γ	Innovation Step Size	1.07	Acemoglu et al. (2023)
α	Input Share	0.4	Standard 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	0.001 0.015	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. (2017). The evidence for this estimate comes from electricity generation, but I use it for transportation as well because it is a similar magnitude to other estimates in the literature that look at the energy sector more broadly (Popp, 2004; Lanzi and Sue Wing, 2011). 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; Fried, 2018).

For the sectoral CES shares, I normalize the sum of the shares to one and select $\{\nu_{car}, \nu_{elec}\}$ to equal the average income shares of transportation (2.8%) and electricity generation (2.3%) from 2000 to 2020. Data on revenue for the electricity generation sector comes from the US Energy Information Agency (EIA).[30] Data on motor vehicle output and GDP come from the US Bureau of Economic Analysis (BEA). I normalize the supply of scientists and workers to $\mathcal{S} = 1$ and $L = 100$. 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.[31] This is a similar value to those 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 $\alpha = 0.4$ to match estimates of the labor share. The standard economy-wide value for the labor share is 0.67, but Barrage (2020) estimates a labor share in electricity

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

[31] 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.

and resource production of 0.403. I set my labor share $1-\alpha$ as a weighted average of the two, with about one quarter weight on the lower estimate from 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.[32] 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 one (Hall, 1993; Hall and Van Reenen, 2000; Bloom et al., 2002; Wilson, 2009). As I discuss in Appendix B.2, 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 one, 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 (Dasgupta, 2008; Barrage, 2018).[33] This debate is not the focus of this paper, but given the long-term 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 clean quantity shares in each sector. I target clean quantity shares, as opposed to income shares, because of data availability, but as I show in Appendix B.2, there is a straightforward mapping between the two. 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).[34] Data on the share of electricity generated from non-fossil sources comes from the EIA.[35] 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,

[32]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.

[33]It is worth noting that moral philosophers universally reject the idea that utility should be discounted simply because it occurs at a later date (Cowen and Parfit, 1992; Ord, 2020). The only justification for discounting future utility that is widely accepted amongst philosophers is existential risk, so it is acceptable to discount future utility in accordance with the probability that it may not happen. Indeed, this is the justification Stern (2007) gives for his discount rate.

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

[35]Specifically, Form EIA-923.

I include clean input subsidies discussed below.

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).[36]

5.4 Model Validation

In this section, I test my model's ability to replicate the progress made by clean technology in both transportation and electricity generation throughout the 2010s. As shown in Figure 3, 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 spite of the fact that no model parameters was selected to target this change. However, when I shut down the spillover network, my model makes the counterfactual prediction that dirty technologies would have become *more* entrenched in both sectors, not less.

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.3. First, I allow for clean input subsidies $\bar{\xi}_{\theta c}$ that reduce effective input prices to $(1-\bar{\xi}_{\theta c})r_{\theta c}$. I set $\bar{\xi}_{car,c} = 0.012$ and $\bar{\xi}_{elec,c} = 0.03$ to match federal spending on the plug-in electric vehicle and energy investment tax credits as a proportion of total spending in the two sectors.[37]

In selecting innovation subsidies, I focus on relative innovation subsidies because the fixed supply of scientists implies that only relative innovation subsidy matter. To this end, I take data on public R&D spending by technology from the International Energy Agency (IEA) and 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 with and without the spillover network.

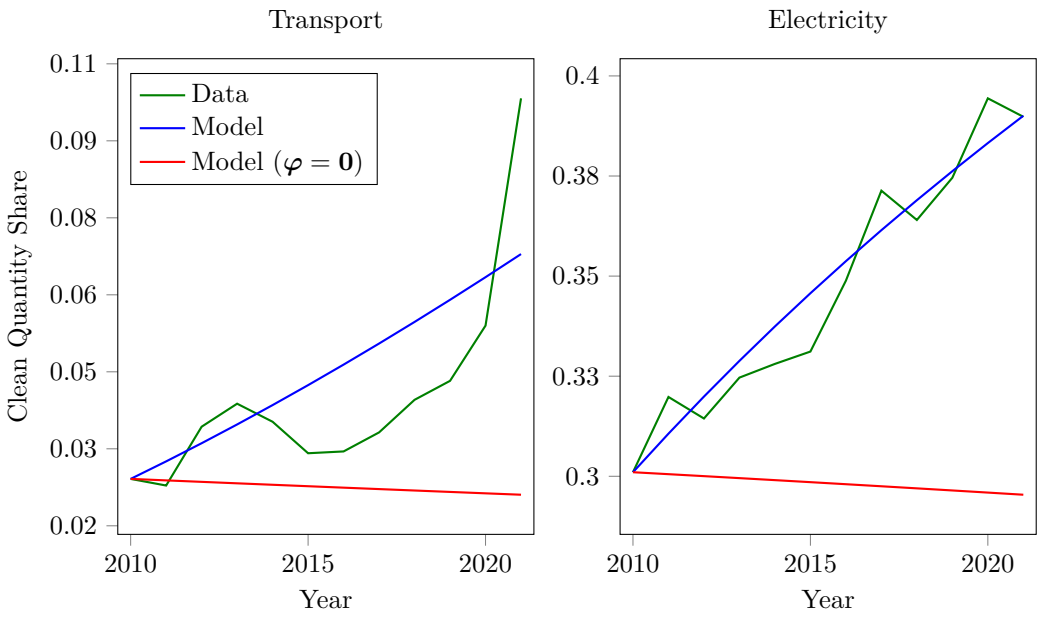
Using the procedure described in Section 5.3, I take 2010 as the start year to set my initial condition for technology. I then simulate the equilibrium path of technology with and without the spillover network and plot the implied dynamics of clean quantity shares in Figure 3.

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 6.8% and 39%. When I shut down the spillover network, the model predicts that clean quantity shares for transportation and electricity generation would have instead

[36] 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.

[37] Data on federal spending on these tax credits during this period comes from US Congressional Research Service Reports IF11017 and IF10479.

Figure 3: 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.

fallen to 2.1% and 29.5% over this period. Therefore, a model without cross-technology spillovers cannot match the qualitative direction of the data.

Using Propositions 2 and 3, we can better understand the role of cross-technology 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, meaning that increasing returns to innovation were not large enough to lock in dirty technology. The model predicts that starting in 2010, transportation and electricity generation will converge halfway to their steady-states in 74 and 101 years, respectively. However, when I shut down the spillover network, the spectral radius rises to 1.01, locking in dirty technology. In that case, US climate policy in the 2010s was too weak to switch these sectors to clean growth, so in the absence of cross-technology spillovers, they would have moved deeper into their dirty basins of attraction. Thus, the model's ability to explain the nascent clean energy transition of the 2010s hinges on the inclusion of cross-technology spillovers.

6 Quantitative Exercises

In this section, I conduct several 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. I will consider sector-specific relative clean technology $\bar{B}_{\theta t} \equiv A_{\theta ct}/A_{\theta dt}$, which is log-linear in relative technology.

Second, I simulate optimal climate innovation policy for various restrictions on the carbon price. I find that optimal clean innovation subsidies are very different from the "big push" of clean innovation

subsidies recommended by models without cross-technology spillovers, and this holds even in the absence of optimal carbon prices.

6.1 Impact of Policy Reform

In this section, I examine the impact of introducing a realistic 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.[38] 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 than 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.

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 2 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. The third panel of Table 2 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.[39]

The first panel of Table 2 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 117.3% for transportation and 120.8% for electricity generation. Together, the carbon price and induced increase in clean technology reduce the economy's steady-state emissions intensity by three quarters.

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, locking in dirty technology. 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 2 shows the policy reform is strong enough to do so for electricity generation, but not for transportation. This difference stems from the fact that clean technology starts out with a greater disadvantage in transportation. 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 emissions intensity.

The second panel of Table 2 considers the transitional impacts of the policy reform. Using Proposition 2, I can characterize the speed of transition to the new steady-state.[40] For my calibrated spillover

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

[39] A common robustness check when using patent citations is to consider only citations added by the patent applicant, rather than the examiner. When I build my spillover network using only applicant citations, the spectral radius becomes 0.989, leaving my conclusions largely unchanged.

[40] Given that these results involve linearizing my model, one may be concerned that my results are not informative

Table 2: Impact of Policy Reform

	No Spillovers	Calibrated Spillovers	Double Spillovers
Long-Run Impacts			
Relative Clean Technology by Sector			
$\%\Delta\bar{B}_{car}$	0%	+117.3%	+33.9%
$\%\Delta\bar{B}_{elec}$	$+\infty$%	+120.8%	+38.7%
Emissions Intensity			
$\%\Delta\bar{\omega}$	-73.1%	-74.4%	-57.2%
Transitional Impacts			
Half-Lives of Convergence by Sector			
$t_{car}^{(1/2)}$	–	119 years	20 years
$t_{elec}^{(1/2)}$	–	131 years	25 years
Carbon Emissions by Year			
$\%\Delta\mathcal{E}_{2035}$	-43.7%	−51%	-53.6%
$\%\Delta\mathcal{E}_{2060}$	-46.8%	-57.1%	-55.1%
Degree of Increasing Returns to Innovation			
Spectral Radius			
$\max\|\kappa_j\|$	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.

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 119 and 131 years, respectively. The policy reform's impact on 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.[41]

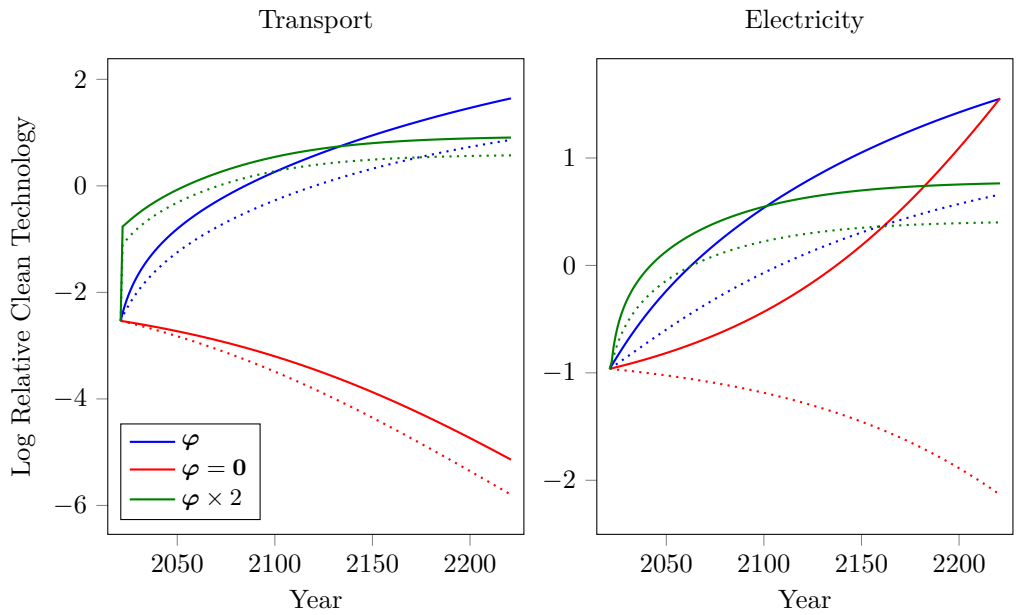
As before, the other two cases illustrate how cross-technology spillovers shape the speed with which policy can influence innovation. When the spillover network is shut down, the economy doesn't converge to an interior steady-state, making the half-lives undefined. Moreover, the long-run emission reductions 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 quick half-lives of 20 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 4 provides 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

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.

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

Figure 4: 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.

reform allows clean technologies to slowly gain a lead over their dirty counterparts as the economy transitions to a new, cleaner steady-state. 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 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.

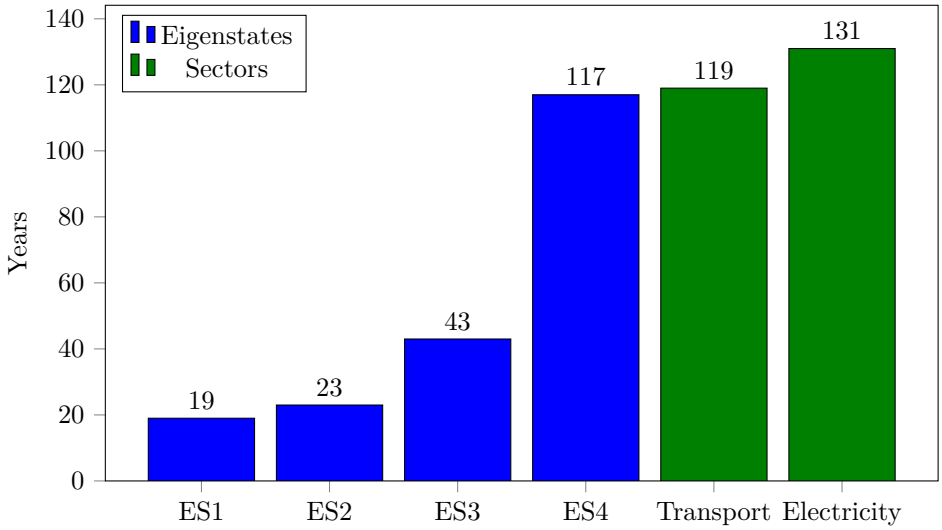
These differences showcase how both the long-run and transitional impacts of a policy reform are shaped by cross-technology spillovers. High cross-technology spillovers allow policy to shape the direction of innovation quickly by allowing less advanced technologies to achieve catchup growth. However, 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.

My theoretical results emphasized the role of both spillovers and substitutability in shaping the impact of policy reforms in both the long-run and transition. To explore this relationship, I first plot in Figure C.4 how the steady-state impact of the policy reform varies in the scale of cross-technology spillovers and the elasticity of substitution between clean and dirty goods σ. As expected, the steady-

state impact reduces with cross-technology spillovers and increases with the elasticity of substitution σ. Next, I plot in Figure C.5 how the spectral radius of the transition matrix varies in the same two variables. Again, the theory is borne out with the degree of increasing returns to innovation decreasing with cross-technology spillovers and increasing with the elasticity of substitution σ. In particular, I find that technology becomes path dependent if cross-technology spillovers are reduced to 66.1% of their calibrated value or if the elasticity of substitution σ rises to 2.29.

The EPA has recently proposed increasing the federal SCC to \$190, so I examine the impact of introducing such a carbon price in Table C.2 and Figure C.6. The results are quantitatively larger but qualitatively similar. 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, so I show in Figures C.7 and C.8 the carbon prices and clean innovation subsidies necessary to shift each technology into its clean basin of attraction in the absence of cross-technology spillovers.

Figure 5: 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 ($\xi_c = 1.43$).

Figure 5 shows the half-lives of convergence for each eigenstate and sector of the economy.[42] 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 convergence speed, and Figure 5 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 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

[42] I perform the same exercise for the case where cross-technology spillovers are doubled in Figure C.9.

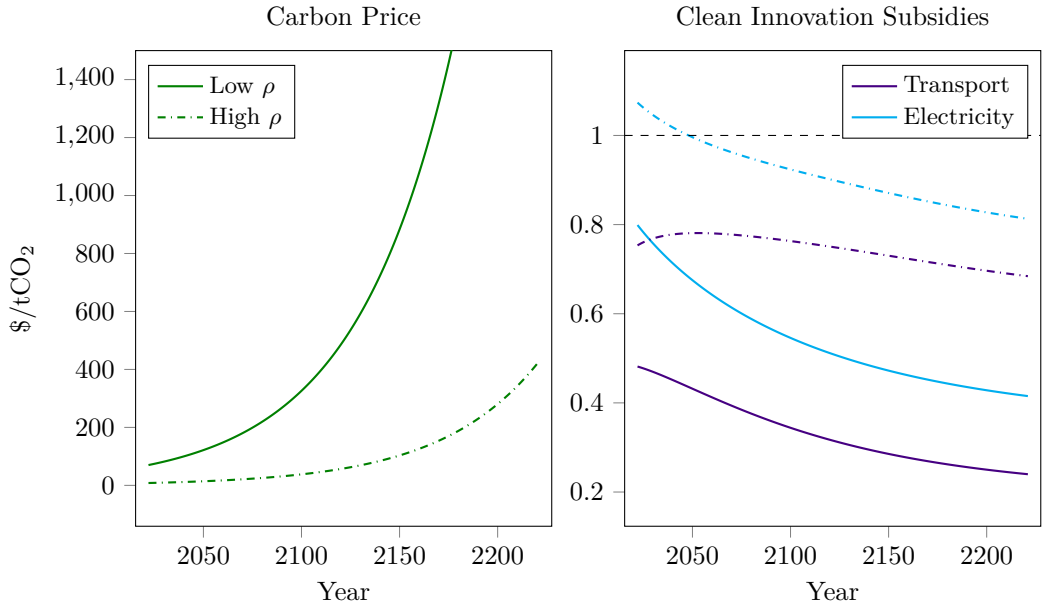
is substantially faster, represents initial conditions where dirty technology is advanced in one sector and clean technology is advanced in another. Moving from dirty to clean technology in both sectors concurrently is especially slow because, as argued in Section 5.1, clean and dirty technologies both 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.[43]

Figure 5 shows that transportation converges a bit more quickly than electricity generation, with a difference in half-lives of about 10.1%. This is due to the higher within-sector spillovers in transportation. To quantify the importance of such spillovers, Figure C.11 shows half-lives when within-sector spillovers are shut down. In that case, the spectral radius rises to 0.998, and the half-lives for transportation and electricity generation rise to 443 and 401 years. The loss of such spillovers substantially slows convergence speed, but more importantly, the sectors switch their speed rankings, signifying that transportation's more rapid transition stems from within-sector spillovers.

6.2 Optimal Policy Path

In this section, I simulate the optimal policy paths starting in 2022.[44] I start by considering the first-best policy path for two values of the discount rate, as shown in Figure 6. 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.

Figure 6: First-Best Policy Path



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

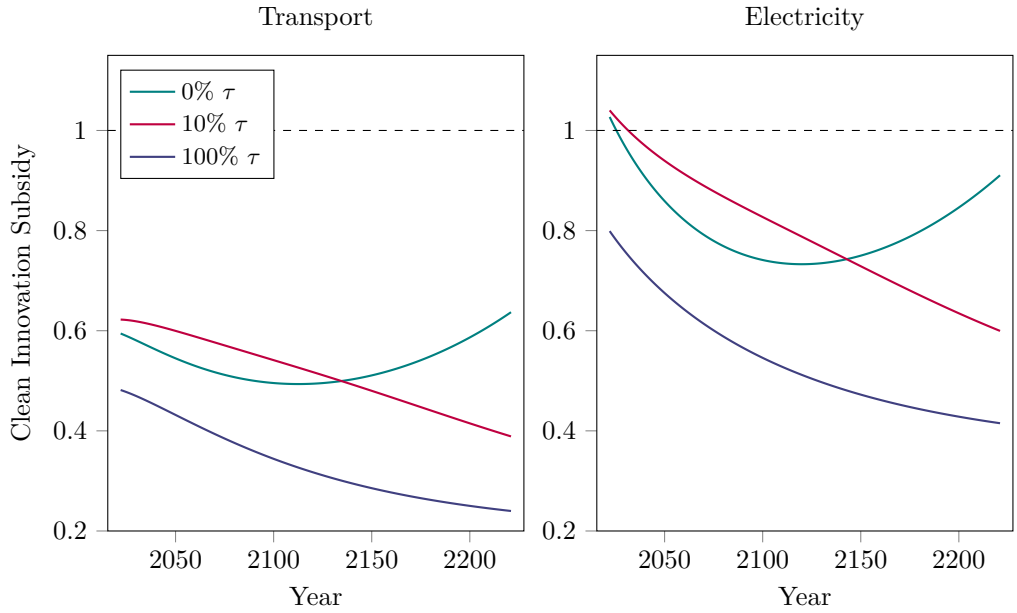
[43] One may wonder whether the degree of increasing returns to innovation depends on the size of the policy reform. Figure C.10 shows that the eigenvalues are not sensitive to the size of the policy reform. Instead, variation in transition speed comes from differences in loading on the eigenstates β.

[44] I provide details on my solution method in Appendix B.4.

The right panel displays clean innovation subsidies as a fraction of the common innovation wedge. Strikingly, clean technologies do not receive favorable innovation subsidies for either discount rate. As shown in Proposition 4, the primary function of first-best innovation subsidies is to reward spillover creation, and on this metric, clean technologies don't score particularly well. Clean transport's eigenvector centrality is barely more than a tenth of its steady-state income share (0.64 vs. 5.1%), while clean electricity's eigenvector centrality is slightly less than a quarter of its steady-state income share (0.91 vs. 3.3%).[45] In other words, clean technologies are useful because they can produce physical goods without pollution, but they are not particularly useful in the production of ideas.

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.

Figure 7: Optimal Innovation Subsidy Path



Notes: Policy paths use the low discount rate. The external carbon price τ is a proportion of the social cost of carbon, so 100% is the first-best. Innovation subsidies are listed as a fraction of the common innovation wedge.

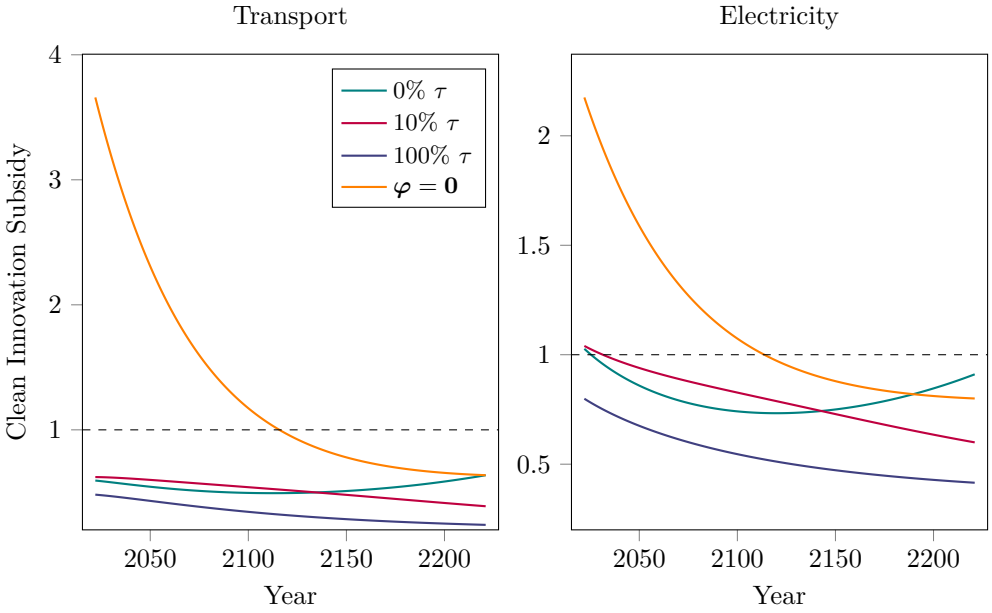
Figure 7 shows how optimal innovation subsidies change when there are restrictions on the carbon price. As shown in Proposition 4, second-best innovation subsidies adjust to accommodate the distortion of incomplete carbon pricing. I consider two cases in addition to the first-best: one where the external carbon price is zero and another where the external carbon price is ten percent of the social cost of carbon. The ten percent case is qualitatively similar to the first-best, with clean innovation subsidies shifted upwards by about 20-30% of the common innovation wedge. When the carbon price is zero,

[45] 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.

clean innovation subsidies start with a similar upward shift as in the small carbon price case and then rise after the first century of policy. These innovation subsidies use my preferred specification of low discounting, but I show in Figure C.12 that the main features of second-best innovation subsidies carry over to the case of high discounting.

To unpack the importance of cross-technology spillovers for optimal policy, Figure 8 plots first-best clean innovation subsidies in the absence of cross-technology spillovers, alongside the previous three cases. When I shut down the spillover network, my model recovers the common prescription in the climate innovation literature of a "big push" of large, temporary clean innovation subsidies (Acemoglu et al., 2012, 2016). Now innovation subsidies for clean transport and electricity start at more than triple and double the common innovation wedge and slowly decrease down to baseline over the first century of policy. Thus, relative to this benchmark case, the inclusion of cross-technology spillovers implies the need for lower, flatter clean innovation subsidies, even when carbon is priced below its social marginal cost. Even in the most distorted case where carbon pricing is altogether absent, the path of clean innovation subsidies looks very different from a big push, with subsidies rising over time, rather than tapering off.

Figure 8: Optimal Innovation Subsidy Path (+No Spillovers)



Notes: Policy paths use the low discount rate. The external carbon price τ is a proportion of the social cost of carbon, so 100% is the first-best. Innovation subsidies are listed as a fraction of the common innovation wedge.

Cross-technology spillovers eliminate the need for a big push because building towards better clean technology in the future no longer requires an exclusive focus on clean innovation today. The need to decarbonize implies that clean technologies will dominate production in the future, so the forward-looking Planner picks today's profile of innovation to build towards this clean future. Without cross-technology spillovers, clean innovation is the only way to build towards clean technology in the future. With such spillovers, clean technologies can build on all of the technologies in the economy, so innovation in every technology can build towards the clean future. Thus, there is no need for large, upfront subsidies

that jolt the direction of innovation towards clean technology.

The above analysis assumes that the spillover network is unchanging over time. While this is consistent with the historical evidence (see Appendix B.1), it is certainly possible that clean technologies will become endogenously more central in the future as they mature. I model this possibility explicitly in Appendix B.5 by allowing the spillover functions to have an elasticity of substitution above one. In that case, clean innovation subsidies start near the common innovation wedge and rise substantially over time. Therefore, the core conclusion of this section, that a big push is not warranted with an empirically disciplined spillover network, carries over to the case where clean technologies become endogenously more important for spillover creation as the clean transition proceeds.

First-best clean innovation policy no longer implies the need for a big push, but why don't inefficient carbon prices do more to increase optimal clean innovation subsidies? There are two reasons. First, clean innovation alone is ineffective at reducing emissions because it generates a rebound effect. This can be seen in the following accounting identity:

$$\frac{\partial \ln(\mathcal{E}_t)}{\partial \ln(A_{jt})} = \frac{\partial \ln(\omega_t)}{\partial \ln(A_{jt})} + \frac{\partial \ln(\mathcal{Y}_t)}{\partial \ln(A_{jt})}. \tag{64}$$

The first term – the change in emissions intensity – will tend to be negative for clean technology as clean innovation generates substitution away from dirty goods. But the second term – the change in output – will always be positive as clean innovation expands the economy's productive capacity. Thus, trying to reduce emissions through clean innovation alone faces an inherent headwind from the rebound effect of greater output. We can see this playing out in the simulation: in the zero carbon price case, the elasticity of contemporaneous emissions with respect to clean transport and electricity generation is only -0.25 and -0.29 on average over the first century of policy.

The second reason is that dirty innovation is less valuable when carbon pollution is under-priced, so the value of spillovers from clean to dirty is reduced in turn. Indeed, in the case of a zero carbon price, dirty innovation is shut down entirely, completely eliminating the value of these spillovers.

These two consequences illustrate the main tradeoff of second-best policy. With the Planner unable to separately correct the two externalities of the economy, emissions are higher and growth is lower. Carbon emissions and gross output growth in the first-best are 0.36 GtC and 2% in the year 2200, whereas they are 3.92 GtC and 1.99% for the ten percent case.[46] When there is no carbon price, these numbers are 9.27 GtC and 1.97%. Therefore, the cost of restricting the carbon price is highly nonlinear. This is because a small, growing carbon price can push dirty technology out of production, allowing some dirty innovation to continue for the sake of spillovers. 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. Put differently, a small carbon price goes a long way in preventing technological advance from translating into increases in emissions.

We can further examine the climate and welfare implications of restrictions on the carbon price. The climate implications can be seen in Figure C.13, where I plot the path of temperature increases associated with each policy path. 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

[46] By gross output I mean output before climate damages. This gives us a measure of technological change that is independent of the different emission paths of these scenarios.

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 in the first-best and ends up roughly a degree higher in the ten percent case. 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.

For welfare, we can again see the highly nonlinear cost of restricting the carbon price. The ten percent case leads to a consumption equivalent loss of 1.44%, 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%. That is, the lion's share of welfare losses come when the carbon price is lost altogether. Welfare losses are lower in the case of high discounting because higher emissions and lower growth are problems that materialize in the future. The less patient Planner has a consumption equivalent loss of 0.3% for the ten percent case and 0.88% 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, but in both cases the cost is nonlinear in the degree of restriction.

Finally, the damages we can expect from a warmer climate are highly uncertain, so I consider the possibility of substantially higher damages in Figure C.14. For that case, I focus on the first-best and 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 actually reduced. 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.

7 Conclusion

One of the main goals of climate policy is to redirect innovation from dirty to clean technology. This paper has argued that cross-technology knowledge spillovers are critical for understanding the role of policy in this transition. Such spillovers matter for both describing the impact of policy reforms and prescribing optimal policy. They enable a more rapid transition by allowing less advanced technologies to achieve catchup growth, but this same mechanism reduces the long-run impact of policy reforms. Formally, I have shown that two forces – cross-technology spillovers and substitution patterns in production – govern both the size and speed of technological transition and that each of these forces can be summarized in terms of a sufficient statistic matrix. Together, these forces govern the degree of increasing returns to innovation, which can be quantified using a spectral radius.

For optimal policy, I have shown that innovation subsidies should focus on spillover creation whenever carbon prices correct the pollution externality. That is, clean technologies should not be subsidized simply because they can create goods without pollution; they should be subsidized insofar as they produce spillovers. However, innovation subsidies should be adjusted to reflect the equilibrium impact of innovation on emissions whenever carbon prices are inefficient.

To take my model to the data, I performed several quantitative exercises, focusing on transportation and electricity generation in the US. First, I showed that cross-technology knowledge spillovers are strong enough to prevent the lock-in of dirty technology, but that the substitutability of clean and dirty goods still leads to increasing to returns to innovation near one. Thus, a realistic carbon price and clean

innovation subsidy bring about a substantial redirection of innovation in the long run, but with a slow transition. Second, I showed that a big push of clean innovation subsidies is not optimal in the presence of a spillover network because such spillovers allow future clean technologies to be built atop a broader range of technologies. This prevents the need to focus exclusively on clean innovation early on and holds even in the absence of efficient carbon prices. Finally, I showed that the welfare cost of restricting the carbon price is highly nonlinear.

Economists have long understood, going back to Hicks (1932), that substitution patterns in production are important whenever the direction of innovation is endogenous. This paper has argued that cross-technology knowledge spillovers are deserving of a similar status. I have made this argument in the context of climate change, but I believe it has more general applicability. Just as the importance of substitutability has been applied to a variety of contexts – including skill-biased technological change, automation, and more – the importance of cross-technology spillovers could be applicable in a wide range of settings as well.

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 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., Barrage, L., Donald, E., Hémous, D., and Liu, E. (2024). Transition to Green Technology along the Supply Chain. *Working Paper*.

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

Akcigit, U., Hanley, D., and Serrano-Velarde, N. (2021). Back to Basics: Basic Research Spillovers, Innovation Policy, and Growth. *The Review of Economic Studies*, 88(1):1–43.

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

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

Arrow, K. (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. (2024). Policies, Projections, and the Social Cost of Carbon: Results from the DICE-2023 Model. *Proceedings of the National Academy of Sciences*, 121(13).

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., Romer, P., Terry, S., 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. (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. and Williams, H. (2021). Innovation: Market Failures and Public Policies. In *Handbook of Industrial Organization*, volume 5, pages 281–388. Elsevier.

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

Cowen, T. and Parfit, D. (1992). Against the Social Discount Rate. In *Philosophy, Politics, and Society: Volume 6, Justice Between Age Groups and Generations*, pages 144–161. Yale University Press.

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

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

Domar, E. (1961). On the Measurement of Technological Change. *The Economic Journal*, 71(284):709–729.

Donald, E., Hanley, D., and Restrepo, P. (2025). Knowledge Spillovers and the Hunt for New Ideas. *Working Paper*.

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.

Gale, D. and Nikaido, H. (1965). The Jacobian Matrix and Global Univalence of Mappings. *Mathematische Annalen*, 159(2):81–93.

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

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

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

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

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. and Ziedonis, R. (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.

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

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

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

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

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

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

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.

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

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

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

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

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

Papageorgiou, C., Saam, M., and Schulte, P. (2017). Substitution Between Clean and Dirty Energy Inputs: A Macroeconomic Perspective. *The Review of Economics and Statistics*, 99(2):281–290.

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

Stern, N. (2007). *The Economics of Climate Change: The Stern Review*. Cambridge 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. (2016). *Climate Shock: The Economic Consequences of a Hotter Planet*. Princeton University Press.

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

Wilson, D. (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

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

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 establish the existence and uniqueness of the economy's equilibrium path, it suffices to show that there exists a unique equilibrium allocation of scientists $\{s_{jt}\}$ in each period. Consider that the research condition (25) and fixed supply of scientists (12) defines, at each point in time, the mapping $\mathbb{T}s$ which follows

$$\mathbb{T}s_j = \frac{\big(\chi_j\xi_{jt}\phi_j(\{A_{it-1}\})\Pi_{jt}(\{s_i, A_{it-1}\},\tau_t)\big)^{\frac{1}{1-\eta}}}{\sum_i(\chi_i\xi_{it}\phi_{it}\Pi_{it})^{\frac{1}{1-\eta}}}\mathcal{S}. \tag{A.6}$$

Note that innovation profits $\{\Pi_{jt}\}$ are a function of realized technology $\{A_{jt}\}$, which can be written as a function of scientists $\{s_{jt}\}$ and inherited technology $\{A_{jt-1}\}$ via the law of motion for technology (8). Now, we need to establish that there exists a unique fixed-point for $\mathbb{T}s$. The fixed supply of scientists (12) establishes a compact convex domain set $\mathbb{X}_s$ for $\mathbb{T}s$, and the definition of $\mathbb{T}s$ implies that it maps into $\mathbb{X}_s$ as well. $\mathbb{T}s$ is continuous in $\{s_{jt}\}$ by assumption, so Brouwer's Fixed Point Theorem establishes existence.

For uniqueness, define the function $\mathbb{U}s \equiv s - \mathbb{T}s$, so a fixed-point of $\mathbb{T}s$ is a root of $\mathbb{U}s$. Theorems 4

and 6 of Gale and Nikaido (1965) establish that a root of $\mathbb{U}s$ is unique if the Jacobian of $\mathbb{U}s$ is either a P-matrix for every element of a closed rectangular region of $\mathbb{R}^J$ that contains $\mathbb{X}_s$ or positive (negative) quasi-definite for every element of $\mathbb{X}_s$. Hence, this condition implies that the equilibrium allocation of scientists is unique as well.

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

$$\max_{\{E_{\theta t}\}} \mathcal{Y}_t - \sum_{\theta} p_{\theta t} E_{\theta t}, \tag{A.7}$$

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

$$\Omega_t^{\frac{\lambda-1}{\lambda}} \nu_\theta^{\frac{1}{\lambda}} \Big(\frac{\mathcal{Y}_t}{E_{\theta t}}\Big)^{\frac{1}{\lambda}} = p_{\theta t}. \tag{A.8}$$

Next, producers at the sector level solve

$$\max_{Y_{\theta dt}, Y_{\theta ct}} p_{\theta t} E_{\theta t} - p_{\theta ct} Y_{\theta ct} - p_{\theta dt} Y_{\theta dt} \tag{A.9}$$

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

$$p_{\theta t} \Big(\frac{E_{\theta t}}{Y_{\theta et}}\Big)^{\frac{1}{\sigma}} = p_{\theta et}, \tag{A.10}$$

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

A.2 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

$$G_i\Big(\big\{\ln\big(Y_j/Y_J\big), \ln\big(p_J/p_j\big)\big\}\Big) = \ln\big(p_J/p_i\big) - \Big[\frac{1}{\lambda}\ln\big(\nu_{\theta(J)}/\nu_{\theta(i)}\big) + \big(\frac{1}{\sigma} - \frac{1}{\lambda}\big)\ln\big(E_{\theta(J)}/E_{\theta(i)}\big) - \frac{1}{\sigma}\ln\big(Y_J/Y_i\big)\Big]. \tag{A.11}$$

That is, G is a function valued at the zero vector when the equilibrium price conditions (A.8) and (A.10) 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

$$\frac{\partial G_i}{\partial \ln\big(Y_j/Y_J\big)} = -\Big[\frac{1}{\sigma}\mathbb{1}(i=j) + \big(\frac{1}{\lambda} - \frac{1}{\sigma}\big)\varepsilon^E_{\theta(i)j}\mathbb{1}(\theta(i) = \theta(j))\Big] \tag{A.12}$$

$$\frac{\partial G_i}{\partial \ln\big(p_J/p_j\big)} = \mathbb{1}(i=j), \tag{A.13}$$

where $\varepsilon^E_{\theta(i)j}$ is the elasticity of sector $\theta(i)$ with respect to good j. Equation (A.12) 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

$$\mathbf{\Sigma} = -\mathbf{D_y}\mathbf{G}^{-1}, \tag{A.14}$$

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

$$\mathbf{\Sigma} = \begin{pmatrix} \tilde{\mathbf{\Sigma}}_1 & 0 & \dots & 0 \\ 0 & \tilde{\mathbf{\Sigma}}_2 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \tilde{\mathbf{\Sigma}}_{\Theta-1} \end{pmatrix}, \tag{A.15}$$

where

$$\tilde{\mathbf{\Sigma}}_\theta = \begin{pmatrix} \lambda + (\sigma - \lambda)\varepsilon^E_{\theta d} & (\lambda - \sigma)\varepsilon^E_{\theta d} \\ (\lambda - \sigma)\varepsilon^E_{\theta c} & \lambda + (\sigma - \lambda)\varepsilon^E_{\theta c} \end{pmatrix}, \tag{A.16}$$

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.3 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

$$\ln\left(s_i/s_J\right) = -\frac{1}{\eta}\ln\left(\chi_i/\chi_J\right) - \frac{1}{\eta}\ln\left(\phi_i/\phi_J\right). \tag{A.17}$$

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

$$(1-\alpha)\ln\left(\bar{A}_i\right) = \ln\left(\Xi_i\right) + \frac{1}{\eta}\ln\left(\chi_i/\chi_J\right) + \frac{1}{\eta}\ln\left(\phi_i/\phi_J\right) + \alpha\ln\left(\mathcal{R}_i\right) + \ln\left(Y_{it}/Y_{Jt}\right), \tag{A.18}$$

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.18).

To consider existence and uniqueness of the steady-state, use Equation (A.18) to define the mapping $\mathbb{T}\ln\left(\bar{A}\right)$ which follows

$$\mathbb{T}\ln\left(\bar{A}_i\right) = \frac{1}{1-\alpha}\Big[\ln\left(\Xi_i\right) + \frac{1}{\eta}\ln\left(\chi_i/\chi_J\right) + \frac{1}{\eta}\ln\left(\phi_i/\phi_J\right) + \alpha\ln\left(\mathcal{R}_i\right) + \ln\left(Y_{it}/Y_{Jt}\right)\Big], \tag{A.19}$$

where the functional dependence on log relative technology comes from relative spillovers and relative quantities. If $\mathbb{T}\ln\left(\bar{A}\right)$ is a contraction, then we have that there exists a unique steady-state. I also constructively establish the existence of a unique steady-state in Appendix A.4 when the production structure is CES and the spillover structure is Cobb-Douglas.

More generally, if there exists a steady-state, Theorems 4w and 6 of Gale and Nikaido (1965) establish

that a fixed-point of $\mathbb{T}\ln(\bar{A})$ is unique if the Jacobian of the function $\mathbb{U}\ln(\bar{A}) \equiv \ln(\bar{A}) - \mathbb{T}\ln(\bar{A})$ is either a P-matrix for every element of $\mathbb{R}^{J-1}$ or positive (negative) quasi-definite for every element of $\mathbb{R}^{J-1}$.

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

$$(1-\alpha)\frac{\partial \ln(\bar{A}_i)}{\partial \ln(\Xi_j)} = \mathbb{1}(i=j) - \frac{1}{\eta}\sum_q \Phi_{iq}\frac{\partial \ln(\bar{A}_q)}{\partial \ln(\Xi_j)} + (1-\alpha)\sum_q \Sigma_{iq}\frac{\partial \ln(\bar{A}_q)}{\partial \ln(\Xi_j)}, \tag{A.20}$$

which in matrix notation gives us

$$\begin{aligned}(1-\alpha)\mathbf{D_{\Xi}}\bar{\mathbf{A}}_{\mathbf{ss}} &= \mathbf{I} - \frac{1}{\eta}\mathbf{\Phi}\mathbf{D_{\Xi}}\bar{\mathbf{A}}_{\mathbf{ss}} + (1-\alpha)\mathbf{\Sigma}\mathbf{D_{\Xi}}\bar{\mathbf{A}}_{\mathbf{ss}} \\ \Rightarrow \mathbf{D_{\Xi}}\bar{\mathbf{A}}_{\mathbf{ss}} &= \eta\big[\mathbf{\Phi} - \eta(1-\alpha)\big(\mathbf{\Sigma}-\mathbf{I}\big)\big]^{-1}.\end{aligned} \tag{A.21}$$

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

$$(1-\alpha)\frac{\partial \ln(\bar{A}_i)}{\partial \ln(\mathcal{R}_j)} = \alpha\mathbb{1}(i=j) - \frac{1}{\eta}\sum_q \Phi_{iq}\frac{\partial \ln(\bar{A}_q)}{\partial \ln(\mathcal{R}_j)} + (1-\alpha)\sum_q \Sigma_{iq}\frac{\partial \ln(\bar{A}_q)}{\partial \ln(\mathcal{R}_j)} - \alpha\Sigma_{ij}, \tag{A.22}$$

which in matrix notation gives us

$$\begin{aligned}(1-\alpha)\mathbf{D_R}\bar{\mathbf{A}}_{\mathbf{ss}} &= \alpha\mathbf{I} - \frac{1}{\eta}\mathbf{\Phi}\mathbf{D_R}\bar{\mathbf{A}}_{\mathbf{ss}} + (1-\alpha)\mathbf{\Sigma}\mathbf{D_R}\bar{\mathbf{A}}_{\mathbf{ss}} - \alpha\mathbf{\Sigma} \\ \Rightarrow \mathbf{D_R}\bar{\mathbf{A}}_{\mathbf{ss}} &= -\eta\alpha\big[\mathbf{\Phi} - \eta(1-\alpha)\big(\mathbf{\Sigma}-\mathbf{I}\big)\big]^{-1}\big(\mathbf{\Sigma}-\mathbf{I}\big).\end{aligned} \tag{A.23}$$

Combining Equations (A.21) and (A.23), we have

$$d\ln(\bar{A}_{ss}) = \eta\big[\mathbf{\Phi} - \eta(1-\alpha)\big(\mathbf{\Sigma}-\mathbf{I}\big)\big]^{-1}\big[d\ln(\Xi) - \alpha\big(\mathbf{\Sigma}-\mathbf{I}\big)d\ln(\mathcal{R})\big], \tag{A.24}$$

which gives us Equations (37) and (38) from Proposition 1.

□

A.4 Steady-State under Parametric Assumptions of Section 2.5

Under the parametric assumptions of Section 2.5, we can further characterize the steady-state for relative technology. Plugging the equilibrium price conditions (A.8) and (A.10) into Equation (A.18), we have

$$\begin{aligned}\sum_{j<J}\Phi_{ij}\ln(\bar{A}_j) = \eta\ln(\Xi_i) + \eta(1-\alpha)(\sigma-1)\ln(\bar{A}_i) \\ + \eta(\sigma-\lambda)\ln\big(p_{\theta(i)t}/p_{\theta(J)t}\big) - \eta\alpha(\sigma-1)\ln(\mathcal{R}_i).\end{aligned} \tag{A.25}$$

Denote relative sector prices by $\mathcal{P}_i \equiv p_{\theta(i)t}/p_{\theta(J)t}$. Writing Equation (A.25) in terms of matrix notation, we have

$$\ln\left(\bar{A}_{ss}\right) = \eta\left[\mathbf{\Phi} - \eta(1-\alpha)(\sigma-1)\mathbf{I}\right]^{-1}\left[\ln\left(\Xi\right) + (\sigma-\lambda)\ln\left(\mathcal{P}\right) - \alpha(\sigma-1)\ln\left(\mathcal{R}\right)\right]. \tag{A.26}$$

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

$$\mathcal{P}_i = \frac{p_{\theta(i)t}}{p_{\theta(J)t}} = \left(\frac{p_{\theta ct}^{1-\sigma} + p_{\theta dt}^{1-\sigma}}{p_{Jt}^{1-\sigma}}\right)^{\frac{1}{1-\sigma}} = \left(\mathcal{R}_{\theta c}^{\alpha(1-\sigma)}\bar{A}_{\theta c}^{(1-\alpha)(\sigma-1)} + \mathcal{R}_{\theta d}^{\alpha(1-\sigma)}\bar{A}_{\theta d}^{(1-\alpha)(\sigma-1)}\right)^{\frac{1}{1-\sigma}}, \tag{A.27}$$

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.26). Note that if the production structure is a simple CES, i.e. $\lambda = \sigma$, then Equation (A.26) constructively provides the economy's unique steady-state.

Furthermore, Equation (A.17) implies that

$$\ln\left(\bar{s}_{ss}\right) = \ln\left(\mathcal{V}\right) + \frac{1}{\eta}\mathbf{\Phi}\ln\left(\bar{A}_{ss}\right), \tag{A.28}$$

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

$$\ln\left(\bar{s}_{ss}\right) = \ln\left(\mathcal{V}\right) + \mathbf{\Phi}\left[\mathbf{\Phi} - \eta(1-\alpha)(\sigma-1)\mathbf{I}\right]^{-1}\left[\ln\left(\Xi\right) + (\sigma-\lambda)\ln\left(\mathcal{P}\right) - \alpha(\sigma-1)\ln\left(\mathcal{R}\right)\right]. \tag{A.29}$$

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

$$\ln\left(\bar{A}_{ss}\right) = \frac{1}{(1-\alpha)(1-\lambda)}\ln\left(\Xi\right) + \frac{\alpha}{1-\alpha}\ln\left(\mathcal{R}\right), \tag{A.30}$$

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

A.5 Proof of Proposition 2

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

$$\ln\left(\bar{A}_t\right) = \mathcal{H}\Big(s_t\big(\ln\left(\bar{A}_{t-1}\right)\big), \ln\left(\bar{A}_{t-1}\right)\Big), \tag{A.31}$$

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

$$\ln\left(\bar{A}_t\right) - \ln\left(\bar{A}_{ss}\right) \approx \boldsymbol{\mathcal{J}}\big(\ln\left(\bar{A}_{t-1}\right) - \ln\left(\bar{A}_{ss}\right)\big), \tag{A.32}$$

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 (40) 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

$$\begin{aligned} \ln\left(\bar{A}_{it}\right) &= \ln\left(\gamma\right)\left[\chi_i s_{it}^{\eta}\phi_{it} - \chi_J s_{Jt}^{\eta}\phi_{Jt}\right] + \ln\left(\bar{A}_{it-1}\right) \\ (1-\eta)\left[\ln\left(s_{it}\right) - \ln\left(s_{Jt}\right)\right] &= \ln\left(\chi_i/\chi_J\right) + \ln\left(\xi_i/\xi_J\right) + \ln\left(\phi_{it}/\phi_{Jt}\right) + \ln\left(p_{it}/p_{Jt}\right) + \ln\left(Y_{it}/Y_{Jt}\right) \\ \sum_i s_{it} &= \mathcal{S}. \end{aligned} \tag{A.33}$$

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

$$\begin{aligned} \varepsilon_{ij}^{\bar{A}} &= g\left[\eta(\varepsilon_{ij}^{s} - \varepsilon_{Jj}^{s}) - \Phi_{ij}\right] + \mathbb{1}(i=j) \\ (1-\eta)(\varepsilon_{ij}^{s} - \varepsilon_{Jj}^{s}) &= -\Phi_{ij} - (1-\alpha)\varepsilon_{ij}^{\bar{A}} + \sum_q \Sigma_{iq}(1-\alpha)\varepsilon_{qj}^{\bar{A}} \\ \Rightarrow \left[(1-\eta) + g\eta(1-\alpha)\right]\varepsilon_{ij}^{\bar{A}} &= (1-\eta)\mathbb{1}(i=j) - g\Phi_{ij} + g\eta(1-\alpha)\sum_q \Sigma_{iq}\varepsilon_{qj}^{\bar{A}}, \end{aligned} \tag{A.34}$$

where the derivative of relative good prices with respect to relative technology comes from Equation (A.4). Transforming into matrix notation, we have Equation (41) from Proposition 2:

$$\begin{aligned} \left[(1-\eta)\mathbf{I} - g\eta(1-\alpha)\left(\boldsymbol{\Sigma} - \mathbf{I}\right)\right]\boldsymbol{\mathcal{J}} &= (1-\eta)\mathbf{I} - g\boldsymbol{\Phi} \\ \Rightarrow \boldsymbol{\mathcal{J}} &= \left[(1-\eta)\mathbf{I} - g\eta(1-\alpha)\left(\boldsymbol{\Sigma} - \mathbf{I}\right)\right]^{-1}\left[(1-\eta)\mathbf{I} - g\boldsymbol{\Phi}\right]. \end{aligned} \tag{A.35}$$

□

A.6 Proof of Corollary 2

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

$$\begin{aligned} \boldsymbol{\mathcal{J}}\boldsymbol{Q} &= \boldsymbol{Q}\boldsymbol{D}(\kappa) \\ \Rightarrow \left[(1-\eta)\mathbf{I} - g\boldsymbol{\Phi}\right]\boldsymbol{Q} &= \left[(1-\eta)\mathbf{I} - g\eta(1-\alpha)\left(\boldsymbol{\Sigma} - \mathbf{I}\right)\right]\boldsymbol{Q}\boldsymbol{D}(\kappa) \\ \Rightarrow \left[(1-\eta)\mathbf{I} - g\eta(1-\alpha)\left(\boldsymbol{\Sigma} - \mathbf{I}\right)\right]\boldsymbol{Q}\boldsymbol{D}(1-\kappa) &= g\left[\boldsymbol{\Phi} - \eta(1-\alpha)\left(\boldsymbol{\Sigma} - \mathbf{I}\right)\right]\boldsymbol{Q} \\ \Rightarrow \boldsymbol{\mathcal{M}} &= g\boldsymbol{Q}\boldsymbol{D}(1-\kappa)^{-1}\boldsymbol{Q}^{-1}\left[(1-\eta)\mathbf{I} - g\eta(1-\alpha)\left(\boldsymbol{\Sigma} - \mathbf{I}\right)\right]^{-1}, \end{aligned} \tag{A.36}$$

which gives us Equation (46) of Corollary 2.

□

A.7 Closed-Form for Steady-State Growth Rate g

Combining Equations (12) and (A.17), we have that the steady-state level of scientists must follow

$$s_j = \frac{\left(\chi_j \phi_j\right)^{-\frac{1}{\eta}}}{\sum_i \left(\chi_i \phi_i\right)^{-\frac{1}{\eta}}} \mathcal{S}. \tag{A.37}$$

Plugging this into the growth rate (9), we have

$$g = \ln\left(\gamma\right)\Big(\sum_j \left(\chi_j \phi_j\right)^{-\frac{1}{\eta}}\Big)^{-\eta} \mathcal{S}^{\eta}. \tag{A.38}$$

That is, the effects of technology-level research productivity aggregate to steady-state growth via a CES with elasticity of substitution $\eta/(1+\eta)$. Because $\eta \in (0,1)$, this CES will be complements with elasticity of substitution between 0 and 1/2.

A.8 Proof of Proposition 4

As described in the Definition 6, the Planner solves

$$\begin{aligned}
\max_{\{c_t, \{A_{jt}, s_{jt}\}\}} \sum_{t \geq 0} \frac{1}{(1+\rho)^t} u(c_t) \;\; s.t. \qquad & \\
\mathcal{Y}_t = c_t + \sum_j r_j \Lambda_{jt} : \;\; & \varkappa_t \\
\ln\left(A_{jt}\right) = \ln\left(\gamma\right) \chi_j s_{jt}^{\eta} \phi_{jt} + \ln\left(A_{jt-1}\right) : \;\; & \epsilon_{jt} \\
\mathcal{S} = \sum_j s_{jt} : \;\; & \varpi_{st} \\
\{\Lambda_{jt}, \{\ell_{j\iota t}\}\} = \text{argmax}\left[\mathcal{Y}_t - \sum_j r_j \Lambda_{jt} - \tau_t \mathcal{E}_t\right] \;\; s.t. \;\; L = \sum_j \int_0^1 \ell_{j\iota t} d\iota, &
\end{aligned} \tag{A.39}$$

where next to each constraint is the corresponding multiplier. The incentive compatibility constraint of the final line defines functions for factors $\{\Lambda_{jt}, \{\ell_{j\iota t}\}\}$ that depend on the choice of technology. Note that I have transformed the law of motion for technology to be in terms of logs and that I am assuming the intermediate subsidy is set according to $\Upsilon = \gamma$, as in the first-best.

Taking the FOC with respect to c_t, we have

$$\frac{u_t'}{(1+\rho)^t} = \varkappa_t, \tag{A.40}$$

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

$$\varpi_{st} = \epsilon_{jt} \ln\left(\gamma\right) \chi_j \eta s_{jt}^{\eta-1} \phi_{jt}. \tag{A.41}$$

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

$$\frac{\epsilon_{jt}}{A_{jt}} = \frac{u'_t}{(1+\rho)^t}\frac{\partial \mathcal{Y}_t}{\partial A_{jt}} + \sum_{\hat{t}\geq t}\frac{u'_{\hat{t}}}{(1+\rho)^{\hat{t}}}\Big[\sum_i \int_0^1 \frac{\partial \mathcal{Y}_{\hat{t}}}{\partial \ell_{i\iota\hat{t}}}\frac{\partial \ell_{i\iota\hat{t}}}{\partial A_{jt}}d\iota + \sum_i \big(\frac{\partial \mathcal{Y}_{\hat{t}}}{\partial \Lambda_{i\hat{t}}} - r_i\big)\frac{\partial \Lambda_{i\hat{t}}}{\partial A_{jt}} - SCC_{\hat{t}}\frac{\partial \mathcal{E}_{\hat{t}}}{\partial A_{jt}}\Big] \tag{A.42}$$
$$+ \frac{\epsilon_{jt+1}}{A_{jt}} + \sum_i \epsilon_{it+1}\ln(\gamma)\chi_i s_{it+1}^{\eta}\frac{\partial \phi_{it+1}}{\partial A_{jt}},$$

where changes in output with respect to changes in inputs should be understood as holding damages constant and SCC_t is the social marginal cost of increasing carbon emissions. This social cost of carbon follows

$$SCC_t = -\sum_{\hat{t}\geq t}\prod_{\hat{s}=1}^{\hat{t}-t}\frac{1}{R_{t+\hat{s}}}\mathcal{Y}_{\hat{t}}\frac{\partial \ln(\Omega_{\hat{t}})}{\partial \mathcal{E}_t}, \tag{A.43}$$

which gives us Equation (58) of Proposition 4. Note that changes in technology at time t generate changes in the equilibrium allocation at all future periods because of the dynamic effect of climate damages. Now consider the producer optimality conditions

$$\begin{aligned}\frac{\partial \mathcal{Y}_t}{\partial \Lambda_{jt}} &= r_j + \omega_j \tau_t \\ \frac{\partial \mathcal{Y}_t}{\partial \ell_{j\iota t}} &= w_{\ell t}.\end{aligned} \tag{A.44}$$

Plugging these conditions into the Planner's optimality condition for technology (A.42), 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

$$\frac{\epsilon_{jt}}{A_{jt}} = \frac{u'_t}{(1+\rho)^t}\frac{\partial \mathcal{Y}_t}{\partial A_{jt}} - \sum_{\hat{t}\geq t}\frac{u'_{\hat{t}}}{(1+\rho)^{\hat{t}}}(SCC_{\hat{t}} - \tau_{\hat{t}})\frac{\partial \mathcal{E}_{\hat{t}}}{\partial A_{jt}} + \frac{\epsilon_{jt+1}}{A_{jt}} + \sum_i \epsilon_{it+1}\ln(\gamma)\chi_i s_{it+1}^{\eta}\frac{\partial \phi_{it+1}}{\partial A_{jt}}. \tag{A.45}$$

Comparing the optimality condition of the research problem (A.5) with the Planner's scientist FOC (A.41), we have

$$\begin{aligned}\chi_j \eta s_{jt}^{\eta-1}\phi_{jt}\xi_{jt}\Pi_{jt} &= w_{st} \\ \chi_j \eta s_{jt}^{\eta-1}\phi_{jt}\epsilon_{jt} &= \frac{\varpi_{st}}{\ln(\gamma)}.\end{aligned} \tag{A.46}$$

The Planner can set the wage for scientists according to

$$w_{st} = \frac{\varpi_{st}\ \ln(\gamma)}{u'_t/(1+\rho)^t}. \tag{A.47}$$

Plugging this into the two research conditions (A.46), we can see that the incentives for innovation will be efficient if the Planner sets innovation subsidies according to

$$\xi_{jt}\Pi_{jt} = \frac{\epsilon_{jt}}{u'_t/(1+\rho)^t}. \tag{A.48}$$

Plugging this into Equation (A.45), we have that innovation subsidies follow the recursion

$$\xi_{jt}\Pi_{jt} = (1-\alpha)S_{jt}\mathcal{Y}_t - \sum_{\hat{t}\geq t}\prod_{\hat{s}=1}^{\hat{t}-t}\frac{1}{R_{t+\hat{s}}}(SCC_{\hat{t}}-\tau_{\hat{t}})\mathcal{E}_{\hat{t}}\frac{\partial \ln(\mathcal{E}_{\hat{t}})}{\partial \ln(A_{jt})} + \frac{1}{R_{t+1}}\Big[\xi_{jt+1}\Pi_{jt+1} + \sum_i \xi_{it+1}\Pi_{it+1}g_{it+1}\varphi_{ijt+1}\Big], \tag{A.49}$$

which gives us Equation (57) of Proposition 4. 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.

I describe first-best policy in detail in Appendix A.10, so one can verify by inspection that the second-best innovation subsidies of Proposition 4 recover those of the first-best whenever $\tau_t = SCC_t$. □

A.9 Proof of Corollary 3

First, denote by $\tilde{\xi}_{jt} \equiv \xi_{jt}S_{jt}$ innovation subsidies multiplied by income shares. Note that with $\Upsilon = \gamma$ average intermediate producer profit follows $\Pi_{jt} = (\gamma-1)(1-\alpha)S_{jt}\mathcal{Y}_t$. Manipulating Equation (57), we have

$$\tilde{\xi}_{jt} = \frac{S_{jt}-\mathcal{T}_{jt}}{\gamma-1} + \frac{\mathcal{Y}_{t+1}/\mathcal{Y}_t}{R_{t+1}}\Big[\tilde{\xi}_{jt+1} + \sum_i \tilde{\xi}_{it+1}g_{it+1}\varphi_{ijt+1}\Big]. \tag{A.50}$$

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.[47] More generally, we would have $g_y \leq g$ as damages could worsen with continued carbon pollution.

Then, evaluating Equation (A.50) in matrix notation at the steady-state, we have

$$\tilde{\xi}' = \frac{1}{\gamma-1}[S'-\mathcal{T}']\Big((1-\tilde{R}^{-1})\mathbf{I} - g\tilde{R}^{-1}\boldsymbol{\varphi}\Big)^{-1}. \tag{A.51}$$

Normalizing by the common innovation wedge gives us Equation (59) of Corollary 3. Another way to gain intuitive understanding of steady-state policy is to write innovation incentives in terms of a Leontief inverse. We can write

$$\hat{\xi}' = [S'-\mathcal{T}']\Big(\mathbf{I} - \frac{g\tilde{R}^{-1}}{1-\tilde{R}^{-1}}\boldsymbol{\varphi}\Big)^{-1} = [S'-\mathcal{T}']\sum_{t\geq 0}\Big(\frac{g\tilde{R}^{-1}}{1-\tilde{R}^{-1}}\boldsymbol{\varphi}\Big)^t. \tag{A.52}$$

The final sum represents the sequence of spillovers that a given act of innovation will create in the future. First, the innovations pass through the spillover network to generate additional innovation in the following period; these downstream innovations then create a second round of spillovers through the network, and so on. All of these effects are discounted by the term $g\tilde{R}^{-1}/(1-\tilde{R}^{-1})$. This full sequence of spillover ripple effects aggregates into a Leontief inverse, which then loads on Domar weights net of the distortion term because these tell us the value of increasing a technology at each point in time.

In interpreting Equation (59), I claim that the research allocation of a myopic Planner will maximize

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

contemporaneous output, while that of a perfectly patient Planner will maximize growth. To see this, first note that Equations (25) and (A.37) imply that $\hat{\xi} \propto s \propto \left(\chi_j \phi_j\right)^{-\frac{1}{\eta}}$. For the myopic Planner, we have innovation incentives set according to the flow value of innovation: $\hat{\xi} = S - \mathcal{T}$. They solve the problem

$$\begin{aligned} \max_{\{A'_j, s_j\}} \mathcal{Y} - \sum_j r_j \Lambda_j \quad s.t. \\ \sum_j s_j = \mathcal{S} \\ \ln\left(A'_j\right) = \ln\left(\gamma\right) z_j + \ln\left(A_j\right), \end{aligned} \tag{A.53}$$

where I have removes time subscripts to reflect the static nature of the myopic Planner's problem. Using similar arguments from Appendix A.8, we that the optimum requires $s \propto S - \mathcal{T}$. Hence, setting $\hat{\xi} = S - \mathcal{T}$ will achieve this condition.

For the patient Planner, we have innovation incentives set according to eigenvector centrality in the gross spillover matrix: $\hat{\xi}' \boldsymbol{\varphi} = \vec{0}'$. Differentiating the steady-state growth rate (A.38) with respect to relative technology gives us that

$$\sum_i \left(\chi_i \phi_i\right)^{-\frac{1}{\eta}} \varphi_{ij} = 0. \tag{A.54}$$

If we have $\hat{\xi}' \boldsymbol{\varphi} = \vec{0}'$, all of these conditions will be satisfied.

□

A.10 First-Best Policy

In this appendix, I will describe 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 first-best allocation as an equilibrium.

Definition A.1 (First-Best Planning Problem)**.** *The Planner solves*

$$\begin{aligned} \max_{\{c_t, \{\Lambda_{jt}, \{\ell_{j\iota t}\}, A_{jt}, s_{jt}\}\}} \sum_{t \geq 0} \frac{1}{(1+\rho)^t} u(c_t) \quad s.t. \\ \mathcal{Y}_t = c_t + \sum_j r_j \Lambda_{jt} : \quad \varkappa_t \\ L = \sum_j \int_0^1 \ell_{j\iota t} d\iota : \quad \varpi_{\ell t} \\ \ln\left(A_{jt}\right) = \ln\left(\gamma\right) \chi_j s_{jt}^{\eta} \phi_{jt} + \ln\left(A_{jt-1}\right) : \quad \epsilon_{jt} \\ \mathcal{S} = \sum_j s_{jt} : \quad \varpi_{st}, \end{aligned} \tag{A.55}$$

That is, the Planner picks the technologically feasible allocation that will maximize household utility. Next to each constraint is the corresponding multiplier.

Taking the FOC with respect to c_t, we have

$$\frac{u_t'}{(1+\rho)^t} = \varkappa_t, \tag{A.56}$$

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

$$\frac{\partial \mathcal{Y}_t}{\partial \Lambda_{jt}} = r_j - \omega_j \sum_{\hat{t} \geq t} \frac{\mathcal{Y}_{\hat{t}}}{(1+\rho)^{\hat{t}-t}} \frac{u_{\hat{t}}'}{u_t'} \frac{\partial \ln(\Omega_{\hat{t}})}{\partial \mathcal{E}_t}. \tag{A.57}$$

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

$$\frac{\partial \mathcal{Y}_t}{\partial \ell_{j\iota t}} = \frac{\varpi_{\ell t}}{u_t'/(1+\rho)^t}. \tag{A.58}$$

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

$$\frac{\epsilon_{jt}}{A_{jt}} = \frac{u_t'}{(1+\rho)^t} \frac{\partial \mathcal{Y}_t}{\partial A_{jt}} + \frac{\epsilon_{jt+1}}{A_{jt}} + \sum_i \epsilon_{it+1} \ln(\gamma) \chi_i s_{it+1}^{\eta} \frac{\partial \phi_{it+1}}{\partial A_{jt}}. \tag{A.59}$$

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

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

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 $\{p_{jt}, \{p_{j\iota t}\}\}$ set equal to marginal products/marginal costs will be efficient.

Next, to align the input demand condition (17) with the Planner's input FOC (A.57), the Planner sets the carbon price according to

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

which is the social cost of carbon.

For innovation subsidies, we can compare the optimality condition of the research problem (A.5) with the Planner's scientist FOC (A.60). We have

$$\begin{aligned} \chi_j \eta s_{jt}^{\eta-1} \phi_{jt} \xi_{jt} \Pi_{jt} &= w_{st} \\ \chi_j \eta s_{jt}^{\eta-1} \phi_{jt} \epsilon_{jt} &= \frac{\varpi_{st}}{\ln(\gamma)}. \end{aligned} \tag{A.62}$$

The Planner can set the wage for scientists according to

$$w_{st} = \frac{\varpi_{st} \ \ln(\gamma)}{u_t'/(1+\rho)^t}. \tag{A.63}$$

Plugging this into the two research conditions (A.62), we can see that the incentives for innovation will be efficient if the Planner sets innovation subsidies according to

$$\xi_{jt}\Pi_{jt} = \frac{\epsilon_{jt}}{u_t'/(1+\rho)^t}. \tag{A.64}$$

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.59), we get the recursion

$$\xi_{jt}\Pi_{jt} = (1-\alpha)S_{jt}\mathcal{Y}_t + \frac{1}{R_{t+1}}\Big[\xi_{jt+1}\Pi_{jt+1} + \sum_i \xi_{it+1}\Pi_{it+1}g_{it+1}\varphi_{ijt+1}\Big]. \tag{A.65}$$

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 Spillover Network Stability

This appendix details my examination of the spillover network's stability over time. To start, I use the application year of citing patents to compute a spillover network that is specific to each of the five year bins from 1970 to 2015. My method for computing the spillover network is otherwise the same as that described in Section 5.1.

My first test of the spillover network's stability considers AR(1)-style regressions where I regress an element of the gross spillover network on its realization from the previous period. The resulting lag coefficient gives us a measure of persistence in that a value very close to one suggests an unchanging spillover network. Formally, I use the estimating equation

$$\tilde{\varphi}_{ijt} = \mu\tilde{\varphi}_{ijt-1} + \delta_{ijt}. \tag{B.1}$$

A specific concern motivating this exercise is that the relative newness of clean technologies may imply that past estimates of the spillovers these technologies create my not be a good guide for the spillovers we should expect them to create in the future. This is especially relevant for optimal policy, where the Planner should pick forward-looking innovation subsidies that reflect spillover creation in future periods. To address this concern, I allow Equation (B.1) to also include time trends that are specific to spillover pairs where the sending technology is clean. I allow for time trends that are shared across both clean technologies as well as specific to transport and electricity generation. The results are presented in Table B.1.

We can see that for every specification the spillover network is highly persistent with a lag coefficient near one. Furthermore, clean technologies do not trend towards greater spillover creation over time. Taken together, these estimates support the use of a Cobb-Douglas spillover network, where spillover

Table B.1: Spillover Network Stability

	Dependent Variable: Spillover Elasticity		
	(1)	(2)	(3)
Lagged Spillover Elasticity	0.982	0.983	0.983
	(0.01)	(0.01)	(0.01)
Clean Technology Trend		-0.0001	
		(0.0002)	
Clean Transport Trend			0.00002
			(0.0002)
Clean Electricity Trend			-0.006
			(0.008)
R^2	0.992	0.992	0.992
Obs	200	200	200

Notes: This table presents estimates that describe the stability of the spillover network. The spillover network is computed in five year bins from 1970 to 2015. Standard errors clustered by spillover recipient are reported in parentheses.

elasticities are highly stable over time.

My second test for spillover stability is motivated by my theoretical finding that spillover centrality is the relevant measure of spillover creation for optimal policy. Quantitatively, I argue in the main text that clean technologies should receive relatively small innovation subsidies because of their low centrality in the spillover network. Again, this claim is vulnerable to the concern that clean technologies may gain in their level of centrality as they become more mature, so to address this concern, I plot the centrality of clean technologies over time in Figure B.1. Neither clean technology has a clear trend in their level of centrality over time, and indeed the clean technology that has seen the greater change in centrality over time – electricity generation – saw a *reduction* in centrality. Of course, it is impossible to completely rule out a structural break in the future spillover network, but the evidence presented here suggests there is no reason to expect such a change. This is consistent with the results of Liu and Ma (2021) who also find a stable citation network over time.

B.2 Details on Numerical Representations

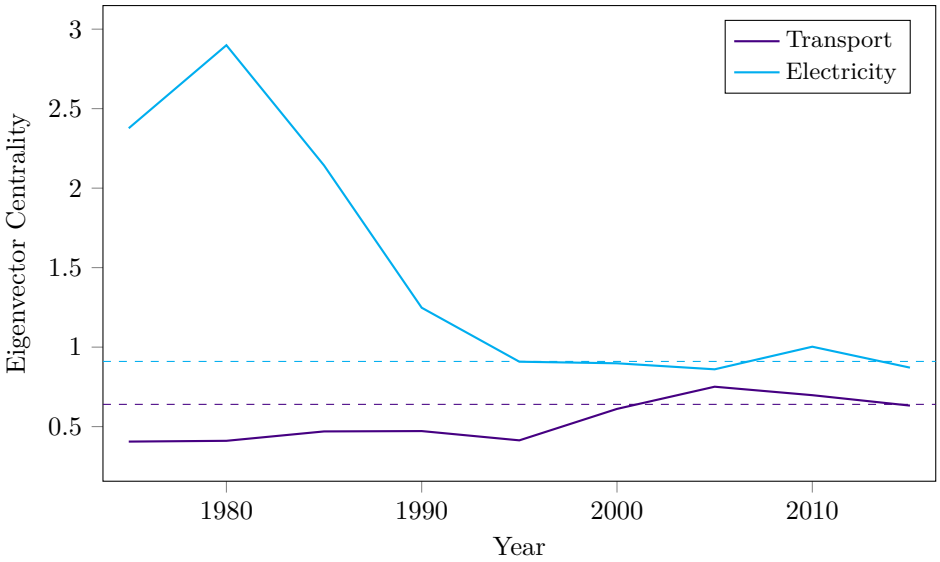
This appendix contains details on how I can represent equilibrium outcomes of my model in terms of fundamentals: technology, policy, and structural parameters. I make the parametric assumptions of Section 2.5. First, I will define what I call pseudo prices. These follow

$$\tilde{p}_{jt} \equiv \left(r_j + \omega_j \tau_t\right)^{\alpha} / A_{jt}^{1-\alpha}. \tag{B.2}$$

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

$$p_{jt} = \left(\frac{1}{\alpha}\right)^{\alpha} \left(\frac{1}{1-\alpha}\right)^{1-\alpha} \left(\frac{\gamma w_{\ell t}}{\Upsilon}\right)^{1-\alpha} \tilde{p}_{jt}. \tag{B.3}$$

Figure B.1: Clean Technology Centrality



Notes: Eigenvector centrality computed for each technology using five year bins. Dashed lines indicate centrality in the sample-wide spillover network.

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

$$\tilde{p}_{\theta t} \equiv \left(\tilde{p}_{\theta ct}^{1-\sigma} + \tilde{p}_{\theta dt}^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \tag{B.4}$$

$$\tilde{P}_t \equiv \Omega_t^{-1}\left(\sum_{\theta} \nu_\theta \tilde{p}_{\theta t}^{1-\lambda}\right)^{\frac{1}{1-\lambda}}, \tag{B.5}$$

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

$$\frac{\gamma w_{\ell t}}{\Upsilon} = (1-\alpha)\frac{\mathcal{Y}_t}{L}. \tag{B.6}$$

Plugging the wage (B.6) into the factor of proportionality and combining the final pseudo price (B.5) with the fact that output is the numeraire, we have for output

$$\mathcal{Y}_t = \alpha^{\frac{\alpha}{1-\alpha}} L \tilde{P}_t^{\frac{-1}{1-\alpha}}. \tag{B.7}$$

Next, using the demand conditions for sectors (A.8) and goods within sectors (A.10), we can derive income shares for sectors and goods within sectors. We have

$$S_{\theta t} \equiv \frac{p_{\theta t} E_{\theta t}}{\mathcal{Y}_t} = \frac{\nu_\theta \tilde{p}_{\theta t}^{1-\lambda}}{\Omega_t^{1-\lambda} \tilde{P}_t^{1-\lambda}} \tag{B.8}$$

$$S_{et}^{\theta} \equiv \frac{p_{\theta et} Y_{\theta et}}{p_{\theta t} E_{\theta t}} = \frac{\tilde{p}_{\theta et}^{1-\sigma}}{\tilde{p}_{\theta t}^{1-\sigma}}, \tag{B.9}$$

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

$$\Lambda_{jt} = \frac{\alpha S_{jt} \mathcal{Y}_t}{r_j + \omega_j \tau_t}, \tag{B.10}$$

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 $q^\theta_{ct} \equiv Y_{\theta ct}/(Y_{\theta ct} + Y_{\theta dt})$. To map between clean income shares and clean quantity shares, note that the demand condition within sectors (A.10) implies that

$$S^\theta_{ct} = \frac{p_{\theta ct} Y_{\theta ct}}{p_{\theta ct} Y_{\theta ct} + p_{\theta dt} Y_{\theta dt}} = \frac{1}{1 + \left((1 - q^\theta_{ct})/q^\theta_{ct}\right)^{\frac{\sigma-1}{\sigma}}}. \tag{B.11}$$

A major advantage of the climate specification of Equation (62) 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

$$\mathcal{C}_{1t} = \psi_p \mathcal{E}_t + \mathcal{C}_{1t-1}. \tag{B.12}$$

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

$$\mathcal{C}_{2t} = (1 - \psi_p)\psi_0 \mathcal{E}_t + \psi \mathcal{C}_{2t-1}. \tag{B.13}$$

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 $\mathcal{C}_{1t_0} = \sum_{\hat{t}=1800}^{t_0} \psi_p \mathcal{E}_{\hat{t}} + \bar{\mathcal{C}}$ 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.[48]

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

$$\begin{aligned} z_{jt} &= \chi_j \left(\frac{R\&D_{jt}}{w_{st}}\right)^\eta \phi_{jt} \\ \Rightarrow \frac{\partial \ln(z_{jt})}{\partial \ln(R\&D_{jt})} &= \eta, \end{aligned} \tag{B.14}$$

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

[48] 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.

condition of the research problem (A.5) implies

$$\eta\chi_j\left(\frac{R\&D_{jt}}{w_{st}}\right)^{\eta}\phi_{jt} = R\&D_{jt} \tag{B.15}$$
$$\Rightarrow \frac{\partial \ln\left(R\&D_{jt}\right)}{\partial \ln\left(w_{st}\right)} = -\frac{\eta}{1-\eta},$$

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

B.3 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 subsidies $\bar{\xi}_{\theta c}$, consider public spending on the subsidy as a proportion of total spending in each sector. We have

$$\frac{\bar{\xi}_{\theta c} r_{\theta c} \Lambda_{\theta ct}}{p_{\theta t} E_{\theta t}} = \frac{\bar{\xi}_{\theta c}}{1-\bar{\xi}_{\theta c}} \frac{\alpha p_{\theta ct} Y_{\theta ct}}{p_{\theta t} E_{\theta t}} = \frac{\bar{\xi}_{\theta c}}{1-\bar{\xi}_{\theta c}} \alpha S^{\theta}_{ct}, \tag{B.16}$$

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

Let $\hat{cred}_{\theta ct}$ denote data on federal spending on clean tax credits as a proportion of total spending in the relevant sector. The numerators come from US Congressional Research Service Reports IF11017 and IF10479, which contain federal outlays on the plug-in electric vehicle and energy investment tax credits, respectively. The former report starts in 2011 and was published in 2019, while the latter starts in 2011 and was published in 2021. Both reports contain projections into later years, but I only include spending that occurred before each report was published. As outlined in the latter report, the vast majority of the expenditures for the energy investment tax credit went to solar. The denominators come from the series on sectoral output discussed in Section 5.3. Next, I take estimates of clean income shares in each sector $\hat{S}^{\theta}_{ct}$ from my data on clean quantity shares using the mapping from Equation (B.11). Plugging these data series into Equation (B.16), we have

$$\frac{\bar{\xi}_{\theta c}}{1-\bar{\xi}_{\theta c}} \sum_{t=2011}^{T_{IF}} \alpha \hat{S}^{\theta}_{ct} = \sum_{t=2011}^{T_{IF}} \hat{cred}_{\theta ct}, \tag{B.17}$$

where T_{IF} is the end year of data for the two congressional reports. This allows me to back out the values $\bar{\xi}_{car,c} = 0.012$ and $\bar{\xi}_{elec,c} = 0.03$.

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

$$\frac{(\xi_j - 1) z_{jt} \Pi_{jt}}{w_{st} \sum_j s_{jt}} = \frac{\xi_j - 1}{\xi_j} \frac{s_{jt}}{\eta \mathcal{S}}. \tag{B.18}$$

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.2 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

$$\frac{\xi_j - 1}{\xi_j} \sum_{t=2010}^{2015} \frac{s_{jt}}{\eta \mathcal{S}} = \sum_{t=2010}^{2015} \hat{pubrd}_{jt}. \tag{B.19}$$

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.19), so I calibrate separate innovation subsidies for the versions of the model with and without spillovers. In the model with spillovers this yields $(\xi_{car,c}, \xi_{car,d}, \xi_{elec,c}, \xi_{elec,d}) = (1.005, 1.015, 1.057, 1.01)$, and in the model without spillovers this yields $(\xi_{car,c}, \xi_{car,d}, \xi_{elec,c}, \xi_{elec,d}) = (1.21, 1.005, 1.272, 1.015)$. 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.2: 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.4 Optimal Policy Simulation Solution Method

To simulate the optimal 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.2, my economy can be reduced to the sequence $\{SCC_{1t}, SCC_{2t}, \mathcal{C}_{1t}, \mathcal{C}_{2t}, \{\tilde{\xi}_{jt}, A_{jt}\}, \varsigma_{1t}, \varsigma_{2t}\}$.

For the social cost of carbon, combining atmospheric carbon concentrations (62) and the damage function (63) gives us

$$SCC_t = \varrho \sum_{\hat{t} \geq t} \prod_{\hat{s}=1}^{\hat{t}-t} \frac{1}{R_{t+\hat{s}}} \mathcal{Y}_{\hat{t}} \big(\psi_p + (1-\psi_p)\psi_0 \psi^{\hat{t}-t}\big). \tag{B.20}$$

This allows us to write the social cost of carbon as a two-dimensional recursion $\{SCC_{1t}, SCC_{2t}\}$, which separately accounts for the permanent and transitory component of carbon pollution. These follow

$$SCC_{1t} = \varrho \psi_p \mathcal{Y}_t + \frac{1}{R_{t+1}} SCC_{1t+1} \tag{B.21}$$

$$SCC_{2t} = \varrho (1-\psi_p)\psi_0 \mathcal{Y}_t + \frac{1}{R_{t+1}} \psi SCC_{2t+1}. \tag{B.22}$$

The sum of these two components gives us the social cost of carbon: $SCC_t = SCC_{1t} + SCC_{2t}$. I can then set the carbon price τ_t to an appropriate fraction of the social cost of carbon. The recursive formulation for $\{\mathcal{C}_{1t}, \mathcal{C}_{2t}\}$ is given by Equations (B.12) and (B.13), 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.50), 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

$$(1-\eta)\ln(\bar{s}_t) = \ln(\tilde{\Xi}_t) - \eta\ln(\mathcal{V}) - \mathbf{\Phi}\ln(\bar{A}_{t-1}), \tag{B.23}$$

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. Because innovation subsidies can be negative, I will add the qualification $s_{jt} = 0$ whenever $\tilde{\xi}_{jt} \leq 0$. Furthermore, as I describe in Appendix B.2, 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 $\{A_{jt}\}$.

For the effect of innovation on equilibrium emissions $\mathcal{T}_{jt}$, we can use the results of Appendix B.2 to write the sum of future emission changes in terms of a two-dimensional recursion $\{\varsigma_{1t}, \varsigma_{2t}\}$. Equation (B.10) defines a mapping from current technology and input prices, as well as past emissions, to current input demand, so for all $\hat{t} > t$ we have

$$\frac{\partial \mathcal{E}_{\hat{t}}}{\partial \ln(A_{jt})} = Z_{\hat{t}}\Big(\sum_{s=0}^{\hat{t}-1-t} \frac{\partial \ln(\Omega_{\hat{t}})}{\partial \mathcal{E}_{t+s}} \frac{\partial \mathcal{E}_{t+s}}{\partial \ln(A_{jt})}\Big), \tag{B.24}$$

where

$$Z_t = \frac{\mathcal{E}_t}{1 - \alpha - \frac{\partial \ln(\Omega_t)}{\partial \mathcal{E}_t}\mathcal{E}_t}. \tag{B.25}$$

This implies that

$$\begin{aligned}\mathcal{T}_{jt} = \frac{\bar{\omega}_t}{1-\alpha}\frac{\partial \ln(\mathcal{E}_t)}{\partial \ln(A_{jt})}\Big((SCC_t - \tau_t) \\ + \frac{1}{R_{t+1}} Z_{t+1}\frac{\partial \ln(\Omega_{t+1})}{\partial \mathcal{E}_t}(SCC_{t+1} - \tau_{t+1}) \\ + \frac{1}{R_{t+1}}\frac{1}{R_{t+2}} Z_{t+2}\Big[\frac{\partial \ln(\Omega_{t+2})}{\partial \mathcal{E}_{t+1}} Z_{t+1}\frac{\partial \ln(\Omega_{t+1})}{\partial \mathcal{E}_t} + \frac{\partial \ln(\Omega_{t+2})}{\partial \mathcal{E}_t}\Big](SCC_{t+2} - \tau_{t+2}) \\ + \dots,\end{aligned} \tag{B.26}$$

which gives us that

$$\begin{aligned}\mathcal{T}_{jt} = \frac{\bar{\omega}_t}{1-\alpha}\frac{\partial \ln(\mathcal{E}_t)}{\partial \ln(A_{jt})}\Big((SCC_t - \tau_t) \\ + \sum_{\tilde{t}\geq 1}\sum_{t'>t}\cdots\sum_{t^{(\tilde{t})}>t^{(\tilde{t}-1)}} Z_{t^{(\tilde{t})}}\frac{\partial \ln(\Omega_{t^{(\tilde{t})}})}{\partial \mathcal{E}_{t^{(\tilde{t}-1)}}}\dots Z_{t'}\frac{\partial \ln(\Omega_{t'})}{\partial \mathcal{E}_t}\prod_{\hat{s}=1}^{t^{(\tilde{t})}-t}\frac{1}{R_{t+\hat{s}}}(SCC_{t^{(\tilde{t})}} - \tau_{t^{(\tilde{t})}})\Big).\end{aligned} \tag{B.27}$$

The sum of the second line, which represents all of the cascading effects of damages on future emissions, can be written as the sum of the following two-dimensional recursion:

$$\varsigma_{1t} = \frac{1}{R_{t+1}}\big(- Z_{t+1}\varrho\psi_p(SCC_{t+1} - \tau_{t+1}) + \varsigma_{1t+1} - Z_{t+1}\varrho\psi_p\big[\varsigma_{1t+1} + \varsigma_{2t+1}\big]\big) \tag{B.28}$$

$$\varsigma_{2t} = \frac{1}{R_{t+1}}\big(- Z_{t+1}\varrho(1-\psi_p)\psi_0\psi(SCC_{t+1} - \tau_{t+1}) + \psi\varsigma_{2t+1} - Z_{t+1}\varrho(1-\psi_p)\psi_0\psi\big[\varsigma_{1t+1} + \varsigma_{2t+1}\big]\big). \tag{B.29}$$

Thus, we can write

$$\mathcal{T}_{jt} = \frac{\bar{\omega}_t}{1-\alpha}\frac{\partial \ln(\mathcal{E}_t)}{\partial \ln(A_{jt})}\big((SCC_t - \tau_t) + \varsigma_{1t} + \varsigma_{2t}\big). \tag{B.30}$$

The contemporaneous effect on emissions follows

$$\begin{aligned}\frac{\bar{\omega}_t}{1-\alpha}\frac{\partial \ln(\mathcal{E}_t)}{\partial \ln(A_{jt})} = \frac{\alpha Z_t}{\mathcal{E}_t}\sum_\theta \frac{\omega_{\theta d}S_{\theta dt}}{r_{\theta d} + \omega_{\theta d}\tau_t}\big[&(1 + (1-\alpha)(1-\lambda))S_{jt} \\ &+ \mathbb{1}(\theta = \theta(j))(1-\alpha)(\lambda - \sigma)S_{jt}^\theta \\ &+ \mathbb{1}(\theta d = j)(1-\alpha)(\sigma - 1)\big].\end{aligned} \tag{B.31}$$

In summary, the root system follows

$$SCC_{1t} = \varrho\psi_p\mathcal{Y}_t + \frac{1}{R_{t+1}}SCC_{1t+1} \tag{B.32}$$

$$SCC_{2t} = \varrho(1-\psi_p)\psi_0\mathcal{Y}_t + \frac{1}{R_{t+1}}\psi SCC_{2t+1}$$

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

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

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

$$\ln(A_{jt}) = \ln(\gamma)\chi\Big(\frac{s_{jt}}{\nu_{\theta(j)}}\Big)^\eta\phi_{jt} + \ln(A_{jt-1})$$

$$\varsigma_{1t} = \frac{1}{R_{t+1}}\big(- Z_{t+1}\varrho\psi_p(SCC_{t+1} - \tau_{t+1}) + \varsigma_{1t+1} - Z_{t+1}\varrho\psi_p\big[\varsigma_{1t+1} + \varsigma_{2t+1}\big]\big)$$

$$\varsigma_{2t} = \frac{1}{R_{t+1}}\big(- Z_{t+1}\varrho(1-\psi_p)\psi_0\psi(SCC_{t+1} - \tau_{t+1}) + \psi\varsigma_{2t+1} - Z_{t+1}\varrho(1-\psi_p)\psi_0\psi\big[\varsigma_{1t+1} + \varsigma_{2t+1}\big]\big).$$

The state variables $\{\mathcal{C}_{1t}, \mathcal{C}_{2t}, \{A_{jt}\}\}$ are backward-looking, so I specify initial conditions using the strategies described in Section 5.3 and Appendix B.2. The policy variables $\{SCC_{1t}, SCC_{2t}, \{\tilde{\xi}_{jt}\}\}$ are forward-looking, so I specify terminal conditions 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 the social cost of carbon relative to output by $S\tilde{C}C_t \equiv SCC_t/\mathcal{Y}_t$. This

achieves a steady-state which follows

$$S\tilde{C}C_1 = \frac{\varrho\psi_p}{1-\tilde{R}^{-1}} \tag{B.33}$$

$$S\tilde{C}C_2 = \frac{\varrho(1-\psi_p)\psi_0}{1-\psi\tilde{R}^{-1}}. \tag{B.34}$$

From this, we can see that the social cost of carbon asymptotes to infinity as the economy continues to grow.

To derive the steady-state for $\{\tilde{\xi}_j\}$, we first need to determine the steady-state behavior of $\mathcal{T}$. First, suppose that $\varsigma_1, \varsigma_2 = 0$. The carbon price τ_t will be either zero or a constant fraction of the social cost of carbon. In the former case, $\mathcal{T}$ will diverge to infinity unless steady-state dirty income shares are zero. Without a growing carbon price, this requires steady-state relative dirty technology to be zero as well. In the latter case, τ_t will grow to infinity, so dirty income shares will again be zero in steady-state. In either case, $\mathcal{T} = \vec{0}$ and emissions will go to zero, but in the case of a zero carbon price, we have the added restriction that steady-state relative dirty technology must also be zero. If emissions go to zero, we will have $Z = 0$, which implies $\varsigma_1, \varsigma_2 = 0$ as conjectured.

To select the steady-state income shares S and growth rate g, we can derive steady-state relative technology by combining Equations (A.28) and (B.23) to achieve

$$\ln\left(\bar{A}_{ss}\right) = \eta\mathbf{\Phi}^{-1}\big[\ln\left(\tilde{\Xi}\right) - \ln\left(\mathcal{V}\right)\big], \tag{B.35}$$

which, combined with the requirement that dirty income shares are zero, defines a fixed-point problem to solve for steady-state $\{\tilde{\xi}_j\}$ using Equation (A.51). Note that the growth rate g comes from Equation (9).

Furthermore, the fact that emissions go to zero implies that the growth rate of output g_y goes to the growth rate of technology g because climate damages asymptote to a constant. Therefore, I assume that both consumption and output grow at rate g in the periods beyond my simulation. The terminal social cost of carbon then follows $SCC_{1T+1} = (1+g)\mathcal{Y}_T S\tilde{C}C_1$ and $SCC_{2T+1} = (1+g)\mathcal{Y}_T S\tilde{C}C_2$.

Finding the steady-state when there are no cross-technology spillovers $\boldsymbol{\varphi} = \mathbf{0}$ is a special case because the inverse of $\mathbf{\Phi}$ 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.28) and scientist supply (12) can be satisfied by setting $s_j = \nu_{\theta(j)}$ for all clean technologies. Plugging this into Equation (B.23) gives us that income shares for clean technologies also follow $S_j = \nu_{\theta(j)}$. From Equation (B.8), this implies that $\bar{A}_j = 1$ for clean technologies and $\bar{A}_j = 0$ for dirty technologies.

B.5 CES Spillovers

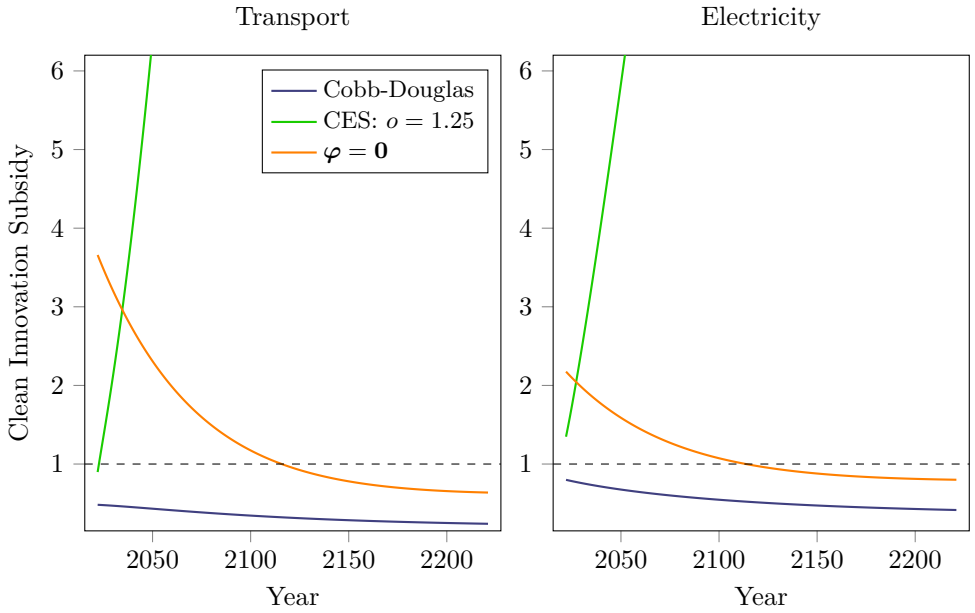
The analysis of the main text assumes that the spillover function is Cobb-Douglas. This assumption is consistent with the historic stability of the spillover network, as described in Appendix B.1, but of course, the future development of technology is famously difficult to predict. In particular, it is plausible that clean technologies will become endogenously more important in the spillover network as they mature in the coming decades. To formalize this possibility, I will consider a CES spillover function

with elasticity of substitution $o > 1$. That is, we have

$$\phi_{it} = \frac{\left(\sum_j \hat{\varphi}_{ij}^{\frac{1}{o}} A_{jt-1}^{\frac{o-1}{o}}\right)^{\frac{o}{o-1}}}{A_{it-1}}. \tag{B.36}$$

This specification explicitly models the potential for technologies to become more important in the spillover network as they become more advanced. I calibrate the matrix of CES shares $\hat{\boldsymbol{\varphi}}$ by matching the gross spillover network to my estimated citation network in the initial period.

Figure B.2: Optimal Innovation Subsidy Path (CES Spillovers)



Notes: Policy paths are all first-best and use the low discount rate. Cobb-Douglas is the base case used in the main text. Innovation subsidies are listed as a fraction of the common innovation wedge.

Figure B.2 shows first-best optimal subsidies for three cases of the spillover network, each of which uses the low discount rate. The Cobb-Douglas and no spillover cases are already described in the main text, and I have also added a CES case with elasticity of substitution $o = 1.25$. The historic stability of the spillover network, as well as ease of comparability, motivates me to pick an elasticity of substitution that is substitutes but still fairly close to one. In the CES case, I assume that the economy converges to a steady-state without dirty technology as this is the hypothetical world following the clean transition that many envision.[49]

As we can see, allowing some degree of substitutability in the spillover functions does substantively change optimal policy, but with very different implications from a big push. When technologies can substitute each other in the creation of spillovers, optimal clean innovation subsidies start out small and rise rapidly over time. This is because the development of clean technologies raises their centrality in the spillover network over time. Indeed, in the CES steady-state, the centrality of clean transport and electricity generation rise to 4.79 and 6.58, respectively. At the same time, relative improvements

[49] I obtain similar paths for optimal policy if I allow for an interior steady-state.

in clean technology lower the income share of transport and electricity (which are eventually produced using only clean technology) because sectors are complements in production. In the CES steady-state, the income shares of clean transport and electricity generation drop to 0.77% and 0.47%, respectively. Therefore, subsidies for clean technology must rise rapidly to cover the growing wedge between spillover creation (summarized by centrality) and private profits (summarized by income shares).

This logic is very different from that of a big push, and as a result, the time paths of optimal subsidies point in opposite directions. A big push calls for an immediate shift in innovation incentives towards clean technology that tapers off over time, whereas a spillover network with substitutable technologies calls for continual support for technologies that have higher centrality than income shares. This will be the case when production and spillover creation have opposing elasticities of substitution because technological leadership will increase weight in the spillover network while lowering it in production. The above figure demonstrates that the undesirability of a big push is robust to different assumption about the elasticity of substitution of the spillover functions, and it suggests that further research on this elasticity is of first order importance for innovation policy.

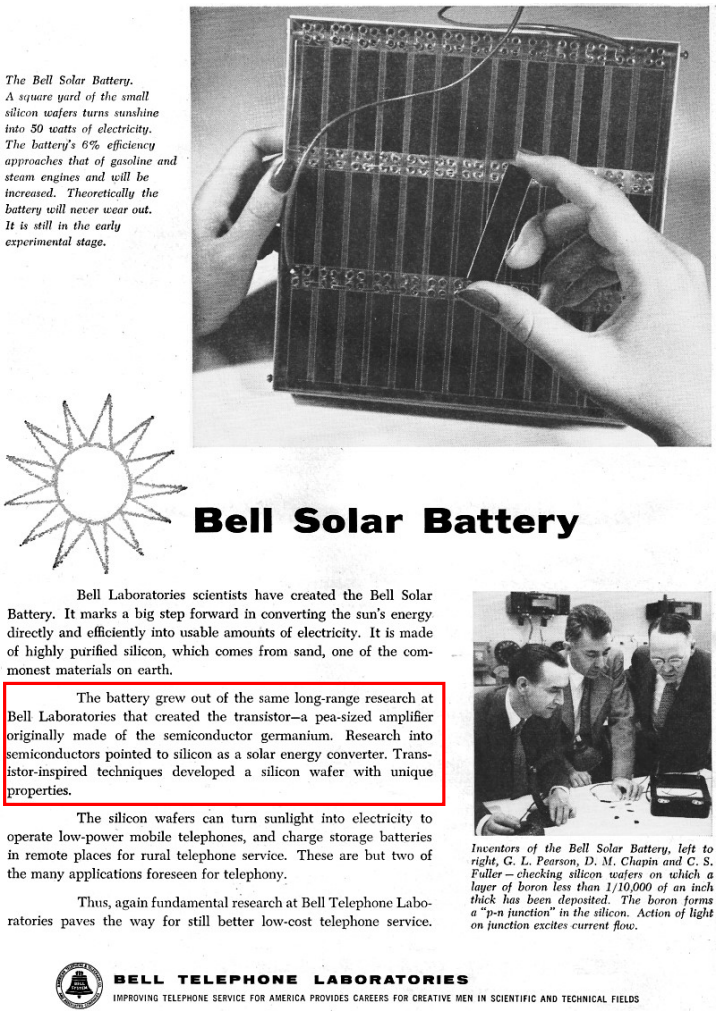
C Additional Tables & Figures

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



The Bell Solar Battery. A square yard of the small silicon wafers turns sunshine into 50 watts of electricity. The battery's 6% efficiency approaches that of gasoline and steam engines and will be increased. Theoretically the battery will never wear out. It is still in the early experimental stage.

Bell Solar Battery

Bell Laboratories scientists have created the Bell Solar Battery. It marks a big step forward in converting the sun's energy directly and efficiently into usable amounts of electricity. It is made of highly purified silicon, which comes from sand, one of the commonest materials on earth.

The battery grew out of the same long-range research at Bell Laboratories that created the transistor—a pea-sized amplifier originally made of the semiconductor germanium. Research into semiconductors pointed to silicon as a solar energy converter. Transistor-inspired techniques developed a silicon wafer with unique properties.

The silicon wafers can turn sunlight into electricity to operate low-power mobile telephones, and charge storage batteries in remote places for rural telephone service. These are but two of the many applications foreseen for telephony.

Thus, again fundamental research at Bell Telephone Laboratories paves the way for still better low-cost telephone service.

Inventors of the Bell Solar Battery, left to right, G. L. Pearson, D. M. Chapin and C. S. Fuller — checking silicon wafers on which a layer of boron less than 1/10,000 of an inch thick has been deposited. The boron forms a "p-n junction" in the silicon. Action of light on function excites current flow.

BELL TELEPHONE LABORATORIES

IMPROVING TELEPHONE SERVICE FOR AMERICA PROVIDES CAREERS FOR CREATIVE MEN IN SCIENTIFIC AND TECHNICAL FIELDS

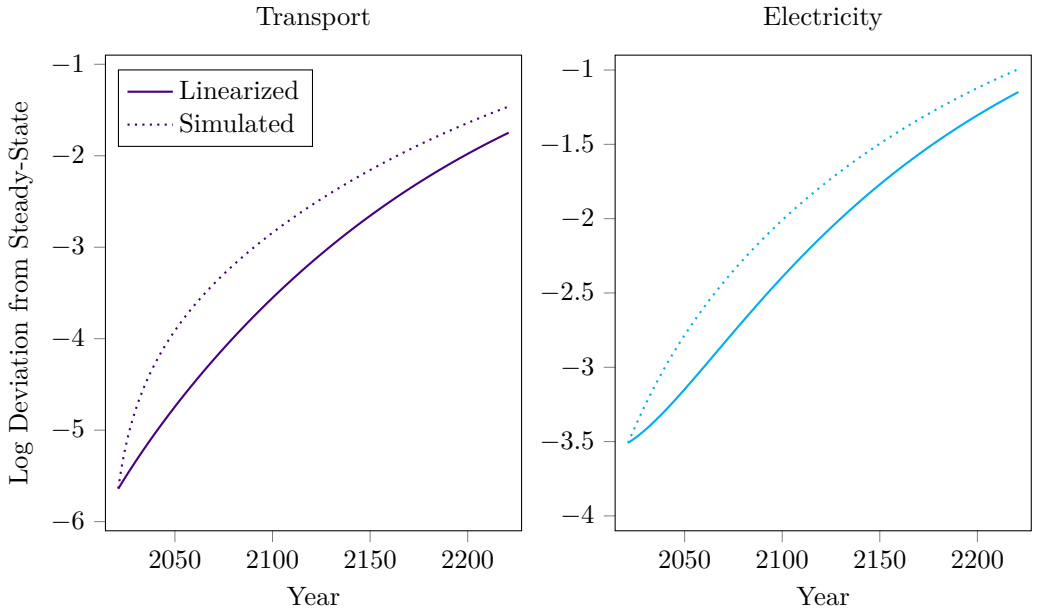
Notes: Bell Telephone Laboratories July 1954 Ad. Red box added to emphasize the description of knowledge spillovers.

Table C.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

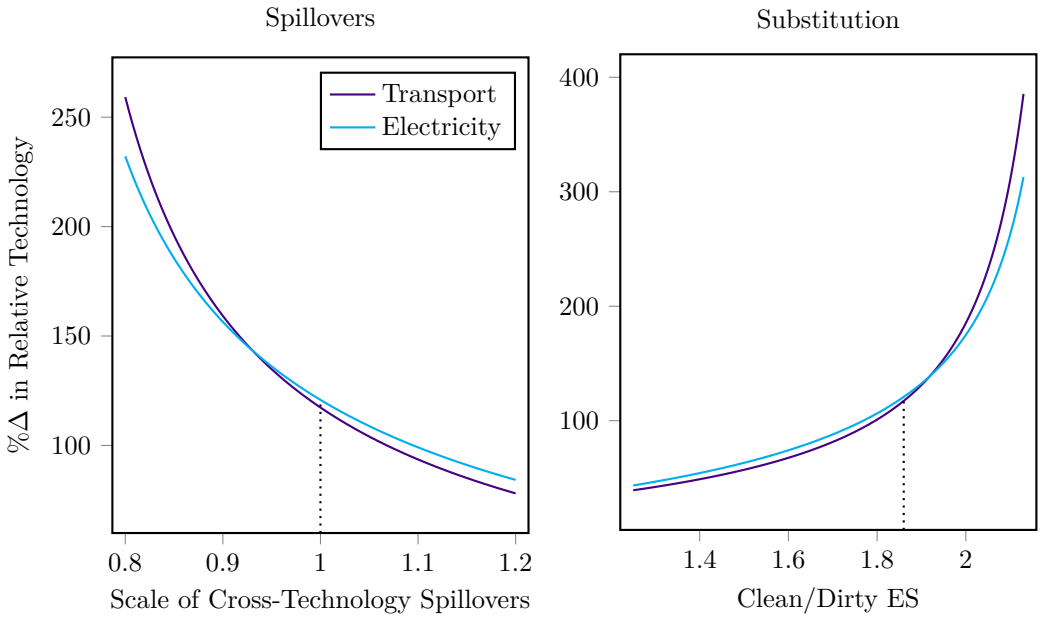
Notes: IPC codes for both clean and dirty transport patents come from Aghion et al. (2016), CPC codes for clean electric patents come directly from CPC subclass Y02E, and IPC codes for dirty electric patents come from Lanzi et al. (2011). I also include CPC codes electromobility (Y02T64-72), efficient charging and discharging systems (Y02T10/92), charging of electric vehicles (Y02T90/10-16), hydrogen technology in transport (Y02T90/40), and hybrids (Y02T10/62) for clean transport; ICE efficiency (Y02T10/12) and engine management (Y02T10/40) for dirty transport; and integration of photovoltaics in buildings (Y02B10/10) for clean electricity. General patents are those not classified as pertaining to either transport or electricity.

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



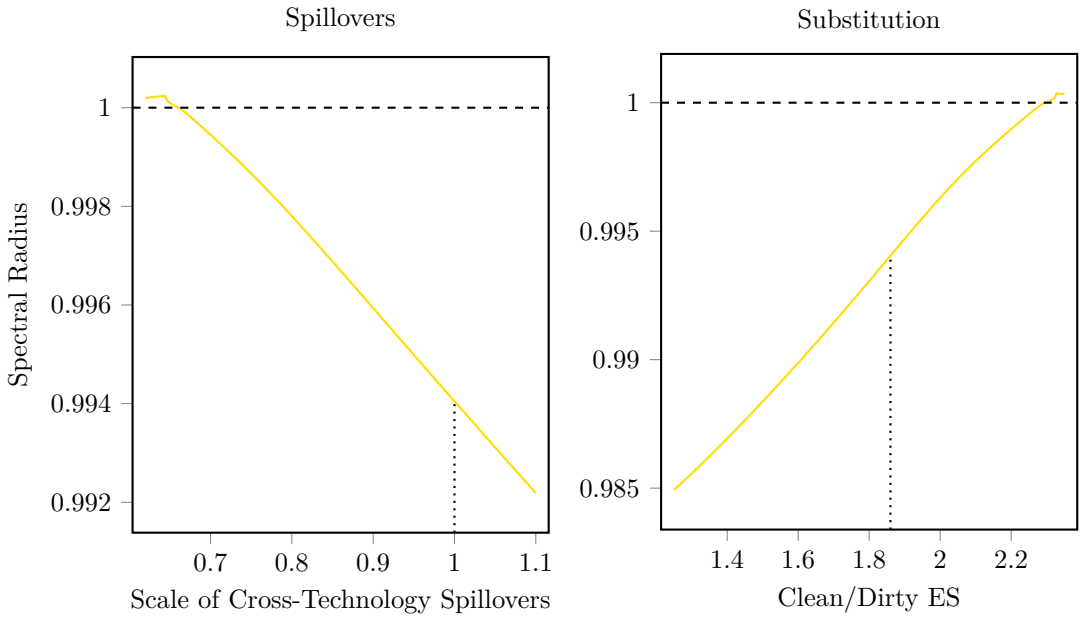
Notes: Technology's transition path following the policy reform of Section 6.1. Solid lines linearized paths using Proposition 2, while dotted lines fully simulated paths.

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



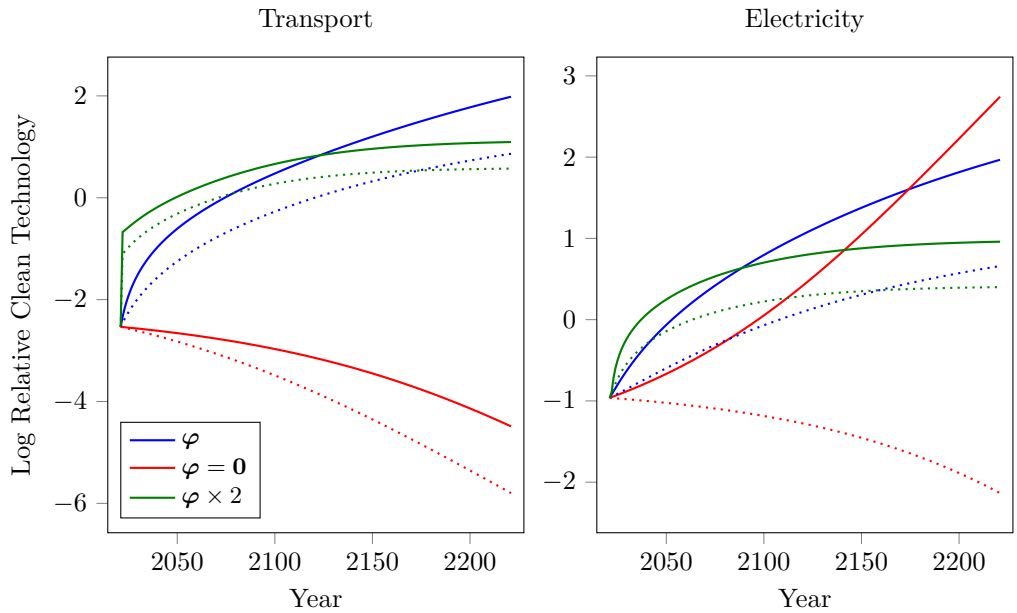
Notes: Spectral radius of the transition matrix as a function of the level of cross-technology knowledge spillovers and the elasticity of substitution between clean and dirty goods. The spectral radius passes one when spillovers reduce to 66.1% of their calibrated level or the elasticity of substitution increases to 2.29. Vertical dotted lines represent the benchmark calibration.

Table C.2: 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%	+220%	+56.9%
$\%\Delta\bar{B}_{elec}$	$+\infty$%	+216.8%	+63.8%
Emissions Intensity			
$\%\Delta\bar{\omega}$	-86.8%	-93.5%	-84.3%
Transitional Impacts			
Half-Lives of Convergence by Sector			
$t_{car}^{(1/2)}$	–	148 years	20 years
$t_{elec}^{(1/2)}$	–	150 years	25 years
Carbon Emissions by Year			
$\%\Delta\mathcal{E}_{2035}$	-73.9%	-80%	-81.6%
$\%\Delta\mathcal{E}_{2060}$	-76.4%	-83.9%	-82.7%
Degree of Increasing Returns to Innovation			
Spectral Radius			
$\max\lvert\kappa_j\rvert$	1.01	0.994	0.978

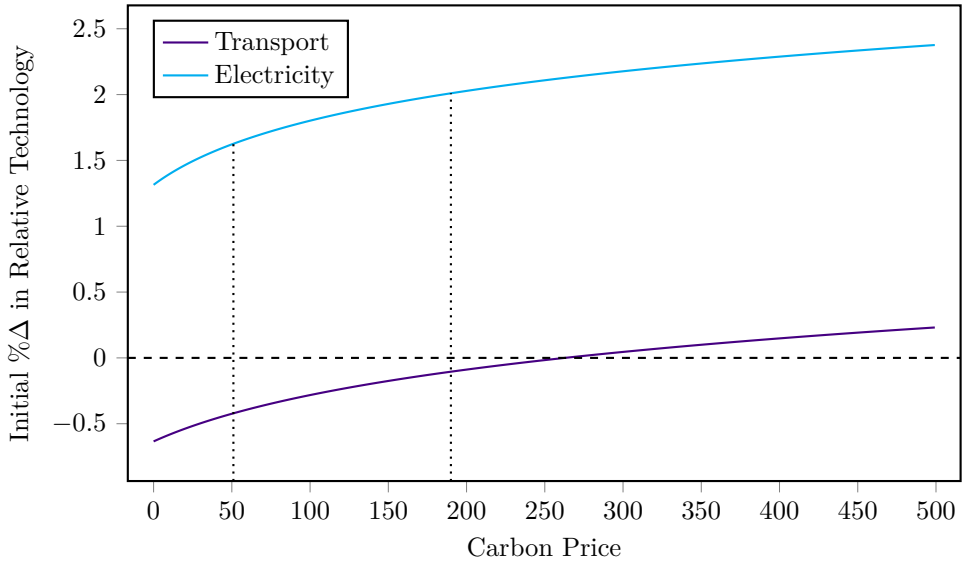
Notes: Impact of introducing a carbon price at the EPA's proposed SCC ($190) and clean innovation subsidy equivalent to a 30% tax credit ($\xi_c = 1.43$). Changes in relative technology are listed in log points. For path dependent economies, long-run impacts refer to corner, rather than interior, steady-states.

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



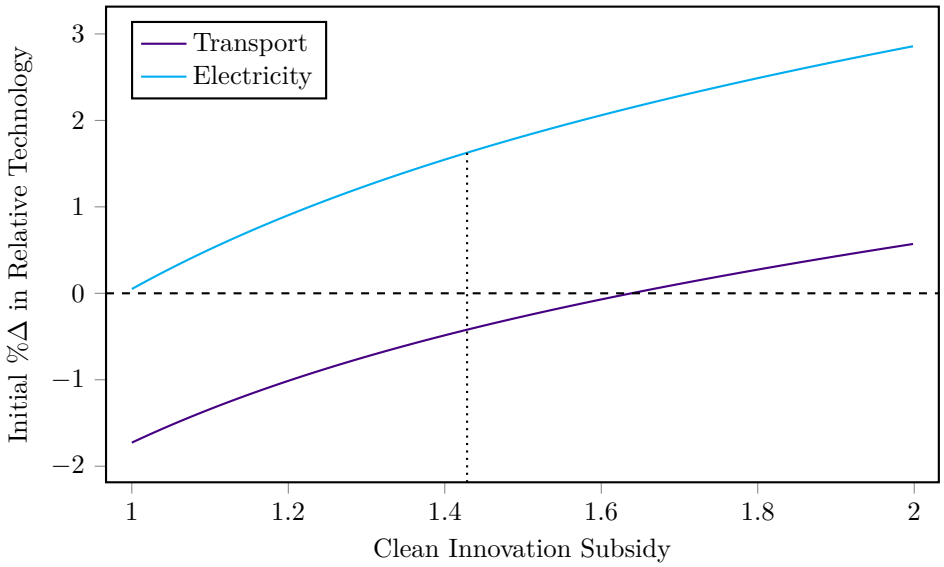
Notes: Impact of introducing a carbon price at the EPA's proposed SCC ($190) and clean innovation subsidy equivalent to a 30% tax credit ($\xi_c = 1.43$). Dotted lines indicate laissez-faire paths.

Figure C.7: Carbon Price Necessary for Clean Growth (No Spillovers)



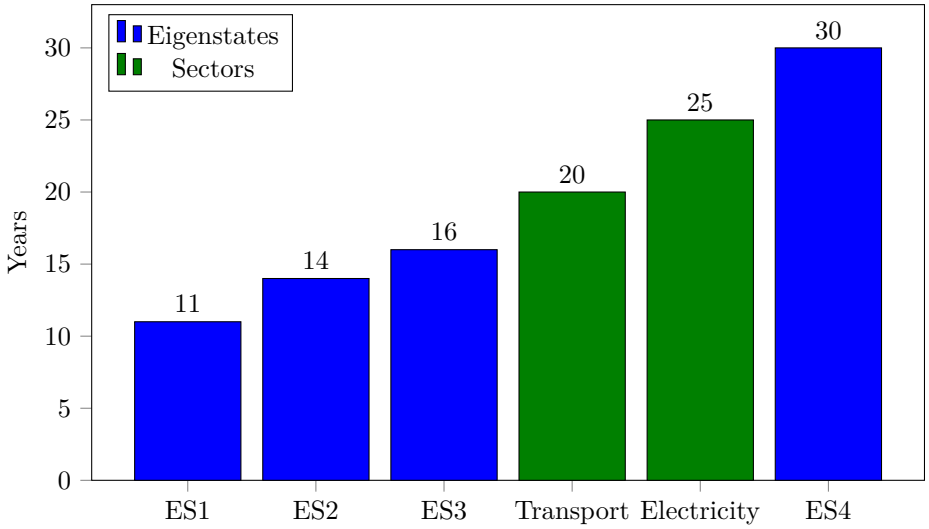
Notes: Direction of innovation as a function of the carbon price when the spillover network has been shut down. Values above zero indicate a sector is in its clean basin of attraction. Carbon prices are in addition to a clean innovation subsidy equivalent to a 30% tax credit ($\xi_c = 1.43$). Vertical 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 $263, while electricity generation switches without a carbon price.

Figure C.8: Clean Innovation Subsidy Necessary for Clean Growth (No Spillovers)



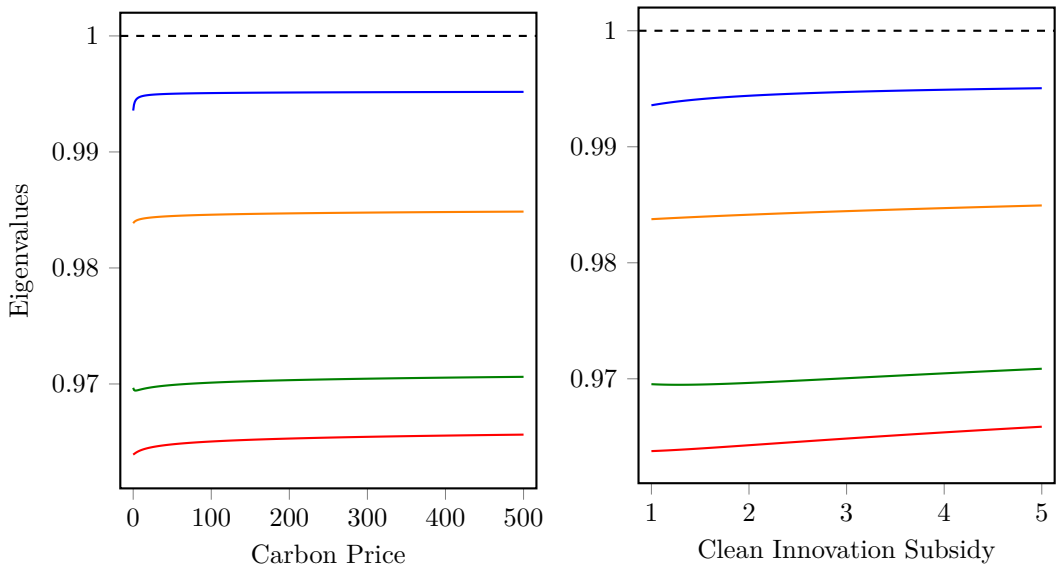
Notes: Direction of innovation as a function of the clean innovation subsidy when the spillover network has been shut down. Values above zero indicate a sector is in its clean basin of attraction. Clean innovation subsidies are in addition to a carbon price at the Biden Administration's SCC ($51). Vertical dotted line references clean innovation subsidy equivalent to a 30% tax credit ($\xi_c = 1.43$). Transportation switches to clean growth with a subsidy of 1.64, while electricity generation switches with a subsidy of 1.

Figure C.9: Half-Lives of Convergence (Double Spillovers)



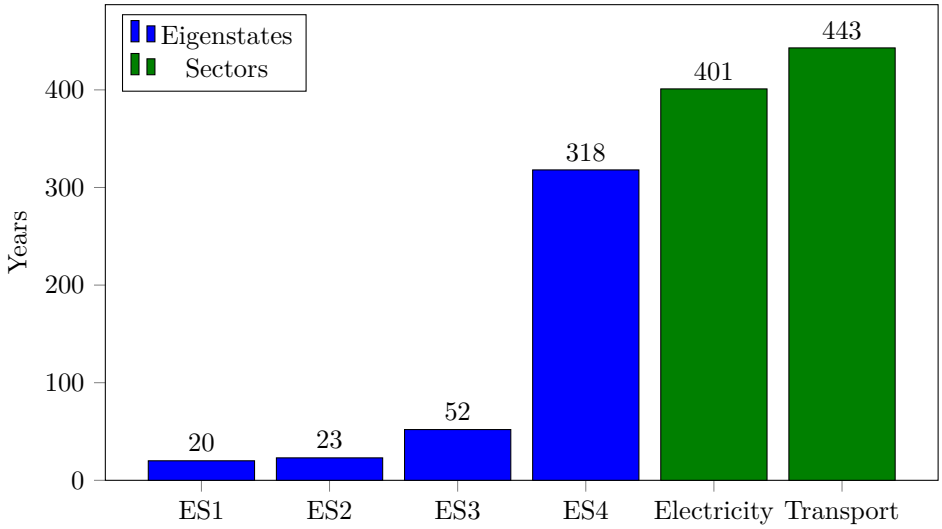
Notes: Transition speeds following the policy reform of Section 6.1. Cross-technology spillovers have been doubled.

Figure C.10: Degree of Increasing Returns Doesn't Depend on Policy Reform



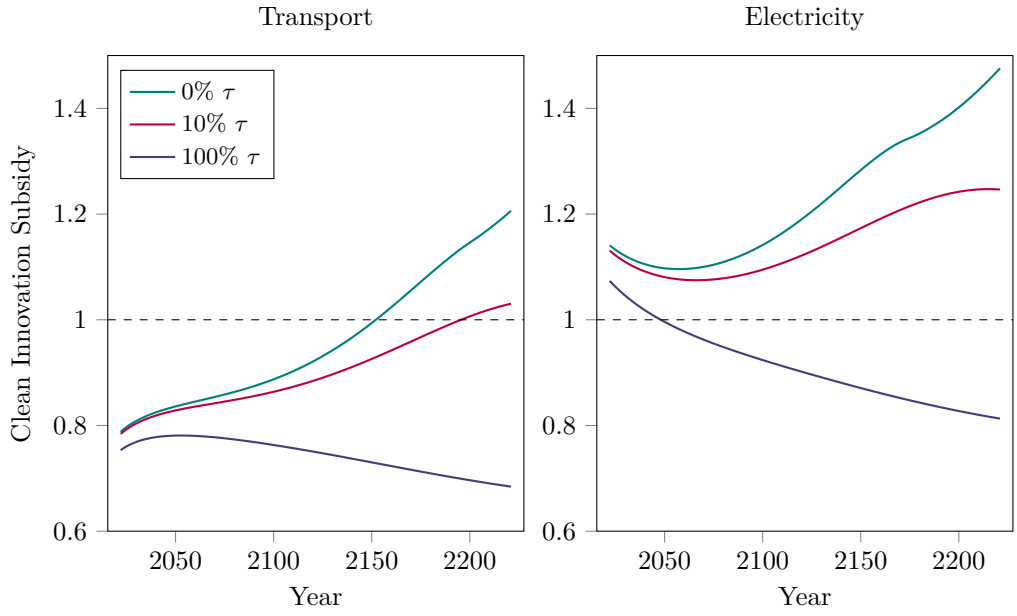
Notes: Eigenvalues of the transition matrix as a function of carbon prices and clean innovation subsidies. In each case, the other policy instrument is kept at its value from the policy reform of Section 6.1.

Figure C.11: Half-Lives of Convergence (No Within-Sector Spillovers)



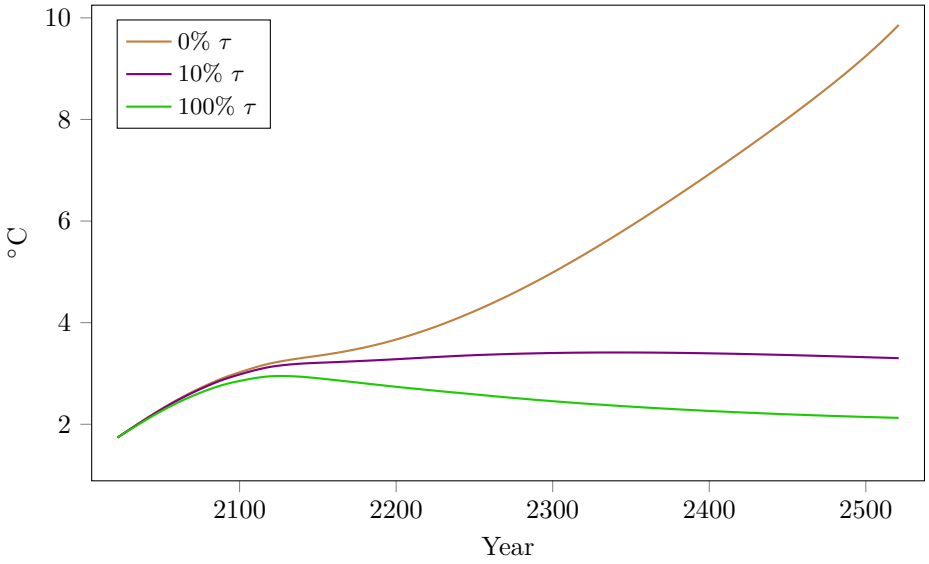
Notes: Transition speeds following the policy reform of Section 6.1. Within-sector spillovers are shut down by transferring them to the diagonal of the gross spillover network.

Figure C.12: Optimal Innovation Subsidy Path (High Discounting)



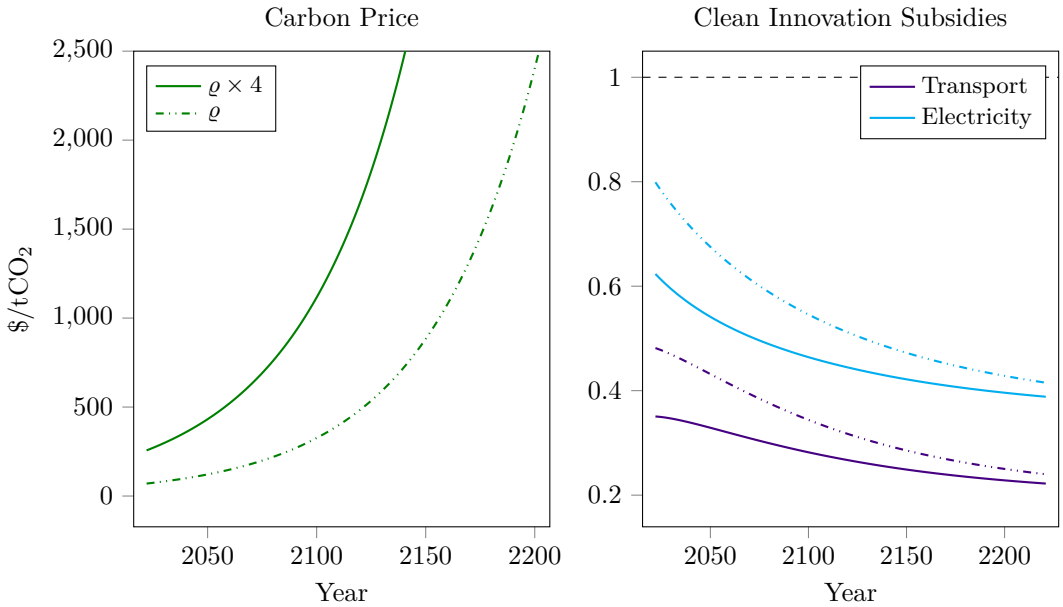
Notes: The optimal path for clean innovation subsidies with the high discount rate. The external carbon price τ is a proportion of the social cost of carbon, so 100% is the first-best. Innovation subsidies are listed as a fraction of the common innovation wedge.

Figure C.13: 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 28 for further discussion. Outside emissions come from the 2010 RICE model. Policy paths use the low discount rate. The external carbon price τ is a proportion of the social cost of carbon, so 100% is the first-best.

Figure C.14: 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 common innovation wedge.