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

Eric Donald†
University of Pittsburgh

November 14, 2025

Click here for latest version

Abstract

Transitioning to clean energy is thought to require massive policy intervention to overcome the incumbency of fossil fuels. This paper shows that cross-technology knowledge spillovers, by allowing clean technologies to achieve catchup growth, can prevent technological lock-in and the need for a "big push". I develop an endogenous growth model with clean and dirty technologies linked by a spillover network and characterize the size and speed of technological transition following a policy reform. Spillovers, together with market size effects, determine whether the economy favors technological leaders or laggards, implying a novel result: policy can generate rapid short-run change or substantial long-run redirection, but not both. I then examine the role of spillovers in optimal innovation policy, deriving subsidies under arbitrary carbon prices. Using patent data to estimate the spillover network, I apply my model to US transport and electricity generation, finding that a big push is neither necessary nor desirable.

*I am deeply grateful to my advisors Pascual Restrepo, David Lagakos, and Yuhei Miyauchi for their guidance and support. I thank Adrien Auclert, Adrien Bilal, Jean-Félix Brouillette, Dani Coen-Pirani, Dave DeJong, John Firth, Masao Fukui, Stefania Garetto, Adam Guren, Doug Hanley, David Hémous, Sungwan Hong, Chad Jones, Bob King, Kevin Kuruc, Nils Lehr, Ernest Liu, Bernardo Ribeiro, Marla Ripoll, John Sturm, Stephen Terry, and Conor Walsh as well as seminar participants at Arizona State, Bank of Canada, Boston University, Global Priorities Institute, Green Line Macro Meeting, Harvard Business School, Midwest Macro, NBER Economics of Decarbonizing Industrial Production, SEA Annual Meeting, SED Annual Meeting, SED Winter Meeting, Stanford, University of Georgia, University of Kentucky, and University of Pittsburgh for their helpful comments.

†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 largely absent, was on track to spend roughly $360 billion on subsidies for clean technology via the Inflation Reduction Act, much of which was explicitly aimed at encouraging clean innovation.

Will such policies generate the radical shift toward clean innovation that their architects envisage? The directed technical change literature takes a binary view. Fossil fuels have been the dominant energy source for centuries, and this legacy makes it difficult for immature clean technologies to compete. The resulting market size effects create strong innovation incentives for dirty technology, further entrenching fossil fuels. Because of this, policy intervention can either overcome the entrenched advantage of fossil fuels with a "big push" and bring about a complete switch to clean energy, or it can accomplish next to nothing – leaving little room for intermediate cases (Acemoglu et al., 2012, 2016).

This paper argues that *cross-technology knowledge spillovers* can undo the historic advantage of fossil fuels, making them crucial for understanding the role of policy in the transition to clean technology. Models of directed innovation typically exclude such spillovers, tacitly assuming that new technologies must start from scratch. The history of many clean technologies suggests that this is not the case. 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 an 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 carbon price and technology-specific innovation subsidies. The impact of these policies on the direction of innovation depends on the interplay of cross-technology spillovers and market size effects, as these forces determine whether the economy invests in technological leaders or laggards. Cross-technology spillovers help laggards by allowing them to build on existing knowledge, stimulating catchup growth. Market size effects help leaders by raising innovation profits for already advanced technologies. Together, these forces determine the level of increasing returns to innovation in a manner analogous to dispersion and agglomeration forces in the spatial literature (Krugman, 1991; Allen and Donaldson, 2020).

[1]Figures D.1 and D.2 provide illustrations of these two examples of cross-technology knowledge spillovers. Looking ahead, novel geothermal techniques "build on developments from the oil and gas industry, including directional drilling and fracking technologies", and cell-cultivated meat production uses bioreactors originally designed for pharmaceuticals (US Congressional Research Service Reports R48090 and R47697, respectively).

Following a policy reform, both the change in technology's long-run steady-state and its transition path can be described in terms of two matrices: one summarizing catchup growth effects from the spillover network and another summarizing market size effects from elasticities of substitution in production. These matrices are sufficient statistics in the sense that they allow one to make first-order predictions about the direction of innovation without additional information about the underlying spillover or production structure.

Combining and eigendecomposing these matrices generates a single, scalar summary statistic of increasing returns, with the novel implication that the size and speed of technological transitions are inversely related. That is, given the spillover and substitution parameters of the model, a policy reform that favors clean technologies may either generate rapid change in the short run or substantial redirection of innovation in the long run, but not both. This is because greater increasing returns – favoring leaders over laggards either through lower spillovers or greater substitutability across technologies – slows down transitions by making it more difficult to dislodge incumbents, but increases long-run redirection by allowing favored technologies to maintain greater leads over their rivals in equilibrium. If increasing returns become strong enough, they will generate technological lock-in, recovering the case where a big push is required for a slow, but total, transition.

A common narrative in the clean transition literature is that learning-by-doing has played an important role in the development of clean technologies, particularly solar (Arkolakis and Walsh, 2023). In an extension, I show that learning-by-doing cannot solve the problem of technological lock-in; in fact, the opposite is true. If productivity growth depends on *doing*, then this favors technologies that already have greater market size, and indeed, I find that learning-by-doing is isomorphic to simply having higher elasticities of substitution. Thus, learning-by-doing allows clean technologies to increase their lead in the long run, but not to initially overcome the historic advantage of fossil fuels. In another extension, I allow for semi-endogenous growth and find that the degree to which ideas get harder to find plays a similar role to cross-technology spillovers. Both of these added mechanisms cash out in terms of their effect on increasing returns to innovation – with learning-by-doing helping leaders and semi-endogeneity helping laggards – so they too have different implications for the long run and transition. Overall, avoiding the standard prediction of technological lock-in requires laggard technologies to have higher research productivity, but this same mechanism makes it more difficult to get rid of dirty technologies in the long run.

Next, given that cross-technology spillovers are an externality, they also play a central role in optimal innovation policy. To this end, 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. Moreover, substituting for the carbon price could provide an alternative rationale for aggressive clean innovation subsidies, so this second-best justification for a big push is worth taking seriously. The subsidy 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 and 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.

To apply these insights, I use the model to measure increasing returns to innovation in US transportation and electricity generation. Whether spillovers or market size effects dominate is an empirical question, so I use the citation network of US patents as a proxy for spillovers and verify that the constructed spillover network strongly predicts increased innovation by technology. I find that spillovers are large enough to prevent lock-in of dirty technologies in both sectors, placing the economy in an intermediate case where incremental reforms can redirect innovation. Yet, because clean and dirty goods are substitutes, increasing returns remain strong enough that reforms produce slow transitions with large long-run effects. I validate the model by examining the progress made in clean transportation and electricity in the 2010s, showing it matches the increasing market share of clean technology over this period. Absent cross-technology spillovers, the model makes the counterfactual prediction that dirty technologies would have become *more* entrenched in both sectors, not less.

I then use the calibrated model for two main quantitative exercises. First, I examine the impact of introducing a \$51 carbon price (the Biden Administration's lower 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 long-run relative quality of clean technology increasing by 92.3% and 98.5% for transportation and electricity generation, respectively. However, the transition to this new steady-state is slow, with half-lives of convergence of 83 and 97 years. In the absence of cross-technology spillovers, the economy is locked in a dirty equilibrium, and the policy reform is only able to switch the direction of innovation in electricity generation, where clean technology is less far behind.

Second, I simulate the optimal path of innovation subsidies for a variety of potential restrictions on the carbon price. In the absence of cross-technology spillovers, my model makes the standard recommendation of a big push: clean innovation subsidies for transport and electricity generation that start out at 347.5% and 210.4% of the common innovation wedge (the gap between private and social returns common to all technologies) 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 38.4% and 66% 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 in the present 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 spillover function elasticity of substitution.

Even in the second-best, where there is an additional justification for clean innovation subsidies, a big push is still not warranted. 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, implying that the clean subsidy only approach is not a viable path to decarbonization.

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; Lemoine, 2024). This implies destabilizing increasing returns to innovation that must be overcome with a big push of clean innovation subsidies. A notable exception is Fried (2018), who studies the impact of carbon pricing in a quantitative framework with an assumed cross-technology spillover parameter. Aghion et al. (2025) studies the adoption of clean technology along the supply chain, finding that strategic complementarity generates lock-in which can be escaped without a big push if policy targets downstream sectors.

My paper also relates to the topic of technological lock-in or path dependence. Several papers in this literature argue for the role of cross-technology knowledge spillovers in stabilizing technology's steady-state (Acemoglu, 2002, 2023; Fried, 2018; Hart, 2019). My formulas generalize these results by providing a test for technological lock-in 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), and Känzig (2023). Examples of theoretical contributions from the skill-biased technological change literature include Acemoglu (1998) and Acemoglu and Restrepo (2018, 2022). 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) in its normative (though not descriptive) implications.[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 of Kleinman et al. (2023). I take a complementary approach to describe transition dynamics in the context of directed innovation.

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 climate policy. Section 7 concludes.

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

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

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. Innovation 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 and 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

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

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

The good specific to technology j is produced according to

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

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

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

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

Dirty inputs can be thought of as fossil fuels whose use generates carbon emissions. Carbon intensities ω_j may differ across dirty technologies due to the different types of fossil fuels they use. 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:

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

There is a fixed endowment of labor that is supplied inelastically, so we have

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

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, where 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, so if intermediate ι for technology j receives an innovation, we have that $a_{j\iota t}$ increases to $\gamma a_{j\iota t}$, where $\gamma > 1$ is the step size of innovation. Define z_{jt} as the mass of intermediates for technology j that receives an innovation. Plugging this into Equation (5), we have that technology evolves according to

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

This formulation implies that technology-specific growth rates follow

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

[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

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

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. This ensures balanced growth in the presence of a fixed supply of scientists. The parameter $\eta \in (0, 1)$ determines the degree of diminishing returns in research, and χ_j is a research productivity shifter specific to each technology.

The matrix of spillover elasticities defines the network of knowledge spillovers.

Definition 1 (Spillover Network $\boldsymbol{\varphi}$)**.** *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. The spillover network is allowed to vary over time, so my only substantive assumption is 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. 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 $\boldsymbol{\varphi}_t = \tilde{\boldsymbol{\varphi}}_t - \mathbf{I}$.

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 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 spillovers play a critical role in shaping both technological transitions and optimal policy.

Finally, there is a fixed endowment of scientists supplied inelastically, so we have

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

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. In a setting where scientists have an outside option such as 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.

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.

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

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

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

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

The producers of technology-specific goods solve

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

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

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

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

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

Intermediate producers internalize their demand curves, so they solve

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

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

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

That is, intermediate prices equal the subsidy-inclusive break-even price of the competitive fringe. Average profit is then equal to the profit of each intermediate producer, so average profit follows

$$\Pi_{jt} = \Upsilon\frac{\gamma-1}{\gamma}(1-\alpha)p_{jt}Y_{jt} \propto S_{jt}\mathcal{Y}_t, \tag{21}$$

where S_{jt} is the income share of technology j. That is, the rents associated with technology j scale with market size. This is a standard result in the directed technological change literature (Acemoglu, 2002; Acemoglu et al., 2012).

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.

The research firm solves

$$\max_{\{s_{jt}\}} \sum_j \xi_{jt} z_{jt} \Pi_{jt} - w_{st}\sum_j s_{jt} \tag{22}$$

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}. The research optimality condition yields

$$\Big(\frac{s_{jt}}{s_{Jt}}\Big)^{1-\eta} = \Big(\frac{\chi_j}{\chi_J}\Big)\Big(\frac{\xi_{jt}}{\xi_{Jt}}\Big)\Big(\frac{\phi_{jt}}{\phi_{Jt}}\Big)\Big(\frac{\Pi_{jt}}{\Pi_{Jt}}\Big), \tag{23}$$

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:

$$c_t = w_{\ell t} L + w_{st} \mathcal{S} + \Pi_t - D_t \tag{24}$$

$$D_t + \tau_t \mathcal{E}_t = \sum_j (\xi_{jt} - 1) z_{jt} \Pi_{jt} + \sum_j \int_0^1 (\Upsilon - 1) p_{j\iota t} y_{j\iota t} d\iota, \tag{25}$$

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* (21) *and solve the profit maximization problem* (22)*; (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* (24) *and* (25) *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. The example economy will nest the environment of Acemoglu et al. (2012) so as to make the comparison of market size and spillover effects explicit. I will assume a multi-sector version of this setup 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.

Final production combines a clean and dirty good according to

$$\mathcal{Y}_t = \Omega_t\Big(Y_{ct}^{\frac{\sigma-1}{\sigma}} + Y_{dt}^{\frac{\sigma-1}{\sigma}}\Big)^{\frac{\sigma}{\sigma-1}}, \tag{26}$$

where Y_{ct} and Y_{dt} are the clean and dirty forms of production. That is, output is a substitutes CES over clean and dirty production with elasticity of substitution $\sigma > 1$.

On the innovation side of the economy, I will assume that both technologies share the same research productivity shifter χ, so the production of innovation follows

$$z_{jt} = \chi s_{jt}^{\eta}\phi_{jt}. \tag{27}$$

The spillover functions follow

$$\phi_{it} = \frac{\prod_j A_{jt-1}^{\tilde{\varphi}_{ij}}}{A_{it-1}}, \tag{28}$$

which implies 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 $\tilde{\boldsymbol{\varphi}}$. The denominator represents the recipient's own knowledge stock.

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 $\{d\tau, \{d\xi_j\}\}$,[5] and I will focus attention on a balanced growth steady-state, though the reader should note that my analysis does not require technology to be in a previous steady-state at the time of the policy reform.

Definition 3 (Balanced Growth Steady-State)**.** *An equilibrium where every technology grows at rate g.*

I will first consider an example economy to express the core insights, then describe in the remainder of the section how those insights generalize. The key questions we will ask are: i) what are the conditions under which the economy is locked-in with dirty technology? And ii) if this is not the case, how do incremental policy reforms shift the direction of innovation? As we will see, the answer to both questions will depend on the interplay of spillover and market size effects and the resulting degree of increasing returns to innovation. Detailed proofs of all propositions can be found in Appendix A.

[5] 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.

3.1 Lock-In in an Example Economy

I will start by illustrating the main arguments of the paper using the example model of Section 2.5. The intention is to express the key insights in a simple setting before explaining how these insights carry over to a more general environment.

To start, we will define total cross-technology spillovers in the economy:

$$\Phi = \varphi_{cd} + \varphi_{dc}. \tag{29}$$

In words, the sum of off-diagonal elements of the spillover network provides a measure for how much the economy's two technologies rely on each other for growth. This object will allow us to determine when dirty technology is locked-in and a big push is required for the energy transition.

It is helpful to think in terms of relative technology $\bar{A}_t \equiv A_{ct}/A_{dt}$ because relative technology achieves a fixed point along a balanced growth path. Then, under a mild regularity condition,[6] technological lock-in occurs whenever

$$\Phi < \eta(1-\alpha)(\sigma-1). \tag{30}$$

From this, we can see that lock-in depends on a horse race between cross-technology spillovers and market size effects. If the two technologies are highly substitutable relative to catchup-inducing spillovers, then the less advanced clean technology will not be profitable enough to attract the innovation effort necessary to overcome their disadvantage. Instead, the more advanced dirty 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 dirty technology. This is the lock-in scenario considered by Acemoglu et al. (2012), and it can only be overcome by a big push to jolt the economy out of its dirty basin of attraction.

But what if that is not the case? How should we handle intermediate cases where incremental policy reforms, such as the introduction of a modest carbon price, change the direction of innovation? To answer these questions, we will consider the steady-state for relative technology:

$$\ln(\bar{A}_{ss}) = \eta \frac{\ln(\Xi) - \alpha(\sigma-1)\ln(\mathcal{R})}{\Phi - \eta(1-\alpha)(\sigma-1)}, \tag{31}$$

where $\Xi \equiv \xi_c/\xi_d$ is the relative innovation subsidy and $\mathcal{R} \equiv r_c/(r_d + \omega_d \tau)$ is the relative input cost. In a lock-in scenario, this demarcates the basin of attraction for clean and dirty growth, so a big push is a policy reform that is large enough to move the economy into its clean basin of attraction.[7] But if the lock-in condition is not met, the economy converges to an interior steady-state. In that world, incremental policy reforms $d\ln(\Xi)$ and $d\ln(\mathcal{R})$ shift the economy's long-run steady-state

[6] 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.

[7] Formally, policy reforms that favor clean technology will shrink the dirty basin of attraction, expanding the set of initial conditions that lead to clean growth.

and induce a technological transition.

To examine the long-run effect of policy, we can differentiate Equation (31) to derive

$$d\ln(\bar{A}_{ss}) = \eta\mathcal{M}\big(d\ln(\Xi) - \alpha(\sigma-1)d\ln(\mathcal{R})\big), \tag{32}$$

where the amplification term $\mathcal{M}$ follows

$$\mathcal{M} = \frac{1}{\Phi - \eta(1-\alpha)(\sigma-1)}. \tag{33}$$

The amplification term $\mathcal{M}$ describes the full general equilibrium impact of a policy reform, which operates through spillover and market size effects. 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 due to cost-disease: complementarity lowers the market size of more advanced technologies.

Next, by linearizing around the steady-state, we have that the transition follows

$$\ln(\bar{A}_{t+1}) - \ln(\bar{A}_{ss}) \approx \mathcal{J}\big(\ln(\bar{A}_t) - \ln(\bar{A}_{ss})\big), \tag{34}$$

where the transition term $\mathcal{J}$ follows

$$\mathcal{J} = \frac{(1-\eta) - g\Phi}{(1-\eta) - g\eta(1-\alpha)(\sigma-1)}. \tag{35}$$

From this, we can see that the speed with which policy can influence the direction of innovation depends on the geometric decay of the transition term $\mathcal{J}$. If $\mathcal{J}$ is close to zero, we will get very rapid convergence, and if $\mathcal{J}$ is close to one from below, we will get slow convergence. Finally, if $\mathcal{J}$ is above one, we have the case of technological lock-in described above because doubling the deviation of relative technology from its steady-state more than doubles it in the following period. In units of time, technology's half-life follows

$$t^{(1/2)} = \left\lceil \frac{\ln(1/2)}{\ln(\mathcal{J})} \right\rceil. \tag{36}$$

The transition term $\mathcal{J}$ again describes the role of spillover and market size effects, 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.[8] Conversely, greater substitutability slows down the transition because it allows more advanced technologies to maintain a larger market size for longer. This makes it more difficult to dislodge incumbents.

[8] The effect of cross-technology spillovers on $\mathcal{J}$ described here ignores changes in g, but in Appendix A.6 I show that changes via g are $\mathcal{O}(1-\mathcal{J})$ and therefore empirically small.

Again, the opposite will be true for complements.

We should therefore interpret $\mathcal{J}$ as a measure of increasing returns to innovation. It aggregates the two forces of spillover and market size effects into a single, scalar measure of the degree to which the economy favors technological leaders or laggards. This gives us the novel result that, depending on the spillover and substitutability parameters of the economy, policy reforms will either have a large impact in the long-run or a rapid impact in the short-run, but not both. This inverse relationship can be formalized as follows:

$$\mathcal{M} = \frac{g(1-\mathcal{J})^{-1}}{(1-\eta) - g\eta(1-\alpha)(\sigma-1)}. \tag{37}$$

With higher increasing returns $\mathcal{J}$, it becomes more difficult to dislodge incumbent technologies and technology becomes slow to react to changes in the policy environment. But at the same time, the impact of favoring some technologies with policy is amplified in the long-run through $\mathcal{M}$ because greater increasing returns allows technologies favored by policy to gain a greater lead over their rivals over time.

3.2 Sufficient Statistic Matrices

In the more general environment, the two forces of spillover and market size effects can no longer be summarized by global, scalar parameters. Instead, this section will define two sufficient statistic matrices that summarize the two main forces of cross-technology knowledge spillovers and market size effects. 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.

First, as an analog to the sum of cross-technology spillovers Φ in the example model, consider the spillover matrix:

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

$$\Phi_{ijt} \equiv \frac{\partial \ln(\phi_{Jt}/\phi_{it})}{\partial \ln(A_{jt-1}/A_{Jt-1})}. \tag{38}$$

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

$$\Phi_{iit} = \sum_{j\neq i} (\varphi_{ijt} - \varphi_{Jjt}), \tag{39}$$

so the diagonal elements of $\boldsymbol{\Phi}_t$ will be large when technologies receive high cross-technology spillovers, relative to the spillovers received by the base technology. The diagonal is thus infor-

[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.

mative 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.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, as an analog to the elasticity of substitution σ in the example model, consider the substitution matrix:

Definition 5 (Substitution Matrix $\mathbf{\Sigma}$)**.** *A $J-1 \times J-1$ matrix with elements*

$$\Sigma_{ijt} \equiv \frac{\partial \ln \left(Y_{it}/Y_{Jt}\right)}{\partial \ln \left(p_{Jt}/p_{jt}\right)}. \tag{40}$$

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, and therefore, how market size responds to changes in prices.[10]

3.3 Steady-State Impact of Policy

Using these sufficient statistic matrices, we can again ask how the balanced growth steady-state changes in response to a policy reform when there is no technological lock-in. However, this analysis will allow us to build to a more general lock-in condition which can be taken to the data as a test for the relative strength of spillover and market size effects. I will again describe the steady-state in terms of relative technology, given by

$$\bar{A}_{jt} \equiv \frac{A_{jt}}{A_{Jt}}, \tag{41}$$

where the base technology is J. In Appendix A.4, I provide conditions for the existence and uniqueness of the economy's steady-state, and in Appendix A.9, I derive a closed-form solution for the steady-state growth rate g.

[10]I show in Appendix A.3 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.

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*

$$d\ln(\bar{A}_{ss}) = \eta\boldsymbol{\mathcal{M}}\big[d\ln(\Xi) - \alpha(\boldsymbol{\Sigma} - \boldsymbol{I})d\ln(\mathcal{R})\big], \tag{42}$$

where $d\ln(\Xi_j) \equiv d\ln(\xi_j/\xi_J)$ 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

$$\boldsymbol{\mathcal{M}} = \big[\boldsymbol{\Phi} - \eta(1-\alpha)(\boldsymbol{\Sigma} - \boldsymbol{I})\big]^{-1}. \tag{43}$$

Proposition 1 generalizes Equation (32) in terms of an amplification matrix $\boldsymbol{\mathcal{M}}$, which describes the role of spillover and market size effects via the two sufficient statistic matrices. In spite of the generality, the interpretation is the same.[11] Greater spillovers reduce the long-run impact of policy by preventing laggards from falling too far behind, while greater substitutability increases the long-run impact of policy by allowing leaders to increase their market size.

The primary difference in this more general environment comes from tracking the full network impact of policy reforms through the spillover network and potentially complicated substitution patterns in production. 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.[12]

3.4 Technology's Transition Path

Having examined how policy affects the long-run steady-state, I can now characterize technology's transition, which I will describe in terms of the log deviation of relative technology from steady-state

$$\bar{\mathcal{A}}_t \equiv \ln(\bar{A}_t) - \ln(\bar{A}_{ss}). \tag{44}$$

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 the linear process*

$$\bar{\mathcal{A}}_t \approx \boldsymbol{\mathcal{J}}\bar{\mathcal{A}}_{t-1}, \tag{45}$$

[11] I will assume that the amplification matrix $\boldsymbol{\mathcal{M}}$ does not flip the sign of policy impacts which, as shown by Corollaries 1 and 2, is consistent with assuming away technological lock-in.

[12] Formally, we have that

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

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

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

$$\boldsymbol{\mathcal{J}} = \big[(1-\eta)\boldsymbol{I} - g\eta(1-\alpha)\big(\boldsymbol{\Sigma} - \boldsymbol{I}\big)\big]^{-1}\big[(1-\eta)\boldsymbol{I} - g\boldsymbol{\Phi}\big]. \tag{46}$$

Proposition 2 generalizes Equation (34) using the transition matrix $\boldsymbol{\mathcal{J}}$, and we can again see that the speed of transition depends on how quickly the transition matrix $\boldsymbol{\mathcal{J}}$ shrinks technology's log deviation from steady-state. 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.

Again, this generalized formula has the same interpretation as in the example economy. The transition matrix $\boldsymbol{\mathcal{J}}$ describes spillover and market size effects using the same two sufficient statistic matrices as the amplification matrix, except now their influence is reversed. Catchup growth from cross-technology spillovers speeds up the transition, while stronger market size effects allow incumbent technologies to hold onto their lead for longer.

The transition matrix $\boldsymbol{\mathcal{J}}$ also allows us to account for network effects from both cross-technology spillovers and market size. The rightmost matrix of Equation (46) – 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.[13]

3.5 Degree of Increasing Returns to Innovation

We have seen that Equations (32) and (34) of the example model generalize to a matrix form where spillover and market size effects are summarized by sufficient statistic matrices. This generalization comes at some cost to interpretability as matrices are generally harder to interpret than scalars, but we can recover scalar interpretation by eigendecomposing the transition matrix. In particular, the spectral radius of the transition matrix provides a scalar summary statistic of increasing returns to innovation. This will allow us to derive a general condition for technological lock-in and again establish 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. Eigendecomposing the transition matrix, we have

$$\boldsymbol{\mathcal{J}} = \boldsymbol{Q}\boldsymbol{D}(\kappa)\boldsymbol{Q}^{-1}, \tag{47}$$

[13] Formally, we have that

$$\big[(1-\eta)\mathbf{I} - g\eta(1-\alpha)\big(\mathbf{\Sigma} - \mathbf{I}\big)\big]^{-1} \propto \mathbf{I} + \Big(\frac{g\eta(1-\alpha)}{1-\eta}(\mathbf{\Sigma}-\mathbf{I})\Big) + \Big(\frac{g\eta(1-\alpha)}{1-\eta}(\mathbf{\Sigma}-\mathbf{I})\Big)^2 + \dots,$$

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

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. Next, iterating forward from technology's initial condition, we have

$$\bar{\mathcal{A}}_t \approx \boldsymbol{Q}\boldsymbol{D}(\kappa)^t \boldsymbol{Q}^{-1} \bar{\mathcal{A}}_0. \tag{48}$$

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 study the importance of each eigenvalue, I use spectral analysis to provide an interpretation of the corresponding eigenvectors in Appendix A.8.

A key advantage of this eigendecomposition is that it describes technology's transition in terms of scalar objects, eigenvalues, that have a clear economic interpretation: they summarize the degree of increasing returns to innovation. Larger eigenvalues imply that it will take more time for advanced incumbent technologies to lose their advantage when the policy environment changes. In particular, the spectral radius, or largest eigenvalue, will have an outsized influence as time progresses, so it can provide us with a single summary statistic of increasing returns to innovation.

Put differently, the two forces of spillovers and substitutability are summarized with high-dimensional matrices, but we can aggregate them into a single, scalar summary statistic using the spectral radius. This measure of increasing returns describes the speed of convergence for systems with stable, interior steady-states, but it also tells us when increasing returns are destabilizing to the point of generating technological lock-in. Technological convergence slows down as the eigenvalues approach one, and once an eigenvalue goes outside of the unit circle, the steady-state becomes unstable and the path of technology will depend discontinuously on its initial state. The following corollary states this point formally.

Corollary 1. *Technological lock-in occurs when the spectral radius of the transition matrix* $\max|\kappa_j|$ *exceeds one.*

If an eigenvalue exceeds one, technology will diverge in the direction of this unstable eigenvector. In that case, the incentives to innovate are so biased towards leading technologies that their lead will become further reinforced over time as they receive greater and greater research effort. This provides a general analog to Inequality (30) of the example model.

My next corollary shows that these eigenvalues matter for the long run as well by establishing a relationship between the amplification and transition matrix that is analogous to Equation (37) of the example model.

Corollary 2. *The amplification matrix follows*

$$\boldsymbol{\mathcal{M}} = g\boldsymbol{Q}\boldsymbol{D}(1-\kappa)^{-1}\boldsymbol{Q}^{-1}\left[(1-\eta)\boldsymbol{I} - g\eta(1-\alpha)(\boldsymbol{\Sigma} - \boldsymbol{I})\right]^{-1}. \tag{49}$$

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, in general, spillover and market size effects have opposite effects in the long run and transition. In Appendix A.8, I show that this inverse relationship between transition speeds and long run effects holds in a large class of dynamic models.

3.6 Extensions

This section considers various extensions to the baseline model and shows how the logic of the above results carries over to these other cases. I reserve discussion of how these extensions affect optimal policy for Section 4.

Semi-Endogenous Growth First, I allow for semi-endogenous growth, where ideas become harder to find (Bloom et al., 2020; Jones, 2022). In my baseline 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. Put differently, assuming that spillover functions are homogeneous of degree zero makes growth endogenous, rather than semi-endogenous.

To consider a semi-endogenous growth model, we can instead assume that spillover functions are homogeneous of degree $-\zeta < 0$. This implies that overall research productivity will decline with the level of technological advance, so ζ determines the degree of fishing out in a semi-endogenous growth model (Jones, 2022). The only other change that must be made to the model is that the stock of scientist $\mathcal{S}_t$ must grow at rate n. As expected, steady-state technology growth requires population growth as $g \approx \frac{\eta}{\zeta} n$.

As I show in Appendix B.1, the formulas of Propositions 1 and 2 apply exactly the same in this case with the only difference being that the spillover matrix $\boldsymbol{\Phi}$ has ζ added along its diagonal. Intuitively, fishing out works in a similar manner as cross-technology spillovers. The key question is whether the economy invests in technological leaders or laggards. Greater substitutability in production pushes innovation investment toward the leading technologies by increasing their market size. Cross-technology spillovers push innovation investment toward the laggard technologies by increasing their relative research productivity, and fishing out has a similar effect because it makes it relatively easier to make productivity gains in backward technologies.

Learning-by-Doing Next, I allow for learning-by-doing where productivity growth for a technology depends on both research effort and production. A common view is that learning-by-doing has played an important role in the declining cost of renewables (Arkolakis and Walsh, 2023), but as I show in Appendix B.2, learning-by-doing cannot solve the problem of technological lock-in.

Formally, one can account for learning-by-doing in my framework with a Leontief inverse that reflects how relative learning changes in response to changes in relative quantities. The Leontief inverse multiplies the substitution matrix, so under the natural case where learning depends on a technology's own relative demand, learning-by-doing is isomorphic to larger elasticities of substitution, and therefore, more powerful market size effects.

Again, the intuition of this result comes from asking whether the economy favors technology leaders or laggards. Like market size effects, learning-by-doing helps leaders because *doing* requires market demand, so those with a larger market size receive a productivity boost from both direct learning and indirect incentives for R&D spending. This clarifies the nuanced sense in which learning-by-doing is "important" for the clean transition. As my results emphasize, market size effects, and therefore also learning-by-doing, make it more difficult to avoid lock-in of dirty technology but allow switches in the direction of innovation to have greater impact in the long run. Put differently, learning-by-doing cannot get clean technologies off the ground, but it can help them to accelerate past dirty technologies once they gain steam.

General Input-Output Structure Finally, I allow for a general input-output structure where, instead of always being denominated in the final good, input costs r_j can be arbitrary functions of the various technology-specific goods in the economy. As I show in Appendix B.3, the only change necessary for this case is that one must account for the fact that changes in technology can also affect input costs through the IO structure. This is captured with a Leontief inverse that reflects the network effect of such changes.

4 Optimal Policy

Up until now, all of the results have been descriptive, examining how cross-technology spillovers influence the impact of an arbitrary policy reform on the direction of innovation. Knowledge spillovers are also a canonical market failure, so I will now turn to the normative and characterize how cross-technology spillovers determine optimal climate innovation policy. To do so, 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.

Cross-technology spillovers intertwine the recursive benefit of innovation across technologies because innovating in one technology today makes it easier to innovate in another technology tomorrow. Because of this, cross-technology spillovers imply that permanent intervention is desirable, rather than the one-off, temporary intervention of a big push.

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 desirable restriction because pricing carbon has proved to be politically difficult and substituting for the carbon price is viewed as another motivation for aggressively subsidizing clean technology.

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*

$$\max_{\{c_t,\{A_{jt},s_{jt}\}\}} \sum_{t\geq 0} \frac{1}{(1+\rho)^t} u(c_t) \;\; s.t. \;\; (7),\; (8),\; (12), \tag{50}$$

$$\{\Lambda_{jt}, \{\ell_{j\iota t}\}\} = \text{argmax}\left[\mathcal{Y}_t - \sum_j r_j \Lambda_{jt} - \tau_t \mathcal{E}_t\right] \;\; s.t. \;\; L = \sum_j \int_0^1 \ell_{j\iota t} d\iota.$$

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 3. *Optimal innovation subsidies follow the recursion*

$$\xi_{jt}\Pi_{jt} = (1-\alpha)S_{jt}\mathcal{Y}_t - \sum_{\hat{t}\geq t} \prod_{\hat{s}=1}^{\hat{t}-t} \frac{1}{R_{t+\hat{s}}} (SCC_{\hat{t}} - \tau_{\hat{t}}) \mathcal{E}_{\hat{t}} \frac{\partial \ln(\mathcal{E}_{\hat{t}})}{\partial \ln(A_{jt})} + \frac{1}{R_{t+1}}\left[\xi_{jt+1}\Pi_{jt+1} + \sum_i \xi_{it+1}\Pi_{it+1} g_{it+1}\varphi_{ijt+1}\right], \tag{51}$$

where the social cost of carbon follows

$$SCC_t = -\sum_{\hat{t}\geq t} \prod_{\hat{s}=1}^{\hat{t}-t} \frac{1}{R_{t+\hat{s}}} \mathcal{Y}_{\hat{t}} \frac{\partial \ln(\Omega_{\hat{t}})}{\partial \mathcal{E}_t}. \tag{52}$$

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

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 (51) 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.[14] Equation (51) 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, the spillover network connects the social value of innovation across technologies because part of the benefit from innovating in one technology today is enabling more innovation in other technologies tomorrow. I show in Appendix B.2 that a similar formula holds in the case of learning-by-doing, the only difference being the inclusion of a Leontief inverse that accounts for the network effect of learning.[15]

A visible implication of Equation (51) is that efficient carbon prices allow the Planner to separately correct the economy's two externalities and achieve the first-best.[16] 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 (52), 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.

[14] 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.

[15] I show in Appendices B.3 and B.4 how Proposition 3 continues to hold with a general input-output structure or endogenous extraction costs for fossil fuels.

[16] I describe the first-best planning problem in detail in Appendix A.12.

4.2 Steady-State Policy

To understand the long-run determinants of optimal innovation subsidies, I will now consider steady-state innovation policy, where I recover the recent results of Liu and Ma (2021) as a special case. As I will show, the presence of cross-technology spillovers determines whether innovation policy should be temporary or permanent.

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 steady-state difference between the private and social benefit of innovation that is common to 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 surplus in perpetuity. In a model with an endogenous supply of scientists (e.g. where workers could choose between science or production work), 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*

$$(1-\tilde{R}^{-1})(\hat{\xi}'-[S'-\mathcal{T}'])-g\tilde{R}^{-1}\hat{\xi}'\boldsymbol{\varphi}=\vec{0}'. \tag{53}$$

As in Liu and Ma (2021), Equation (53) 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 $\tilde{R}^{-1}$.[17]

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 optimal 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

[17] Steady-state policy can also be written in terms of a Leontief inverse, which I describe in Appendix A.11. One should also note that greater spillover creation per se does not lead to a greater subsidy; what matters is spillover creation relative to market size (net of externalities). For example, in a knife-edge steady-state without pollution where spillover centrality happens to equal income shares, the laissez-faire composition of research effort is optimal.

focus on the spillover benefits of innovation.[18] They set innovation incentives according to

$$\hat{\xi}'\varphi = \hat{\xi}'(\tilde{\varphi} - \mathbf{I}) = \vec{0}', \tag{54}$$

where $\tilde{\varphi}$ is the gross spillover network discussed in Section 2.2. That is, $\hat{\xi}$ solves $\hat{\xi}'\tilde{\varphi} = \hat{\xi}'$, 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. I show in Appendix B.1 that this logic carries over to the case of semi-endogenous growth with the only difference being that the Planner maximizes the level, rather than growth, effect of more productive researchers.

Therefore, the permanence of directional innovation policy depends on the presence of cross-technology spillovers. Without such spillovers (and an efficient carbon price), the centrality term goes away and policy converges to a uniform subsidy. This is why the big push only requires temporary intervention. But in the presence of cross-technology spillovers, the Planner must consider that ideas are used to produce new ideas as well as physical goods, so innovation policy converges to a permanent intervention focused on centrality in the spillover network.

5 Calibration

In this section, I describe my calibration of the model's structural parameters. I use a multi-sector version of the model outlined in Section 2.5, the details of which are in Appendix A.2. The main difference is that we can now have multiple sectors, indexed by θ, that aggregate to final output via a CES with elasticity of substitution $\lambda < 1$ and CES shares ν_θ. My application is transport and electricity generation in the US. These are two sectors that can be clearly delineated into "clean" and "dirty", and they together account for a little more than half of US emissions. I also allow for a general sector composed of only a clean technology that accounts for the rest of the economy. As I describe in Appendix C.1, I use the setup of Golosov et al. (2014) to describe the climate.

A time period represents one year, but I show in Appendix C.2 that the model's predictions are similar when patents are expected to last five or ten years instead. Finally, all dollar values referenced throughout the paper are in 2021 dollars.

5.1 Spillover Network

In this section, I describe how I calibrate the spillover network, following Liu and Ma (2021) in using the citation network of patents as a proxy for the spillover network.[19] For my sample of patents, I take the universe of granted US patents from PatentsView. First, I assign patents to my five technology classes: clean/dirty transportation, clean/dirty electricity generation, and general

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

[19] Dechezleprêtre et al. (2017) and Jee and Srivastav (2024) also use patent citations to study knowledge spillovers across clean and dirty technologies.

technology. To do so, I use the International Patent Classification (IPC) and Cooperative Patent Classification (CPC) systems.[20] IPC codes pertaining to clean and dirty transportation come from Aghion et al. (2016), who identify IPC codes in their empirical study of directed innovation in the auto industry. 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. Finally, I define general technology patents as those patents that are not in any of the above categories. Table D.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:

$$\tilde{\varphi}_{ij} = \frac{cites_{ij}}{\sum_k cites_{ik}}, \tag{55}$$

where $cites_{ij}$ is the number of citations from patents in technology class i that reference patents in technology class j. This approach leverages the widely accepted view that patent citations are a metric of knowledge spillovers (Jaffe et al., 1993; Hall et al., 2005).

My model predicts that receiving spillovers from other technologies increases the research productivity of one's own scientists, and Table 1 shows that this is indeed the case. Following a similar approach to Bloom et al. (2013) and Liu and Ma (2021), I find that my constructed spillover network is strongly predictive of innovation output for a given level of innovation effort. I proxy for innovation output using both citation-weighted and raw patent counts while also controlling for innovation effort using R&D spending. The details of this exercise are in Appendix C.3.

Columns (1) and (4) show my baseline specification, while Columns (2) and (5) test the prediction that research productivity is enhanced by spillovers a technology *receives*, rather than *sends*, through the network. Such downstream spillovers are insignificant in my preferred specification of citation-weighted patents, and they are less important and noisier than upstream spillovers for raw patents. Finally, to address the endogeneity concern that cross-technology spillovers are confounded with serially correlated shocks that benefit multiple technologies, I follow Bloom et al. (2013) by instrumenting for spillovers using variation in innovation activity driven by state-level R&D prices. Columns (3) and (6) show the results of this IV specification, with cross-technology spillovers remaining highly predictive of innovation productivity.

Another potential concern with using citation shares 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 they will create in the future. To address this concern, I compute the spillover network in five year bins between 1975 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

[20] 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 IPC codes that I use map to the same CPC codes with the same meaning.

Table 1: Predicted Innovation from Cross-Technology Spillovers

	Dependent Variable:					
	ln(Citations)			ln(Patents)		
	(1)	(2)	(3)	(4)	(5)	(6)
ln(R&D Spending)	0.800***	0.795***	0.813***	0.997***	1.046***	0.976***
	(0.08)	(0.09)	(0.06)	(0.10)	(0.08)	(0.08)
ln(Cross-Technology Spillovers)	3.111***	3.079***	2.924***	2.552***	2.678***	2.369***
	(0.37)	(0.33)	(0.35)	(0.55)	(0.30)	(0.48)
ln(Downstream Spillovers)		-0.269			1.668**	
		(0.81)			(0.81)	
Specification	OLS	OLS	IV	OLS	OLS	IV
IV 1st Stage F-Stat			983.84			907.66
R^2	0.886	0.886	0.969	0.907	0.913	0.976
Obs	176	176	144	176	176	144
Fixed Effects	Technology, Year			Technology, Year		

Notes: Standard errors clustered by technology are reported in parentheses. OLS regressions go from years 1980 to 2023, while IV regressions go from 1980 to 2015. All specifications exclude the general technology, focusing on the four climate-relevant technologies. Details on the construction of regression variables can be found in Appendix C.3.

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 C.4.

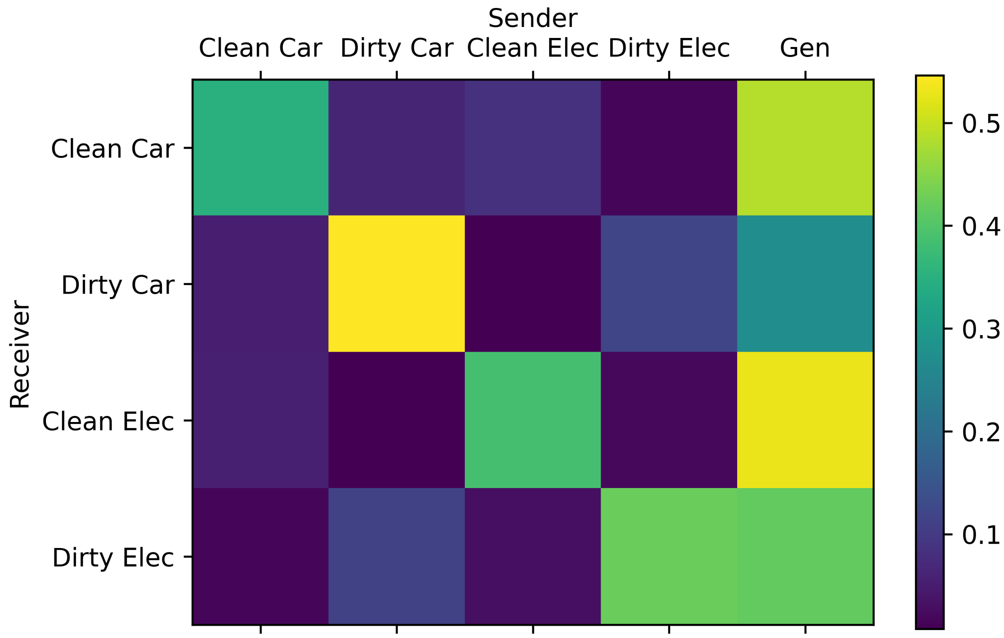
Figure 1 contains a heat map that illustrates the gross spillover network $\tilde{\varphi}$. It excludes the row pertaining to the general technology as almost all of the general technology's patent citations reference another general technology patent.[21] Figure 1 allows us to see several intuitive 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 discuss in Appendix C.8.

5.2 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 $\{A_{jt_0}\}$.

[21] Specifically, 98.5% 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.

In Appendix C.5, 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 2 provides a summary of my parameter choices.

For the within-sector elasticity of substitution between clean and dirty production, I set $\sigma = 1.86$ in line with the average estimate from Papageorgiou et al. (2017). This estimate comes from electricity generation, but I use it for transportation as well because it is a similar magnitude to other estimates 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, energy is typically a complement to other factors (Fried, 2018).

For the sectoral CES shares, I normalize their sum to one and select $\{\nu_{car}, \nu_{elec}\}$ to equal the average income shares of transport (2.8%) and electricity generation (2.2%) from 2001 to 2020. Data on revenue for electricity generation come from the US Energy Information Agency (EIA). Data on motor vehicle output and aggregate 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.[22] This is a similar value to those found in other models of step-ladder innovation (Acemoglu et al., 2016; Akcigit and Kerr,

[22] 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.

Table 2: 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 0.022	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)

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 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. 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.[23] Therefore, I set the dirty input costs to $r_{\theta d} = 2.25 \approx 1/0.45$.

For my choice of the research elasticity 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 (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 and Van Reenen, 2000; Bloom et al., 2002; Wilson, 2009). As I discuss in Appendix C.5, both sets of findings correspond to a value of $\eta = 0.5$.

[23] This efficiency factor is consistent with the US Department of Energy estimate that EVs 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 EVs likely reflects an intrinsic difference between the two technologies.

For the preference parameters, I use CRRA utility and set the inverse intertemporal elasticity of substitution ϑ equal to the standard value of one, which implies a logarithmic 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).[24] 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 C.5, 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).[25] Data on the share of electricity generated from non-fossil sources comes from the EIA. Next, I set cross-sector relative technology $\{A_{\theta dt_0}/A_{\Theta ct_0}\}$ to match the income share of each sector. In both cases, I include 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 72% 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).[26]

5.3 Model Validation

Whether spillovers or market size effects dominate is fundamentally an empirical question, so now that we have a calibrated model, we can use it to measure increasing returns to innovation. In doing so, we will also validate the model by testing whether the model-implied level of increasing returns

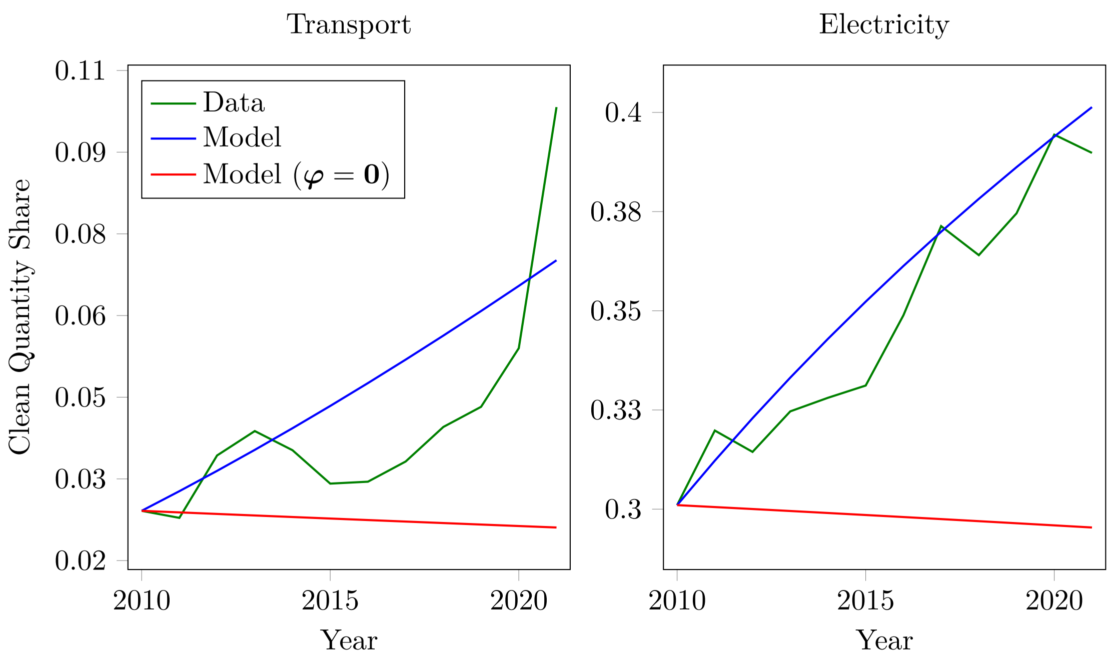
[24]It is worth noting that moral philosophers universally reject that utility should be discounted simply because it occurs at a later date (Cowen and Parfit, 1992; Ord, 2020). The only justification for discounting future utility that is widely accepted amongst philosophers is existential risk, i.e.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

[25]In particular, see Table 6.02 of the Transportation Energy Data Book. I consider hybrids as a form of clean transportation for two reasons. First, this definition is consistent with the previous literature (Aghion et al., 2016). Second, and more importantly, very few electric vehicles were sold in the US in the first half of the 2010s, so including hybrids allows me to extend the available data back further and test my model over a longer time horizon.

[26]To be exact, I subtract 56.8% of US emissions as this is the average contribution of transportation and electricity generation to US emissions from 2000 to 2020.

can replicate the progress made by clean technology in both transport and electricity generation throughout the 2010s. As shown in Figure 2, 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%.

Figure 2: 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.

Using the procedure described in Section 5.2, I take 2010 as the starting year for technology. Comparing the simulations with the data in Figure 2, one can see that my calibrated model rejects lock-in and in doing so matches the increasing market share of clean technology over this period. In particular, the spectral radius of the transition matrix equals 0.991, so the spillover effect was strong enough to prevent the market size effect from generating lock-in. The close match of the model to the data is noteworthy because no parameters were selected to target this change. 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. With market size effects alone, the spectral radius rises to 1.01, locking in dirty technology. In that case, US climate policy in the 2010s would have been 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.

To account for US climate policy during this time, I allow for subsidies to both innovation and clean inputs. Details on the selection of these subsidies are in Appendix C.6. In Figures D.3 and D.4, I examine the carbon prices and clean innovation subsidies that would have been required to shift each technology into its clean basin of attraction in the absence of cross-technology spillovers.

Because clean transportation was so far behind it's rival in 2010, it would have required massive intervention from either policy instrument. Clean electricity generation had made more progress by 2010 (which I take as given here), but it still would have required a carbon price of more than \$100; a stringent intervention even by today's standards.

I focus here on US climate policy, but the concern remains that a foreign big push could have brought about the clean progress observed domestically. To address this concern, I show in Figure C.4 that solar prices have seen log-linear declines since their invention in the 1950s. Such a steady accumulation process is more consistent with the catchup growth story of this paper than the abrupt redirection of a big push. Furthermore, the German and Chinese feed-in tariffs that are often discussed in this literature began after solar prices had already declined for half a century.

6 Quantitative Exercises

We have seen that US transportation and electricity generation is not locked in with dirty technology, so the economy is in an intermediate case where incremental reforms can redirect innovation. In this section, I examine the impact of introducing a realistic 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.

Next, 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 recommended by models without cross-technology spillovers and this holds even in the absence of 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 lower estimate of the social cost of carbon, as well as a uniform clean innovation subsidy equivalent to a 30% tax credit, consistent with the investment tax credits in the Inflation Reduction Act.[27] As shown by 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 3 summarizes the impact of the policy reform for three cases that differ in their level of cross-technology spillovers. First, I shut down the spillover network. Second, I take my calibrated spillover network. Third, I double the level of cross-technology spillovers in the economy. The third panel of Table 3 summarizes the difference between the three cases in terms of their degree of increasing returns to innovation, as measured by the spectral radius of the transition matrix.[28] Appendix C.8 uses spectral analysis to

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

[28] 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 0.996, leaving my conclusions largely unchanged.

examine the features of the spillover network that determine convergence speed by sector.

The first panel of Table 3 considers the long-run impacts of the policy reform. The first two rows make use of Proposition 1 and show the first-order change in steady-state relative clean technology by sector. With calibrated spillovers, we have an increase of 92.3% for transportation and 98.5% for electricity generation. Together, the carbon price and induced increase in clean technology reduce the economy's steady-state emissions intensity by two thirds.

Table 3: Impact of Policy Reform

	No Spillovers	Calibrated Spillovers	Double Spillovers
Long-Run Impacts			
Relative Clean Technology by Sector			
$\%\Delta\bar{B}_{car}$	0%	+92.3%	+29.6%
$\%\Delta\bar{B}_{elec}$	$+\infty$%	+98.5%	+33.8%
Emissions Intensity			
$\%\Delta\bar{\omega}$	-71.5%	-67.7%	-52.3%
Transitional Impacts			
Half-Lives of Convergence by Sector			
$t_{car}^{(1/2)}$	–	83 years	18 years
$t_{elec}^{(1/2)}$	–	97 years	22 years
Carbon Emissions by Year			
$\%\Delta\mathcal{E}_{2035}$	-39.4%	-47%	-49.7%
$\%\Delta\mathcal{E}_{2060}$	-42.4%	-53%	-50.8%
Degree of Increasing Returns to Innovation			
Spectral Radius			
$\max\lvert\kappa_j\rvert$	1.01	0.992	0.974

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 locked in economies, long-run impacts refer to corner, rather than interior, steady-states.

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 0.992 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 3 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 0.992 to 0.974, 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 3 considers the transitional impacts of the policy reform. Using Proposition 2, I can characterize the speed of transition to the new steady-state.[29] For my calibrated spillover network, increasing returns to innovation are strong enough to create a slow transition but are not so strong that they generate lock-in. Following the policy reform, transportation and electricity generation converge halfway to their steady-states in 83 and 97 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.

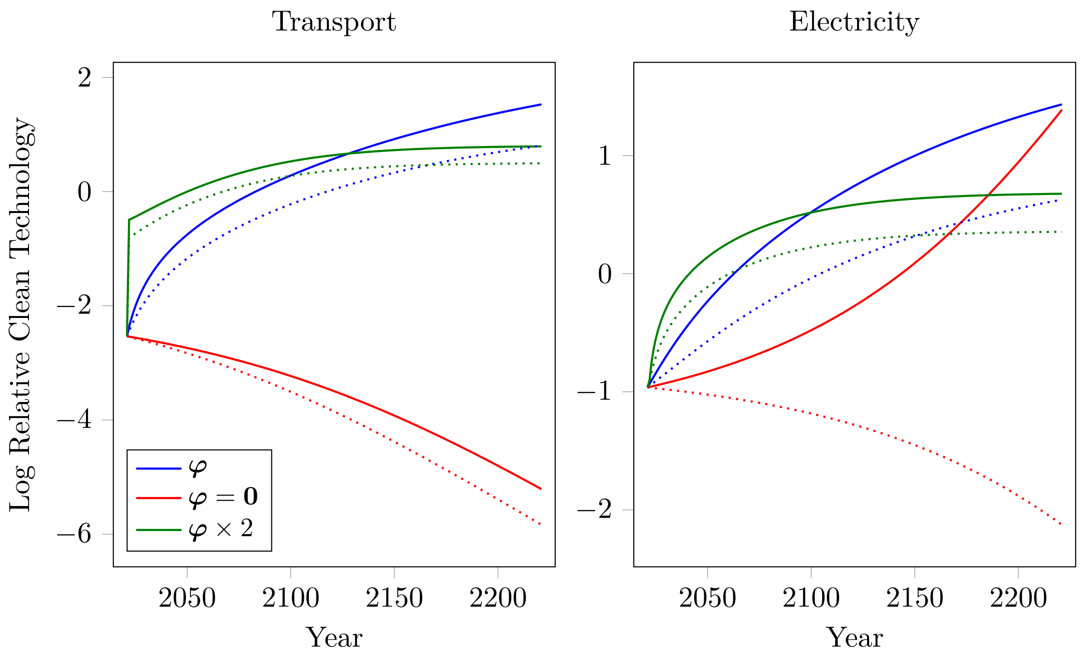
As before, the other two cases illustrate how cross-technology spillovers shape the speed with which policy can influence innovation. When the spillover network is shut down, the economy doesn't converge to an interior steady-state, making the half-lives undefined. Moreover, the long-run emission reductions 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 18 and 22 years. In the near term, the more rapid technological transition allows for greater emission reductions, but this lead is eventually lost as high spillovers allow dirty technologies to stay relatively advanced.

Figure 3 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 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 lock-in, 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

[29] Given that these results involve linearizing my model, one may be concerned that my results are not informative about such a large policy reform. To address this concern, I compare the linearized transition path to the full simulation in Figure D.5 and show that the two do not differ significantly.

Figure 3: 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.

of policy by preventing any technology from gaining a significant advantage.

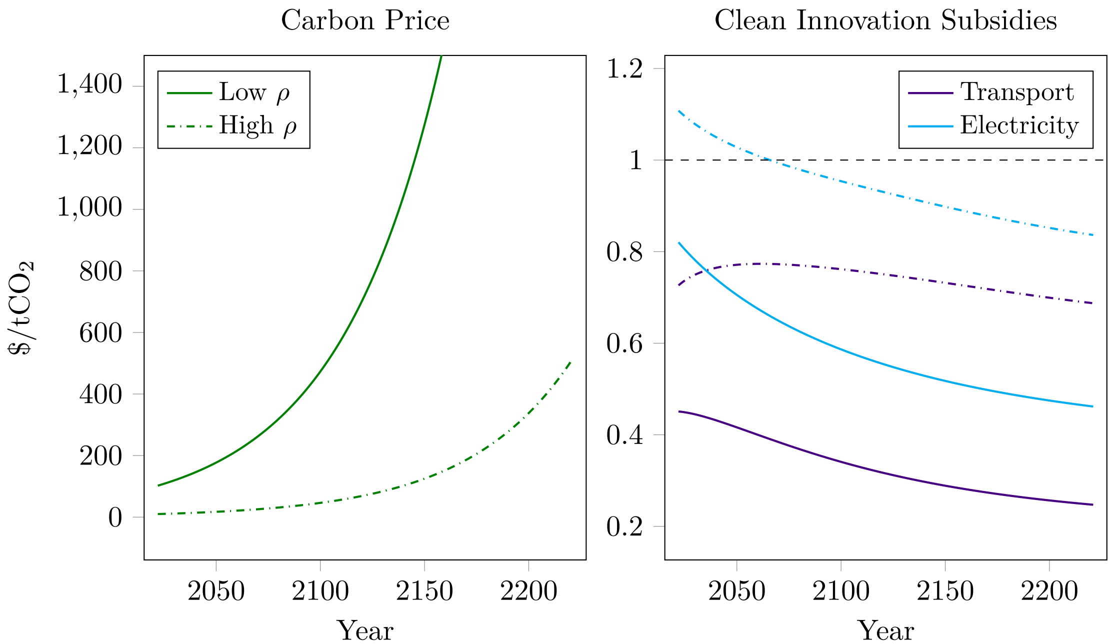
My theoretical results emphasized the role of both spillover and market size effects in shaping the impact of policy reforms in both the long-run and transition. To explore this relationship, I first plot in Figure D.6 how the steady-state impact of the policy reform varies in the scale of cross-technology spillovers and the elasticity of substitution between clean and dirty goods. As expected, the steady-state impact reduces with cross-technology spillovers and increases with the elasticity of substitution. Next, I plot in Figure D.7 how the spectral radius of the transition matrix varies in the same two variables. Again, the theory is borne out with increasing returns to innovation decreasing with cross-technology spillovers and increasing with the elasticity of substitution. In particular, I find that technology becomes locked in if cross-technology spillovers are reduced to 60.1% of their calibrated value or if the elasticity of substitution rises to 2.43.

Finally, the EPA has recently proposed increasing the federal SCC to $190, so I examine the impact of introducing such a carbon price in Table D.2 and Figure D.8. 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.

6.2 Optimal Policy Path

In this section, I simulate the optimal policy paths starting in 2022. I start by considering the first-best policy path for two values of the discount rate, as shown in Figure 4. 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 about a century to reach the more patient Planner's initial carbon price.

Figure 4: First-Best Policy Path



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

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 3, 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 (0.67 vs. 4.9%), while clean electricity's eigenvector centrality is slightly less than a third of its steady-state income share (1 vs. 3.1%).[30] In other words, clean technologies are useful because they can produce goods without pollution, but they are not as useful in the production of ideas.

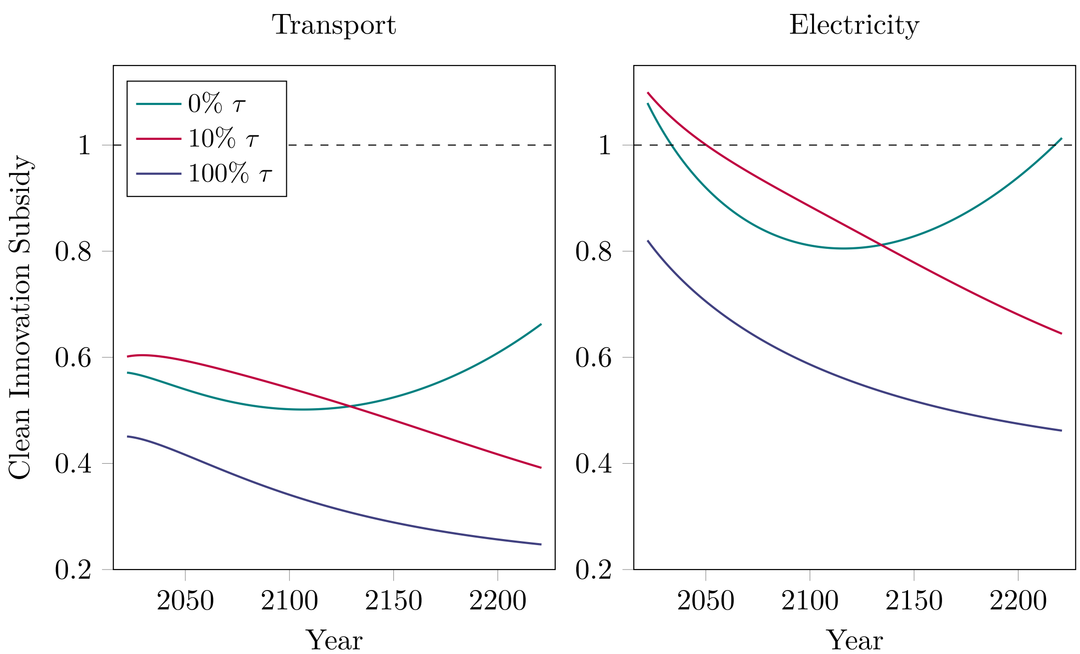
My definition of the general technology is highly aggregated, so the centrality of clean technologies could depend on the way the non-climate-related technologies of the economy are aggregated. But as I show in Appendix C.4, the centrality of clean technologies is robust to disaggregating the general technology into 3-digit CPC codes.

Interestingly, 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

[30] 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.

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 5: Optimal Innovation Subsidy Path

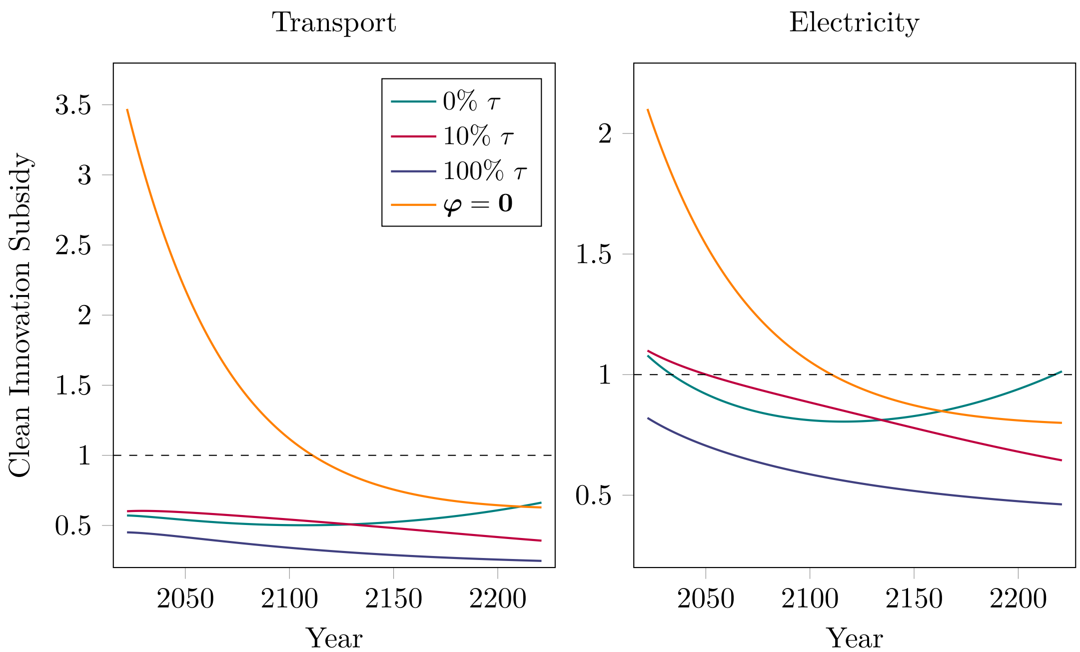


Notes: Policy paths use the low discount rate. The external carbon price τ is a proportion of the social cost of carbon, so 100% is the first-best. Innovation subsidies are listed as a fraction of the common innovation wedge.

Figure 5 shows how optimal innovation subsidies change when there are restrictions on the carbon price. As shown in Proposition 3, 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, 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 D.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 6 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 6: 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 C.4), 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 C.10 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 spillovers become endogenously more important 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:

$$\frac{\partial \ln(\mathcal{E}_t)}{\partial \ln(A_{jt})} = \frac{\partial \ln(\bar{\omega}_t)}{\partial \ln(A_{jt})} + \frac{\partial \ln(\mathcal{Y}_t)}{\partial \ln(A_{jt})}. \tag{56}$$

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 0.27 GtC and 2% in the year 2200, whereas they are 3.08 GtC and 1.99% for the ten percent case.[31] When there is no carbon price, these numbers are 8.46 GtC and 1.96%. 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 D.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 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 1.49%, 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

[31] 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.

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 0.29% for the ten percent case and 0.87% for the zero case. 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 potentially very large (Bilal and Känzig, 2024), so I consider the possibility of substantially higher damages in Figure D.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 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. Without such spillovers, new technologies must be built up from scratch and market size effects will generate lock-in for incumbent dirty technologies. This leads to the incorrect conclusion that only a big push can redirect innovation, and once this view is overturned, a big push is neither necessary nor desirable.

Such spillovers matter for both describing the impact of policy reforms and prescribing optimal policy. Spillovers compete with market size effects in a horse race that determines the degree of increasing returns to innovation, which can be quantified using a spectral radius. Quantitatively, cross-technology spillovers prevent US transport and electricity generation from being locked in with dirty technology, allowing incremental policy reforms to shift the direction of innovation

For optimal policy, spillovers link the social benefit of innovation across technologies and imply the need for steady, permanent intervention, rather than a temporary big push. For the clean transition, cross-technology 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.

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., 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. (2022). Demographics and Automation. *The Review of Economic Studies*, 89(1):1–44.

Aghion, P., Barrage, L., Donald, E., Hémous, D., and Liu, E. (2025). Transition to Green Technology along the Supply Chain. Working Paper 33934, National Bureau of Economic Research.

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.

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

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

Arkolakis, C. and Walsh, C. (2023). Clean Growth. Working Paper 31615, National Bureau of Economic Research.

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

Bilal, A. and Känzig, D. (2024). The Macroeconomic Impact of Climate Change: Global vs. Local Temperature. Working Paper 32450, National Bureau of Economic Research.

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

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

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

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

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.

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.

Dechezleprêtre, A., Martin, R., and Mohnen, M. (2017). Knowledge Spillovers from Clean and Dirty Technologies. *Working Paper*.

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

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

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

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

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.

Hart, R. (2019). To Everything There Is a Season: Carbon Pricing, Research Subsidies, and the Transition to Fossil-Free Energy. *Journal of the Association of Environmental and Resource Economists*, 6(2):349–389.

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.

Jee, S. J. and Srivastav, S. (2024). Knowledge Spillovers Between Clean and Dirty Technologies: Evidence from the Patent Citation Network. *Ecological Economics*, 224:108310.

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.

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

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. (2024). Innovation-Led Transitions in Energy Supply. *American Economic Journal: Macroeconomics*, 16(1):29–65.

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.

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.

Online Appendix

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

Eric Donald

Contents

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

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

Turning to the innovation side of the economy, the research firm solves the problem described in Equation (22). 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 (23) 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 (23) 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.

A.2 Nested-CES Parametric Model

This appendix outlines the multi-sector version of the example model of Section 2.5 that I use for my quantitative exercises. Production is divided into Θ-many sectors, each of which has a clean and dirty form of production. Final output now follows

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

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 is as before:

$$E_{\theta t} = \Big(Y_{\theta ct}^{\frac{\sigma-1}{\sigma}} + Y_{\theta dt}^{\frac{\sigma-1}{\sigma}} \Big)^{\frac{\sigma}{\sigma-1}}, \tag{A.8}$$

where $Y_{\theta c}$ and $Y_{\theta d}$ are the clean and dirty forms of production for sector θ.

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.

For the production of ideas, I will now allow for differences in productivity shifters according to $\chi_j = \chi \nu_{\theta(j)}^{-\eta}$, where $\theta(j)$ is the sector associated with technology j. 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, 2022). Thus, the relevant notion of effective scientific effort is the number of scientists per economic task associated with a technology.

Finally, household utility follows the common CRRA specification

$$u(c_t) = \frac{c_t^{1-\vartheta} - 1}{1 - \vartheta}, \tag{A.9}$$

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

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

$$\max_{\{E_{\theta t}\}} \mathcal{Y}_t - \sum_\theta p_{\theta t} E_{\theta t}, \tag{A.10}$$

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

$$\Omega_t^{\frac{\lambda-1}{\lambda}} \nu_\theta^{\frac{1}{\lambda}} \Big(\frac{\mathcal{Y}_t}{E_{\theta t}}\Big)^{\frac{1}{\lambda}} = p_{\theta t}. \tag{A.11}$$

Next, producers at the sector level solve

$$\max_{Y_{\theta dt}, Y_{\theta ct}} p_{\theta t} E_{\theta t} - p_{\theta ct} Y_{\theta ct} - p_{\theta dt} Y_{\theta dt} \tag{A.12}$$

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

$$p_{\theta t}\Big(\frac{E_{\theta t}}{Y_{\theta e t}}\Big)^{\frac{1}{\sigma}} = p_{\theta e t}, \tag{A.13}$$

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

A.3 Substitution Matrix for Nested-CES Production

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

$$G_i\Big(\{\ln(Y_j/Y_J), \ln(p_J/p_j)\}\Big) = \ln(p_J/p_i) - \Big[\frac{1}{\lambda}\ln(\nu_{\theta(J)}/\nu_{\theta(i)}) + (\frac{1}{\sigma} - \frac{1}{\lambda})\ln(E_{\theta(J)}/E_{\theta(i)}) - \frac{1}{\sigma}\ln(Y_J/Y_i)\Big]. \tag{A.14}$$

That is, G is a function valued at the zero vector when the equilibrium price conditions (A.11) and (A.13) 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

$$\frac{\partial G_i}{\partial \ln(Y_j/Y_J)} = -\Big[\frac{1}{\sigma}\mathbb{1}(i=j) + (\frac{1}{\lambda} - \frac{1}{\sigma})\varepsilon^E_{\theta(i)j}\mathbb{1}(\theta(i) = \theta(j))\Big] \tag{A.15}$$

$$\frac{\partial G_i}{\partial \ln(p_J/p_j)} = \mathbb{1}(i=j), \tag{A.16}$$

where $\varepsilon^E_{\theta(i)j}$ is the elasticity of sector $\theta(i)$ with respect to good j. Equation (A.15) relies on the fact that $E_{\theta(J)} = Y_J$ because the final sector Θ only has a clean form of production.

Applying the Implicit Function Theorem gives us

$$\mathbf{\Sigma} = -\mathbf{D_y G}^{-1}, \tag{A.17}$$

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

$$\mathbf{\Sigma} = \begin{pmatrix} \tilde{\mathbf{\Sigma}}_1 & 0 & \dots & 0 \\ 0 & \tilde{\mathbf{\Sigma}}_2 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \tilde{\mathbf{\Sigma}}_{\Theta-1} \end{pmatrix}, \tag{A.18}$$

where

$$\tilde{\mathbf{\Sigma}}_\theta = \begin{pmatrix} \lambda + (\sigma - \lambda)\varepsilon^E_{\theta d} & (\lambda - \sigma)\varepsilon^E_{\theta d} \\ (\lambda - \sigma)\varepsilon^E_{\theta c} & \lambda + (\sigma - \lambda)\varepsilon^E_{\theta c} \end{pmatrix}, \tag{A.19}$$

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

A.4 Proof of Proposition 1

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

$$\ln(s_i/s_J) = -\frac{1}{\eta}\ln(\chi_i/\chi_J) - \frac{1}{\eta}\ln(\phi_i/\phi_J). \tag{A.20}$$

Plugging this into the research condition (23) and using the no-arbitrage condition (21) and relative good prices (A.4), we have

$$(1-\alpha)\ln(\bar{A}_i) = \ln(\Xi_i) + \frac{1}{\eta}\ln(\chi_i/\chi_J) + \frac{1}{\eta}\ln(\phi_i/\phi_J) + \alpha\ln(\mathcal{R}_i) + \ln(Y_{it}/Y_{Jt}), \tag{A.21}$$

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.21).

To consider existence and uniqueness of the steady-state, use Equation (A.21) to define the mapping $\mathbb{T}\ln(\bar{A})$ which follows

$$\mathbb{T}\ln(\bar{A}_i) = \frac{1}{1-\alpha}\Big[\ln(\Xi_i) + \frac{1}{\eta}\ln(\chi_i/\chi_J) + \frac{1}{\eta}\ln(\phi_i/\phi_J) + \alpha\ln(\mathcal{R}_i) + \ln(Y_{it}/Y_{Jt})\Big], \tag{A.22}$$

where the functional dependence on log relative technology comes from relative spillovers and relative quantities. If $\mathbb{T}\ln(\bar{A})$ 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.5 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 $\mathbf{D_\Xi \bar{A}_{ss}}$, we can differentiate

$$(1-\alpha)\frac{\partial \ln(\bar{A}_i)}{\partial \ln(\Xi_j)} = \mathbb{1}(i=j) - \frac{1}{\eta}\sum_q \Phi_{iq}\frac{\partial \ln(\bar{A}_q)}{\partial \ln(\Xi_j)} + (1-\alpha)\sum_q \Sigma_{iq}\frac{\partial \ln(\bar{A}_q)}{\partial \ln(\Xi_j)}, \tag{A.23}$$

which in matrix notation gives us

$$\begin{aligned}(1-\alpha)\mathbf{D_\Xi \bar{A}_{ss}} &= \mathbf{I} - \frac{1}{\eta}\mathbf{\Phi}\mathbf{D_\Xi \bar{A}_{ss}} + (1-\alpha)\mathbf{\Sigma}\mathbf{D_\Xi \bar{A}_{ss}} \\ \Rightarrow \mathbf{D_\Xi \bar{A}_{ss}} &= \eta\big[\mathbf{\Phi} - \eta(1-\alpha)(\mathbf{\Sigma}-\mathbf{I})\big]^{-1}.\end{aligned} \tag{A.24}$$

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

$$(1-\alpha)\frac{\partial \ln(\bar{A}_i)}{\partial \ln(\mathcal{R}_j)} = \alpha\mathbb{1}(i=j) - \frac{1}{\eta}\sum_q \Phi_{iq}\frac{\partial \ln(\bar{A}_q)}{\partial \ln(\mathcal{R}_j)} + (1-\alpha)\sum_q \Sigma_{iq}\frac{\partial \ln(\bar{A}_q)}{\partial \ln(\mathcal{R}_j)} - \alpha\Sigma_{ij}, \tag{A.25}$$

which in matrix notation gives us

$$\begin{aligned}(1-\alpha)\mathbf{D_R \bar{A}_{ss}} &= \alpha\mathbf{I} - \frac{1}{\eta}\mathbf{\Phi}\mathbf{D_R \bar{A}_{ss}} + (1-\alpha)\mathbf{\Sigma}\mathbf{D_R \bar{A}_{ss}} - \alpha\mathbf{\Sigma} \\ \Rightarrow \mathbf{D_R \bar{A}_{ss}} &= -\eta\alpha\big[\mathbf{\Phi} - \eta(1-\alpha)(\mathbf{\Sigma}-\mathbf{I})\big]^{-1}(\mathbf{\Sigma}-\mathbf{I}).\end{aligned} \tag{A.26}$$

Combining Equations (A.24) and (A.26), we have

$$d\ln(\bar{A}_{ss}) = \eta\big[\mathbf{\Phi} - \eta(1-\alpha)(\mathbf{\Sigma}-\mathbf{I})\big]^{-1}\big[d\ln(\Xi) - \alpha(\mathbf{\Sigma}-\mathbf{I})d\ln(\mathcal{R})\big], \tag{A.27}$$

which gives us Equations (42) and (43) from Proposition 1.

□

A.5 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.11) and (A.13) into Equation (A.21), we have

$$\begin{aligned}\sum_{j<J}\Phi_{ij}\ln(\bar{A}_j) = {} & \eta\ln(\Xi_i) + \eta(1-\alpha)(\sigma-1)\ln(\bar{A}_i) \\ & + \eta(\sigma-\lambda)\ln\big(p_{\theta(i)t}/p_{\theta(J)t}\big) - \eta\alpha(\sigma-1)\ln(\mathcal{R}_i).\end{aligned} \tag{A.28}$$

Denote relative sector prices by $\mathcal{P}_i \equiv p_{\theta(i)t}/p_{\theta(J)t}$. Writing Equation (A.28) in terms of matrix notation, we have

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

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

$$\mathcal{P}_i = \frac{p_{\theta(i)t}}{p_{\theta(J)t}} = \Big(\frac{p_{\theta ct}^{1-\sigma} + p_{\theta dt}^{1-\sigma}}{p_{Jt}^{1-\sigma}}\Big)^{\frac{1}{1-\sigma}} = \Big(\mathcal{R}_{\theta c}^{\alpha(1-\sigma)}\bar{A}_{\theta c}^{(1-\alpha)(\sigma-1)} + \mathcal{R}_{\theta d}^{\alpha(1-\sigma)}\bar{A}_{\theta d}^{(1-\alpha)(\sigma-1)}\Big)^{\frac{1}{1-\sigma}}, \tag{A.30}$$

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.29). Note that if the production structure is a simple CES, i.e. $\lambda = \sigma$, then Equation (A.29) constructively provides the economy's unique steady-state.

Furthermore, Equation (A.20) implies that

$$\ln(\bar{s}_{ss}) = \ln(\mathcal{V}) + \frac{1}{\eta}\mathbf{\Phi}\ln(\bar{A}_{ss}), \tag{A.31}$$

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

$$\ln(\bar{s}_{ss}) = \ln(\mathcal{V}) + \mathbf{\Phi}\big[\mathbf{\Phi} - \eta(1-\alpha)(\sigma-1)\mathbf{I}\big]^{-1}\big[\ln(\Xi) + (\sigma-\lambda)\ln(\mathcal{P}) - \alpha(\sigma-1)\ln(\mathcal{R})\big]. \tag{A.32}$$

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

$$\ln(\bar{A}_{ss}) = \frac{1}{(1-\alpha)(1-\lambda)}\ln(\Xi) + \frac{\alpha}{1-\alpha}\ln(\mathcal{R}), \tag{A.33}$$

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

A.6 Proof of Proposition 2

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

$$\ln(\bar{A}_t) = \mathcal{H}\Big(s_t\big(\ln(\bar{A}_{t-1})\big), \ln(\bar{A}_{t-1})\Big), \tag{A.34}$$

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

$$\ln(\bar{A}_t) - \ln(\bar{A}_{ss}) \approx \boldsymbol{\mathcal{J}}\big(\ln(\bar{A}_{t-1}) - \ln(\bar{A}_{ss})\big), \tag{A.35}$$

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 (45) 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 (23), the no-arbitrage condition (21), and the market clearing condition for scientists (12). I will omit time subscripts on policy variables. This yields

$$\begin{aligned} \ln(\bar{A}_{it}) &= \ln(\gamma)\big[\chi_i s_{it}^{\eta}\phi_{it} - \chi_J s_{Jt}^{\eta}\phi_{Jt}\big] + \ln(\bar{A}_{it-1}) \\ (1-\eta)\big[\ln(s_{it}) - \ln(s_{Jt})\big] &= \ln(\chi_i/\chi_J) + \ln(\xi_i/\xi_J) + \ln(\phi_{it}/\phi_{Jt}) + \ln(p_{it}/p_{Jt}) + \ln(Y_{it}/Y_{Jt}) \\ \sum_i s_{it} &= \mathcal{S}. \end{aligned} \tag{A.36}$$

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

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

$$\begin{aligned} \varepsilon_{ij}^{\bar{A}} &= g\big[\eta(\varepsilon_{ij}^{s} - \varepsilon_{Jj}^{s}) - \Phi_{ij}\big] + \mathbb{1}(i=j) \\ (1-\eta)(\varepsilon_{ij}^{s} - \varepsilon_{Jj}^{s}) &= -\Phi_{ij} - (1-\alpha)\varepsilon_{ij}^{\bar{A}} + \sum_q \Sigma_{iq}(1-\alpha)\varepsilon_{qj}^{\bar{A}} \\ \Rightarrow \big[(1-\eta) + g\eta(1-\alpha)\big]\varepsilon_{ij}^{\bar{A}} &= (1-\eta)\mathbb{1}(i=j) - g\Phi_{ij} + g\eta(1-\alpha)\sum_q \Sigma_{iq}\varepsilon_{qj}^{\bar{A}}, \end{aligned} \tag{A.37}$$

where the derivative of relative good prices with respect to relative technology comes from Equation (A.4). Transforming into matrix notation, we have Equation (46) from Proposition 2:

$$\begin{aligned} \big[(1-\eta)\mathbf{I} - g\eta(1-\alpha)(\boldsymbol{\Sigma} - \mathbf{I})\big]\boldsymbol{\mathcal{J}} &= (1-\eta)\mathbf{I} - g\boldsymbol{\Phi} \\ \Rightarrow \boldsymbol{\mathcal{J}} &= \big[(1-\eta)\mathbf{I} - g\eta(1-\alpha)(\boldsymbol{\Sigma} - \mathbf{I})\big]^{-1}\big[(1-\eta)\mathbf{I} - g\boldsymbol{\Phi}\big]. \end{aligned} \tag{A.38}$$

□

As I discuss in Footnote 8, my description of the effect of cross-technology spillovers on the transition term in the example model of Section 3.1 ignores changes in g. In particular, the full derivatives of $\mathcal{J}$ with respect to the cross-technology spillover parameters (but holding fixed the steady-state) follow

$$\frac{\partial \mathcal{J}}{\partial \varphi_{cd}} \propto -1 + \left[\Phi - \mathcal{J}\eta(1-\alpha)(\sigma-1)\right]\phi_c^{-\frac{1}{\eta}} \ln\left(\bar{A}_{ss}\right) \tag{A.39}$$

$$\frac{\partial \mathcal{J}}{\partial \varphi_{dc}} \propto -1 - \left[\Phi - \mathcal{J}\eta(1-\alpha)(\sigma-1)\right]\phi_d^{-\frac{1}{\eta}} \ln\left(\bar{A}_{ss}\right), \tag{A.40}$$

where the second term in both equations comes from changes in the growth rate. Also, note that the factor of proportionality is positive and so does not change the sign of the derivative. The term in square brackets is $\mathcal{O}(1-\mathcal{J})$ because $\mathcal{J}$ will be close to one if $\Phi \approx \eta(1-\alpha)(\sigma-1)$. Section 5 shows that this is the empirically relevant case, which is why one can ignore changes in the growth rate when changing the cross-technology spillover parameters.

A.7 Proof of Corollary 2

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

$$\boldsymbol{\mathcal{J}}\boldsymbol{Q} = \boldsymbol{Q}\boldsymbol{D}(\kappa) \tag{A.41}$$

$$\Rightarrow \left[(1-\eta)\mathbf{I} - g\boldsymbol{\Phi}\right]\boldsymbol{Q} = \left[(1-\eta)\mathbf{I} - g\eta(1-\alpha)(\boldsymbol{\Sigma}-\mathbf{I})\right]\boldsymbol{Q}\boldsymbol{D}(\kappa)$$

$$\Rightarrow \left[(1-\eta)\mathbf{I} - g\eta(1-\alpha)(\boldsymbol{\Sigma}-\mathbf{I})\right]\boldsymbol{Q}\boldsymbol{D}(1-\kappa) = g\left[\boldsymbol{\Phi} - \eta(1-\alpha)(\boldsymbol{\Sigma}-\mathbf{I})\right]\boldsymbol{Q}$$

$$\Rightarrow \boldsymbol{\mathcal{M}} = g\boldsymbol{Q}\boldsymbol{D}(1-\kappa)^{-1}\boldsymbol{Q}^{-1}\left[(1-\eta)\mathbf{I} - g\eta(1-\alpha)(\boldsymbol{\Sigma}-\mathbf{I})\right]^{-1},$$

which gives us Equation (49) of Corollary 2.

□

A.8 Spectral Analysis

The eigendecomposition of Section 3.5 shows that 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

$$\beta = \boldsymbol{Q}^{-1}\bar{\mathcal{A}}_0. \tag{A.42}$$

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.[A.1] 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.[A.2] Since $\boldsymbol{Q}$ forms a basis, any state of technology can be expressed as a linear combination of these eigenstates. Moreover, each eigenstate

[A.1] 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.

[A.2] 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.

has a convergence rate determined by its corresponding eigenvalue. For instance, if the initial state of technology is proportional to the jth eigenstate, then β will load exclusively on the jth dimension, and technology will geometrically converge to the steady state with a decay rate of κ_j. More generally, the loading of the initial state on the eigenstates determines the speed of convergence as a weighted average of the eigenvalues. The following proposition provides a formal statement of how the spectral properties of the transition matrix determine technology's transition.

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

$$\bar{\mathcal{A}}_t \approx \sum_j \kappa_j^t \beta_j Q_j, \tag{A.43}$$

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 A.1 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.

Finally, to measure convergence speeds in units of time, one can 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

$$t_{ES_j}^{(1/2)} = \left\lceil \frac{\ln(1/2)}{\ln(\kappa_j)} \right\rceil, \tag{A.44}$$

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 A.1, each technology will have its own half-life governed by the initial state's loading on the eigenstates.

Corollary 2 argues for an inverse relationship between the speed of transition and long-run effect of policy due to increasing returns to innovation. I will now show why this is a general phenomenon that holds in a large class of dynamic models. Imagine a dynamic system

$$x_{t+1} = F(x_t; \theta), \tag{A.45}$$

where $x_t \in \mathbb{R}^J$ is the state, and θ is some set of fundamental parameters. These parameters could be policies, productivities, etc. Then a steady-state $\bar{x}$ satisfies

$$\bar{x} = F(\bar{x}; \theta). \tag{A.46}$$

Linearizing around a steady-state, we have

$$x_{t+1} - \bar{x} \approx \boldsymbol{\mathcal{J}}(x_t - \bar{x}), \tag{A.47}$$

where $\boldsymbol{\mathcal{J}}$ is the Jacobian of F with respect to x (evaluated at the steady-state). Furthermore, if we perturb θ and consider the long-run effect, we have

$$d\bar{x} = (\mathbf{I} - \boldsymbol{\mathcal{J}})^{-1}\boldsymbol{G}d\theta, \tag{A.48}$$

where $\boldsymbol{G}$ is the Jacobian of F with respect to θ. The direct effect of shocking θ is given by $\boldsymbol{G}d\theta$ and this feeds into long-run effects through the Leontief inverse $(\mathbf{I} - \boldsymbol{\mathcal{J}})^{-1}$. Denote this amplification term by $\tilde{\boldsymbol{\mathcal{M}}}$.

Eigendecomposing $\boldsymbol{\mathcal{J}}$, we have

$$\boldsymbol{\mathcal{J}} = \boldsymbol{Q}\boldsymbol{D}(\kappa)\boldsymbol{Q}^{-1}, \tag{A.49}$$

where $\boldsymbol{D}(\kappa)$ is a diagonal matrix whose diagonal elements are the eigenvalues, and $\boldsymbol{Q}$ is a matrix whose columns are the eigenvectors. This also implies that

$$\tilde{\boldsymbol{\mathcal{M}}} = \boldsymbol{Q}\boldsymbol{D}(1-\kappa)^{-1}\boldsymbol{Q}^{-1}. \tag{A.50}$$

So if κ is an eigenvalue of $\boldsymbol{\mathcal{J}}$, then it must be that $(1-\kappa)^{-1}$ is an eigenvalue of $\tilde{\boldsymbol{\mathcal{M}}}$. Therefore, the inverse relationship expressed in Corollary 2 is far more general than the direction of innovation.

The above result implies that changes to fundamental parameters θ, such as policy reforms, can either have impacts on the state x_t that are occur quickly or have a large impact in the long-run, but not both. To see this explicitly, consider an economy that is near an initial steady-state $\bar{x}_0$. We will then shock the new economy, so the new steady-state will follow

$$\bar{x} \approx \bar{x}_0 + (\mathbf{I} - \boldsymbol{\mathcal{J}})^{-1}\boldsymbol{G}d\theta. \tag{A.51}$$

Denote by $\tilde{\beta} \equiv \boldsymbol{Q}^{-1}\boldsymbol{G}d\theta$ the direct policy effect projected into the eigenspace. To then consider the transition, we have

$$\begin{aligned} x_t - \bar{x} &\approx \boldsymbol{\mathcal{J}}^t(x_0 - \bar{x}) \\ &= -\boldsymbol{\mathcal{J}}^t(\mathbf{I} - \boldsymbol{\mathcal{J}})^{-1}\boldsymbol{G}d\theta \\ &= -\sum_j \frac{\kappa_j^t}{1-\kappa_j} Q_j \beta_j, \end{aligned} \tag{A.52}$$

where the first line comes from the fact that the economy was initially near its old steady-state, $x_0 \approx \bar{x}_0$, and the second comes from the two eigendecompositions. Therefore, we can decompose the initial effect of the policy reform into the eigenvectors $\boldsymbol{G}d\theta = \boldsymbol{Q}\tilde{\beta}$, and the components of the policy reform that have the largest long-run effect from $(1-\kappa)^{-1}$ will also have the largest half-lives of transition.

Intuitively, we have that for a general class of model, big long-run changes in the steady-state are inherently slow because reforms that change the steady-state in directions where reforms have the most leverage are also those acting on the components of state space where responses are the most sluggish. Note here that "slow" is defined proportionally, so this does not just follow from the fact that it takes longer to travel a given distance (when travel is a linear process).

While, the math here is fairly straightforward, there is a substantive economic implication. Namely, any structural parameter that slows the speed of convergence must also increase the long-run effect of policy reforms. Transition speeds and long-run comparative statics are generally considered as independent features of a dynamic system, but the above result implies this is not actually possible in general. Once you have established something about one, you have inescapably said something about the other. Put differently, persistence of initial conditions is structurally tied to the long-run leverage of policy as they are two sides of same eigenproperty coin.

A.9 Closed-Form for Steady-State Growth Rate g

Combining Equations (12) and (A.20), we have that the steady-state level of scientists must follow

$$s_j = \frac{\left(\chi_j \phi_j\right)^{-\frac{1}{\eta}}}{\sum_i \left(\chi_i \phi_i\right)^{-\frac{1}{\eta}}} \mathcal{S}. \tag{A.53}$$

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

$$g = \ln(\gamma) \Big(\sum_j \left(\chi_j \phi_j\right)^{-\frac{1}{\eta}} \Big)^{-\eta} \mathcal{S}^{\eta}. \tag{A.54}$$

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.10 Proof of Proposition 3

As described in the Definition 6, the Planner solves

$$\max_{\{c_t, \{A_{jt}, s_{jt}\}\}} \sum_{t \geq 0} \frac{1}{(1+\rho)^t} u(c_t) \;\; s.t. \tag{A.55}$$

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

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

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

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

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

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

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

$$\varpi_{st} = \epsilon_{jt} \ln(\gamma)\chi_j \eta s_{jt}^{\eta-1}\phi_{jt}. \tag{A.57}$$

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

$$\frac{\epsilon_{jt}}{A_{jt}} = \frac{u_t'}{(1+\rho)^t}\frac{\partial \mathcal{Y}_t}{\partial A_{jt}} + \sum_{\hat{t}\geq t}\frac{u_{\hat{t}}'}{(1+\rho)^{\hat{t}}}\Big[\sum_i \int_0^1 \frac{\partial \mathcal{Y}_{\hat{t}}}{\partial \ell_{i\iota\hat{t}}}\frac{\partial \ell_{i\iota\hat{t}}}{\partial A_{jt}}d\iota + \sum_i \big(\frac{\partial \mathcal{Y}_{\hat{t}}}{\partial \Lambda_{i\hat{t}}} - r_i\big)\frac{\partial \Lambda_{i\hat{t}}}{\partial A_{jt}} - SCC_{\hat{t}}\frac{\partial \mathcal{E}_{\hat{t}}}{\partial A_{jt}}\Big] \tag{A.58}$$

$$+ \frac{\epsilon_{jt+1}}{A_{jt}} + \sum_i \epsilon_{it+1}\ln(\gamma)\chi_i s_{it+1}^{\eta}\frac{\partial \phi_{it+1}}{\partial A_{jt}},$$

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

$$SCC_t = -\sum_{\hat{t}\geq t}\prod_{\hat{s}=1}^{\hat{t}-t}\frac{1}{R_{t+\hat{s}}}\mathcal{Y}_{\hat{t}}\frac{\partial \ln(\Omega_{\hat{t}})}{\partial \mathcal{E}_t}, \tag{A.59}$$

which gives us Equation (52) of Proposition 3. 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

$$\frac{\partial \mathcal{Y}_t}{\partial \Lambda_{jt}} = r_j + \omega_j \tau_t \tag{A.60}$$
$$\frac{\partial \mathcal{Y}_t}{\partial \ell_{j\iota t}} = w_{\ell t}.$$

Plugging these conditions into the Planner's optimality condition for technology (A.58), 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

$$\frac{\epsilon_{jt}}{A_{jt}} = \frac{u_t'}{(1+\rho)^t}\frac{\partial \mathcal{Y}_t}{\partial A_{jt}} - \sum_{\hat{t}\geq t}\frac{u_{\hat{t}}'}{(1+\rho)^{\hat{t}}}(SCC_{\hat{t}} - \tau_{\hat{t}})\frac{\partial \mathcal{E}_{\hat{t}}}{\partial A_{jt}} + \frac{\epsilon_{jt+1}}{A_{jt}} + \sum_i \epsilon_{it+1}\ln(\gamma)\chi_i s_{it+1}^{\eta}\frac{\partial \phi_{it+1}}{\partial A_{jt}}. \tag{A.61}$$

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

(A.57), we have

$$\chi_j \eta s_{jt}^{\eta-1}\phi_{jt}\xi_{jt}\Pi_{jt} = w_{st} \quad \text{(A.62)}$$
$$\chi_j \eta s_{jt}^{\eta-1}\phi_{jt}\epsilon_{jt} = \frac{\varpi_{st}}{\ln(\gamma)}.$$

The Planner can set the wage for scientists according to

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

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

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

Plugging this into Equation (A.61), we have that innovation subsidies follow the recursion

$$\xi_{jt}\Pi_{jt} = (1-\alpha)S_{jt}\mathcal{Y}_t - \sum_{\hat{t}\geq t}\prod_{\hat{s}=1}^{\hat{t}-t}\frac{1}{R_{t+\hat{s}}}(SCC_{\hat{t}}-\tau_{\hat{t}})\mathcal{E}_{\hat{t}}\frac{\partial\ln(\mathcal{E}_{\hat{t}})}{\partial\ln(A_{jt})} + \frac{1}{R_{t+1}}\left[\xi_{jt+1}\Pi_{jt+1} + \sum_i \xi_{it+1}\Pi_{it+1}g_{it+1}\varphi_{ijt+1}\right], \quad \text{(A.65)}$$

which gives us Equation (51) of Proposition 3. 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 (25) in each period.

I describe first-best policy in detail in Appendix A.12, so one can verify by inspection that the second-best innovation subsidies of Proposition 3 recover those of the first-best whenever $\tau_t = SCC_t$. □

A.11 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 (51), we have

$$\tilde{\xi}_{jt} = \frac{S_{jt}-\mathcal{T}_{jt}}{\gamma-1} + \frac{\mathcal{Y}_{t+1}/\mathcal{Y}_t}{R_{t+1}}\left[\tilde{\xi}_{jt+1} + \sum_i \tilde{\xi}_{it+1}g_{it+1}\varphi_{ijt+1}\right]. \quad \text{(A.66)}$$

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.[A.3] More generally, we would have $g_y \leq g$ as damages could worsen with continued carbon pollution.

[A.3]This is the steady-state under the parametric assumptions of Sections 2.5 and C.1.

Then, evaluating Equation (A.66) in matrix notation at the steady-state, we have

$$\tilde{\xi}' = \frac{1}{\gamma - 1}[S' - \mathcal{T}']\Big((1 - \tilde{R}^{-1})\mathbf{I} - g\tilde{R}^{-1}\boldsymbol{\varphi}\Big)^{-1}. \tag{A.67}$$

Normalizing by the common innovation wedge gives us Equation (53) 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

$$\hat{\xi}' = [S' - \mathcal{T}']\Big(\mathbf{I} - \frac{g\tilde{R}^{-1}}{1 - \tilde{R}^{-1}}\boldsymbol{\varphi}\Big)^{-1} = [S' - \mathcal{T}']\sum_{t \geq 0}\Big(\frac{g\tilde{R}^{-1}}{1 - \tilde{R}^{-1}}\boldsymbol{\varphi}\Big)^{t}. \tag{A.68}$$

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 (53), I claim that the research allocation of a myopic Planner will maximize contemporaneous output, while that of a perfectly patient Planner will maximize growth. To see this, first note that Equations (23) and (A.53) imply that $\hat{\xi} \propto s \propto \big(\chi_j \phi_j\big)^{-\frac{1}{\eta}}$. 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

$$\begin{aligned} \max_{\{A'_j, s_j\}} \mathcal{Y} - \sum_j r_j \Lambda_j \;\; s.t. \\ \sum_j s_j = \mathcal{S} \\ \ln(A'_j) = \ln(\gamma) z_j + \ln(A_j), \end{aligned} \tag{A.69}$$

where I have removes time subscripts to reflect the static nature of the myopic Planner's problem. Using similar arguments from Appendix A.10, 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.54) with respect to relative technology gives us that

$$\sum_i \big(\chi_i \phi_i\big)^{-\frac{1}{\eta}} \varphi_{ij} = 0. \tag{A.70}$$

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

□

A.12 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*

$$\max_{\{c_t,\{\Lambda_{jt},\{\ell_{j\iota t}\},A_{jt},s_{jt}\}\}} \sum_{t\geq 0} \frac{1}{(1+\rho)^t} u(c_t) \;\; s.t. \tag{A.71}$$

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

$$L = \sum_j \int_0^1 \ell_{j\iota t} d\iota : \;\; \varpi_{\ell t}$$

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

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

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

$$\frac{u_t'}{(1+\rho)^t} = \varkappa_t, \tag{A.72}$$

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

$$\frac{\partial \mathcal{Y}_t}{\partial \Lambda_{jt}} = r_j - \omega_j \sum_{\hat{t}\geq t} \frac{\mathcal{Y}_{\hat{t}}}{(1+\rho)^{\hat{t}-t}} \frac{u_{\hat{t}}'}{u_t'} \frac{\partial \ln(\Omega_{\hat{t}})}{\partial \mathcal{E}_t}. \tag{A.73}$$

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

$$\frac{\partial \mathcal{Y}_t}{\partial \ell_{j\iota t}} = \frac{\varpi_{\ell t}}{u_t'/(1+\rho)^t}. \tag{A.74}$$

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

$$\frac{\epsilon_{jt}}{A_{jt}} = \frac{u_t'}{(1+\rho)^t} \frac{\partial \mathcal{Y}_t}{\partial A_{jt}} + \frac{\epsilon_{jt+1}}{A_{jt}} + \sum_i \epsilon_{it+1} \ln(\gamma)\chi_i s_{it+1}^{\eta} \frac{\partial \phi_{it+1}}{\partial A_{jt}}. \tag{A.75}$$

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

$$\varpi_{st} = \epsilon_{jt} \ln(\gamma)\chi_j \eta s_{jt}^{\eta-1} \phi_{jt}. \tag{A.76}$$

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.73), the Planner sets the carbon price according to

$$\tau_t = -\sum_{\hat{t}\geq t} \prod_{\hat{s}=1}^{\hat{t}-t} \frac{1}{R_{t+\hat{s}}} \mathcal{Y}_{\hat{t}} \frac{\partial \ln\left(\Omega_{\hat{t}}\right)}{\partial \mathcal{E}_t}, \tag{A.77}$$

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.76). We have

$$\begin{aligned} \chi_j \eta s_{jt}^{\eta-1} \phi_{jt} \xi_{jt} \Pi_{jt} &= w_{st} \\ \chi_j \eta s_{jt}^{\eta-1} \phi_{jt} \epsilon_{jt} &= \frac{\varpi_{st}}{\ln\left(\gamma\right)}. \end{aligned} \tag{A.78}$$

The Planner can set the wage for scientists according to

$$w_{st} = \frac{\varpi_{st} \ \ln\left(\gamma\right)}{u_t'/(1+\rho)^t}. \tag{A.79}$$

Plugging this into the two research conditions (A.78), we can see that the incentives for innovation will be efficient if the Planner sets innovation subsidies according to

$$\xi_{jt}\Pi_{jt} = \frac{\epsilon_{jt}}{u_t'/(1+\rho)^t}. \tag{A.80}$$

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.75), we get the recursion

$$\xi_{jt}\Pi_{jt} = (1-\alpha) S_{jt} \mathcal{Y}_t + \frac{1}{R_{t+1}} \left[\xi_{jt+1}\Pi_{jt+1} + \sum_i \xi_{it+1}\Pi_{it+1} g_{it+1} \varphi_{ijt+1}\right]. \tag{A.81}$$

Finally, the Planner sets the lump-sum tax D_t to balance the government's budget (25) in each period.

B Extensions

B.1 Semi-Endogenous Growth

This appendix describes how the results of the main text carry over to the case of semi-endogenous growth. The model of Section 2 will remain the same but for two changes. First, rather than homogeneous of degree zero, the spillover functions will be homogeneous of degree $-\zeta < 0$.[B.1] Second, the supply of scientists will grow at rate n to satisfy the well-known property of semi-endogenous growth models that economic growth requires population growth (Jones, 2022). From here on, it will be useful to refer to the proportion of scientists allocated to a technology $\bar{s}_{jt} \equiv s_{jt}/\mathcal{S}_t$.

The homogeneity assumption implies that we can write the spillover functions as

$$\phi_{jt} = A_{Jt-1}^{-\zeta} \phi_{jt}(\bar{A}_{1t-1}, \bar{A}_{2t-1}, \dots, 1). \tag{B.1}$$

I will denote the spillover function evaluated in relative technology by $\bar{\phi}_{jt}$. This implies that relative spillovers can always be evaluated in terms of relative technology: $\phi_{it}/\phi_{jt} = \bar{\phi}_{it}/\bar{\phi}_{jt}$.

Along a balanced growth path, all technologies grow at the same rate g, so relative technologies achieve a fixed point. Therefore, we have that

$$z = \chi_j \bar{s}_j^{\eta} \mathcal{S}_t^{\eta} A_{Jt-1}^{-\zeta} \bar{\phi}_j, \tag{B.2}$$

so it must be that $(1+n)^{\eta} = (1+g)^{\zeta}$ or $g \approx \frac{\eta}{\zeta} n$.

As before, we will define sufficient statistic matrices using Definitions 4 and 5 to summarize the economy's spillover structure and substitution patterns in production. Proposition 1 follows immediately from the fact that relative spillovers only depend on relative technologies. Proposition 2, and therefore Corollaries 1 and 2, also follow, but this requires some additional work as I describe below.

Proof. The complication, relative to the endogenous growth case, is that there is now a new state variable describing the absolute level of fishing out. I will denote this additional state variable by $\delta_t \equiv \mathcal{S}_t^{\eta}(1+n)^{\eta} A_{Jt}^{-\zeta}$. Fortunately, we can still write the economy as a dynamic system that achieves a fixed point along an interior balanced growth path:

$$\big(\ln(\bar{A}_t), \ln(\delta_t)\big) = \mathcal{H}\Big(\bar{s}_t\big(\ln(\bar{A}_{t-1})\big), \ln(\bar{A}_{t-1}), \ln(\delta_{t-1})\Big). \tag{B.3}$$

[B.1] The ongoing work of Donald and Restrepo (2025) examines the implications and empirical support for this assumption.

This dynamic system is described by the following system of equations:

$$\begin{aligned} \ln(\bar{A}_{it}) &= \ln(\gamma)\left[\chi_i \bar{s}_{it}^{\eta}\bar{\phi}_{it} - \chi_J \bar{s}_{Jt}^{\eta}\bar{\phi}_{Jt}\right]\delta_{t-1} + \ln(\bar{A}_{it-1}) \qquad \text{(B.4)} \\ (1-\eta)\left[\ln(\bar{s}_{it}) - \ln(\bar{s}_{Jt})\right] &= \ln(\chi_i/\chi_J) + \ln(\xi_i/\xi_J) + \ln(\bar{\phi}_{it}/\bar{\phi}_{Jt}) + \ln(p_{it}/p_{Jt}) + \ln(Y_{it}/Y_{Jt}) \\ \sum_i \bar{s}_{it} &= 1 \\ \ln(\delta_t) &= \eta \ln(1+n) + \ln(\delta_{t-1}) - \zeta \ln(\gamma)\chi_J \bar{s}_{Jt}^{\eta}\delta_{t-1}\bar{\phi}_{Jt}. \end{aligned}$$

When evaluating elasticities at the steady-state, we have that $\frac{\partial \ln(\bar{A}_{it})}{\partial \ln(\delta_{t-1})} = 0$, so local to the steady-state, the linear system can ignore the level of fishing out and be described with the same formula as in the endogenous growth model. Note that this also depends on the fact that the equilibrium allocation of scientists only depends on relative technology.

□

So how does semi-endogeneity matter? We have seen that the *absolute* level of fishing out does not matter (locally), but the *relative* level of fishing out still has an impact that operates through the spillover network. To see this, we can write spillover functions as $\phi_{jt} = \tilde{\phi}_{jt}/A_{jt-1}^{1+\zeta}$, where $\tilde{\phi}_{jt}$ denotes a constant returns gross spillover function. This function in turn generates a gross spillover network $\tilde{\boldsymbol{\varphi}}_t$.

As a thought exercise, suppose we read Bloom et al. (2020) and learn that $\zeta > 0$ but do not change our beliefs about the gross spillover network $\tilde{\boldsymbol{\varphi}}_t$. Then our updated spillover network would be $\boldsymbol{\varphi}_t = \tilde{\boldsymbol{\varphi}}_t - (1+\zeta)\mathbf{I}$. That is, we would add negative mass to the diagonal of the spillover network.

For the above results, we need to look at the spillover matrix. By Equation (38), we have that $\Phi_{ijt} = \varphi_{Jjt} - \varphi_{ijt}$, so the off-diagonal components would not change with our semi-endogenous update. The diagonal components would equal $\tilde{\varphi}_{Jit} + \sum_{j \neq i} \tilde{\varphi}_{ijt} + \zeta$, so we would now be adding positive mass to the diagonal. That is, we would add $\zeta \mathbf{I}$ to our previous estimate of the spillover matrix.

For optimal policy, Proposition 3 stills applies because it does not depend on the homogeneity of the spillover functions. The steady-state composition of research effort is where semi-endogeneity begins to matter. For expositional simplicity, I will assume an efficient price on carbon. Following the same steps as in Appendix A.11,[B.2] we can derive

$$((1-\tilde{R}^{-1}) + g\tilde{R}^{-1}\zeta)(\hat{\xi}' - S') - g\tilde{R}^{-1}\hat{\xi}'(\boldsymbol{\varphi} + \zeta\mathbf{I}) = \vec{0}', \qquad \text{(B.5)}$$

which has the same interpretation as in Section 4.2, with some adjustment for the semi-endogenous setting. With a myopic Planner, i.e. $\tilde{R}^{-1} \to 0$, innovation incentives are set according to Domar weights: $\hat{\xi} = S$. This maximizes the contemporaneous benefit of innovation by focusing exclusively on the immediate impact on output.

[B.2] A slight difference is that the growth adjusted discount rate $\tilde{R}_t$ now includes $(1+n)$ in the denominator if L_t is also growing by n. The baseline innovation wedge also adjusts to $(\gamma-1)\big((1-\tilde{R}^{-1}) + g\tilde{R}^{-1}\zeta\big)$.

With a longtermist Planner, i.e. $\tilde{R}^{-1} \to 1$, innovation incentives satisfy $\zeta(\hat{\xi}' - S') - \hat{\xi}'(\boldsymbol{\varphi} + \zeta \mathbf{I}) = \vec{0}'$. With $\zeta = 0$, this would set innovation incentives according to eigenvector centrality, which would maximize the growth rate. But with $\zeta > 0$, the growth rate is pinned down by population growth. Instead, the longtermist Planner maximizes the level of output along the balanced growth path, which is consistent with the core implication of semi-endogenous growth that policy has *level*, rather than *growth*, effects. This is why income shares show up with semi-endogenous growth; they contain information about the best technology portfolio to maximize the level effect.

To see why this innovation policy maximizes the level of output along the balanced growth path, first note that $\hat{\xi} = \bar{s}$ in the steady-state. Then constant growth implies that steady-state scientist shares follow

$$\bar{s}_j = \frac{\left(\chi_j \bar{\phi}_j\right)^{-\frac{1}{\eta}}}{\sum_i \left(\chi_i \bar{\phi}_i\right)^{-\frac{1}{\eta}}}. \tag{B.6}$$

Plugging this into Equation (B.2) gives us that

$$A_{Jt} = \Big(\frac{\mathcal{S}_1^{\eta}}{z}\Big)^{\frac{1}{\zeta}} \Big(\sum_j \left(\chi_j \bar{\phi}_j\right)^{-\frac{1}{\eta}}\Big)^{-\frac{\eta}{\zeta}} (1+n)^{\frac{\eta}{\zeta} t}, \tag{B.7}$$

and of course $A_{jt} = \bar{A}_j A_{Jt}$. The question is what choice of relative technology maximizes the level of output along the balanced growth path. The above equation gives us a mapping from relative technology to output, so differentiating (and removing time subscripts), we have

$$\frac{\partial \ln(\mathcal{Y})}{\partial \ln(\bar{A}_j)} = \sum_i \frac{\partial \ln(\mathcal{Y})}{\partial \ln(A_i)} \frac{\partial \ln(A_i)}{\partial \ln(\bar{A}_j)} = S_j + \frac{1}{\zeta} \sum_i \bar{s}_i \varphi_{ij} = 0. \tag{B.8}$$

All of these conditions will be satisfied if innovation incentives are set according to $\zeta(\hat{\xi}' - S') - \hat{\xi}'(\boldsymbol{\varphi} + \zeta \mathbf{I}) = \vec{0}'$. Thus, the longtermist Planner maximizes the level of output along the balanced growth path.

B.2 Learning-by-Doing

In the main text, I focus on productivity growth through research, but in this appendix, I will allow for learning-by-doing, where productivity growth also comes from simply producing a good, or "doing". It is commonly thought that learning by doing will play a central role in the clean transition and that it explains much of the cost declines of batteries and renewables observed thus far (Nemet, 2019; Arkolakis and Walsh, 2023; Hong, 2025). By including this mechanism in my model, I clarify the subtle sense in which this is true. Most notably, learning-by-doing cannot solve the problem of technological lock-in.

To start, I will modify the idea production function as follows

$$z_{jt} = \chi_j s_{jt}^{\eta} \phi_j(\{A_{it-1}\}) \mathcal{Q}_j(\{Y_{it}\}), \tag{B.9}$$

where $\mathcal{Q}_j(.)$ is a learning function that stipulates how productivity growth depends on the physical production of goods. Here I am making two substantive assumptions. First, learning can only take place with some amount of research. That is, allocating zero scientists to a technology implies zero productivity growth, independent of learning. Second, I assume that the learning functions are homogeneous of degree zero. This assumption is analogous to the homogeneity assumption for the spillover functions in that it allows for balanced growth with a fixed supply of scientists. Beyond these assumptions, this approach is very general.

Now to consider how learning-by-doing affects the direction of innovation, we can define the following matrix:

Definition B.1 (Learning Matrix $\boldsymbol{\mathcal{L}}$). *A $J-1 \times J-1$ matrix with elements*

$$\mathcal{L}_{ijt} \equiv \frac{\partial \ln\left(\mathcal{Q}_{it}/\mathcal{Q}_{Jt}\right)}{\partial \ln\left(Y_{jt}/Y_{Jt}\right)}. \tag{B.10}$$

This matrix tells us how relative learning changes in response to changes in relative demand. Using this matrix, we can modify the amplification and transition matrices as follows:

$$\boldsymbol{\mathcal{M}} = \left[\boldsymbol{\Phi} - \eta(1-\alpha)\left((\mathbf{I} + \frac{1}{\eta}\boldsymbol{\mathcal{L}})\boldsymbol{\Sigma} - \mathbf{I}\right)\right]^{-1} \tag{B.11}$$

$$\boldsymbol{\mathcal{J}} = \left[(1-\eta)\mathbf{I} - g\eta(1-\alpha)\left((\mathbf{I} + \frac{1}{\eta}\boldsymbol{\mathcal{L}})\boldsymbol{\Sigma} - \mathbf{I}\right)\right]^{-1}\left[(1-\eta)\mathbf{I} - g\boldsymbol{\Phi}\right]. \tag{B.12}$$

Therefore, the effect of learning-by-doing, as summarized by the learning matrix $\boldsymbol{\mathcal{L}}$, loads entirely on the substitution matrix. This is because the substitution matrix tells us how relative quantities change in response to changes in relative prices, which now has an effect on both learning and market size. This implies that, to a first order, learning-by-doing is isomorphic to market size effects because one could devise a set of substitution patterns that replicates the first-order impact of learning-by-doing.

To make this isomorphism more clear, consider the natural case where learning-by-doing depends on demand relative to output: $\mathcal{Q}_{jt} = Y_{jt}/\mathcal{Y}_t$. In words, a technology learns more when it produces more relative to the size of the economy. This implies a learning matrix equal to the identity: $\boldsymbol{\mathcal{L}} = \mathbf{I}$. Therefore, in this example, learning-by-doing is as if there are simply higher elasticities of substitution because the diagonal of the substitution matrix will expand by the factor $(\eta+1)/\eta$.

For optimal policy, learning-by-doing creates a new source of externality. Now that production itself creates externalized productivity gains, the Planner needs to both directly subsidize production and modify innovation subsidies to reflect this. Denote production subsidies by $\check{\xi}_{jt}$. To ease exposition, I will assume that carbon is priced efficiently, so the Planner is able to achieve the first best. Furthermore, learning-by-doing generates a new innovation network that we can summarize with the following matrix:

Definition B.2 (Learning Network $\boldsymbol{\mathcal{G}}$)**.** *A $J \times J$ matrix with elements*

$$\mathcal{G}_{ijt} \equiv \frac{\partial \ln \left(\mathcal{Q}_{it}\right)}{\partial \ln \left(Y_{jt}\right)}. \tag{B.13}$$

This matrix tells us how learning in each technology depends on production of each technology-specific good. To implement the first-best as an equilibrium, production subsidies must follow

$$\check{\xi}_{jt} = 1 + (\gamma - 1)(1 - \alpha)\frac{\sum_i \xi_{it} S_{it} g_{it} \mathcal{G}_{ijt}}{S_{jt}}. \tag{B.14}$$

That is, production is subsidized so as to internalize the productivity benefit of learning, which is reflected in the second term. Next, innovation subsidies must follow

$$\left(\xi_t \circ \Pi_t\right)' = \left[(1-\alpha)S_t' + \frac{1}{R_{t+1}}\left(\xi_{t+1} \circ \Pi_{t+1}\right)'\left(\mathbf{I} + \boldsymbol{D}(g_{t+1})\boldsymbol{\varphi}_{t+1}\right)\right]\left(\mathbf{I} - (1-\alpha)\boldsymbol{D}(g_t)\boldsymbol{\mathcal{G}}_t\right)^{-1}, \tag{B.15}$$

which modifies the formula of Proposition 3 to include learning-by-doing. The difference comes in the form of a Leontief inverse which accounts for the full network effect of additional learning. That is, an improvement in one technology sets off a cascade of learning by increasing its own production, generating learning and production increases in other technologies, and so on. These effects accumulate in a Leontief inverse that then multiplies the direct benefit of technology from contemporaneous output growth and future spillovers.

B.3 General Input-Output Structure

My baseline model allows for a very general input-output structure in terms of the mapping from technology-specific goods to final output. However, I have assumed that technology-specific inputs are produced using the final good. In this appendix, I will relax this assumption and allow for inputs to be produced using arbitrary functions of the technology-specific goods. I will then show how one needs to adjust the formulas of Propositions 1 and 2 to accommodate this case.

Each input is produced by a competitive firm who solves the problem

$$\max_{\{Y_{ijt}\}} r_{it}\Lambda_{it}\left(\{Y_{ijt}\}\right) - \sum_j p_{jt} Y_{ijt}, \tag{B.16}$$

where $\Lambda_{it}(.)$ is a constant returns function of the technology-specific goods, and Y_{ijt} is the quantity of good j used in the production of inputs for good i. Therefore, there are an additional set of resource constraints which require

$$Y_{jt} = Y_{yjt} + \sum_i Y_{ijt}, \tag{B.17}$$

where Y_{yjt} is the quantity of good j used in the production of the final good.[B.3] For our purposes,

[B.3] With this formulation, all production of the final good is consumed.

we will focus on the cost functions implied by this problem, denoted by

$$r_{it} = C_{it}(\{p_{jt}\}). \tag{B.18}$$

That is, the input cost for good i can be written in terms of the prices of the various technology-specific goods. This nests the baseline setup where inputs are produced using only the final good. In that case, the cost functions only differ up to a scale constant. For expositional convenience, I will assume that policy makers levy ad valorem carbon prices $\tilde{\tau}_{jt}$, so that relative prices follow

$$\frac{p_{jt}}{p_{Jt}} = \Big(\frac{C_{jt}}{C_{Jt}}\Big)^{\alpha}\Big(\frac{1+\tilde{\tau}_{jt}}{1+\tilde{\tau}_{Jt}}\Big)^{\alpha}\Big(\frac{A_{jt}}{A_{Jt}}\Big)^{\alpha-1}. \tag{B.19}$$

To consider how this more general input-output setup affects the direction of innovation, we can define the following matrix:

Definition B.3 (IO Perturbation Matrix $\boldsymbol{\mathcal{P}}$)**.** *A $J-1 \times J-1$ matrix with elements*

$$\mathcal{P}_{ijt} \equiv \frac{\partial \ln\left(C_{it}/C_{Jt}\right)}{\partial \ln\left(p_{jt}/p_{Jt}\right)} \tag{B.20}$$

This matrix tells us how relative input costs change in response to relative prices, and it allows us to characterize the Jacobian of log relative prices with respect to log relative technology as $-(1-\alpha)\big(\mathbf{I}-\alpha\boldsymbol{\mathcal{P}}\big)^{-1}$.

Following the same arguments as Propositions 1 and 2, we now have

$$\boldsymbol{\mathcal{M}} = \big[\boldsymbol{\Phi} - \eta(1-\alpha)\big(\boldsymbol{\Sigma}-\mathbf{I}\big)\big(\mathbf{I}-\alpha\boldsymbol{\mathcal{P}}\big)^{-1}\big]^{-1} \tag{B.21}$$

$$\boldsymbol{\mathcal{J}} = \big[(1-\eta)\mathbf{I} - g\eta(1-\alpha)\big(\boldsymbol{\Sigma}-\mathbf{I}\big)\big(\mathbf{I}-\alpha\boldsymbol{\mathcal{P}}\big)^{-1}\big]^{-1}\big[(1-\eta)\mathbf{I} - g\boldsymbol{\Phi}\big]. \tag{B.22}$$

Therefore, the insights of the main text extend to the case of a general input-output structure. The only necessary adjustment is that market size effects now include changes in relative input prices, the full network effect of which is reflected in a Leontief inverse multiplying the substitution matrix net of the identity.

Finally, the optimal policy formula of Proposition 3 does not change with a general input-output structure. Due to Hulten's Theorem, the contemporaneous benefit of innovation is summarized by market size, and this remains true independent of the complexity of the underlying input-output structure.

B.4 Endogenous Extraction Costs

In the main text, I assumed that input costs r_j are exogenous, but given that I interpret dirty inputs as fossil fuels, it is useful to consider a case where input costs are endogenously determined by extraction costs. This appendix derives this case and shows that endogenous extraction costs do not change the optimal policy formulas of Proposition 3.

I will allow for a very general specification of endogenous input costs, where the quantity of the final good required to produce an array of inputs depends on the past sequence of input production. That is, the resource constraint becomes

$$\mathcal{Y}_t = c_t + \mathcal{D}_t(\{\Lambda_{j\hat{t}}\}_{\hat{t}\leq t}). \tag{B.23}$$

This clearly nests the main specification of Equation (7), but it also allows for an extraction cost schedule where each additional unit of fossil fuels becomes more difficult to extract.

Inputs are produced by a unit interval of competitive, identical firms,[B.4] so these firms solve

$$\max_{\{\Lambda_{jt}\}} \sum_{t\geq 0} \prod_{\hat{s}=1}^{t} \frac{1}{R_{\hat{s}}} \Big[\sum_j r_{jt}\Lambda_{jt} - \mathcal{D}_t \Big]. \tag{B.24}$$

These firms are price takers, so their optimality conditions follow

$$r_{jt} = \sum_{\hat{t}\geq t} \prod_{\hat{s}=1}^{\hat{t}-t} \frac{1}{R_{t+\hat{s}}} \frac{\partial \mathcal{D}_{\hat{t}}}{\partial \Lambda_{jt}}. \tag{B.25}$$

Note that this, plus the social cost of carbon, is the Planner's shadow price for inputs. Therefore, by Pigou Principle reasoning, endogenous input costs do not add anything to the logic of Proposition 3. Put differently, it is not the underlying structure generating private input costs that matters for policy; it is just the wedge between private and social input costs that matters.

C Details on Calibration & Simulation

C.1 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

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

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

[B.4] I abstract from market power or extraction externalities because these would introduce additional market failures that are orthogonal to the main arguments of the paper.

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 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.[C.1] 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.

A major advantage of the climate specification of Equation (C.1) 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

$$\mathcal{C}_{1t} = \psi_p \mathcal{E}_t + \mathcal{C}_{1t-1}. \tag{C.2}$$

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

$$\mathcal{C}_{2t} = (1-\psi_p)\psi_0 \mathcal{E}_t + \psi \mathcal{C}_{2t-1}. \tag{C.3}$$

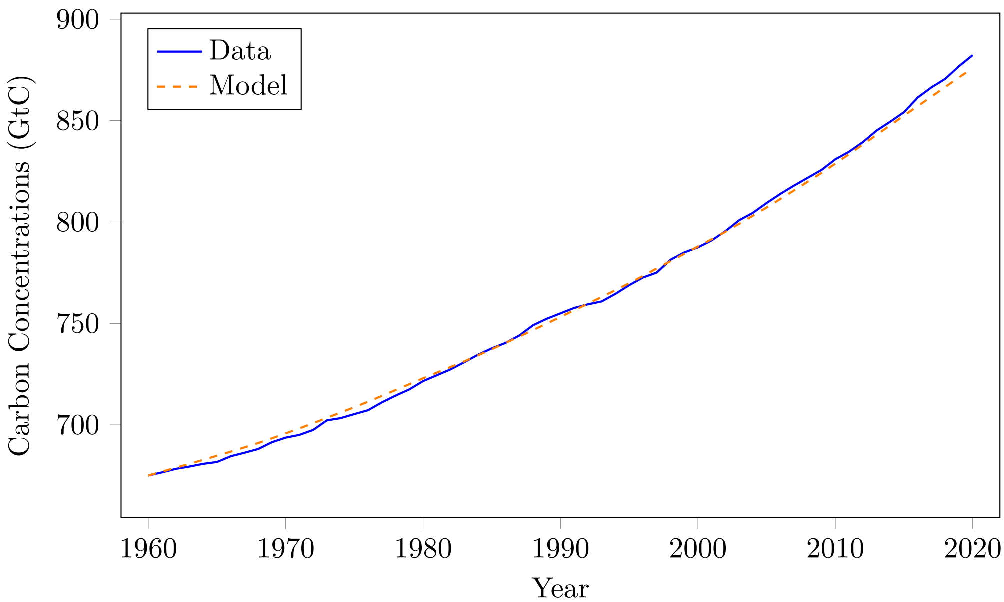
The sum of these two components is therefore atmospheric carbon concentration: $\mathcal{C}_t = \mathcal{C}_{1t} + \mathcal{C}_{2t}$. Writing the state of the climate recursively allows for a simpler representation of the climate system as one does not need to keep track of the entire history of emissions. 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.

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 match the impulse response functions of a carbon pulse reported in Joos et al. (2013). Using a model intercomparison, they find that 41% of the original carbon pulse remains in the atmosphere after 100 years and 25% remains after 1000 years. This implies values of $\psi_0 = 0.308$ and $\psi = 0.998$ in the context of my parametric model. Figure C.1 tests the model's ability to match the relationship between carbon emissions and atmospheric carbon concentrations throughout the 1960-2020 period and finds a close correspondence between the model's predictions and the data. 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

[C.1] 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.

Loa Observatory.[C.2]

Figure C.1: Climate Model Matches Historic Relationship of Emissions and Carbon Concentrations



Notes: Match of model predicted atmospheric carbon concentrations with data when $\psi_p = 0.2$, $\psi_0 = 0.308$, and $\psi = 0.998$. 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

$$\Omega_t = \exp\big(-\varrho(\mathcal{C}_t - \bar{\mathcal{C}})\big), \tag{C.4}$$

where $\varrho > 0$ is a scale parameter stipulating the semi-elasticity of final output with respect to atmospheric carbon concentrations.[C.3] 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 (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 (Bilal and Känzig, 2024), 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).[C.4]

[C.2] 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.

[C.3] 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).

[C.4] 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; Donald and Hanley, 2025). A thorough consideration of climate uncertainty is beyond the scope of this paper, but I will capture some

C.2 Variable Patent Lengths

In the quantitative results of the main text, a model period is one year, implying that purchasing an innovation grants a one year long monopoly right. In reality, patents typically grant monopoly rights for longer periods, so this raises a concern that the model may not match the data when patent protection lasts longer than one period. In this appendix, I address that concern.

To do so, I extend the model to allow for patents that expire randomly at the end of a period with probability $1 - \Psi$. This implies the following recursive pricing formula for innovation:

$$V_{jt} = \Pi_{jt} + \frac{\Psi}{R_{t+1}} V_{jt+1}. \tag{C.5}$$

Note that this nests the baseline model when $\Psi = 0$, and varying Ψ allows for straightforward variation in expected patent duration. In this more general setting, the price of a technology-specific innovation is given by the expected present value of flow profits, rather than just the contemporaneous flow of profit. The model is otherwise the same, and in particular, the steady-state will be exactly the same as relative innovation prices will equal relative income shares. This formulation brings us closer to a Schumpeterian growth model like that of Aghion and Howitt (1992), but here I am not allowing the loss of a patent to be endogenous. That is, I abstract from the possibility that future innovators steal your position as the productivity leader.

Figure C.2 plots the validation exercise of Section 5.3 for three expected patent durations: one, five, and ten years. Parameter values are otherwise the same as in Section 5.3, with US climate policy fixed at its 2010s level forever. Because their is now a private dynamic decision, I take the Nordhaus rate of pure time preference $\rho = 0.015$ because this is a purely descriptive exercise. For the case of technological lock-in, I assume that agents expect the economy to continue moving towards a dirty steady-state.

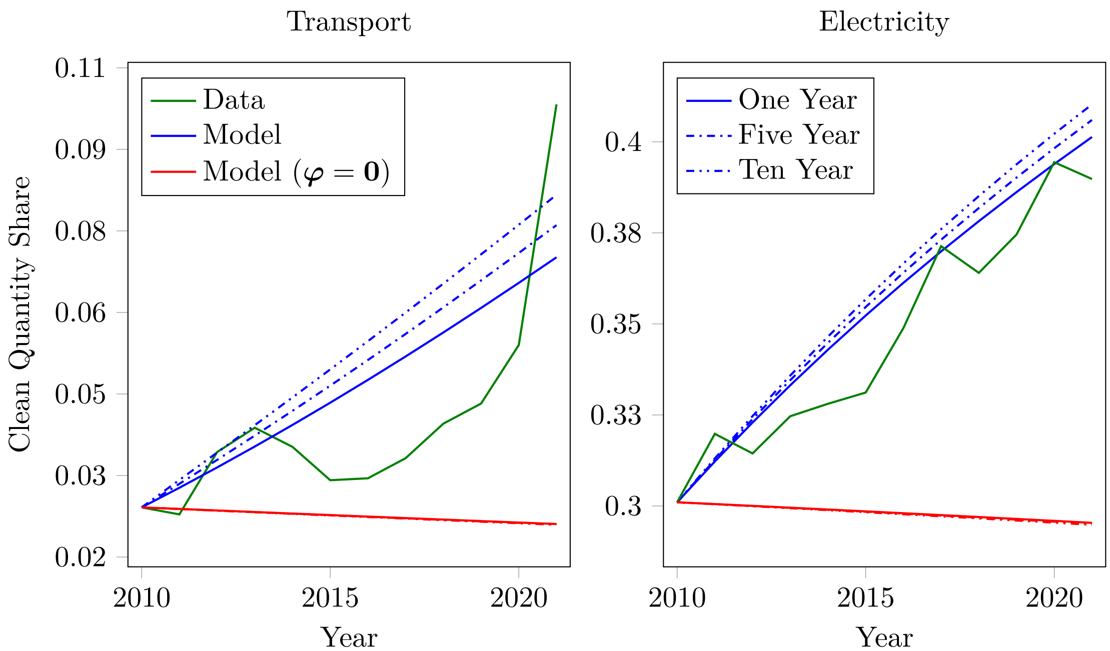
As we can see, changing the expected patent length does not substantively change the model's ability to match clean progress in the 2010s as an untargeted moment. Longer patent lengths slightly speed up clean convergence in both sectors, but all three lines are within the bounds of empirical reasonableness. Notably, shutting down the spillover network still leads to technological lock-in, independent of the patent duration. In fact, the model's predicted path for the direction of innovation in the absence of spillovers is so similar across the three patent durations that the reader cannot see them distinctly. This exercise shows that assuming a year long model period, and therefore patent length, is simply a matter of convenience that does not substantively change the quantitative implications.

C.3 Reduced Form Evidence for Cross-Technology Spillovers

This appendix explains the details of the empirical exercise described in Table 1, which shows that cross-technology spillovers lead to increased innovation, as predicted by the theory. To start, note

elements of insuring against climate risk by considering the possibility of catastrophic damages.

Figure C.2: Model Validation with Variable Patent Lengths



Notes: The solid, dash-dotted, and dash-double-dotted line represent expected patent durations of one, five and ten years. These durations correspond to Ψ equal 0, 0.8, and 0.9 respectively.

that the production function for innovation (10) implies the following estimating equation:

$$\ln\left(pats_{jt}\right) = \alpha_j + \alpha_t + \mu_1 \ln\left(R\&D_{jt}\right) + \mu_2 \ln\left(spill_{jt}\right) + \varepsilon_{jt}. \tag{C.6}$$

That is, new ideas in the form of patents are log-linear in research effort and cross-technology spillovers. Note that the model describes research effort in terms of scientists s_{jt}, but this is proportional to $R\&D_{jt} \equiv w_{st}s_{jt}$, so the difference is absorbed in the time fixed effect.

As the model makes explicit, shocks to other technologies will affect total research output through several channels, including both spillover and market size effects. Because of this, advances in other, substitute technologies could lower innovation effort on net even in the presence of spillovers. However, market size effects influence the profitability of innovation, which is reflected in the equilibrium demand for scientists. Therefore, by controlling for research effort, we can isolate spillover effects and measure how advances in other technologies increase the research productivity of a given level of scientific effort.

I measure innovation output using patents, so in all of the exercises below, I use both citation weighted and raw counts. To account for the fact that patents from earlier periods should mechanically have more citations, I follow Arora et al. (2023) and set the citation weight as the number of citations relative to the application year by 3-digit CPC code average.[C.5] Furthermore, I add one to every patent's citation count, so that patents that do not receive any citations are still included

[C.5] Patents with multiple CPC codes are normalized by the average of averages across their CPC codes.

in the sample.

For my measure of knowledge stocks, I follow Bloom et al. (2013) and Aghion et al. (2016) in using the perpetual inventory method. That is,

$$A_{jt} = pat_{it} + (1-\delta)A_{jt-1}, \tag{C.7}$$

where $\delta = 0.15$ is the decay rate of patent knowledge. I begin these technology stocks by assuming a steady-state in 1972, so $A_{j,72} = (1+g)pat_{j,72}/(\delta+g)$, where $g = 0.02$. My assumed steady-state takes place before the unexpected rise in energy prices from the 1970s oil shocks, after which there was more concerted effort to invest in energy-saving technology (Hassler et al., 2021). My regressions also only consider years 1980 onward, further reducing the possibility of biasing the results with a steady-state assumption.

These knowledge stocks generate cross-technology spillovers through the network, so that measured cross-technology spillovers follow

$$\ln(spill_{it}) = \sum_{j \neq i} \tilde{\varphi}_{ij} \ln(A_{jt-1}). \tag{C.8}$$

This excludes own-technology spillovers both to avoid endogeneity from serially correlated shocks and to focus on the model mechanism of enhanced research productivity from building on the knowledge of *other* technologies. To focus on the role of cross-technology spillovers for the development of climate-relevant technologies, I do not include the general technology in the regressions of Table 1, though it is of course included in the computation of spillovers.

For R&D spending, I use firm-level R&D expenditures from Compustat. To assign these expenditures to technologies, I first assign patents to firms using the crosswalks of both Arora et al. (2021) and Kogan et al. (2017). Then, I split firm's R&D expenditures across technology classes using ten-year moving averages of the share of a firm's patents that are in each technology class. Aggregating theses expenditures across firms for each technology class gives a measure of total R&D spending by year. However, the theoretically relevant measure is research effort towards innovation at a point in time, so because there can be a long and variable lag between R&D spending and a patent application, the R&D measure of Equation (C.6) includes lags of R&D spending with a decay rate of δ.[C.6]

Columns (2) and (5) of Table 1 include the spillovers that a technology sends through the network, in addition to those it receives. This tests the directionality of the spillover network. According to the theory, the research productivity of a technology depends on the technology stocks of those upstream in the network, not downstream. To test this hypothesis, I include in Equation (C.6)

$$\ln(down_{it}) = \sum_{j \neq i} \tilde{\varphi}_{ji} \ln(A_{jt-1}), \tag{C.9}$$

[C.6]To project backwards indefinitely, I also assume that R&D spending is in a steady-state in 1972.

where the difference from Equation (C.8) comes from flipping the direction of the spillover elasticity. Note that this variable is still built with lagged technology stocks. Had I used lead technology stocks, the theory would predict a mechanical correlation from reverse causality. The intention here is to test whether my estimated spillover network gets the direction of spillover creation correct, and as shown in Table (1), this is largely the case.

Finally, as discussed extensively in Bloom et al. (2013), there are endogeneity concerns in this setting stemming from the fact that multiple technologies could receive beneficial shocks that are serially correlated. This would show up in my specification as a spillover when the real reason was the confounded shock. In my setting, explicit policy goals surrounding climate change could cause such a problem. For example, if a new discovery in dirty technology raises demand for fossil fuels, policy makers may respond with favorable policies for clean technology. The inclusion of fixed effects and downstream spillovers already goes some way in addressing this common shock concern, but to address it further, I follow Bloom et al. (2013) in instrumenting knowledge stocks with innovation driven by exposure to state-level R&D taxes.

Table C.1: Firm-Level Patenting from State-Level R&D Price Exposure

	Dependent Variable:	
	ln(Citations)	ln(Patents)
	(1)	(2)
ln(State-Level R&D Price)	-0.649	-0.956**
	(0.41)	(0.45)
R^2	0.0002	0.0006
Obs	60,206	60,206
Fixed Effects	Firm × Technology, Year	

Notes: Standard errors clustered by firm and technology are reported in parentheses.

To this end, I take state-specific R&D prices $\hat{\rho}_{st}$ from Lucking et al. (2019). Variation here is induced by differences in corporate tax rates and R&D tax credits across states. To derive firm-technology-level idiosyncratic exposure to state-level R&D prices, I take the 10-year moving average of the proportion of firm f's technology j inventors located in state s, $\hat{\theta}_{fjst}$, a proxy for the geographic distribution of R&D spending.[C.7] Firm-technology exposure is then $\hat{\rho}_{fjt} \equiv \sum_s \hat{\theta}_{fjst} \hat{\rho}_{st}$. We can then run the following regression to predict firm-level patenting in technology j, shown in Table C.1:

$$\ln(pat_{fjt}) = \alpha_{fj} + \alpha_t + \mu_1 \ln(\hat{\rho}_{fjt}) + \varepsilon_{fjt}. \tag{C.10}$$

I only include a firm-technology-year observation if a firm has a patent in the relevant technology class in that year. Furthermore, the availability of state-level R&D prices cuts the IV sample to

[C.7] Patents with multiple inventors are attributed fractionally across inventors, so a patent with two inventors counts as half for each of the two inventor states.

end in the year 2015.

To construct an instrument, I sum over predicted patenting by firm $\hat{pat}_{fjt}$ and construct a knowledge stock instrument using the same method as before. In this case, knowledge stocks only reflect innovations that were driven by exogenous exposure to state-level R&D prices, satisfying the exclusion restriction. I then build an instrument for cross-technology spillovers by considering the spillovers these knowledge stocks send through the network in the same manner as in Equation (C.8). As shown in Table C.2, this instrument has a strong first stage, and the resulting spillover variation remains highly predictive of research productivity, as shown in Columns (3) and (6) of Table 1.

Table C.2: First-Stage for Cross-Technology Spillover Instrument

	Dependent Variable:	
	ln(Citation-Based Spillovers)	ln(Patent-Based Spillovers)
	(1)	(2)
ln(R&D Spending)	0.005	-0.007
	(0.02)	(0.01)
ln(Spillover Instrument)	0.969***	1.022***
	(0.04)	(0.04)
R^2	0.892	0.917
Obs	144	144
Fixed Effects	Technology, Year	

Notes: Standard errors clustered by technology are reported in parentheses.

C.4 Details on Spillover Network

This appendix details some of the properties of the spillover network.[C.8] I start by examining its stability over time. To do so, I use the application year of citing patents to compute a spillover network that is specific to each of the five year bins from 1975 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

$$\tilde{\varphi}_{ijt} = \mu\tilde{\varphi}_{ijt-1} + \varepsilon_{ijt}. \tag{C.11}$$

[C.8] I abstract 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.

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 (C.11) 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 C.3.

Table C.3: Spillover Network Stability

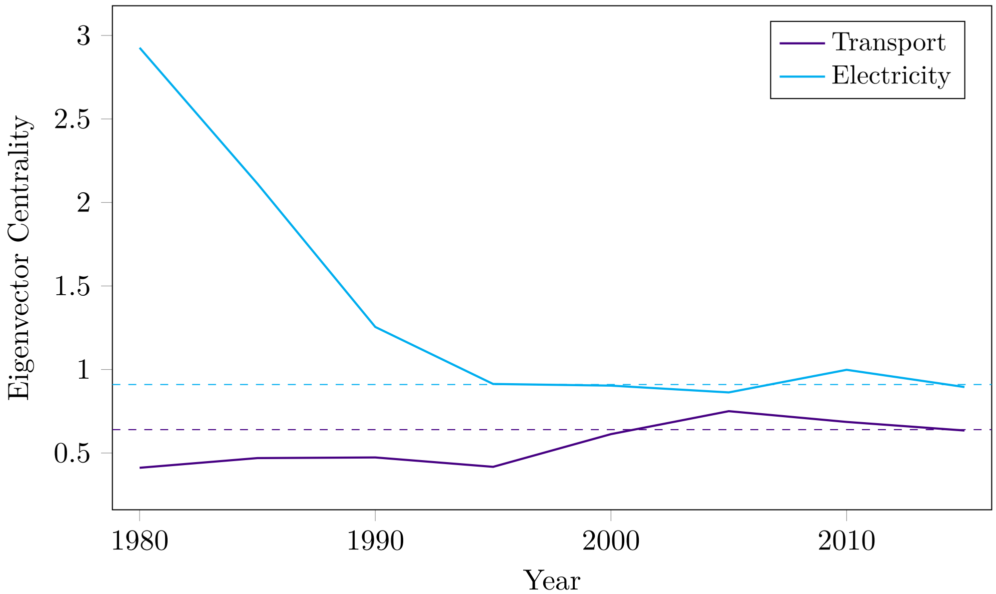
	Dependent Variable: Spillover Elasticity		
	(1)	(2)	(3)
Lagged Spillover Elasticity	0.982***	0.983***	0.983***
	(0.01)	(0.01)	(0.01)
Clean Technology Trend		-0.0001	
		(0.0002)	
Clean Transport Trend			0.00002
			(0.0002)
Clean Electricity Trend			-0.0002
			(0.0003)
R^2	0.992	0.992	0.992
Obs	200	200	200

Notes: This table presents estimates that describe the stability of the spillover network. The spillover network is computed in five year bins from 1975 to 2015. Standard errors clustered by spillover recipient are reported in parentheses.

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 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 C.3. 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.

Figure C.3: Clean Technology Centrality



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

As I show in Section 4, the centrality of technologies in the spillover network is one of the main determinants of optimal innovation policy. However, my general technology is a highly aggregated residual in that it includes all patents that are not related to transportation or electricity generation. This is a concern because, in general, the aggregation of technologies will affect the computation of centrality scores. For instance, if clean transport is rarely cited by the general technology, but the citations it does receive are from sub-technologies that are highly central, then the centrality of clean transport could be underestimated in the aggregated network.

To address this concern, I consider a spillover network where the general technology is disaggregated into 3-digit CPC codes. The definitions of the climate-related technologies are as before, and to maintain consistency with the previous approach, citations that are sent to general patents with multiple CPC codes are fractionally allocated across these codes. That is, if a general patent with two assigned CPC codes is cited, then this citation will be counted as half a citation from the perspective of the two CPC codes. This guarantees that the spillovers sent across my four climate-related technologies will be the same as in the aggregated network.

When I compute centrality in the disaggregated network, I find very little difference from the base case. In particular, clean transportation has a centrality score of 0.67 in the aggregated network and 0.73 in the disaggregated network, while clean electricity generation has centrality scores of 1 and 1.19, respectively. Therefore, the centrality scores of my clean technologies do not seem to depend on how I aggregate the non-climate-related technologies of the economy.

C.5 Details on Numerical Representations

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

$$\tilde{p}_{jt} \equiv \big(r_j + \omega_j \tau_t\big)^{\alpha} / A_{jt}^{1-\alpha}. \tag{C.12}$$

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

$$p_{jt} = \Big(\frac{1}{\alpha}\Big)^{\alpha}\Big(\frac{1}{1-\alpha}\Big)^{1-\alpha}\Big(\frac{\gamma w_{\ell t}}{\Upsilon}\Big)^{1-\alpha}\tilde{p}_{jt}. \tag{C.13}$$

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

$$\tilde{p}_{\theta t} \equiv \big(\tilde{p}_{\theta ct}^{1-\sigma} + \tilde{p}_{\theta dt}^{1-\sigma}\big)^{\frac{1}{1-\sigma}} \tag{C.14}$$

$$\tilde{P}_t \equiv \Omega_t^{-1}\big(\sum_{\theta} \nu_\theta \tilde{p}_{\theta t}^{1-\lambda}\big)^{\frac{1}{1-\lambda}}, \tag{C.15}$$

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

$$\frac{\gamma w_{\ell t}}{\Upsilon} = (1-\alpha)\frac{\mathcal{Y}_t}{L}. \tag{C.16}$$

Plugging the wage (C.16) into the factor of proportionality and combining the final pseudo price (C.15) with the fact that output is the numeraire, we have for output

$$\mathcal{Y}_t = \alpha^{\frac{\alpha}{1-\alpha}} L \tilde{P}_t^{\frac{-1}{1-\alpha}}. \tag{C.17}$$

Next, using the demand conditions for sectors (A.11) and goods within sectors (A.13), we can derive income shares for sectors and goods within sectors. We have

$$S_{\theta t} \equiv \frac{p_{\theta t} E_{\theta t}}{\mathcal{Y}_t} = \frac{\nu_\theta \tilde{p}_{\theta t}^{1-\lambda}}{\Omega_t^{1-\lambda}\tilde{P}_t^{1-\lambda}} \tag{C.18}$$

$$S_{et}^{\theta} \equiv \frac{p_{\theta et} Y_{\theta et}}{p_{\theta t} E_{\theta t}} = \frac{\tilde{p}_{\theta et}^{1-\sigma}}{\tilde{p}_{\theta t}^{1-\sigma}}, \tag{C.19}$$

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

$$\Lambda_{jt} = \frac{\alpha S_{jt} \mathcal{Y}_t}{r_j + \omega_j \tau_t}, \tag{C.20}$$

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

I define clean quantity shares by sector as $q_{ct}^{\theta} \equiv Y_{\theta ct}/(Y_{\theta ct} + Y_{\theta dt})$. To map between clean income shares and clean quantity shares, note that the demand condition within sectors (A.13) implies that

$$S_{ct}^{\theta} = \frac{p_{\theta ct} Y_{\theta ct}}{p_{\theta ct} Y_{\theta ct} + p_{\theta dt} Y_{\theta dt}} = \frac{1}{1 + \left((1 - q_{ct}^{\theta})/q_{ct}^{\theta}\right)^{\frac{\sigma-1}{\sigma}}}. \tag{C.21}$$

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

$$\begin{aligned} z_{jt} &= \chi_j \Big(\frac{R\&D_{jt}}{w_{st}} \Big)^{\eta} \phi_{jt} \\ \Rightarrow \frac{\partial \ln (z_{jt})}{\partial \ln (R\&D_{jt})} &= \eta, \end{aligned} \tag{C.22}$$

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

$$\begin{aligned} \eta \chi_j \Big(\frac{R\&D_{jt}}{w_{st}} \Big)^{\eta} \phi_{jt} &= R\&D_{jt} \\ \Rightarrow \frac{\partial \ln (R\&D_{jt})}{\partial \ln (w_{st})} &= -\frac{\eta}{1-\eta}, \end{aligned} \tag{C.23}$$

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

C.6 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 model validation exercise of Section 5.3. First, I allow for clean input subsidies $\bar{\xi}_{\theta c}$ that reduce effective input prices to $(1 - \bar{\xi}_{\theta c}) r_{\theta c}$. To determine their level, consider public spending on the subsidy as a proportion of total spending in each sector. We have

$$\frac{\bar{\xi}_{\theta c} r_{\theta c} \Lambda_{\theta ct}}{p_{\theta t} E_{\theta t}} = \frac{\bar{\xi}_{\theta c}}{1 - \bar{\xi}_{\theta c}} \frac{\alpha p_{\theta ct} Y_{\theta ct}}{p_{\theta t} E_{\theta t}} = \frac{\bar{\xi}_{\theta c}}{1 - \bar{\xi}_{\theta c}} \alpha S_{ct}^{\theta}, \tag{C.24}$$

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 5.2. 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 (C.21). Plugging these data series into Equation (C.24), we have

$$\frac{\bar{\xi}_{\theta c}}{1-\bar{\xi}_{\theta c}} \sum_{t=2011}^{T_{IF}} \alpha \hat{S}^{\theta}_{ct} = \sum_{t=2011}^{T_{IF}} \hat{cred}_{\theta ct}, \tag{C.25}$$

where T_{IF} is the end year of data for the two congressional reports. This allows me to back out the values $\bar{\xi}_{car,c} = 0.011$ and $\bar{\xi}_{elec,c} = 0.03$.

To select 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

$$\frac{(\xi_j - 1) z_{jt} \Pi_{jt}}{w_{st} \sum_j s_{jt}} = \frac{\xi_j - 1}{\xi_j} \frac{s_{jt}}{\eta \mathcal{S}}. \tag{C.26}$$

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 C.4 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

$$\frac{\xi_j - 1}{\xi_j} \sum_{t=2010}^{2015} \frac{s_{jt}}{\eta \mathcal{S}} = \sum_{t=2010}^{2015} \hat{pubrd}_{jt}. \tag{C.27}$$

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 (C.27), so I calibrate separate innovation subsidies for the versions of the model with and without spillovers. In the model with spillovers this yields $(\xi_{car,c}, \xi_{car,d}, \xi_{elec,c}, \xi_{elec,d}) = (1.005, 1.016, 1.053, 1.01)$, and in the model without spillovers this yields $(\xi_{car,c}, \xi_{car,d}, \xi_{elec,c}, \xi_{elec,d}) = (1.211, 1.005, 1.272, 1.015)$. 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.

C.7 Did a Foreign Big Push cause Solar Prices to Fall?

Throughout the paper, I focus on the United States, but this opens the concern that a foreign big push may have redirected innovation towards clean technology. To address this concern, I plot the reduction in solar photovoltaic prices per watt in Figure C.4, starting immediately after the invention of solar in the 1950s. Both series come from Our World in Data[C.9]

[C.9] The main solar price series starts in 1975, but I add a data point in 1956 from Roser (2020) to emphasize that log-linear price declines stretch all the way back to the invention of solar.

Table C.4: 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.

The most striking feature of Figure C.4 is that the price reductions of solar have been almost perfectly log-linear over the entire period, a point emphasized in Arkolakis and Walsh (2023). To start, such a steady process of improvement is more consistent with the story of this paper – that cross-technology spillovers allow for catchup growth – than a big push abruptly shifting research resources towards solar. Put differently, there is no obvious point in the last 75 years to date a foreign big push. To further argue this point, I have also plotted the share of global electricity generated by solar, a proxy for global market size. For almost the entire period, solar's share was essentially zero, meaning that market size could not have not have been incentivizing the observed improvements in solar.

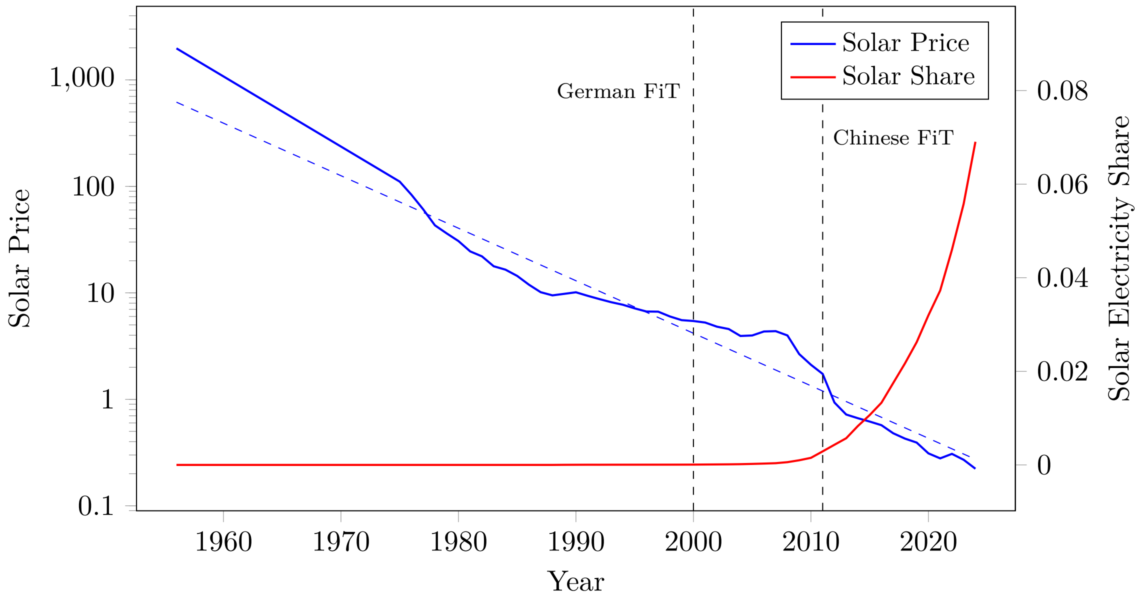
If market size effects were the only operative force, the observed improvements would have had to have been underpinned by massive policy intervention. To consider this possibility, I mark with vertical dash lines the introduction of two policies that are emphasized in this literature: the introduction of feed-in tariffs for solar electricity in Germany in 2000 and China in 2011 (Nemet, 2019). In both cases, the policy intervention took place after half a century of steady productivity improvement. Note that I am not arguing that these policies were irrelevant; the opposite is true. The argument of this paper is that cross-technology spillovers *enabled* these policies to have force because, in a world with only market size effects, such incremental policies would not have been able to spur solar adoption.

C.8 Sectoral Determinants of Convergence Speed

This appendix uses the spectral analysis of Appendix A.8 to unpack the determinants of variation in convergence speed by sector and how they relate to the properties of the spillover network. Figure C.5 shows the half-lives of convergence for each eigenstate and sector of the economy.[C.10] Proposition A.1 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

[C.10] I perform the same exercise for the case where cross-technology spillovers are doubled in Figure D.9.

Figure C.4: Solar Prices & Electricity Shares



Notes: Solar photovoltaic prices per watt are in log scale, with a dashed line of best fit. Both the price and share series come from Our World in Data. The two vertical dashed lines represent the introduction of feed-in tariffs (FiT) for solar electricity. Germany introduced their FiT in 2000, while China introduced theirs in 2011.

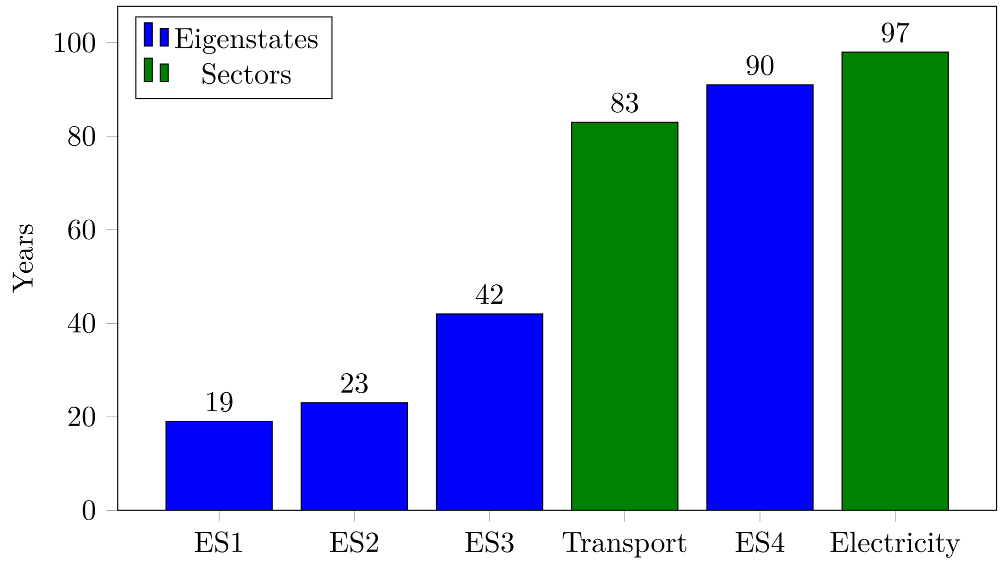
space with a distinct convergence speed, and Figure C.5 shows that the slowest eigenstate drives the speed of convergence for both sectors. That is, technology's initial condition loads primarily on the eigenstate representing the region of state space where convergence is especially slow. This eigenstate 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 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.[C.11]

Figure C.5 shows that transportation converges more quickly than electricity generation, with a difference in half-lives of about 16.9%. This is due to higher within-sector spillovers in transportation. To quantify the importance of such spillovers, Figure D.11 shows half-lives when within-sector spillovers are shut down. In that case, the spectral radius rises to 0.996, and the half-lives for transportation and electricity generation rise to 224 and 199 years. The loss of such spillovers sub-

[C.11] One may wonder whether the degree of increasing returns to innovation depends on the size of the policy reform. Figure D.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.

Figure C.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 ($\xi_c = 1.43$).

stantially slows convergence speed, but more importantly, the sectors switch their speed rankings, signifying that transportation's more rapid transition stems from within-sector spillovers.

C.9 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 C.1. Using the results described in Appendix C.5, 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 (C.1) and the damage function (C.4) gives us

$$SCC_t = \varrho \sum_{\hat{t} \geq t} \prod_{\hat{s}=1}^{\hat{t}-t} \frac{1}{R_{t+\hat{s}}} \mathcal{Y}_{\hat{t}}(\psi_p + (1-\psi_p)\psi_0 \psi^{\hat{t}-t}). \tag{C.28}$$

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

$$SCC_{1t} = \varrho \psi_p \mathcal{Y}_t + \frac{1}{R_{t+1}} SCC_{1t+1} \tag{C.29}$$

$$SCC_{2t} = \varrho(1-\psi_p)\psi_0 \mathcal{Y}_t + \frac{1}{R_{t+1}} \psi SCC_{2t+1}. \tag{C.30}$$

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 (C.2) and (C.3), 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.66), and technology $\{A_{jt}\}$ follows the law of motion (8). Note that the research condition (23) 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

$$(1-\eta)\ln(\bar{s}_{jt}) = \ln(\tilde{\Xi}_{jt}) - \eta\ln(\mathcal{V}_j) + \ln(\phi_{jt}/\phi_{Jt}), \tag{C.31}$$

where $\bar{s}_{jt} \equiv s_{jt}/s_{Jt}$ are relative scientists, $\tilde{\Xi}_{jt} \equiv \tilde{\xi}_{jt}/\tilde{\xi}_{Jt}$ 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 C.5, 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 C.5 to write the sum of future emission changes in terms of a two-dimensional recursion $\{\varsigma_{1t}, \varsigma_{2t}\}$. Equation (C.20) 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

$$\frac{\partial \mathcal{E}_{\hat{t}}}{\partial \ln(A_{jt})} = Z_{\hat{t}}\Big(\sum_{s=0}^{\hat{t}-1-t} \frac{\partial \ln(\Omega_{\hat{t}})}{\partial \mathcal{E}_{t+s}} \frac{\partial \mathcal{E}_{t+s}}{\partial \ln(A_{jt})}\Big), \tag{C.32}$$

where

$$Z_t = \frac{\mathcal{E}_t}{1-\alpha-\frac{\partial \ln(\Omega_t)}{\partial \mathcal{E}_t}\mathcal{E}_t}. \tag{C.33}$$

This implies that

$$\begin{aligned}\mathcal{T}_{jt} = \frac{\bar{\omega}_t}{1-\alpha}\frac{\partial \ln(\mathcal{E}_t)}{\partial \ln(A_{jt})}\Big(&(SCC_t - \tau_t) \\ &+ \frac{1}{R_{t+1}} Z_{t+1}\frac{\partial \ln(\Omega_{t+1})}{\partial \mathcal{E}_t}(SCC_{t+1} - \tau_{t+1}) \\ &+ \frac{1}{R_{t+1}}\frac{1}{R_{t+2}} Z_{t+2}\Big[\frac{\partial \ln(\Omega_{t+2})}{\partial \mathcal{E}_{t+1}} Z_{t+1}\frac{\partial \ln(\Omega_{t+1})}{\partial \mathcal{E}_t} + \frac{\partial \ln(\Omega_{t+2})}{\partial \mathcal{E}_t}\Big](SCC_{t+2} - \tau_{t+2}) \\ &+ \dots,\end{aligned} \tag{C.34}$$

which gives us that

$$\begin{aligned}\mathcal{T}_{jt} = \frac{\bar{\omega}_t}{1-\alpha}\frac{\partial \ln(\mathcal{E}_t)}{\partial \ln(A_{jt})}\Big(&(SCC_t - \tau_t) \\ &+ \sum_{\tilde{t}\geq 1}\sum_{t'>t}\dots\sum_{t^{(\tilde{t})}>t^{(\tilde{t}-1)}} Z_{t^{(\tilde{t})}}\frac{\partial \ln(\Omega_{t^{(\tilde{t})}})}{\partial \mathcal{E}_{t^{(\tilde{t}-1)}}}\dots Z_{t'}\frac{\partial \ln(\Omega_{t'})}{\partial \mathcal{E}_t}\prod_{\hat{s}=1}^{t^{(\tilde{t})}-t}\frac{1}{R_{t+\hat{s}}}(SCC_{t^{(\tilde{t})}} - \tau_{t^{(\tilde{t})}})\Big).\end{aligned} \tag{C.35}$$

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:

$$\varsigma_{1t} = \frac{1}{R_{t+1}}\big(- Z_{t+1}\varrho\psi_p(SCC_{t+1} - \tau_{t+1}) + \varsigma_{1t+1} - Z_{t+1}\varrho\psi_p[\varsigma_{1t+1} + \varsigma_{2t+1}]\big) \tag{C.36}$$

$$\varsigma_{2t} = \frac{1}{R_{t+1}}\big(- Z_{t+1}\varrho(1-\psi_p)\psi_0\psi(SCC_{t+1} - \tau_{t+1}) + \psi\varsigma_{2t+1} - Z_{t+1}\varrho(1-\psi_p)\psi_0\psi[\varsigma_{1t+1} + \varsigma_{2t+1}]\big). \tag{C.37}$$

Thus, we can write

$$\mathcal{T}_{jt} = \frac{\bar{\omega}_t}{1-\alpha}\frac{\partial \ln(\mathcal{E}_t)}{\partial \ln(A_{jt})}\big((SCC_t - \tau_t) + \varsigma_{1t} + \varsigma_{2t}\big). \tag{C.38}$$

The contemporaneous effect on emissions follows

$$\begin{aligned}\frac{\bar{\omega}_t}{1-\alpha}\frac{\partial \ln(\mathcal{E}_t)}{\partial \ln(A_{jt})} = \frac{\alpha Z_t}{\mathcal{E}_t}\sum_\theta \frac{\omega_{\theta d} S_{\theta dt}}{r_{\theta d} + \omega_{\theta d}\tau_t}\big[&(1+(1-\alpha)(1-\lambda))S_{jt} \\ &+ \mathbb{1}(\theta = \theta(j))(1-\alpha)(\lambda-\sigma)S^\theta_{jt} \\ &+ \mathbb{1}(\theta d = j)(1-\alpha)(\sigma-1)\big].\end{aligned} \tag{C.39}$$

In summary, the root system follows

$$\begin{aligned}SCC_{1t} &= \varrho\psi_p\mathcal{Y}_t + \frac{1}{R_{t+1}}SCC_{1t+1} \\ SCC_{2t} &= \varrho(1-\psi_p)\psi_0\mathcal{Y}_t + \frac{1}{R_{t+1}}\psi SCC_{2t+1} \\ \mathcal{C}_{1t} &= \psi_p\mathcal{E}_t + \mathcal{C}_{1t-1} \\ \mathcal{C}_{2t} &= (1-\psi_p)\psi_0\mathcal{E}_t + \psi\mathcal{C}_{2t-1} \\ \tilde{\xi}_{jt} &= \frac{S_{jt} - \mathcal{T}_{jt}}{\gamma - 1} + \frac{\mathcal{Y}_{t+1}/\mathcal{Y}_t}{R_{t+1}}[\tilde{\xi}_{jt+1} + \sum_i \tilde{\xi}_{it+1}g_{it+1}\varphi_{ijt+1}] \\ \ln(A_{jt}) &= \ln(\gamma)\chi\Big(\frac{s_{jt}}{\nu_{\theta(j)}}\Big)^\eta\phi_{jt} + \ln(A_{jt-1}) \\ \varsigma_{1t} &= \frac{1}{R_{t+1}}\big(- Z_{t+1}\varrho\psi_p(SCC_{t+1} - \tau_{t+1}) + \varsigma_{1t+1} - Z_{t+1}\varrho\psi_p[\varsigma_{1t+1} + \varsigma_{2t+1}]\big) \\ \varsigma_{2t} &= \frac{1}{R_{t+1}}\big(- Z_{t+1}\varrho(1-\psi_p)\psi_0\psi(SCC_{t+1} - \tau_{t+1}) + \psi\varsigma_{2t+1} - Z_{t+1}\varrho(1-\psi_p)\psi_0\psi[\varsigma_{1t+1} + \varsigma_{2t+1}]\big).\end{aligned} \tag{C.40}$$

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 5.2 and Appendix C.5. 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 $\tilde{SCC}_t \equiv SCC_t/\mathcal{Y}_t$.

This achieves a steady-state which follows

$$S\tilde{C}C_1 = \frac{\varrho\psi_p}{1 - \tilde{R}^{-1}} \tag{C.41}$$

$$S\tilde{C}C_2 = \frac{\varrho(1-\psi_p)\psi_0}{1 - \psi\tilde{R}^{-1}}. \tag{C.42}$$

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 from Equation (A.21) 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.67). 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.31) and scientist supply (12) can be satisfied by setting $s_j = \nu_{\theta(j)}$ for all clean technologies. Plugging this into Equation (C.31) gives us that income shares for clean technologies also follow $S_j = \nu_{\theta(j)}$. From Equation (C.18), this implies that $\bar{A}_j = 1$ for clean technologies and $\bar{A}_j = 0$ for dirty technologies.

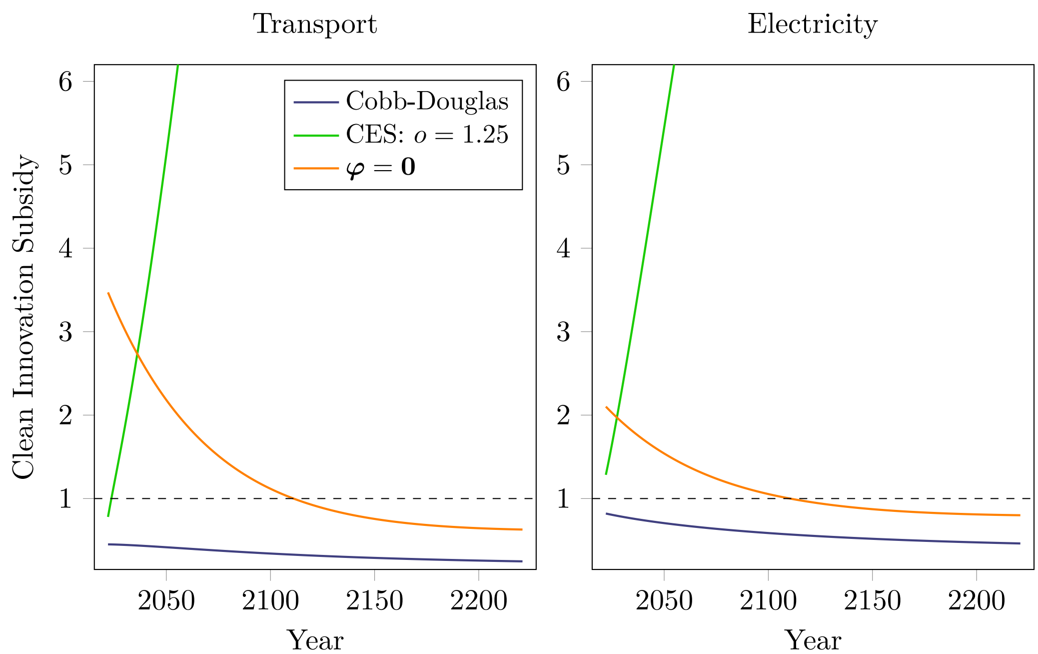
C.10 CES Spillovers

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 C.4, 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

$$\phi_{it} = \frac{\left(\sum_j \hat{\varphi}_{ij}^{\frac{1}{o}} A_{jt-1}^{\frac{o-1}{o}}\right)^{\frac{o}{o-1}}}{A_{it-1}}. \tag{C.43}$$

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 C.6: 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 C.6 shows first-best optimal subsidies for three cases of the spillover network, each of which use 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 post-clean transition world that many envision.[C.12]

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

[C.12] I obtain similar paths for optimal policy if I allow for an interior steady-state.

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 4.9 and 6.62, respectively. At the same time, relative improvements 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 0.78% and 0.5%, 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).

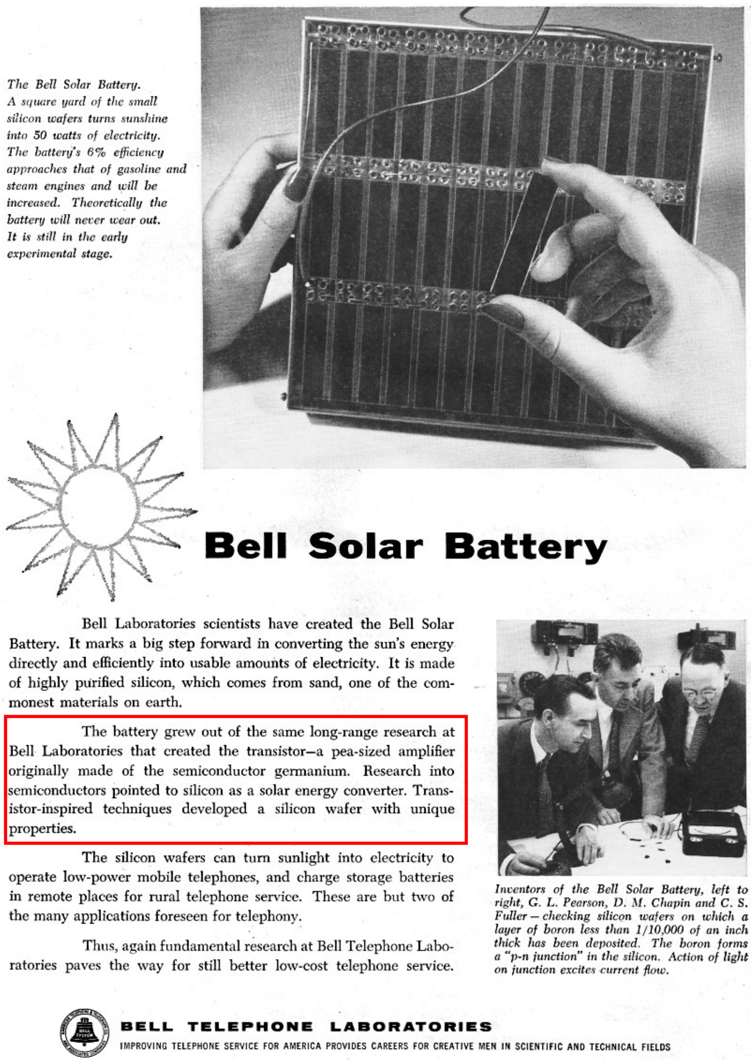
This logic is very different from that of a big push, and as a result, the time paths of optimal subsidies point in opposite directions. A big push calls for an immediate shift in innovation incentives towards clean technology that tapers off over time, whereas a spillover network with substitutable 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 assumptions 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.

D Additional Tables & Figures

Figure D.1: Tesla Mule 1

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

Figure D.2: Bell Labs Solar Cell



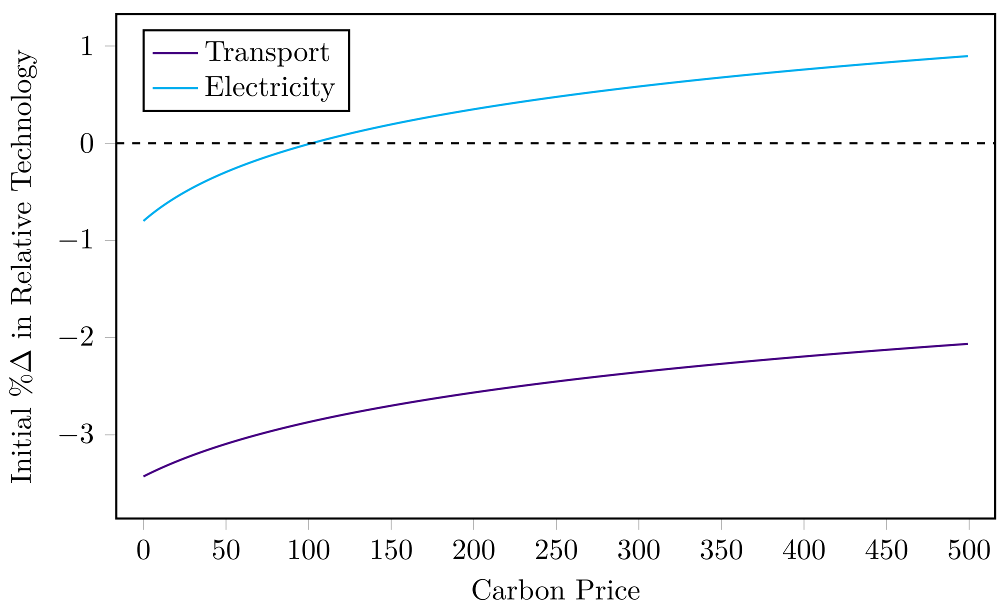
Notes: Bell Telephone Laboratories July 1954 Ad. Red box added to emphasize the description of knowledge spillovers.

Table D.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

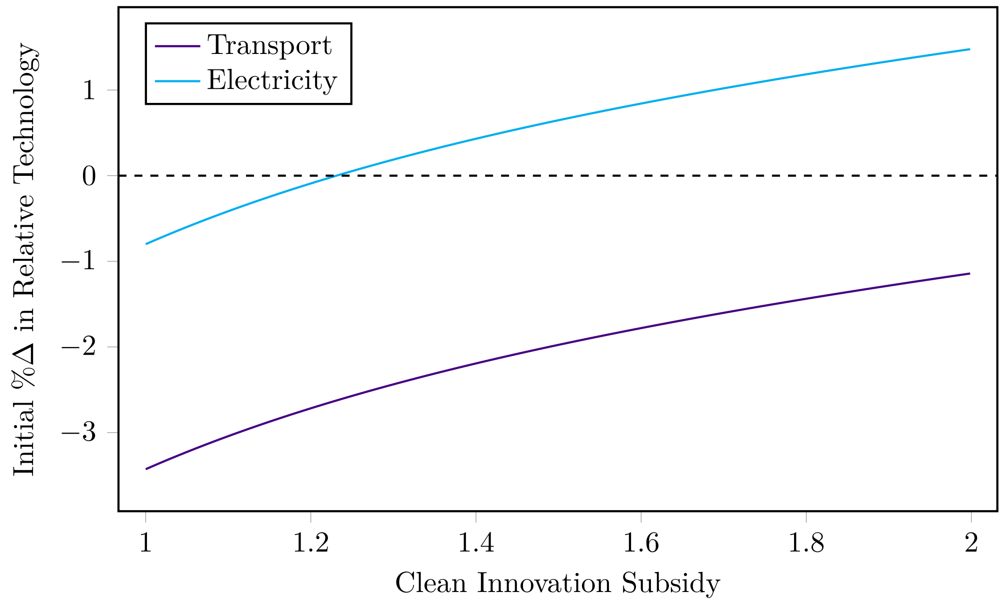
Notes: IPC codes for both clean and dirty transport patents come from Aghion et al. (2016), CPC codes for clean electric patents come directly from CPC subclass Y02E, and IPC codes for dirty electric patents come from Lanzi et al. (2011). I also include CPC codes electromobility (Y02T64-72), efficient 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 D.3: Carbon Price Necessary for Clean Growth (No Spillovers)



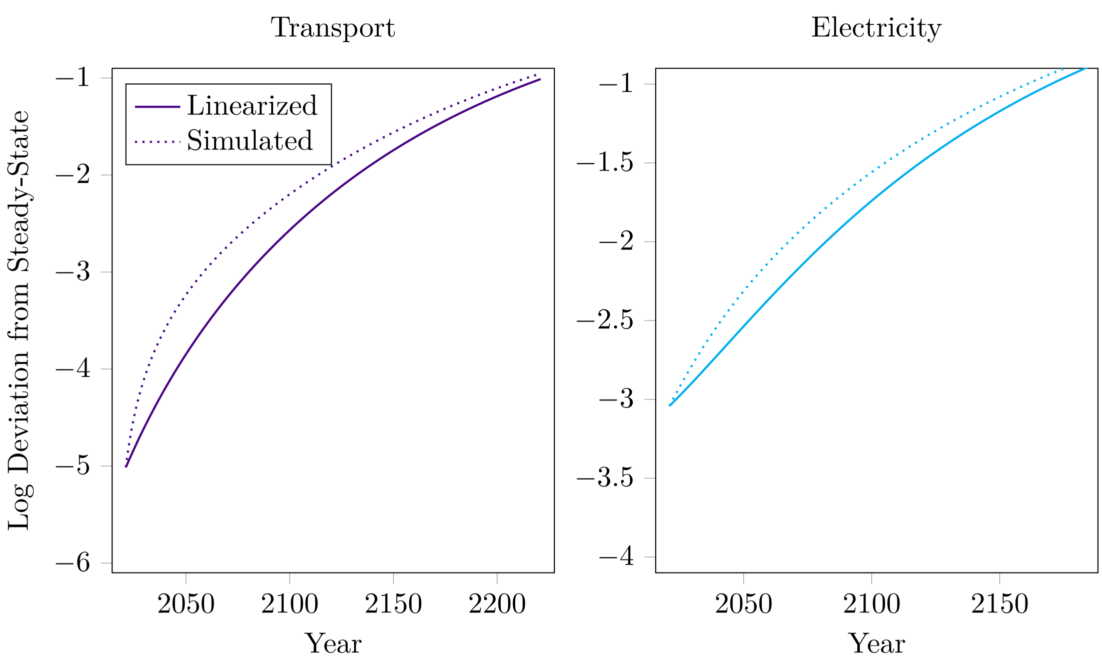
Notes: Direction of innovation as a function of the carbon price when the spillover network has been shut down, starting in 2010. Values above zero indicate a sector is in its clean basin of attraction. Electricity generation switches to clean growth with a carbon price of about $102, while transportation requires a carbon price of more than $10,000.

Figure D.4: Clean Innovation Subsidy Necessary for Clean Growth (No Spillovers)



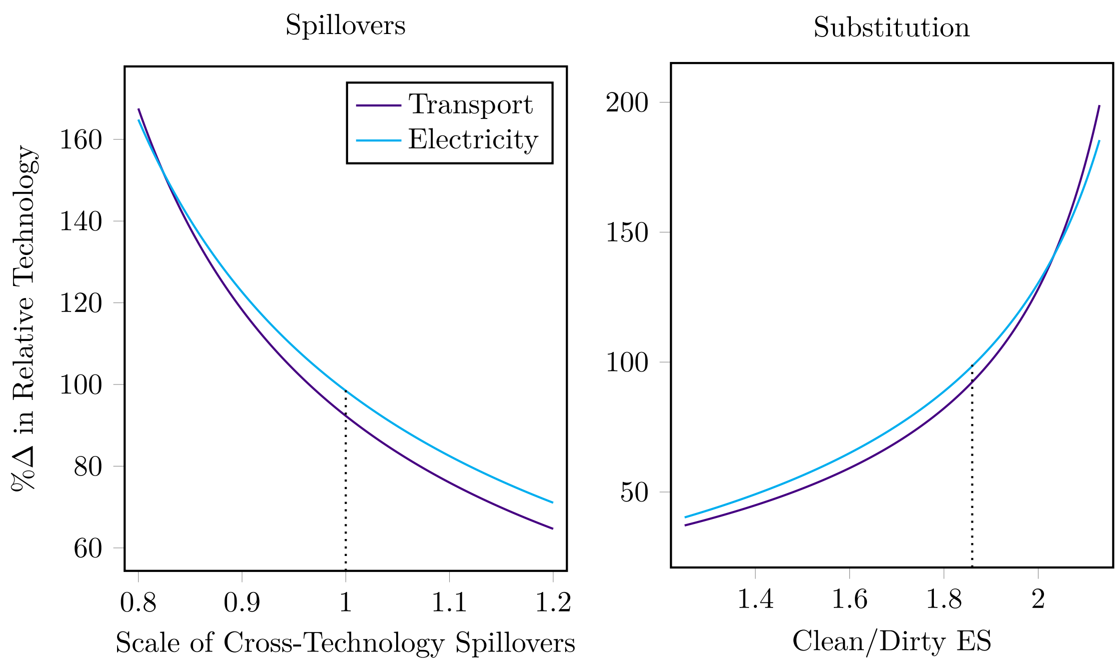
Notes: Direction of innovation as a function of the clean innovation subsidy when the spillover network has been shut down, starting in 2010. Values above zero indicate a sector is in its clean basin of attraction. Transportation switches to clean growth with a subsidy of 2.99, while electricity generation switches with a subsidy of 1.23.

Figure D.5: Linearized Transition Path Accurately Approximates Full Simulation



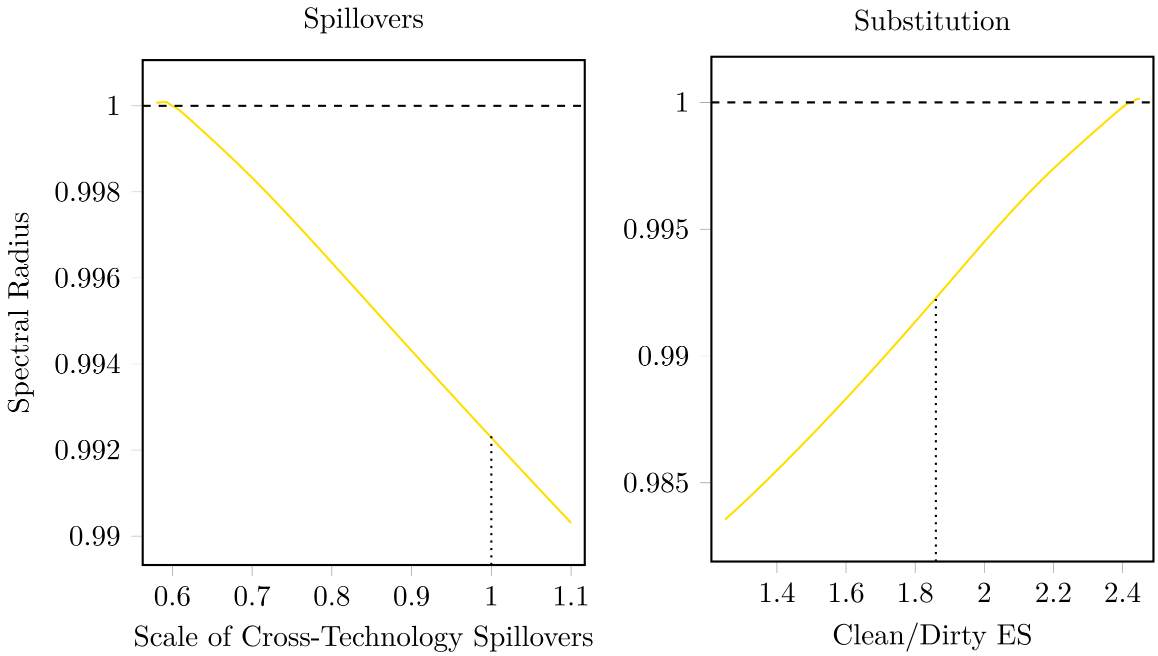
Notes: Technology's transition path following the policy reform of Section 6.1. Solid lines linearized paths using Proposition 2, while dotted lines fully simulated paths.

Figure D.6: Determinants of Steady-State Policy Impact



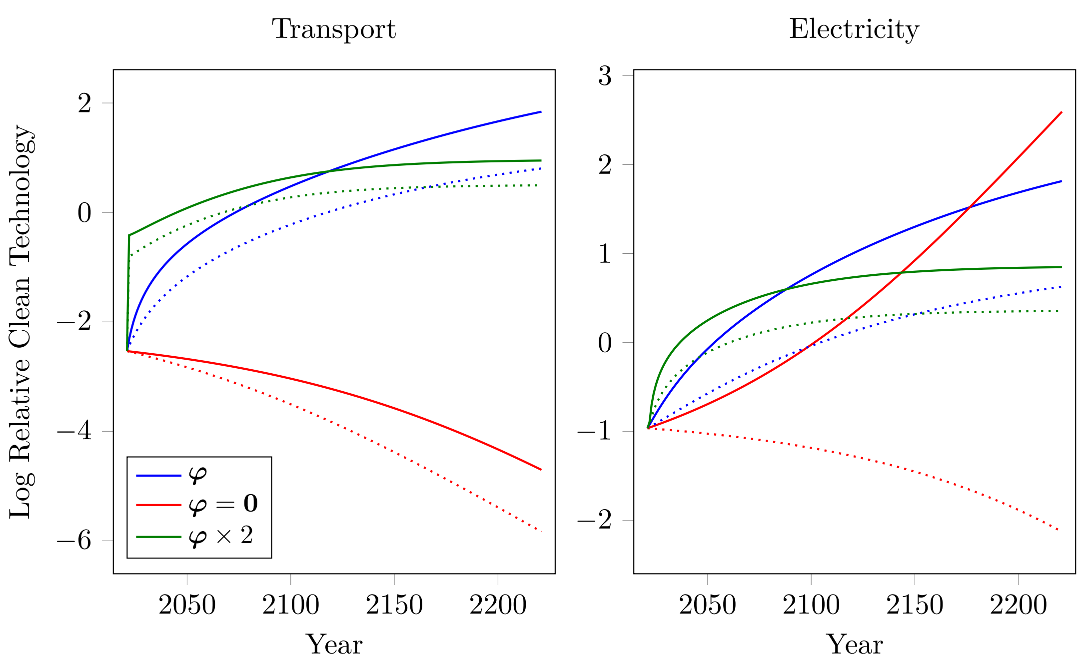
Notes: Impact of the policy reform of Section 6.1 as a function of the level of cross-technology knowledge spillovers and the elasticity of substitution between clean and dirty goods. Vertical dotted lines represent the benchmark calibration.

Figure D.7: 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 60.1% of their calibrated level or the elasticity of substitution increases to 2.43. Vertical dotted lines represent the benchmark calibration.

Figure D.8: Technology Path ($190 Carbon Price)



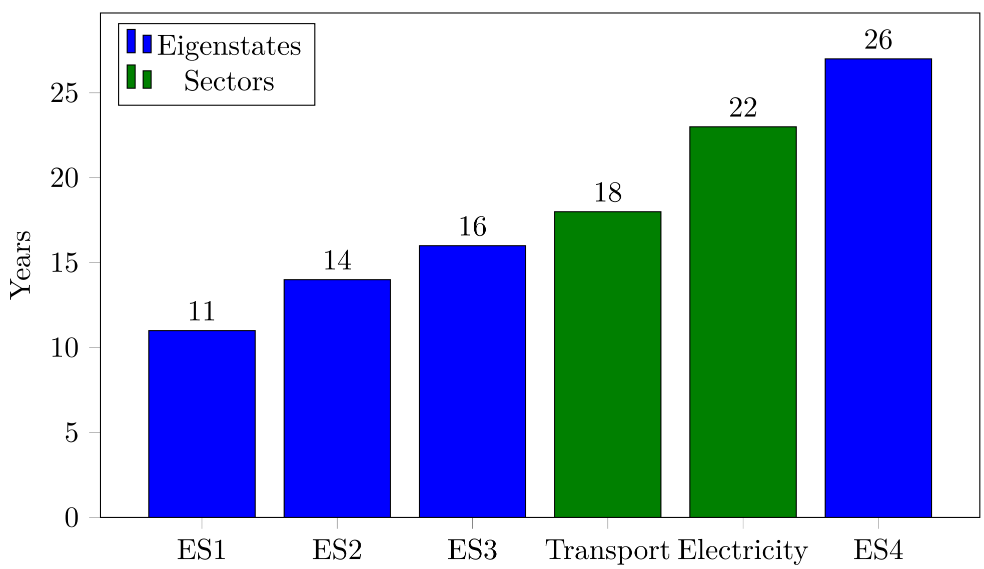
Notes: Impact of introducing a carbon price at the EPA's proposed SCC ($190) and clean innovation subsidy equivalent to a 30% tax credit ($\xi_c = 1.43$). Dotted lines indicate laissez-faire paths.

Table D.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%	+167.7%	+48.9%
$\%\Delta\bar{B}_{elec}$	$+\infty\%$	+173.3%	+55.5%
Emissions Intensity			
$\%\Delta\bar{\omega}$	-85.1%	-90.1%	-81%
Transitional Impacts			
Half-Lives of Convergence by Sector			
$t_{car}^{(1/2)}$	–	99 years	18 years
$t_{elec}^{(1/2)}$	–	107 years	22 years
Carbon Emissions by Year			
$\%\Delta\mathcal{E}_{2035}$	-70.1%	-76.8%	-78.7%
$\%\Delta\mathcal{E}_{2060}$	-72.7%	-81%	-79.6%
Degree of Increasing Returns to Innovation			
Spectral Radius			
$\max\lvert\kappa_j\rvert$	1.01	0.993	0.974

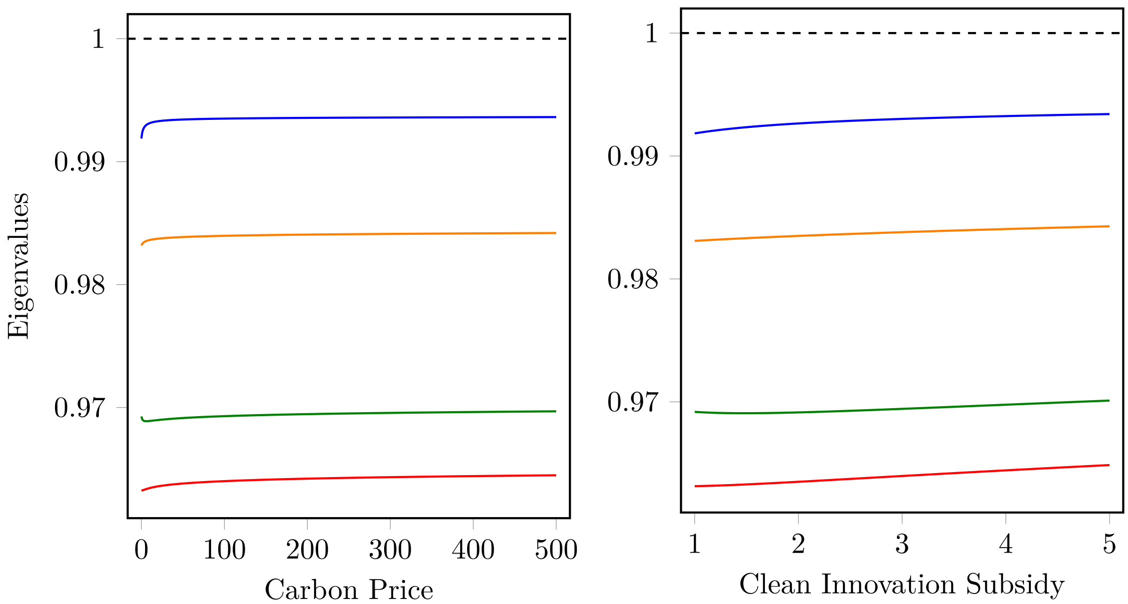
Notes: Impact of introducing a carbon price at the EPA's proposed SCC ($190) and clean innovation subsidy equivalent to a 30% tax credit ($\xi_c = 1.43$). Changes in relative technology are listed in log points. For locked in economies, long-run impacts refer to corner, rather than interior, steady-states.

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



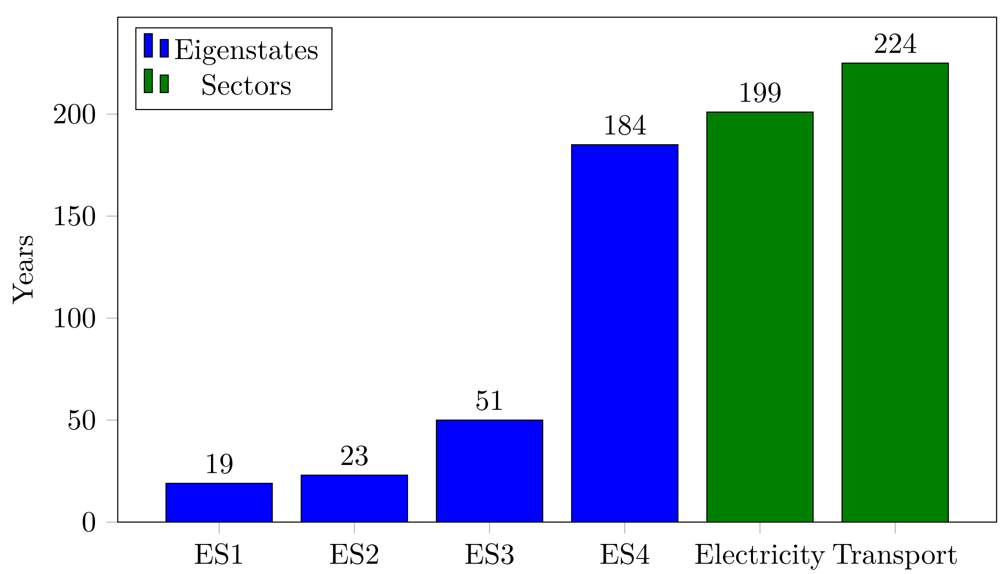
Notes: Transition speeds following the policy reform of Section 6.1. Cross-technology spillovers have been doubled.

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



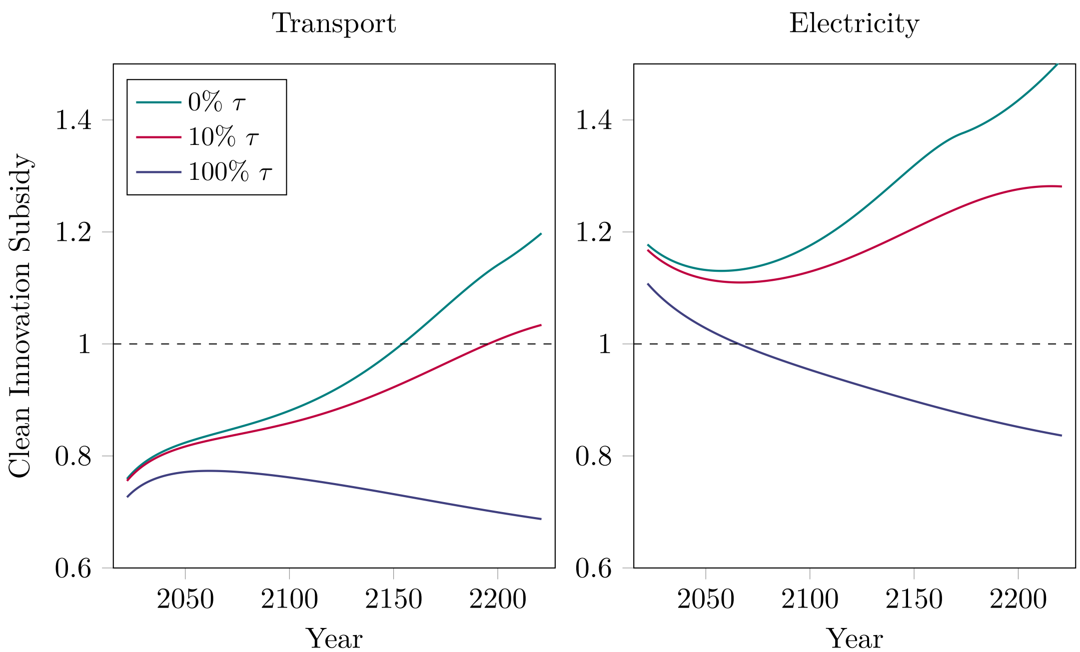
Notes: Eigenvalues of the transition matrix as a function of carbon prices and clean innovation subsidies. In each case, the other policy instrument is kept at its value from the policy reform of Section 6.1.

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



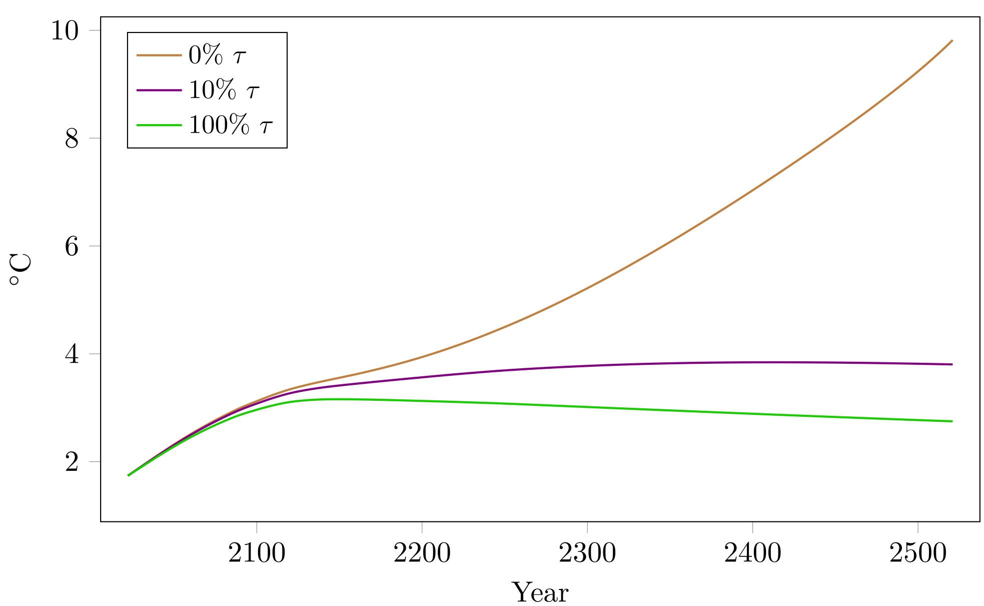
Notes: Transition speeds following the policy reform of Section 6.1. Within-sector spillovers are shut down by transferring them to the diagonal of the gross spillover network.

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



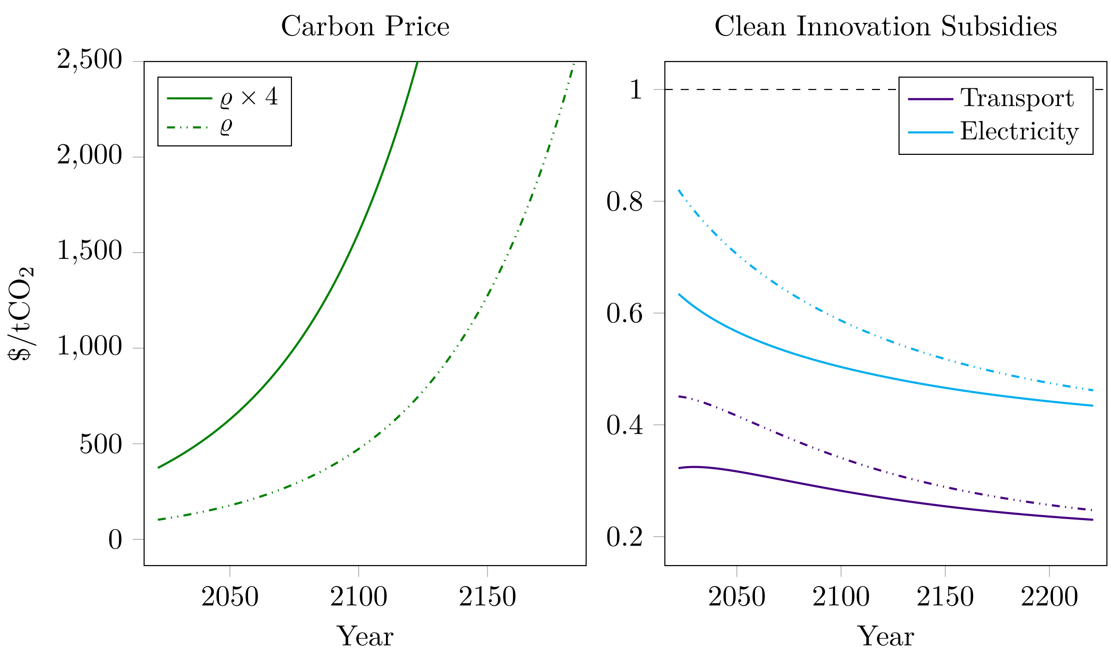
Notes: The optimal path for clean innovation subsidies with the high discount rate. The external carbon price τ is a proportion of the social cost of carbon, so 100% is the first-best. Innovation subsidies are listed as a fraction of the common innovation wedge.

Figure D.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 C.3 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 D.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.

References

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

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.

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

Arkolakis, C. and Walsh, C. (2023). Clean Growth. Working Paper 31615, National Bureau of Economic Research.

Arora, A., Belenzon, S., and Sheer, L. (2021). Knowledge Spillovers and Corporate Investment in Scientific Research. *American Economic Review*, 111(3):871–98.

Arora, A., Cohen, W., Lee, H., and Sebastian, D. (2023). Invention Value, Inventive Capability, and the Large Firm Advantage. *Research Policy*, 52(1):104650.

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).

Bilal, A. and Känzig, D. (2024). The Macroeconomic Impact of Climate Change: Global vs. Local Temperature. Working Paper 32450, National Bureau of Economic Research.

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

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

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

Donald, E. and Hanley, D. (2025). The Social Cost of Carbon with Existential Risk. *Working Paper*.

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

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

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.

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.

Hong, S. (2025). Green Industrial Policy in a Globalized Economy. *Working Paper*.

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

Joos, F., Roth, R., Fuglestvedt, J., Peters, G., Enting, I., von Bloh, W., Brovkin, V., Burke, E., Eby, M., Edwards, N., Friedrich, T., Frölicher, T., Halloran, P., Holden, P., Jones, C., Kleinen, T., Mackenzie, F., Matsumoto, K., Meinshausen, M., Plattner, G.-K., Reisinger, A., Segschneider, J., Shaffer, G., Steinacher, M., Strassmann, K., Tanaka, K., Timmermann, A., and Weaver, A. (2013). Carbon Dioxide and Climate Impulse Response Functions for the Computation of Greenhouse Gas Metrics: A Multi-Model Analysis. *Atmospheric Chemistry and Physics*, 13(5):2793–2825.

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

Kogan, L., Papanikolaou, D., Seru, A., and Stoffman, N. (2017). Technological Innovation, Resource Allocation, and Growth. *The Quarterly Journal of Economics*, 132(2):665–712.

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.

Lucking, B., Bloom, N., and Van Reenen, J. (2019). Have R&D Spillovers Declined in the 21st Century? *Fiscal Studies*, 40(4):561–590.

Nemet, G. (2019). *How Solar Energy Became Cheap: A Model for Low-Carbon Innovation*. Routledge.

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

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.