Transition to Clean Technology*

Daron Acemoglu Ufuk Akcigit Douglas Hanley William Kerr

November 30, 2012

[Preliminary]

Abstract

We develop a microeconomic model of endogenous growth where clean and dirty technologies compete in production and innovation—in the sense that research can be directed to either clean or dirty technologies. If dirty technologies are more advanced to start with, the potential transition to clean technology can be difficult both because clean research must climb several steps to catch up with dirty technology and because this gap discourages research effort directed towards clean technologies. Carbon taxes and research subsidies may nonetheless encourage production and innovation in clean technologies, though the transition will typically be slow. We characterize certain general properties of the transition path from dirty to clean technology. We then estimate the model using a combination of regression analysis on the relationship between R&D and patents, and simulated method of moments using microdata on employment, production, R&D, firm growth, entry and exit from the US energy sector. The model's quantitative implications match a range of moments not targeted in the estimation quite well. We then characterize the optimal policy path implied by the model and our estimates. Optimal policy heavily relies on research subsidies as well as carbon taxes. We use the model to evaluate the welfare consequences of a range of alternative policy structures. For example, just relying on carbon taxes or delaying intervention both have significant welfare costs—though their implications for medium run temperature increases are quite different.

JEL Classification: O30, O31, O33, C65.
Keywords: carbon cycle, directed technological change, environment, innovation, optimal policy.

---

*Authors' addresses (respectively): MIT and CIFAR, University of Pennsylvania, University of Pennsylvania, Harvard University. We thank our discussant John Van Reenen and the participants at the European Economic Association Meetings in Malaga and the Conference on Climate and the Economy in Stockholm for valuable comments. We thank David Popp for very helpful guidance on technology codes. Acemoglu and Akcigit also thank Bilkent University Economics Department for the great hospitality during this project. The research in this paper was conducted while the authors were Special Sworn Status researchers of the US Census Bureau at the Boston Census Research Data Center (BRDC). Support for this research from the NSF grant ITR-0427889 [BRDC] is gratefully acknowledged. Research results and conclusions expressed are the authors' and do not necessarily reflect the views of the Census Bureau or the NSF. This paper has been screened to ensure that no confidential data are revealed.

1 Introduction

Recent economic research has recognized the importance of transition to clean technology in controlling and reducing fossil fuel emissions and potentially limiting climate change.[1] Recent empirical work has also shown that innovation may switch away from dirty to clean technologies in response to changes in prices and policies. For example, Newell, Jaffe and Stavins (1999) show that following the oil price hikes, innovation in air-conditioners turned towards producing more energy-efficient units compared to the previous focus on price reduction; Popp (2002) finds that higher energy prices are associated with a significant increase in energy-saving innovations; Hassler, Krusell and Olovsson (2011) estimate a trend break in factor productivities in the energy-saving direction following the era of higher oil prices; and Aghion et al. (2012) find a significant impact of carbon taxes on the direction of innovation in the automobile industry and further provide evidence that clean innovation has a self-perpetuating nature feeding on its own past success. Based on this type of evidence, Acemoglu et al. (2012a) suggest that a combination of (temporary) research subsidies and carbon taxes can successfully redirect technological change towards cleaner technologies. Several conceptual and quantitative questions remain, however. The first is whether, in the context of a micro-founded quantitative model, reasonable policies can secure a transition to clean technology. The second is whether, in the presence of carbon taxes, there is still any role for significant research subsidies. The third concerns how rapidly the transition to clean technology should take place under optimal policy.

A systematic investigation of these questions necessitates a micro model of innovation and production where clean and dirty technologies can compete given the prevailing policies and research incentives (and the direction of technological change) are also endogenously determined as a function of these policies.[2] It also necessitates a combination of micro data for the modeling of competition in production and innovation, and a quantitative model flexible enough to represent realistic dynamics of carbon emissions and potential climate change. This paper is an attempt in this direction.

Our first contribution is to develop a tractable and parsimonious microeconomic model for this purpose. In our model, which we view as an abstract representation of the energy

[1] On climate change, see, e.g., Stott et al. (2004) on the contribution of human activity to the European heatwave of 2003, Emanuel (2005) and Landsea (2005) on the increased impact and destructiveness of tropical cyclones and Atlantic hurricanes over the last decades; and Nicholls and Lowe (2006) on sea-level rise. On economic costs of climate change, see Mendelsohn et al. (1994), Pizer (1998), and Weitzman (2009). On economic analyses of climate change, see, e.g., Golosov et al. (2011), Hassler and Krusell (2012), Krusell and Smith (2009), MacCracken et al. (1999), Nordhaus (1994), Nordhaus and Boyer (2000), Nordhaus (2008), and Stern (2007). On endogenous technology and climate change, see, Acemoglu et al. (2012a), Bovenberg and Smulders (1995, 1996), Goulder and Mathai (2000), Goulder and Schneider (1999), Grimaud et al. (2011), Hartley et al. (2011), Hassler, Krusell and Olovsson (2011), Popp (2002, 2004), and Van der Zwaan et al. (2002).

[2] Acemoglu et al. (2012a) assume that clean and dirty inputs are combined with a constant elasticity of substitution, which allows for limited form of competition between clean and dirty technologies.

production and delivery sectors, each one of a continuum of intermediate goods can be produced either using a dirty or clean technology, each of which has a knowledge stock represented by a (separate) quality ladder. Given production taxes (which are differential by type of technology), profit-maximizing final good producers choose which technology to utilize. Profit-maximizing firms also decide whether to conduct research to improve clean or dirty technologies. Clean research, for example, leads to an improvement over an existing clean technology, though there is also a small probability of a breakthrough which will build on and surpass the dirty technology when the dirty technology is the frontier in the relevant product line. Research and innovation decisions are impacted both by policies and the current state of technology in the two sectors. For example, when clean technology is far behind, most research directed to that sector will generate incremental innovations that cannot be profitably produced (unless there are very high levels of carbon taxes). However, if clean research can be successfully maintained for a while, it slowly becomes self-sustaining as the range of clean technologies that can compete with dirty ones expands as a result of a series of incremental innovations.

Our second contribution is to estimate parameters of this model using microdata on R&D expenditures, patents, sales, employment and firm entry and exit from a sample of US firms in the energy sector. The data we use for this exercise are from the Census Bureau's Longitudinal Business Database and Economic Censuses, the National Science Foundation's Survey of Industrial Research and Development, and the NBER Patent Database. We design our sample around innovative firms in the energy sector that are in operation during the 1975-2004 period. We use our sample to directly estimate some key parameters of the model and the initial distributions of dirty- and clean-energy product lines.[3] In particular, we estimate two of the key parameters of the model with regression analysis using R&D and patents. We also estimate the initial distribution of productivity gaps between clean and dirty technologies in the economy by allocating the patent stocks of firms innovating in these technology areas across the three-digit industries in which the firms are operating. The remaining four crucial parameters are estimated using simulated method of moments (we impose the discount rate and the fraction of scientists in the labor force from the data rather than estimating these from the model). We show that, despite its parsimony, the fit of the model to a rich and diverse set of moments not targeted in the estimation is fairly good.

We then combine this structure with a parsimonious model of the carbon cycle. Our modeling of the carbon cycle follows Golosov et al. (2011) and is fairly flexible despite its simplicity. Our final contribution is to use this estimated quantitative model for the analysis of optimal policy, in particular optimal carbon taxes and research subsidies,[4] and a range of counterfactual policy experiments.

[3] See Popp (2006) and Jaffe et al. (2010) for background on technology, R&D and innovation in the energy sector.

[4] We do not allow additional tax instruments to remove the monopoly distortions in the economy.

Our main results are as follows. Though it is intuitive to expect that carbon taxes should do most of the work in the optimal allocation—because they both reduce current emissions and encourage R&D directed to clean technologies—quantitatively we find a major role for research subsidies. For example, with an annual discount rate of 1% (similar to the number favored by Nordhaus, 2007) and focusing on constant policies, the optimal research subsidy is 61% (meaning that the government pays for 61 cents out of every dollar of R&D expenditure for clean technology) while the carbon tax is 16%. The numbers are more extreme with a discount rate of 0.1% for the social planner (similar to the number favored by Stern, 2007) but with a similarly major role for research subsidies: a research subsidy of 95% and a carbon tax of 44%. When we allow time-varying policies, the overall pattern is broadly similar and still heavily relies on research subsidies, but with some notable differences: first, the research subsidy is initially slightly more aggressive and then declines somewhat over time; second, with a discount rate of 1%, carbon taxes are backloaded (low, in fact zero, for an extended period of time and then high); and third, with a discount rate of 0.1%, carbon taxes are frontloaded (starting out higher and declining over time).[5] Despite the differences between the models, the reason for the major role for research subsidies is related to the one emphasized in Acemoglu et al. (2012a).[6] Research subsidies are powerful in redirecting technological change, and given this, it is not worth distorting the initial production too much by introducing heavy carbon taxes. It is important to emphasize that research subsidies are not being used just because there is a market failure (and an uninternalized externality) in research. In fact, in our model, in the absence of externalities from carbon, or in the special case in which there is only a dirty or a clean sector, the social planner would have no reason to use research subsidies—because a scarce factor, skilled labor, is being used for research and no other purpose, and thus the social planner cannot increase the growth rate by subsidizing research. The reason why the social planner heavily uses research subsidies is because when carbon creates negative externalities, inducing a transition to clean technology is an effective way of reducing future carbon emissions by changing the path of technological progress.

Another useful comparison is to current US policies. We estimate the effective research

[5] Our time-varying optimal policy results need to be interpreted with caution, since the resulting optimal policy sequence is not time consistent.

[6] Major differences between the models include: (1) here the damage from atmospheric carbon is modeled as impacting production along the lines of previous literature rather than directly utility; (2) here there is no "environmental disaster" threshold, making it possible for us to calibrate the parameters more closely to data and without taking a position on carbon emissions in the rest of the world; (3) in contrast to the constant elasticity of substitution formulation, dirty and clean sectors are not complements in our model, but explicitly compete in each product line. This last one is the most important distinction, enabling us to use microdata on innovation and production. It also implies a different pattern of production distortions from carbon taxes. In Acemoglu et al. (2012a), carbon taxes are particularly distortionary when the dirty sector is behind (and thus its relative prices high because of the imperfect substitutability). In contrast, in our model the carbon tax is least distortionary when the clean technology has already taken over or is about to take over almost all product lines.

subsidy from the differential between clean and dirty firms in our sample in the use of federally funded R&D expenditure. Utilizing this estimate and different values of effective carbon tax at the moment and its likely values in the future, our estimated optimal policies are quite different from their US counterparts, and we show that under US policies, climate change dynamics will be significantly different (and worse).

In terms of counterfactual policies, we investigate the welfare costs of just relying on carbon taxes and delaying intervention. The most notable result here is that the welfare costs of delaying the optimal policy by 50 years (laissez faire) is very significant. With a discount rate of 1%, delaying optimal policy by 50 years has a welfare cost equivalent to a permanent 8% drop in consumption. With a discount rate of 0.1%, the consumption-equivalent welfare cost is 16.6%. The costs of relying just on carbon tax (without any research subsidy) are more modest but still significant, 4.2% and 3.4%, with the same two discount rates, respectively.

We also consider several variations and robustness checks to show which aspects of the model are important for our main theoretical and quantitative results. In particular, we investigate the implications of using different discount rates and estimates of the damage of carbon concentration on economic activity, allowing different degrees of distortions from research subsidies, different estimates of the microeconomic elasticities in the R&D technology, and different distributions of productivity gaps between clean and dirty technologies. Overall, most of the main qualitative and quantitative features of optimal policy appear to be fairly robust to a range of plausible variations.

Our model combines elements from four different lines of research (and is thus related to each of these four lines). First, we build on the growing literature on quantitative general equilibrium models of climate change, such as Golosov et al. (2011), Hassler and Krusell (2012), Krusell and Smith (2009), Nordhaus (1994), Nordhaus and Boyer (2000), Nordhaus (2008), and Stern (2007). We follow these papers in introducing a simple model of the carbon cycle and the economic costs of carbon emissions in a general equilibrium model, and then characterizing optimal policy. Second, we introduce endogenous and directed technological change along the lines of Acemoglu (1998, 2002) in a model where producers have a choice between clean and dirty production methods. In combining these two first lines of research, we are following Acemoglu et al. (2012a) as well as several other papers listed in footnote 1 above. Third, we develop a tractable but rich model of competition between dirty and clean technologies building on the literature on step-by-step competition as in Harris and Vickers (1995), Aghion et al. (2001), and Acemoglu and Akcigit (2012). Fourth, we model the microeconomics of innovation, employment and output dynamics building on Klette and Kortum (2004), where each firm consists of a number of products and technologies (different from other applications, technologies here are different from products because of the competition between clean and dirty sectors).

In estimating a general equilibrium model of firm-level innovation and employment dynamics, we follow Lentz and Mortensen (2008) and Acemoglu et al. (2012b). We differ from existing work in this area in three important respects, however. First, we combine this type of estimation strategy with a model of clean and dirty technologies and estimate some of the parameters of the R&D technology directly from microdata. Second, rather than focusing on steady-state comparisons, we study non-steady-state dynamics, which is crucial for the question of transitioning to clean technology. Third, we characterize optimal policies in such a framework.

The remainder of the paper is organized as follows. Section 2 introduces our model and characterizes the equilibrium. Section 3 describes the dataset we will use for estimation and quantitative evaluation, outlines the different components of our estimation strategy, and presents the estimates of some of the parameters we obtain from our micro data. Section 4 presents the simulated method of moments estimates of our parameters and discusses the fit of the model. Section 5 quantitatively characterizes the structure of optimal environmental policy. In this section, we also conduct a range of counterfactual exercises. Section 6 discusses a range of robustness exercises intended to convey which sorts of assumptions and parameters are important for the qualitative and quantitative results of the paper. Section 7 concludes.

2 Model

In this section, we present our baseline model. This is a simple dynamic general equilibrium model, where final output combines intermediates produced either using a clean or dirty technology. The productivity of the dirty and clean technology for each intermediate is represented by a quality ladder. Production is also subject to taxes, so profit-maximizing final good producers choose whether to use clean or dirty intermediates as a function of the productivity gap between the two and taxes. Research is directed towards clean or dirty technology, and progresses both with incremental research increasing productivity by one rung on the quality ladder and with occasional breakthrough research which enables the firm to surpass the current frontier technology. Research is conducted both by entrants and incumbent firms which already hold a portfolio of products and technologies. Finally, dirty technology contributes to carbon emissions, which create potential economic damage. We next describe each module of the model in turn.

2.1 Preferences and Endowments

We model an infinite-horizon closed economy in continuous time. Since the consumer side is not our focus, we simplify the discussion by modeling it with a representative household with a logarithmic instantaneous utility function. The lifetime utility is then

$$U_0 = \int_0^\infty e^{-\rho t} \ln C_t dt, \tag{1}$$

where C_t is the household's consumption at time t and $\rho > 0$ is the discount rate. We assume that the representative household consists of mass one of production workers and mass L^s of "scientists" who will be employed in R&D activities. All workers supply one unit of labor inelastically. The representative household owns all the firms in the economy, so its problem will be to maximize (1) subject to the following budget constraint

$$w_t^u + w_t^s L^s + \Pi_t \geq C_t,$$

and the usual no Ponzi-game condition. Here Π_t is the total sum of corporate profits net of R&D expenses, w_t^u and w_t^s are the wage rates (and thus wage incomes) of the production and R&D workers.

Since the economy is closed, there is no physical capital, and intermediates and the R&D sector use labor, aggregate consumption is equal to the production of the final good:

$$C_t = Y_t.$$

2.2 Final Good Technology, Intermediate Production and Pricing

The final good is produced by combining a measure one of intermediates with an elasticity of substitution equal to one. In addition, its production is negatively affected by the amount of atmospheric carbon concentration, which we denote by S_t. We follow the formulation suggested by Golosov et al. (2011), which builds on earlier work by Mendelsohn et al. (1994) and Nordhaus (1994, 2008), and assume

$$\ln Y_t = -\gamma \left(S_t - \bar{S} \right) + \int_0^1 \ln y_{i,t} di, \tag{2}$$

where $\bar{S} > 0$ is the pre-industrial level of the atmospheric carbon concentration, $\gamma \geq 0$ is a scale parameter, and $y_{i,t}$ is the quantity of intermediate good i. When $\gamma = 0$, (2) gives the standard (unitary elasticity of substitution) production function for combining intermediates to produce a final good. When $\gamma > 0$, levels of atmospheric carbon concentration above the pre-industrial level reduce productivity with elasticity γ, for reasons discussed in Mendelsohn et al. (1994), Nordhaus (1994, 2008) and Stern (2007).

A feature of (2), which will play a central role in our quantitative exercise, is worth noting: the proportional cost of a unit increase in atmospheric carbon concentration is independent of its current level. Though nonlinearities, or even major threshold effects, are likely to be present in the impact of atmospheric concentration on economic activity, this functional form is not only in line with assumptions made by other economic approaches to climate change (e.g., Nordhaus, 1994, 2002, 2007, Nordhaus and Boyer, 2000, Stern, 2007, Golosov et al., 2011), but also enables us to study the implications of carbon emissions from our economy, with parameters estimated from the US and calibrated to US aggregates, without taking a position on the path of carbon emissions on the rest of the world. Without this assumption, the marginal cost of carbon emissions, and thus optimal policy, would strongly depend on assumptions on the evolution of emissions from other countries.

Each intermediate $i \in [0,1]$ can be produced with either a dirty or a clean technology, and when it is produced with the clean (dirty) technology we denote it by $y^c_{i,t}$ ($y^d_{i,t}$). We will sometimes refer to clean and dirty technologies as clean and dirty "sectors," and we also use the terms "intermediaries" and "product lines" interchangeably.

Firm f can produce intermediate i with either a clean or dirty technology ($j \in \{c,d\}$) with the following production function $y^j_{i,t}(f) = q^j_{i,t}(f)\, l^j_{i,t}(f)$, where $l^j_{i,t}(f)$ is the employment of (production workers) by this firm and $q^j_{i,t}(f)$ is the labor productivity of the technology that this firm has access to for producing with clean or dirty technology j in product line i. In equilibrium, only firms with the highest technology either in the clean or dirty sector will produce, so we simplify this equation by suppressing firm indices and with the implicit convention that the labor productivity q always refers to the most advanced clean or dirty technology, thus writing:

$$y^j_{i,t} = q^j_{i,t} l^j_{i,t}.$$

Though only firms with the most advanced technology for intermediate i within the clean or dirty sector can ever produce it, because of taxes it is not necessarily the most advanced technology between these two sectors that will always be active. In particular, there is a tax at the rate τ^j_t on sector (technology) j at time t, which implies that the marginal cost of production is

$$MC^j_{i,t} = \frac{\left(1+\tau^j_t\right) w^u_t}{q^j_{i,t}}, \;\; j \in \{c,d\} \text{ and } i \in [0,1],$$

where w^u_t is the wage rate of production workers. We define tax-adjusted labor productivity as

$$\tilde{q}^j_{i,t} \equiv \frac{q^j_{i,t}}{1+\tau^j_t}.$$

In equilibrium, only the technology with the lower marginal cost (inclusive of taxes)—or equivalently the one with the higher tax-adjusted labor productivity—will produce. Summarizing this, we have

$$\text{produce intermediate } i \text{ with technology } j \text{ if } \tilde{q}^j_{i,t} > \tilde{q}^{-j}_{i,t} \text{ where } j \neq -j \in \{c,d\}.$$

We assume that if clean and dirty technologies have equal tax-adjusted labor productivities, each produces with probability 50% at any point in time.[7] Thus, the tax-adjusted technology level use in the production of intermediate i at time t can be written as

$$\bar{q}_{i,t} = \begin{cases} \tilde{q}^{d}_{i,t} & \text{if } \tilde{q}^{d}_{i,t} \geq \tilde{q}^{c}_{i,t} \\ \tilde{q}^{c}_{i,t} & \text{otherwise} \end{cases} .$$

Finally, we also assume that at the initial date $t = 0$, for each leading technology of quality $q^{j}_{i,0}$, there also exists an intermediate good of quality $q^{j}_{i,0}/\lambda$, which ensures that markups in the initial date will not exceed λ (this will be guaranteed endogenously in subsequent dates).

2.3 Innovation, the Quality Ladder and Dynamics

Labor productivity for each intermediate (for each technology) evolves as a result of innovation. Research is directed towards clean or dirty technologies. A successful innovation leads to one of two types of innovation. The first is an *incremental* innovation, which takes place with probability $1 - \alpha$; and the second is a *breakthrough* innovation, which takes place with probability α (independently of all other events).

If research directed to sector $j \in \{c, d\}$ leads to an incremental innovation, then the innovator improves over the sector j technology of a randomly chosen intermediate. This is incremental innovation in the sense that it enables the innovator to go up by one rung in the quality ladder over producing technology, and we assume that each rung corresponds to an improvement of $\lambda > 1$. Consequently, labor productivity of technology j in intermediate i at time t can be written as

$$q^{j}_{i,t} = \lambda^{n^{j}_{i,t}},$$

where $n^{j}_{i,t} \in \mathbb{Z}_{+}$ is the effective number of steps that this technology has taken since time $t = 0$ (when all technologies are, by assumption, normalized to $q^{j}_{i,0} = 1$).

Relative productivity of dirty to clean technology in intermediate i at time t can be written as

$$\frac{q^{d}_{i,t}}{q^{c}_{i,t}} = \lambda^{n_{i,t}}$$

where

$$n_{i,t} \equiv n^{d}_{i,t} - n^{c}_{i,t} \in \mathbb{Z}$$

is defined as the technology gap between dirty and clean sectors in product line i at time t. In what follows we will need to keep track of the share of intermediates with technology gap $n \in \mathbb{Z}$, and we denote this by $\mu_{n,t} \in [0, 1]$ at time t.

[7] In other models of this type, e.g., Acemoglu and Akcigit (2012), which of two firms produces is immaterial. But here, since one of them uses the dirty technology and thus will contribute to carbon emissions, we need to specify exactly who produces in this case.

Breakthrough innovations, on the other hand, enable the successful innovator to improve by one rung over the frontier technology, even if this frontier is set by the alternative technology—i.e., a breakthrough clean innovation will improve over the dirty technology even if the latter is far ahead of the clean sector, thus allowing the clean sector to leapfrog the dirty one.

Therefore, conditional on an innovation in technology j for intermediate i between times t and $t+\Delta t$, the evolution of $q_{i,t}^{j}$ can be written as

$$q_{i,t+\Delta t}^{j} = \begin{cases} \lambda q_{i,t}^{j} & \text{with probability} & 1 & \text{if } q_{i,t}^{j} \geq q_{i,t}^{-j} & \text{(incremental)} \\ \lambda q_{i,t}^{j} & \text{with probability} & 1-\alpha & \text{if } q_{i,t}^{j} < q_{i,t}^{-j} & \text{(incremental)} \\ \lambda q_{i,t}^{-j} & \text{with probability} & \alpha & \text{if } q_{i,t}^{j} < q_{i,t}^{-j} & \text{(breakthrough)} \end{cases}.$$

Let z_t^j denote the aggregate innovation rate which is the sum of incumbents' and entrants' innovation rates in technology j. The law of motion for the technology gap $n_{i,t}$ can then be expressed as follows:

$$n_{i,t+\Delta t} = \begin{cases} n_{i,t}-1 & \text{with probability} & (1-\alpha)\, z_t^c \Delta t & \forall n_{i,t} \\ n_{i,t}+1 & \text{with probability} & (1-\alpha)\, z_t^d \Delta t & \forall n_{i,t} \\ -1 & \text{with probability} & \alpha z_t^c \Delta t & \text{if } n_{i,t} > 0 \\ n_{i,t}-1 & \text{with probability} & \alpha z_t^c \Delta t & \text{if } n_{i,t} \leq 0 \\ 1 & \text{with probability} & \alpha z_t^d \Delta t & \text{if } n_{i,t} \leq 0 \\ n_{i,t}+1 & \text{with probability} & \alpha z_t^d \Delta t & \text{if } n_{i,t} > 0 \\ n_{i,t} & \text{otherwise} & & \end{cases}$$

Note that innovations here have a creative destruction element (e.g., Aghion and Howitt, 1992, Grossman and Helpman, 1991) because, by improving over an existing product typically operated by another firm, they transfer the leading-edge technology to the current innovator.

In what follows, for notational and computational tractability, we assume that the gross tax rates are multiples of λ such that $1+\tau_t^j = \lambda^{m_t^j}$. Since taxes are chosen by the social planner, especially when λ is not too large, this is without much loss of generality. Given this assumption, we can write

$$\frac{1+\tau_t^d}{1+\tau_t^c} = \lambda^{m_t},$$

where

$$m_t \equiv m_t^d - m_t^c,$$

and thus tax-adjusted technologies can be written as

$$\frac{\tilde{q}_{i,t}^d}{\tilde{q}_{i,t}^c} = \frac{q_{i,t}^d}{1+\tau_t^d}\frac{1+\tau_t^c}{q_{it}^c} = \lambda^{n_{i,t}-m_t}.$$

We will say that dirty is the leading (tax-adjusted) technology if $n_{i,t} > m_t$; the two technologies are neck and neck if $n_{i,t} = m_t$; and clean is the leading technology otherwise.

2.4 Firms, R&D and Free Entry

Following Klette and Kortum (2004), we define a firm as a collection of leading-edge technologies. Let u_f^j denote the number of intermediates where firm f has the leading-edge technologies in sector $j \in \{c, d\}$ (but these are not necessarily more advanced than the technologies available in the other sector $-j$). Again following Klette and Kortum (2004), we assume that u_f^j captures the stock of knowledge of the firm for further innovations with technology $j \in \{c, d\}$. In particular, we assume that firms combine their knowledge stock u_f^j with scientists (R&D workers) H^j in order to generate a Poisson flow rate of X^j new innovations (in continuous time) according to the following production function

$$X^j = \theta \left(H^j\right)^{\eta} \left(u^j\right)^{1-\eta}, \tag{3}$$

where $\eta \in (0, 1)$ is the R&D elasticity with respect to scientists and $\theta > 0$ is a scale parameter. Thus the *variable* cost of generating a flow rate of X^j is simply $w_t^s u \left(x^j\right)^{\frac{1}{\eta}} \theta^{-\frac{1}{\eta}}$ where $x^j \equiv X^j/u^j$ is the innovation intensity per product line and w_t^s is the wage rate of scientists. In addition, R&D activities also require each firm to hire a number of scientists per product line (as fixed cost). We assume that, per product line, firm f will need to hire $F_{I,i}u$ scientists where $F_{I,i,t} \in [(1-\xi) F_I, (1+\xi) F_I]$ is an iid (across firms and over time) draw with mean F_I and $\xi \in (0, 1)$.[8] Hence, the total cost of R&D for firm i performing R&D directed at technology $j \in \{c, d\}$ at time t is

$$\begin{aligned} C_t\left(u, x^j\right) &= w_t^s u \left(h^j + F_{I,i,t}\right) \\ &= w_t^s u \left(\left(x^j\right)^{\frac{1}{\eta}} \theta^{-\frac{1}{\eta}} + F_{I,i,t}\right), \end{aligned}$$

where $h^j \equiv H^j/u$ is the average scientists hired per product line and the cost function is indexed by time because of the wage rate of scientists.

Entrants can also undertake R&D directed to either sector. We assume that to do this they need to hire $F_E \geq F_I$ scientists, and this will lead to a flow rate of innovation equal to one. We denote the endogenously determined mass of entrants performing R&D directed to technology j at time t by E_t^j.

On the policy side, incumbents performing R&D for sector j receive a proportional government subsidy at the rate $s_{I,t}^j \in [0, 1]$, and entrants performing R&D for sector j receive a subsidy at the rate $s_{E,t}^j \in [0, 1]$.

[8] This heterogeneity in fixed costs is necessary to make the dynamics (computationally) well behaved. Because of "creative destruction" in these types of models, equilibrium path in which some types of firms stop doing R&D (as clean firms will do without policy and dirty firms under our optimal policy), there will be a discontinuous behavior shortly before this point because creative destruction is expected to cease. Heterogeneity in fixed costs smooths this transition.

2.5 The Carbon Cycle

While clean intermediate production $y^c_{i,t}$ creates no carbon emission, dirty production $y^d_{i,t}$ emits κ units of carbon per intermediate output. This implies that total amount of carbon emission at time t is

$$K_t = \kappa \int_0^1 y^d_{i,t} di. \tag{4}$$

We follow Golosov et al. (2011) in assuming that the atmospheric carbon concentration S_t is determined as follows

$$S_t = \int_0^{t-T} (1 - d_l) K_{t-l} dl, \tag{5}$$

where $t = T$ is the first date when emission started and

$$d_l = (1 - \varphi_P) \left[1 - \varphi_0 e^{-\varphi l}\right]$$

is the amount of carbon emitted l years ago still left in the atmosphere. In addition, $\varphi_P \in (0,1)$ is the share of emission that remains permanently in the atmosphere, $(1 - \varphi_P)\varphi_0 \in (0,1)$ is the fraction of the transitory component that remains in the first period, and $\varphi \in (0,1)$ is the rate of decay of carbon concentration over time. As explained in Golosov et al. (2011), this is a flexible specification that approximates the more complex dynamics of carbon concentration in the atmosphere used by Nordhaus (2008). Though considerably simpler, this specification fairly closely approximates the observed dynamics of atmospheric carbon concentration as we show below.

2.6 Equilibrium

In this section, we characterize certain properties of the equilibrium path of this economy.

The economy at time $t = 0$ is characterized by a distribution of technology gaps between clean and dirty sectors $\mu_{n,0}$ for $n \in \mathbb{Z}$, and the equilibrium path will be defined for a given sequence of taxes and subsidies. Then a dynamic equilibrium path is a sequence of intermediate outputs, prices, innovation rates by incumbents and entrants, skilled and unskilled wages, measures of entrants, growth rate of aggregate output, interest rate, and atmospheric concentration, i.e., $\left[y^j_{i,t}, p^j_{i,t}, x^j_{I,t}, x^j_{E,t}, w^s_t, w^u_t, E^j_t, r_t, S_t\right]_{t=0}^{\infty}$, such that, given sequences of policies, all firms maximize profits, skilled and unskilled labor markets clear, free entry conditions hold (with complementary slackness), consumers optimize dynamically, and atmospheric carbon evolves according to the carbon cycle model presented above (i.e., (5)). To determine this dynamic time path we also have to keep track of the distribution of sectors by technology gaps, $\{\mu_{n,t}\}_{n=-\infty}^{\infty}$.

2.7 Prices and Profits

Given the aggregate production function (2), which implies unit elastic demand for intermediates, the demand for intermediates at time t is

$$y_{i,t} = \frac{\tilde{Y}_t}{p_{i,t}}, \quad \forall i \in [0,1], \tag{6}$$

where $\tilde{Y}_t \equiv Y_t \exp\left(\gamma\left(S_t - \bar{S}\right)\right)$ is net aggregate output (net of environmental damage).

We now characterize equilibrium prices. As explained in the previous section, if the leading technology for intermediate i at time t is $q^j_{i,t}$, another firm will have access to technology $q^j_{i,t}/\lambda$ for free. This clearly also applies to tax adjusted labor productivity, which is what is relevant for production decisions: when the leading technology is $\tilde{q}^j_{i,t}$, there is always a follower with technology with $\tilde{q}^j_{i,t}/\lambda$. Thus, equilibrium markups can never exceed λ. However, in equilibrium, there may not be any markup in some of the intermediates because of the competition between clean and dirty technologies. In particular, intermediate i will be produced using technology $j \in \{c,d\}$ only if there $\tilde{q}^{-j}_{i,t} \leq \tilde{q}^j_{i,t}$. If $\tilde{q}^{-j}_{i,t} < \tilde{q}^j_{i,t}$, the equilibrium markup will be λ, and if $\tilde{q}^{-j}_{i,t} = \tilde{q}^j_{i,t}$, there will be zero markup. Therefore:

$$p^j_{i,t} = \begin{cases} \frac{w^u_t}{\tilde{q}^j_{i,t}} & \text{if} \quad \tilde{q}^j_{i,t} = \tilde{q}^{-j}_{i,t} \\ \frac{\lambda w^u_t}{\tilde{q}^j_{i,t}} & \text{if} \quad \tilde{q}^j_{i,t} > \tilde{q}^{-j}_{i,t} \end{cases}. \tag{7}$$

Now, combining (6) and (7), the output of intermediate i as a function of tax-adjusted labor productivities output can be written as

$$y^j_{i,t} = \begin{cases} \frac{\tilde{Y}\tilde{q}^j_{i,t}}{w^u_t} & \text{if} \quad \tilde{q}^c_{i,t} = \tilde{q}^d_{i,t} \\ \frac{\tilde{Y}\tilde{q}^j_{i,t}}{\lambda w^u_t} & \text{if} \quad \tilde{q}^j_{i,t} > \tilde{q}^{-j}_{i,t} \end{cases}. \tag{8}$$

Then, the equilibrium profits (gross of R&D expenditures), as a function of m and n, can be expressed follows

$$\begin{array}{llll} \pi^c_{n,t} = \tilde{Y}_t \frac{\lambda-1}{\lambda} & \pi^d_{n,t} = 0 & \text{if} & n < m \\ \pi^c_{n,t} = 0 & \pi^d_{n,t} = \tilde{Y}_t \frac{\lambda-1}{\lambda} & \text{if} & n > m \\ \pi^c_{n,t} = 0 & \pi^d_{n,t} = 0 & \text{if} & n = m \end{array}. \tag{9}$$

2.8 Innovation Incentives

We now characterize innovation incentives, which are the only forward-looking part of firm behavior in our model. To simplify the exposition, we first assume that firms are myopic and maximize instantaneous (one-step ahead profits) rather than discounted sum of profits. This enables us to provide analytical expressions for R&D decisions, clarifying the basic economic forces. We will then turn to forward-looking maximization by firms and show that exactly the same expressions and intuitions apply, with the only exception that one term will then be replaced by a solution to a Hamilton-Jacobi-Bellman (HJB) equation rather than being explicitly given as in this subsection.

Not every successful innovation leads to profitable production for two reasons. First, the innovation might be in technology j which is behind technology $-j$, and thus may still not be active even after the improvement in labor productivity. Second, even if it leads to production, this might happen at zero markup if the tax-adjusted labor productivities are the same with the two technologies. Clearly, innovation incentives will be determined by the probability of generating positive profits following innovation. We denote this probability for innovation directed at sector j by $\Gamma_t^j \in [0,1]$. For dirty sector, this is

$$\begin{aligned}\Gamma_t^d &\equiv (1-\alpha)\sum\nolimits_{n\geq m}\mu_{n,t} + \alpha\left(\mathbb{I}_{(m\leq 0)} + \mathbb{I}_{(m>0)}\sum\nolimits_{n\geq m}\mu_{n,t}\right) \\ &= \sum\nolimits_{n\geq m}\mu_{n,t} + \alpha\left(1-\sum\nolimits_{n\geq m}\mu_{nt}\right)\mathbb{I}_{(m\leq 0)},\end{aligned}$$

where $\mathbb{I}_{(m\leq 0)}$ is the indicator function for the event $m \leq 0$.

The interpretation of this expression is as follows: If the innovation is incremental (which, conditional on successful innovation, has probability $1-\alpha$), then it will only be profitable if it builds on an intermediate technology where the dirty sector is ahead or neck and neck with the clean sector which, given uniform random draws from the set of all intermediates, has probability $\sum_{n\geq m}\mu_{n,t}$. Alternatively, with probability α, the innovating firm will necessarily be at least one step ahead of the competing technology (either the dirty sector is ahead or with the breakthrough technology, it leapfrogs the clean sector). However, in this case, it may leapfrog the clean technology but still not compete with it on the basis of tax-adjusted productivity because of the higher tax on dirty production (i.e., because of a "carbon tax"). In particular, if $m \leq 0$ (so that $\mathbb{I}_{(m\leq 0)} = 1$), then there is no carbon tax (if anything there might be a carbon subsidy), then it will certainly be at least one step ahead of the clean technology and will be able to charge a markup. If, on the other hand, $m > 0$, then the innovation will be profitable only for intermediates where the technology gap is already sufficiently large for the dirty sector to have higher tax-adjusted technology, which is in the sectors with $n \geq m$.

A similar reasoning leads to a probability of positive profit following clean innovation of

$$\begin{aligned}\Gamma_t^c &\equiv (1-\alpha)\sum\nolimits_{n\leq m}\mu_{n,t} + \alpha\left(\mathbb{I}_{(m\geq 0)} + \mathbb{I}_{(m<0)}\sum\nolimits_{n\leq m}\mu_{n,t}\right) \\ &= \sum\nolimits_{n\leq m}\mu_{n,t} + \alpha\left(1-\sum\nolimits_{n\leq m}\mu_{n,t}\right)\mathbb{I}_{(m\geq 0)}. \qquad (10)\end{aligned}$$

Let us denote the expected value of a successful innovation in technology j by $\bar{v}_t^j$. Since in this subsection we are assuming myopic behavior on the sides of firms, this is equal to the expected immediate (rather than discounted) profits from a successful innovation given by

$$\bar{v}_t^j = \frac{\Gamma_t^j(\lambda-1)\tilde{Y}_t}{\lambda}. \qquad (11)$$

Then, dropping the firm subscript i (in $F_{I,i,t}$), the maximization problem of a firm with the leading-edge technology in u^j intermediates in sector $j \in \{c,d\}$ can be written as:

$$\max_{X_{I,t}^j \geq 0} \left\{ X_{I,t}^j \bar{v}_t^j - \left(1 - s_{I,t}^j\right) w_t^s \left[H\left(X_{I,t}^j, u^j\right) + \mathbb{I}_{(X_t^j > 0)} u^j F_{I,t} \right] \right\}, \tag{12}$$

where $H\left(X_{I,t}^j, u^j\right)$ denotes the number of scientist hired by a firm that has u^j product lines and innovates at the rate $X_{I,t}^j$. In this expression, the indicator function allows us to turn off the fixed costs of R&D when the firm chooses not to perform any R&D activities. Dividing this objective function by u^j, the maximization problem of a firm "per leading-edge technology" (i.e., expression (12) divided by the number of products in which the firm has the leading-edge technology in sector j) is

$$\max_{x_{I,t}^j \geq 0} \left\{ x_{I,t}^j \bar{v}_t^j - \left(1 - s_{I,t}^j\right) w_t^s \left[h\left(x_{I,t}^j\right) + \mathbb{I}_{(x_{I,t}^j > 0)} F_{I,t} \right] \right\}. \tag{13}$$

where $h\left(x_{I,t}^j\right) \equiv H\left(X_{I,t}^j, u^j\right)/u^j$ is defined as the average number of scientists hired and $x_{I,t}^j \equiv X_{I,t}^j/u^j$ is the average innovation intensity. Using the R&D production function defined in (3), equilibrium innovation rate for $j \in \{c,d\}$ can be expressed as

$$x_{I,t}^j = \mathbb{I}_{(x_{I,t}^j > 0)} \left(\frac{\bar{v}_t^j \eta \theta^{\frac{1}{\eta}}}{\left(1 - s_{I,t}^j\right) w_t^s} \right)^{\frac{\eta}{1-\eta}}. \tag{14}$$

A number of important conclusions follow from (14):

1. Higher net output, higher markups and lower scientist wages increase research effort as should be expected.

2. Subsidies to research increase research effort. This will be important in encouraging clean innovation by means of research subsidies.

3. Through the Γ_t^j's, carbon taxes increase clean research effort (and reduce dirty research effort). This can be seen by considering higher values of m in (10), which given the distribution of technology gaps, increases Γ_t^c, because production shifts from dirty to clean technologies (and neck-and-neck sectors shift to positive markups for clean technologies). This shows that just carbon taxes may be sufficient to encourage clean innovation and thus a transition to clean technology. Whether they will in fact be sufficient is an empirical and quantitative question we will try to address below.

4. Again through the Γ_t^j's, we can also see the path-dependent nature of innovation. When there are large technology gaps between dirty and clean, $\sum_{n \leq m} \mu_{n,t}$ will be very small, and thus Γ_t^c will be small (and Γ_t^d will be high), discouraging clean innovation and

encouraging dirty innovation. But if clean innovation can be maintained for a while, then $\sum_{n\leq m}\mu_{nt}$ will increase, and so will Γ_t^c (while Γ_t^d declines). Thus clean innovation will naturally self-reinforce over time. To the extent that $\sum_{n\leq m}\mu_{n,t}$ changes only slowly, this will also be a slow process.

2.9 Free Entry and Labor Market Clearing

The previous subsection characterized the R&D decisions of the incumbents (as a function of the state of the economy and policies). The other component of R&D, creating demand for scientists, is from entrants. With a similar reasoning to the profitability of the R&D of incumbents, the free entry condition for entrants for technology $j \in \{c,d\}$ can be written as

$$\max_{x_{E,t}^j \geq 0}\left\{x_{E,t}^j \bar{v}_t^j - \left(1-s_{E,t}^j\right) w_t^s \left[h\left(x_{E,t}^j\right)+F_E\right]\right\} \leq 0, \tag{15}$$

with this condition holding as equality if $E_t^j > 0$. Hence, the innovation rate by entrants is

$$x_{E,t}^j = \mathbb{I}_{(x_{E,t}^j>0)}\left(\frac{\bar{v}_t^j \eta \theta^{\frac{1}{\eta}}}{\left(1-s_{E,t}^j\right) w_t^s}\right)^{\frac{\eta}{1-\eta}} \quad \text{for } j\in\{c,d\}. \tag{16}$$

Inspection of (15) establishes that at time t, there can be positive entry into technology j only if the "policy-adjusted" value of innovation is higher in sector j than in sector $-j$. In other words, entrants will direct their R&D to the clean technology if $\bar{v}_t^c/\left(1-s_{E,t}^c\right) > \bar{v}_t^d/\left(1-s_{E,t}^d\right)$ and to the dirty technology if the reverse inequality holds. We also adopt the tiebreaking rule that if $\bar{v}_t^c/\left(1-s_{E,t}^c\right) = \bar{v}_t^d/\left(1-s_{E,t}^d\right)$, then half of the entrants will direct R&D to each sector. Therefore, denoting the total number (measure) of entrants at time t by E_t, we have that the number of entrants with technology directed to sector j is given by

$$E_t^j = \begin{cases} E_t & \text{if } \bar{v}_t^j/\left(1-s_{E,t}^j\right) > \bar{v}_t^{-j}/\left(1-s_{E,t}^{-j}\right) \\ 0 & \text{if } \bar{v}_t^j/\left(1-s_{E,t}^j\right) < \bar{v}_t^{-j}/\left(1-s_{E,t}^{-j}\right) \\ E_t/2 & \text{if } \bar{v}_t^j/\left(1-s_{E,t}^j\right) = \bar{v}_t^{-j}/\left(1-s_{E,t}^{-j}\right) \end{cases}.$$

A comparison of equations (14) and (16) shows, conditional on entry an entrant's innovation rate (directed to sector $j\in\{c,d\}$) will only be different from an incumbent's in (14) because of differential subsidies.

It is also useful to inspect the R&D to sales relationship implied by our model. Suppose that free entry condition holds for entry directed at technology $k\in\{c,d\}$. Conditional on investing in R&D, $x_{I,t}^j > 0$, the equilibrium R&D to sales ratio (for $j\in\{c,d\}$) would be:

$$\frac{R\&D_{i,t}^j}{Sales_{i,t}^j} = \eta^\eta (1-\eta)^{1-\eta}\frac{\lambda-1}{\lambda}\frac{\Gamma_t^k}{1-s_{E,t}^k}\theta F_E^{-(1-\eta)}\left[\left(\frac{\Gamma_t^j\left(1-s_{E,t}^k\right)}{\Gamma_t^k\left(1-s_{I,t}^j\right)}\right)^{\frac{1}{1-\eta}}\frac{\eta}{1-\eta}F_E + F_{I,i,t}\right].$$

Note that higher profitability of R&D in the sector for which the free entry condition holds increases the R&D to sales ratio of that sector, but may reduce it in the other sector. The impact of fixed cost requirements of incumbents on R&D to sales ratio result is positive. However, the impact of the fixed cost of entry is ambiguous. On the one hand, it reduces the equilibrium wage, and thus R&D expenditure. On the other, it increases labor requirements, increasing R&D expenditures. The interplay of these two forces makes R&D to sales ratio non-monotonic in the fixed cost for entrance. These different impacts of the fixed cost for incumbents and entrants will enable us to identify both parameters in the estimation.

The labor market clearing condition for scientists can be written as

$$L^s = \sum_{j\in\{c,d\}} \left[\begin{array}{c} \left(\left(\frac{\bar{v}_t^j \theta\eta}{(1-s_{E,t}^j) w_t^s} \right)^{\frac{1}{1-\eta}} + F_E \right) E_t^j \\ + \int_0^1 \mathbb{I}_{(x_{it}^j>0)} \left(\left(\frac{\bar{v}_t^j \theta\eta}{(1-s_{E,t}^j) w_t^s} \right)^{\frac{1}{1-\eta}} + F_{I,i,t} \right) di \end{array} \right]. \tag{17}$$

This equation shows that the demand for scientists is decreasing in the skilled wage w_t^s and will be higher when R&D is more profitable and is subsidized more heavily.

We next characterize labor market clearing for production workers. From the equilibrium production decision in (8) the unskilled labor demand is

$$l_{i,t} = \begin{cases} \frac{\tilde{Y}_t}{(1+\tau_i^j) w_t^u} & \text{if } \tilde{q}_{i,t}^j = \tilde{q}_{i,t}^{-j} \\ \frac{\tilde{Y}_t}{(1+\tau_i^j)\lambda w_t^u} & \text{if } \tilde{q}_{i,t}^j \neq \tilde{q}_{i,t}^{-j} \end{cases}$$

Substituting the optimal quantities (8) into the final good production function (2),

$$w_t^u = \bar{Q}_t \Lambda_t^\mu, \tag{18}$$

where

$$\bar{Q}_t \equiv \exp\left(\int \ln \bar{q}_{it} di \right)$$

is the quality index of active tax-adjusted labor productivities, and

$$\Lambda_t^\mu = \lambda^{-\left(1-\mu_{m,t}\right)}$$

is an inverse function of equilibrium markups (where $\mu_{m,t}$ refers to the fraction of product lines where the lead of dirty is exactly equal to m steps, so that clean and dirty are neck and neck in tax-adjusted productivity). In particular, Λ_t^μ takes the value λ^{-1} when all intermediates charge a markup (which is the case when $\mu_{m,t} = 0$) and the value 1 when no intermediates charger markup (which is the case when $\mu_{m,t} = 1$). The labor market clearing for production workers can then be expressed as

$$1 = \frac{\tilde{Y}_t}{w_t^u} \left\{ \frac{\mu_{m,t}}{2} \left(\frac{1}{1+\tau_t^d} + \frac{1}{1+\tau_t^c} \right) + \frac{1}{\lambda} \left(\frac{\sum_{n<m} \mu_{n,t}}{1+\tau_t^c} + \frac{\sum_{n>m} \mu_{n,t}}{1+\tau_t^d} \right) \right\}.$$

This equation shows both the impact of taxes on labor demand (both types of taxes reduce labor demand and thus wages) and the distribution of technology gaps (because these affect markups). It also shows that if there were only one type of technology, an increase in the tax rate would have no impact on production, just reducing the unskilled wage rate. This is no longer true, however, with two types of technologies, because a tax on dirty technology, for example, would also change the prices of intermediates produced by dirty technology relative to those produced by clean technology, thus impacting production.

This equation also enables us to express aggregate output as a function of the quality index of active tax-adjusted labor productivities as follows

$$Y_t = \exp\left(-\gamma\left(S_t - \bar{S}\right)\right)\tilde{Y}_t = \frac{\bar{Q}_t \Lambda_t^{\mu}}{\Omega_t^{\mu}\exp\left(\gamma\left(S_t - \bar{S}\right)\right)}, \tag{19}$$

where

$$\Omega_t^{\mu} \equiv \frac{\mu_{m,t}}{2}\left(\frac{1}{1+\tau_t^d} + \frac{1}{1+\tau_t^c}\right) + \frac{1}{\lambda}\left(\frac{\sum_{n<m}\mu_{n,t}}{1+\tau_t^c} + \frac{\sum_{n>m}\mu_{n,t}}{1+\tau_t^d}\right)$$

is an adjustment for labor demand coming both from taxes and markups.

2.10 Dynamics and Equilibrium Redux

Equilibrium dynamics are determined by changes in the interest rate and the evolution of technologies and technology gaps. Household maximization leads to the usual Euler equation

$$g_t = r_t - \rho, \tag{20}$$

where g_t is the growth rate of consumption and r_t is the interest rate at time t (and in addition we impose the usual transversality condition).

The evolution of technology gaps $\mu_{n,t}$ can be derived as follows. Let us denote the aggregate innovation rate in technology j as $z_t^j \equiv \left(1 + E_t^j\right) x_t^j$ and the total innovation rate as $z_t \equiv z_t^d + z_t^c$. Then, the flow equations for the distribution of technology gap $n > 1$ can be expressed as

$$\dot{\mu}_{n>1,t} = z_t^d \mu_{n-1,t} + (1-\alpha)\, z_t^c \mu_{n+1,t} - z_t \mu_{n,t}.$$

The change in the share depends on the difference between inflows and outflows. There will be inflows into state n from $n-1$ when a dirty innovation occurs and from $n+1$ when a clean innovation occurs without leapfrogging. On the other hand, an outflow will happen with both clean or dirty innovation as it will bring the state into $n+1$, $n-1$ or -1 depending on the innovation type. We repeat the same reasoning for $n \leq 1$ below:

$$\begin{aligned}
\dot{\mu}_{1,t} &= z_t^d \mu_{0,t} + (1-\alpha)\, z_t^c \mu_{2,t} + \alpha z^d \mu_{-t}^c - z\mu_{1,t} \\
\dot{\mu}_{0,t} &= (1-\alpha)\, z^d \mu_{-1,t} + (1-\alpha)\, z_t^c \mu_{1,t} - z\mu_{0,t} \\
\dot{\mu}_{-1,t} &= z^c \mu_{0,t} + (1-\alpha)\, z_t^d \mu_{-2,t} + \alpha z_t^c \sum_{n>0} \mu_{n,t} - z\mu_{-1,t} \\
\dot{\mu}_{n<-1t} &= z^c \mu_{n+1,t} + (1-\alpha)\, z_t^d \mu_{n-1,t} - z\mu_{n,t}.
\end{aligned} \tag{21}$$

Total dirty intermediate production at time t, Y_t^d, which creates pollution is given as

$$\begin{aligned} Y_t^d &= \int y_{i,t}^d di = \int_{i\in\mu_m} \frac{y_{i,t}^d}{2} di + \sum_{n>m} \int_{i\in\mu_n} y_{i,t}^d di \\ &= \frac{\tilde{Y}_t}{\left(1+\tau_t^d\right) w_t^u} \left[\frac{1}{2} Q_{m,t}^d + \frac{1}{\lambda} \sum_{n>m} Q_{n,t}^d\right], \end{aligned} \tag{22}$$

where we break up the productivity aggregates by step size differential n defining (with a slight abuse of notation where $i \in \mu_n$ denotes intermediates where the technology gap is n steps):

$$Q_{n,t}^d \equiv \int_{i\in\mu_n} q_{i,t}^d di.$$

We now summarize the dynamic equilibrium path using the equations we have derived in this section. For any given time path of policies $\left[\tau_t^j, s_{I,t}^j, s_{E,t}^j\right]_{t=0}^{\infty}$, a dynamic equilibrium path is characterized by time path of

$$\left[y_{i,t}^j, p_{i,t}^j, x_{I,t}^j, x_{E,t}^j, w_t^s, w_t^u, E_t^j, \{\mu_{n,t}\}_{n=-\infty}^{\infty}, \{Q_{n,t}^d\}_{n=-\infty}^{\infty}, g_t, r_t, S_t\right]_{t=0}^{\infty}$$

such that $[i]$ $y_{i,t}^j$ and $p_{i,t}^j$ maximize profits as in (7) and (8); $[ii]$ $x_{I,t}^j$ and $x_{E,t}^j$ solve incumbent's and entrant's R&D decision as in (14) and (16); $[iii]$ w_t^u clears unskilled labor market as in (18); $[iv]$ w^s is determined from the free entry condition (15) when there is positive entry and from skilled labor market clearing (17) when there is no positive entry; $[v]$ E_t^j is determined from the skilled labor market clearing (17) when there is positive entry; $[vi]$ technology gap shares $\{\mu_{n,t}\}_{n=-\infty}^{\infty}$ satisfy the set of flow equations (21); $[vii]$ total productivity of the sectors with n-step gap $Q_{n,t}^d$ evolves according to the innovation rates in (14) and (16), $[viii]$ the growth rate is consistent with the innovation rates $x_{I,t}^j$ and $x_{E,t}^j$; and $[ix]$ the interest rate satisfies the Euler equation (20), and $[x]$ S_t is given by (5).

2.11 Full Model

We now relax the assumption of myopic firms and assume that firms maximize their discounted profits (and this full model will be used in our quantitative analysis also).

Let $\vec{n}^j \equiv \left[n_1^j, ..., n_u^j\right]$ denote the vector of product lines where the firm in question holds the leading-edge technology (a total of $u = u_t^j$ of them for this firm) and n_i^j the technology gap compared to technology $-j$ within the same product line. Let $\vec{n}_{-i}^j$ denote the same vector $\vec{n}^j$ without its ith element n_i^j. Then the value of a firm with a portfolio of products given by $\vec{n}^j$ then satisfies the HJB equation:

$$\begin{aligned} & rV_{\vec{n}^j,t}^j - \dot{V}_{\vec{n}^j,t}^j \\ = & \sum\nolimits_{i=1}^{u} \left[\pi_{n_i,t}^j + z_t^j \left(V_{\vec{n}_{-i}^j}^j - V_{\vec{n}^j,t}^j\right) + z_t^{-j}(1-\alpha)\left(V_{\vec{n}_{-i}^j \cup \{n_i^j - 1\},t}^j - V_{\vec{n}_i^j,t}^j\right) + z_t^{-j}\alpha\left(V_{\vec{n}_{-i}^j,t}^j - V_{\vec{n}_i^j,t}^j\right)\right] \\ & + \int \max_{x_t^j \geq 0} \left[u_t^j x_t^j \left(V_{\vec{n}^j \cup \{n_{u+1}^j\},t}^j - V_{\vec{n}^j,t}^j\right) - \left(1 - s_{I,t}^j\right) u_t^j w_t^s \left(\left(x_t^j\right)^{\frac{1}{\eta}} \theta^{-\frac{1}{\eta}} + \mathbb{I}_{(x_{n,t}^d > 0)} F_{I,t}\right)\right] dF_{I,t}. \end{aligned} \tag{23}$$

The interpretation is straightforward. The right-hand side includes the profits generated from u product lines, which is given by the first term. In addition, at the flow rate rate z_t^j, each product line i will experience an innovation by another firm from the same technology j in which case i is taken out of firm's portfolio (so that the firm's portfolio becomes $\vec{n}_{-i}^j$). If instead production line i experiences an innovation from the alternative technology $-j$, which happens at the rate z_t^{-j}, then there are two possibilities: either the innovation is incremental (probability $(1-\alpha)$) and the current incumbent will still continue with its production in which case the technology gap will be narrowed by one step (so that $n_i^j = n_i^j - 1$) or the innovation might be drastic (probability α), in which case the firm will lose this product line (again reducing its portfolio to $\vec{n}_{-i}^j$). Finally, the firm invests in R&D itself and innovates at the flow rate $X_t^j = u_t^j x_t^j$, and the option value of this R&D (inclusive of costs) is added as the second line of the right-hand side, with the integral taking care of the fact that fixed costs are stochastic. (Note in particular that when it is successful, the firm adds a new product line so that its portfolio becomes $\vec{n}^j \cup \{n_{u+1}^j\}$).

The next lemma provides a convenient re-expression of this Bellman equation in per product terms:

Lemma 1 *Equation* (23) *can be re-expressed as* $V_{\vec{n}^j,t}^j = \tilde{Y}_t \sum_{i=1}^u v_{n_i,t}^j$ *where*

$$\begin{aligned} \rho v_{n_i,t}^j - \dot{v}_{n_i,t}^j &= \pi_{n_i}^j - z_t^j v_{n_i,t}^j + z_t^{-j}(1-\alpha)\left(v_{n_i-1,t}^j - v_{n_i,t}^j\right) + z_t^{-j}\alpha\left(v_{-1,t}^j - v_{n_i,t}^j\right) \\ &+ \int \max_{x_t^j \geq 0}\left[x_t^j \bar{v}_t^j - \left(1 - s_{I,t}^j\right)\tilde{w}_t^s\left(\left(x_t^j\right)^{\frac{1}{\eta}}\theta^{-\frac{1}{\eta}} + \mathbb{I}_{(x_{n,t}^d>0)}F_{I,t}\right)\right] dF_{I,t}. \end{aligned} \tag{24}$$

and $\bar{v}_t^j \equiv \mathbb{E}_i v_{n_i,t}^j$.

Proof of Lemma 1. Substituting (24) into (23), we obtain

$$\begin{aligned} & r\tilde{Y}_t \sum_{i=1}^u v_{n_i,t}^j - \frac{d}{dt}\tilde{Y}_t \sum_{i=1}^u v_{n_i,t}^j - \tilde{Y}_t \sum_{i=1}^u \dot{v}_{n_i,t}^j \\ = \; & \sum_{i=1}^u \left[\tilde{Y}_t \pi_{n_i}^j - z_t^j \tilde{Y}_t v_{n_i,t}^j + z_t^{-j}(1-\alpha)\left(\tilde{Y}_t v_{n_i-1,t}^j - \tilde{Y}_t v_{n_i,t}^j\right) + z_t^{-j}\alpha\left(\tilde{Y}_t v_{-1,t}^j - \tilde{Y}_t v_{n_i,t}^j\right)\right] \\ & + \int \max_{x_t^j \geq 0}\left[u x_t^j \tilde{Y}_t \bar{v}_t^j - \left(1 - s_{I,t}^j\right) u w_t^s\left(\left(x_t^j\right)^{\frac{1}{\eta}}\theta^{-\frac{1}{\eta}} + \mathbb{I}_{(x_{n_i,t}^d>0)}F_{I,t}\right)\right] dF_{I,t}. \end{aligned}$$

where

$$\tilde{w}_t^s \equiv \frac{w_t^s}{\tilde{Y}_t}$$

and

$$\begin{array}{llll} \pi_n^c = \frac{\lambda-1}{\lambda} & \pi_n^d = 0 & \text{if} & n < m \\ \pi_n^c = 0 & \pi_n^d = \frac{\lambda-1}{\lambda} & \text{if} & n > m \\ \pi_n^c = 0 & \pi_n^d = 0 & \text{if} & n = m \end{array} .$$

Then this can further be simplified to

$$
\begin{aligned}
& r \sum\nolimits_{i=1}^{u} v_{n,t}^{j} - g_t \sum\nolimits_{i=1}^{u} v_{n,t}^{j} - \sum\nolimits_{i=1}^{u} \dot{v}_{n,t}^{j} \\
= & \sum\nolimits_{i=1}^{u} \left[\pi_{n_i}^{j} - z_t^{j} v_{n,t}^{j} + z_t^{-j} (1-\alpha) \left(v_{n-1,t}^{j} - v_{n,t}^{j} \right) + z_t^{-j} \alpha \left(v_{-1,t}^{j} - v_{n,t}^{j} \right) \right] \\
& + u \int \max_{x_t^j \geq 0} \left[x_t^j \bar{v}_t^j - \left(1 - s_{I,t}^{j} \right) \tilde{w}_t^s \left(\left(x_t^j \right)^{\frac{1}{\eta}} \theta^{-\frac{1}{\eta}} + \mathbb{I}_{(x_{n,t}^d > 0)} F_{I,t} \right) \right] dF_{I,t}.
\end{aligned}
$$

where we used the fact that $\frac{d}{dt}\tilde{Y}_t = g_t \tilde{Y}_t$ in the first line. Next, we eliminate $\tilde{Y}_t$ throughout and use the Euler equation (20), $r_t - g_t = \rho$, to establish the desired result. ■

An important implication is that incumbent innovation rates per product line will be independent of the portfolio of the incumbent and given by

$$
x_t^j = \mathbb{I}_{(x_{I,t}^j > 0)} \left(\frac{\bar{v}_t^j \eta \theta^{\frac{1}{\eta}}}{\left(1 - s_{I,t}^j \right) \tilde{w}_t^s} \right)^{\frac{\eta}{1-\eta}} \quad \text{for } j \in \{c, d\},
$$

which can be easily verified to be identical to (14) except that now $\bar{v}_t^j$ is given as a solution to (24).

We can now describe the full dynamic equilibrium path of this economy, which will be essentially identical to the equilibrium path with myopic firms, with $\bar{v}_t^j$ given as the solution to the HJB equation (24).

For any given time path of policies $\left[\tau_t^j, s_{I,t}^j, s_{E,t}^j \right]_{t=0}^{\infty}$, a dynamic equilibrium path is characterized by time path of

$$
\left[\bar{v}_t^j, y_{i,t}^j, p_{i,t}^j, x_{I,t}^j, x_{E,t}^j, w_t^s, w_t^u, E_t^j, \{\mu_{n,t}\}_{n=-\infty}^{\infty}, \{Q_{n,t}^d\}_{n=-\infty}^{\infty}, g_t, r_t, S_t \right]_{t=0}^{\infty}
$$

such that $\bar{v}_t^j \equiv \mathbb{E}_i v_{n_i,t}^j$, and each $v_{n,t}^j$ satisfies (24). In addition: [*i*] $y_{i,t}^j$ and $p_{i,t}^j$ maximize profits as in (7) and (8); [*ii*] $x_{I,t}^j$ and $x_{E,t}^j$ solve incumbent's and entrant's R&D decision as in (14) and (16); [*iii*] w_t^u clears unskilled labor market as in (18); [*iv*] w^s is determined from the free entry condition (15) when there is positive entry and from skilled labor market clearing (17) when there is no positive entry; [*v*] E_t^j is determined from the skilled labor market clearing (17) when there is positive entry; [*vi*] technology gap shares $\{\mu_{n,t}\}_{n=-\infty}^{\infty}$ satisfy the set of flow equations (21); [*vii*] total productivity of the sectors with n-step gap $Q_{n,t}^d$ evolves according to the innovation rates in (14) and (16), [*viii*] the growth rate is consistent with the innovation rates $x_{I,t}^j$ and $x_{E,t}^j$; and [*ix*] the interest rate satisfies the Euler equation (20), and [*x*] S_t is given by (5).

3 Empirical Strategy and Data

Our model has 14 parameters/variables to be determined:

$$
\left\{ \rho, \bar{S}, \gamma, \varphi, \varphi_0, \varphi_P, \kappa, L^s, \alpha, \eta, \theta, \lambda, F_I, F_E \right\}.
$$

In addition, the initial distribution of technology gaps between clean and dirty technologies, $\{\mu_{0t}\}_{n=-\infty}^{\infty}$, needs to be determined. It is useful to note that, as will become clearer below, given $\{\mu_{nt}\}_{n=-\infty}^{\infty}$, estimation of the remaining parameters can be done without knowledge of taxes and subsidies, and also without any information on $\bar{S}, \gamma, \varphi, \varphi_0$, and φ_P. These become relevant only for our policy analysis. Nevertheless, here we specify our choices for all these parameters.

We proceed in four steps. First, we externally calibrate some of the parameters, in particular the parameters of the carbon cycle and the discount rate. In all, the parameters $\rho, \bar{S}, \gamma, \varphi, \varphi_0, \varphi_P$, and κ are taken from external sources. Second, we directly estimate L^s, α, and η from microdata. Third, we choose the initial distribution of technology gaps to match the distribution of patents between firms innovating mostly with clean and mostly with dirty technologies as we explain below. Finally, we estimate the remaining parameters θ, λ, F_I and F_E using simulated method of moments, with moments being selected to model the firm-level R&D behavior, growth rates, and entry/exit rates for the energy sector as we describe below. The model performs well and is able to replicate these moments reasonably closely.

Throughout our focus is on the energy sector, the behavior of which has motivated our theoretical model. The energy sector is defined as firms involved in the sourcing, refinement and delivery of energy inputs for residential and industrial applications (e.g., oil and gas, electricity), firms that provide complementary inputs and equipment into this energy-production process (e.g., drilling equipment, power plant technologies), and firms that interface with the energy inputs for residential and industrial use (e.g., motor manufacturers). As such, our group of 1576 firms that make up our sample includes oil and gas producers, mining and exploration firms, engine manufacturers, power companies building upon multiple techniques, energy equipment manufacturers, and similar.[9]

The data we use for estimation comes from the Census Bureau's Longitudinal Business Database and Economic Censuses, the National Science Foundation's Survey of Industrial Research and Development, and the NBER Patent Database. We design our sample around innovative firms in the energy sector that are in operation during the 1975-2004 period.

3.1 External Calibration

We choose $\bar{S}, \gamma, \varphi, \varphi_0, \varphi_P$, and κ to link our model to the carbon cycle and its impact on aggregate output following Golosov et al.'s (2011) approach. This approach takes into account the current level of carbon stock and its increase since pre-industrial times; the rate at which new emissions enter the atmosphere, the terrestrial biosphere or shallow oceans, and the deep oceans; how that movement and the various reservoirs of carbon influence the earth's temperature; and how higher temperatures and environmental damage hurt the economy. This

[9] We exclude approximately 50 non-profit research centers and similar entities to match our model's focus on profit-seeking firms. Our estimations below are robust to retaining these entities.

work builds upon prior work in environmental economics (e.g., Nordhaus 2008, Nordhaus and Boyer 2000), but is more flexible in allowing non-linear absorption of atmospheric carbon, but does not allow any delay on the impact of this carbon content on economic outcomes and temperature changes (which result from different rates at which oceans change temperature, for example) and does not separately keep track of the dynamics of the atmospheric concentration of carbon dioxide (CO_2), methane (CH_4) and nitrous oxide (N_2O).

Our value of the pre-industrial stock of carbon dioxide in the atmosphere $\bar{S}$ is 581 GtC (gigatons of carbon). To model how emission increases the atmospheric stock of CO2, we define the three parameters φ, φ_0, and φ_P as follows. First, φ_P is the portion of new emissions that will remain in the atmosphere for a very long time, likely for thousands of years, and estimates of this parameter from IPCC (2007) and Archer (2005) are about 20%. The other two parameters, φ and φ_0, govern the short- and medium-term movement of the emitted carbon that will not become part of this very long duration stock in the atmosphere. These emissions influence the earth's temperature over short horizons, but they are ultimately absorbed into the deep oceans. To identify these parameters, we utilize the 30 year half-life of carbon and match the carbon stock evolution under emissions during the 1900-2008 period. We use the following formula to determine the atmospheric carbon concentration S_t in every year during 1900-2008 period

$$S_t = \int_0^{t-1900} (1 - d_l)\, K_{t-l} dl + S_{1900}, \tag{25}$$

where

$$d_l = (1 - \varphi_P) \left[1 - \varphi_0 e^{-\varphi l}\right]$$

The emission data for $\{K_t\}_{t=1900}^{2008}$ is shown in Figure 1 below.

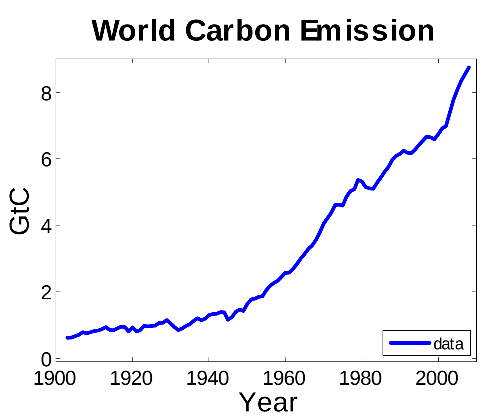

FIGURE 1

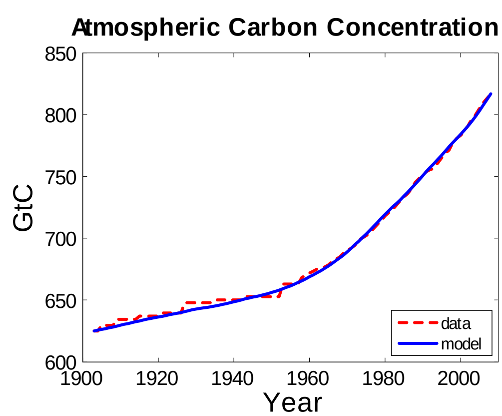

FIGURE 2

Figure 2 shows the fit of (25) which yields $\varphi = 0.0313$ and $\varphi_0 = 0.7661$. The dynamics implied by equation (5) at these parameter values match the actual evolution of atmospheric carbon over the past century very well as shown by the close correspondence between the solid blue line representing the data and the dashed red line corresponding to the model-implied atmospheric concentration in Figure 2.

Our damage function also follows Golosov et al. (2011) and we choose the γ parameter in our baseline policy analysis to be the same as theirs, 5.3 x 10^{-5} GtC^{-1},[10] though this number may be too low, particularly because, in contrast to their paper, we are not allowing policy to adjust to new information about damages as this becomes available, so the certainty equivalent average of estimates rather than arithmetic average might be more appropriate. Section 6.1 provides robustness checks with higher values of γ.

The κ parameter is chosen to link current emissions levels to the baseline output level of the model. In doing so, we are making a simplifying assumption that the emission of our economy can proxy for the emission of the entire US economy. As suggested in Golosov et al. (2011), the modeling of the carbon cycle and its impact on production has the attractive feature that the social value of marginal emissions is the same (relative to output). This implies that our results would be essentially unchanged if we take a future path of emissions from the rest of the world, with the only difference being that the implied temperature changes we depict below would no longer apply (and we would need to talk about incremental temperature changes due to the US energy sector). It is also worth noting that our model and this modeling strategy certainly abstract from several important aspects of international cooperation or competition that impact climate change outcomes (e.g., Hassler and Krusell 2012).

Finally, we report all of our results for a single private discount rate $\rho = 1\%$ and two values of the social discount rate 1% and 0.1%. The first is $\rho = 1\%$, which is close to the 1.5% chosen by Nordhaus in his models, and the second is $\rho = 0.1\%$ used by Stern (2007), on the basis that with a higher discount rate there is almost no weight put on the welfare of future generations.

3.2 Sample Construction and Data Sources

We combine several datasets for this study. The NBER Patent Database and the NSF Survey of Industrial Research and Development are the backbones for our study, with additional information and details being collected from the Longitudinal Business Database and the various Economic Censuses conducted by the Census Bureau. We introduce each dataset as we describe the steps in our sample construction.

3.2.1 Patent Data for Energy Sector

Our first data source is the individual records of all patents granted by the United States Patent and Trademark Office (USPTO) from January 1975 to May 2009. This database was

[10] This approach provides a fairly good approximation of the damage function developed in Nordhaus (2007), who incorporates a typical estimate that a doubling of the stock of atmospheric carbon leads to a 3°C increase and then a proportional damage function of how global temperature increases affect the economy. Golosov et al. (2011) show a close correspondance of these functions over relevant temperature ranges.

first developed by the NBER and was subsequently extended by HBS Research to include patenting in recent years. Each patent record provides information about the invention and the inventors submitting the application. Hall et al. (2001) provide extensive details about these data, and Griliches (1990) surveys the use of patents as economic indicators of technology advancement. We collect from this database the patents that are 1) filed by inventors living in the United States at the time of the patent application, and 2) assigned to industrial firms. In a representative year, 1997, this group comprised about 77 thousand patents (about 40% of the total USPTO patent count in 1997).

We then identify patents that are related to the energy sector. This is a key step for our study, and we outline our approach in detail. We use patent technology codes to identify patents related to the energy sector. The technology codes are the most disaggregated level of the USPTO's classification scheme and number over 150,000. This is important as energy-sector patents are spread out over multiple patent classes (the next higher level of the classification system with about 450 groups). Two examples related to solar energy are "Power Plants/utilizing natural heat/Solar" and "Stoves and Furnaces/Solar heat collector". Moreover, we describe later how these patenting technologies are also used to classify firms as being primarily clean- or dirty-energy firms. This separation can only be done at the technology level as the patent class level includes both types (e.g., "Power Plants" includes technologies for coal-powered plants too).

We identify relevant technology codes through four steps. First, we adopt the prior classifications developed in a study of alternative energy by Popp (2002) and Popp and Newell (2012). Popp and Newell's (2012) work is particularly helpful in that they provide classifications into various types of energy technologies that we discuss in greater detail below. Given this authoritative prior work, we also report results below for our key parameters that just use their classification system.

We are interested, however, in several technologies (e.g., nuclear power) not considered by Popp and Newell (2012). We thus extend their list through three additional steps. Our second step utilizes resources from the OECD's work to identify environment-related technologies. OECD (2011) lists such technologies using the International Patent Classification (IPC) scheme, which some observers believe is better designed to identify and group environment-related technologies than the USPTO classification framework. We use concordances between the IPC and USPTO framework to identify additional technologies to be included.

The third step utilizes information on energy-related R&D data from the NSF Survey of Industrial Research and Development that we describe in greater detail next. For the firms identified in this survey to be conducting energy-related R&D, we list their patent technology codes and frequency. We then manually search the 600 most-frequent codes to identify if they are energy related. In a related fourth step, we also manually search the USPTO database using key words like "Coal" and "Solar" to determine relevant technologies. This identification process constructs a pool of patents related to the energy sector. As a representative year, 1997, our energy-related patents comprised 7.6% of the total US patent count.[11]

3.2.2 Operating Data for Energy Sector

Our next step links our energy patent data to firm-level operating data collected by the US Census Bureau. The Longitudinal Business Database is a business registry that contains annual employment levels for every private-sector establishment in the United States with payroll from 1976 onward. We also employ Economic Censuses that are conducted every five years by sector of the economy; these censuses provide additional plant and firm operation data (e.g., sales). Sourced from US tax records and Census Bureau surveys, these micro-records document the universe of establishments and firms, making them an unparalleled laboratory for studying our model of firm dynamics in the US energy sector.

We match the patent data to these operating data using firm-name and geographic-location matching algorithms. The basic concept in these algorithms is to identify Census Bureau firms that have similar names to USPTO patent assignees and that have establishments in the same geographic area as where inventors of the patents are located.[12] This linkage also accomplishes a related step of aggregating patent assignees to firms, as some firms file patents through multiple patent assignee codes. This aggregation is due to the Census Bureau's establishment-firm hierarchy, as we observe establishment-level names within multi-unit firms that help identify subsidiaries and major corporate restructurings like mergers and acquisitions, and through the name-matching process that consolidates slight name variants over patent assignees.

We focus our sample on the years in which Economic Censuses are conducted, specifically every five years starting in 1977 and ending in 2002. We adopt this focus for several reasons: 1) the operating data are often best measured around these years due to heightened Census Bureau resources, 2) some specific variables from the Economic Censuses are only available at those five-year marks, and 3) our innovation data are most appropriately considered over short time periods. The third rationale is due, for example, to lumpiness in firm applications for patents; as we describe next, our R&D expenditures data are also often collected biannually. We thus measure variables using the average of observed values for firms in five-year windows surrounding these Economic Census years (e.g., 1985-1989 for the 1987-centered period). We have six time periods covering the 30 year interval of 1975-2004.

[11] Energy-related patents account for 5%-15% of US patents over our sample period, with a declining share in recent years; in absolute terms, patent counts for the energy sector are stable or growing throughout the period. The declining share is partly due to the sector not growing as fast as others, and partly due to external changes over time to allow for patents to be made in sectors that traditionally did not patent, especially software patents.

[12] The algorithms are described in detail in an internal Census Bureau report by Ghosh and Kerr (2010). This patent matching builds upon the prior work of Balasubramanian and Sivadasan (2011) and Kerr and Fu (2008).

3.2.3 R&D Data for Energy Sector

We next utilize the Census Bureau's internal linkages to collect information on R&D expenditures from the NSF's Survey of Industrial Research and Development (R&D Survey). The R&D Survey is the US government's primary instrument for surveying the R&D expenditures and innovative efforts of US firms. This is an annual or biannual survey conducted jointly by the Census Bureau and NSF. The survey includes with certainty all public and private firms, as well as foreign-owned firms, undertaking over a minimum threshold of R&D expenditure in the United States. For most of our sample period, this expenditure threshold is one million dollars of R&D within the US. The survey frame also sub-samples firms conducting less than the certain expenditure threshold. These micro-records begin in 1972 and provide the most detailed statistics available on firm-level R&D efforts. In 1997, 3,741 firms reported positive R&D expenditures that sum to $158 billion. Foster and Grim (2010) and Akcigit and Kerr (2010) discuss these data in greater detail.

The R&D Survey provides us with information on many firms' R&D expenditures and employments of science and engineering workers. We use the data, along with the patenting of the firm, to calculate the innovation production function for the sector (e.g., the η and α parameters). We describe these calculations below, and for these calculations we only utilize firm observations for which we always observe reported data on R&D expenditures or scientist counts—meaning that these calculations use only firms that conduct more than the minimum threshold of one million dollars in R&D or are sub-sampled completely. While this might present an issue for sectors like consumer internet start-ups, this is not very restrictive for the supply side of the energy sector given the large amounts of R&D expenditures required by many start-ups.

For our broader moments on firm dynamics, this minimum threshold creates a challenge, however, for the consistent calculation of the entry margin and growth rates. Our model requires that firms be innovative from the start of their lives, and thus an innovative firm that falls below threshold value in its first period would be inappropriately dropped if we restricted the sample only to firms for which we always observe R&D expenditure. By contrast, the patent data are universally observed. To ensure a complete distribution, we thus use patents to impute R&D values for firms that are less than the certainty threshold and not sub-sampled. Overall, our moments combine the R&D and patent data into a single measure of innovation (in R&D terms) that accords with the model for the characterization of firm dynamics in the energy sector. As the R&D expenditures in these sub-sampled cases are low (by definition), this imputation choice versus treating unsurveyed R&D expenditures as zero expenditures conditional on patenting is not very important. The firm does not need to conduct R&D or patent in every year of the initial five-year window, but the firm must do one of the two activities at least once.[13]

To summarize, the key idea is that our sample requires that a firm either patent or have measured R&D in the first period of its life. If the firm is an incumbent in the initial 1975-1979 period, it must have either a patent or measured R&D. Our sample does not condition on innovative activity before 1975-1979. Thus, these incumbents may have had some point in their past when they did not conduct R&D or patent. Our model allows for firms to transition out of R&D, and thus we include firms that quit being innovative. On the other hand, we do not consider non-innovative incumbent firms starting to do innovation. As the probability that an existing, non-innovative firm commences R&D or patenting over the ensuing five years (conditional on survival) is only about 1%, this exclusion is reasonable.[14] As one exception to this sample construction, we only estimate the key innovation production function over firms that have continually observed R&D expenditures (so that imputation procedures are not required).

3.2.4 Sample Inclusion Rules and Sample Size

These procedures define the base pool of innovative firms in the energy sector. To be retained in our final sample, the firm must meet two additional requirements. The first is that the firm has positive employment and obtains three or more patents in the energy sector during the 1975-2004 period. These are not very high hurdles given our purpose, and we thus exclude entities that only obtain one or two energy-related patents over their lifetime. Second, and more important, we also require that 10% of the firm's total patenting be in energy-related fields. This is an important hurdle as it excludes innovative firms that are not very active in the sector. The 10% bar is more substantial than it may initially appear as we have been fairly conservative in terms of defining energy-sector patents.

Thus, our compiled dataset includes innovative firms in the energy sector from 1975-2004. A record in our dataset is a firm-period observation that aggregates over the firm's different establishments. We have 6228 observations from 1576 firms. While focused on a single sector, our firm panel contains 19% of all US R&D industrial expenditures during the 1975-2004 period. The panel accounts for about 70% of industrial patents for the energy sector in the United States. Across all activity in the economy, our sample typically account for 1% of establishments, 5% of employment, and 10% of sales. In the 1997 period, our sample accounts for over a trillion dollars in sales, 3.9 million employees, and 25 billion dollars in R&D expenditures, and the firms obtain 56 thousand patents during 1995-1999.

[13] In a small number of cases where we have scientists counts from the R&D Survey but lack R&D expenditures, we similarly use the scientist counts to impute R&D values for firms below the certainty threshold.

[14] Note that it would have been impossible to build a consistent sample that would also include incumbents switching into innovation. To see why, consider keeping all of the past records for incumbent firms that start conducting R&D in 1987. In the prior periods, this approach would induce a mismeasurement of exit propensities and growth dynamics because a portion of the sample will include firms conditioned on survival until 1987.

Our sample is very important for studying emissions in two ways. First, we account for a substantial amount of activity in several of the main sectors responsible for emissions (e.g., Mueller et al. 2011). In the 1992-1997 period, for example, we account for 59% of sales in industries related to coal and oil extraction, refinement, and shipment; 33% of sales in industries related to electricity production; and 21% of manufacturing sales. Among manufacturing industries, our sample contains higher shares in industries more closely linked with emissions (e.g., 64% in petroleum refinement, 31% in primary metals). Second, while our sample does not include many firms directly from two high-emission sectors, agriculture and transportation, our sample does include many of the manufacturers of products that are key inputs to these sectors.

3.2.5 Designation of Firms as Clean or Dirty Energy

Beyond the development of the firm panel, our leap-frogging calculations below require us to identify whether patents are related to "dirty" or "clean" energy. We do so through the field of patent technologies. We identify patents as related to dirty technologies if they are connected to the extraction, refinement or use of fossil-fuel based energy, including oil, coal, natural gas, and shale technologies. We group into clean-energy patents fields that are related to geothermal, hydroelectric, nuclear, solar, and wind energy. We also include in the clean-energy group identified patents for conservation and utilization of energy. The results below are robust to reclassifications of border group types.

For our model's initial conditions, we also need to identify whether firms are primarily operating in dirty- or clean-energy applications. We do so through a simple rule that has two steps. We first classify a firm-period observation as focused on clean energy if 25% or more of its energy-sector patents are devoted to clean-energy fields; otherwise the firm is classified as a dirty-energy firm in the period. We use the 25% threshold as our assignments of clean-energy fields are conservative compared to dirty-energy fields. We then describe the firm overall as a clean-energy firm if half or more of its time periods achieve this clean-energy focus. The distribution between clean and dirty uses at the firm level is fairly bimodal—96% of observations have 75% or more of their patents in one technology—making the exact details of these procedures less important. In total, 11% of our firms are classified in the clean-energy sub-sector; 14% of energy-sector patents are classified as clean energy.

Several points are worth noting at this stage. First, we generally include technologies that are designed to make fossil fuels cleaner in the dirty-energy group. In this one regard, we deviate from the classifications developed by Popp and Newell (2012) where coal liquefaction and gasification are included in alternative energy, for example. When we directly use Popp and Newell's (2012) technology scheme as a robustness check, we classify the technologies as in their original work. Second, we have not built our sample selection or grouping procedures around technologies related to pollution abatement. We retain all patents for included firms, and thus they are part of our overall technology description, but we only use energy-directed patents to classify patents and firms into dirty- or clean-energy groups. Finally, we also use the more detailed information the R&D Survey collects from selected major R&D producers. We specifically utilize information collected from about 100 firms on R&D expenditures related to specific energy applications like coal or solar energy. We earlier identified one application of this extra information in that we manually searched the major patenting technology codes from these R&D entities to identify energy-sector patenting groups that we should be including. A second application is to verify our data development procedures, for example by assigning firms based upon the types of R&D they conduct rather than observed patents. This group from the R&D Survey also confirms the bimodal nature of our firm groupings. While the group of firms asked these questions is too small and selected to serve as the backbone of our sample, these checks are comforting. While Census Bureau disclosure prevents us from listing firms, we overlap well with Popp and Newell's (2012) listed producers as one example.

3.3 Estimation and Choice of Parameters from Microdata

We first estimate the η parameter from our innovation production function, $X_f = \theta(H_f)^{\eta}(u_f)^{1-\eta}$, which can be rewritten as $\ln(X/u_f) = \ln(\theta) + \eta \cdot \ln(H/u_f)$. We measure X by the firm's count of patents, H by the firm's R&D expenditures or scientist counts, and u by the number of four-digit SIC industries in which a firm is operating. We also check the robustness of our results to using three-digit SIC industry counts, sales and establishment counts as our proxies for firm size u. Our patent count measure is weighted by citations, with citation counts normalized by the average citations achieved by other patents in the same patent class and application year.

To estimate the η elasticity as accurately as possible, we use the panel nature of our data and later return to estimating the θ parameter. As noted earlier, we only use for this exercise firms that have a full panel of reported R&D data. To focus on higher quality data for our differenced estimations, we also require that the firm be present in at least three periods. We first estimate a linear regression with year fixed effects δ_t, yielding

$$\ln(Patents/product_{f,t}) = \delta_t + 0.625\ (0.043) \cdot \ln(R\&D/product_{f,t}) + \epsilon_{f,t}, \tag{26}$$

with standard errors clustered by firm. We then extend the estimation to allow for firm fixed effects, and we estimate the panel elasticity in a first-differenced format, yielding

$$\Delta\ln(Patents/product_{f,t}) = \delta_t + 0.353\ (0.057) \cdot \Delta\ln(R\&D/product_{f,t}) + \epsilon_{f,t}, \tag{27}$$

The range of these point estimates is representative of a broader set of estimates for the η parameter. Table 1a summarizes eight variants of the OLS levels regressions. The rows indicate four measures of firm size u_f: SIC3 industry counts, SIC4 industry counts, sales, and establishment counts. Column headers indicate whether R&D inputs are being measured through expenditures or counts of scientists. The eight coefficients are from eight separate estimations of regression (26). The average of these eight estimations is 0.69, and the estimates are consistently within the range of 0.63-0.76.

Table 1a. OLS Levels Estimates for η Parameter

	R&D Input Measure H_f	
Firm Size Measure u_f:	**R&D Expenditure**	**Scientist Counts**
SIC3 Counts	0.632 (0.042)	0.653 (0.048)
SIC4 Counts	0.625 (0.043)	0.644 (0.048)
Sales	0.761 (0.053)	0.751 (0.048)
Establishments	0.714 (0.039)	0.732 (0.041)

Notes: Table presents eight variants of regression (26).

Table 1b similarly summarizes eight estimation variants of the first-differenced regression (27). The average across these variants is lower at 0.37, with a wider range from 0.29 to 0.51.

Table 1b. OLS First-Differenced Estimates for η Parameter

	R&D Input Measure H	
Firm Size Measure u_f:	**R&D Expenditure**	**Scientist Counts**
SIC3 Counts	0.342 (0.056)	0.286 (0.052)
SIC4 Counts	0.353 (0.057)	0.296 (0.053)
Sales	0.405 (0.075)	0.348 (0.065)
Establishments	0.505 (0.058)	0.455 (0.054)

Notes: Table presents variants of regression (27).

Our baseline value for η is 0.5, taking a mid point within the range of estimates in Tables 1A and 1B.

We also find comparable η parameters in robustness checks off of this sample platform. For example, restricting the sample to firms with energy patents as more than 30% of their innovations yields levels and first-differences estimates of 0.744 (0.065) and 0.384 (0.100), respectively. Restricting our sample to firms that would have been defined for the sector using Popp and Newell's (2012) codes yields levels and first-differences estimates of 0.704 (0.049) and 0.301 (0.071), respectively. Relaxing our requirement that the firm be present in three periods yields levels and first-differences estimates of 0.614 (0.043) and 0.338 (0.056), respectively. We likewise find similar outcomes when incorporating industry-year fixed effects, using unweighted patent counts, or similar adjustments. In addition, Acemoglu et al. (2012b) describe a related instrumental variable elasticity of patenting to science and engineering workers of 0.694 (0.295) across the whole economy developed through H-1B visa reforms estimated by Kerr and Lincoln (2010).

Finally, Table 1C shows estimates from Poisson models that allow for zero patenting outcomes. We report both random effects and fixed effects formats; to conserve space, we only provide two choices of firm size that mostly bound the other variants. Standard errors are bootstrapped. Using four-digit industry counts to measure size consistently delivers elasticities around 0.33, while using establishment counts delivers elasticities around 0.57. Our baseline estimate of $\eta = 0.5$ falls again within these ranges.

Table 1c. Poisson Estimates for η Parameter

	R&D Input Measure H_f	
Technique, Firm Size Measure u_f:	**R&D Expenditure**	**Scientist Counts**
Random Effects, SIC4 Counts	0.326 (0.122)	0.361 (0.079)
Fixed Effects, SIC4 Counts	0.321 (0.106)	0.357 (0.089)
Random Effects, Establishments	0.567 (0.108)	0.584 (0.064)
Fixed Effects, Establishments	0.565 (0.103)	0.583 (0.076)

Notes: Table presents fixed and random effects Poisson estimates similar to Tables 1A and 1B.

We next turn to the α parameter, which in our model describes technology leap-frogging. This process is challenging to model empirically, and we are unfortunately unable to identify exact races between clean and dirty technologies directly within the patent data. This limitation is due to the extreme narrowness of the technology codes that are entirely clean or dirty in application, while patent class divisions are too broad and few in number. We thus instead identify the rate at which patents with exceptional quality emerge using patent citations. We specifically quantify the rate at which patents enter and establish quickly high levels of citations compared to their incumbent peers.

We start with our dataset of all energy-sector patents granted to US inventors during the post-1975 period. We calculate among these energy-sector patents the citation count distribution among incumbent patents by year, excluding within-firm citations. Incumbent patents are defined to be those that were applied for 5-10 years before the focal year; we cap at 10 years prior so that we can have a stable window across a time period from 1985 onwards for analysis. Citations are coming from patents being applied for in the focal year. By conditioning the citation distribution upon a patent receiving a citation in a given year, we are effectively looking at technologies that are being actively used, with many incumbent patents dropping out as no one is building on them.

We then calculate for new patents the citations they receive by year. We designate a major entrant as any patent that has a citation count that exceeds the 90th percentile of the incumbent distribution in any of its first three years. This evaluation approach is designed to keep the incumbent groups (5-10 years earlier) separate from the entrant groups (max of three years earlier). 4.2% of entrants achieve this level of major entrant. We find a slightly lower estimate at 4.0% using Popp and Newell's (2012) definitions, and a rate of 4.1% when making the citation distributions specific to each patent class. Based upon these findings, we set $\alpha = 4\%$.

Finally, for L^s, which is the supply of scientists and engineers involved in R&D-type activities in the model (relative to production workers), we use 5.5%. We calculate this share from Census IPUMS using the 2000 5% sample. We keep non-group quartered workers who are aged 20-65 years old and working in industries closely related to the energy sector. We also require 20 weeks worked within the year and a usual hours worked of 20 or more during each week. 5.5% is the share of these workers with bachelor's educations and higher employed in occupations related to science and engineering.

3.4 Initial Technology Gaps

To provide the initial distributions of the model, we develop estimates of the cumulative stock of technologies invented by clean- and dirty-energy firms using three-digit SIC industries as approximations of product lines. We develop these distributions in three steps. The first step is to calculate the sum of patents made by each firm during the 1975-2004 period and the firm's distribution of employment across SIC3 industries in these sectors over the same period. We then apportion the firm's cumulative patent stock across SIC3 industries using the firm's employment distributions. For each SIC3 industry, we finally sum the apportioned patents made by clean- and dirty-energy firms. This sum of patents across all firms, active or inactive, reflects the quality ladders structure of our model.

These calculations provide us over 400 estimates of comparative clean- and dirty-energy stocks. Across these SIC3 industries, clean-energy firms are estimated to have a higher cumulative patent stock in 13.1% of industries. For data quality and Census Bureau disclosure restrictions, we focus on the upper half of the industry distribution in terms of cumulative clean and dirty patent counts, which has 13.0% of industries being led by the clean-energy stock; within manufacturing and energy production specifically, this share is 12.5%. The following table summarizes some details of these lines:

Table 2. Initial Condition Distributions SIC3

Metric:	**Clean Energy**	**Dirty Energy**
Mean Patent Total	260	1029
Standard Deviation	515	1500
Share: [0,20]	37%	0%
Share: [21,100]	25%	6%
Share: [101,500]	22%	50%
Share: [500+]	16%	44%

The average gap to the frontier for dirty-patents stocks in the 13% of cases where clean patents have the lead is 424 patents, or in relative terms, 39% of the total patenting in that line to date. The average gap to the frontier for clean-patent stocks in the 87% of cases where dirty patents have the lead is 947 patents and 76% in relative terms. To convert the empirical gap into the units of the model, we use the following reasoning. In our model, the annual patent flows of incumbents is 16.1% per product line (the sum of $x^c = 3.9\%$ and $x^d = 12.2\%$). In the data, the median annual flow of patents is approximately 17.1 per line. Hence we divide the empirical patent distribution of clean and dirty (which consists of patents registered between 1975-2004) by a conversion factor 17.1/0.161 and round it to the closest integer. This gives us the initial number of improvements $n^d_{j,0}$ and $n^c_{j,0}$. Then we compute the initial productivities as $q^d_{j,0} = \lambda^{n^d_{j,0}}$ and $q^c_{j,0} = \lambda^{n^c_{j,0}}$ to provide the initialization values. The following graph plots the density of this distribution with gaps between dirty and clean technologies on the horizontal axis:

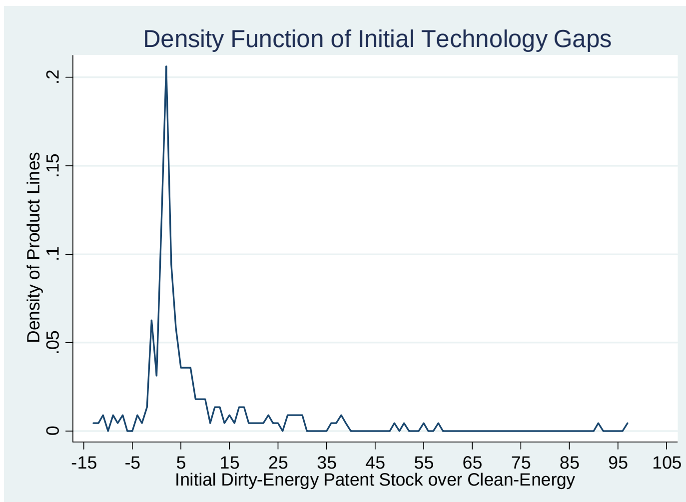

FIGURE 3

This graph shows that in most product lines the dirty technology is only a few steps ahead of clean technology, but there is a long tail of product lines with a large gap between dirty and clean, and a small set in which clean is ahead of dirty. The fraction of product lines with a non-zero gap in terms of step sizes is 90%. Clean energy leads by one or more step sizes in 9% of cases. Dirty energy has a lead of 20 and 50 steps sizes or more in 11% and 2% of technologies, respectively. We later consider an alternative initial distributions that modifies several of the modelling choices made here.

3.5 Simulated Method of Moments

The remaining parameters θ, λ, F_I and F_E are estimated using simulated method of moments (SMM). McFadden (1989), Pakes and Pollard (1989) and Gourieroux and Monfort (1996) characterize the statistical properties of the SMM estimator. This quantitative approach takes a set of key moments from our model, and then chooses the parameter vector so as to minimize the distance between these moments as implied by our model and as computed from the Census

Bureau data,

$$\min \sum_{i=1}^{4} |\text{model}\,(i) - \text{data}\,(i)|\,,$$

where we index each moment by i. SMM iteratively searches repeatedly across sets of parameter values for θ, λ, F_I and F_E in the model until the model's moments are as close as possible to the empirical moments (see Adda and Cooper, 2003, for further details). We also choose the heterogeneity parameter, ξ, as 10% and verify that our results are not sensitive to this parameter.

We use three moments from the microdata—firm entry rates, firm exit rates, and the average R&D/sales ratio of firms—together with the growth rate of the sector to identify these parameters. The entrant's labor share and exit rates are calculated across the five-year intervals of our Census Bureau data and then transformed into annualized net rates of 1.3% and 1.75%, respectively. We match the construction of these entry and exit rate moments in the model. The weighted average R&D-to-sales ratio is 6.57%, using log sales as weights and capping the R&D/sales ratio at 10x to reduce outliers (approximately the 99.8th percentile of the distribution). The aggregate annual sales growth per worker is 1.23% for the sector across the 1975-2004 period. After identifying these parameters in the estimation section, we investigate the fit of the model by comparing the implications of the model with a battery of other non-targeted moments from the microdata.

3.6 Computational Algorithm

Our theoretical analysis shows that microeconomic behavior is independent of climate dynamics. We therefore solve for value functions, innovation rates, and distributions first, then use those to find the time path for the carbon stock, temperature and other variables of interest.

The solution algorithm for this model involves finding the transition dynamics as the fixed point of a forward-backward iteration process, as in Conesa and Krueger (1998). See Zangwill and Garcia (1981) for further references. If the state space were of a more manageable size, one could simply solve the value function over this space and characterize the dynamics given arbitrary initial conditions. However, in this case the state space is the distribution of product lines over the technology gaps between the clean and dirty technologies. For any reasonable approximation, this results in a very high dimensional state space. Therefore, we solve each element of the model as a function of time given the initial conditions from the patenting data. The value function in early periods will thus depend on value functions in later periods. These later period value functions will in turn depend on the later period product distribution, which depends on early period value functions and innovation rates.

To solve for the fixed point of the sequence of value functions, we first discretize time into $N = 2048$ steps and set a terminal period $T = 2000$. Due to the symmetry between technology types inherent in this model, when a single type of technology is dominant—in the sense that the technology gap distribution is heavily skewed to either clean or dirty technology—one can analytically characterize value functions $v^j_{n,\infty}$ and innovation rates x^j_∞ and z^j_∞. We use these values as terminal conditions, though we do not know in advance which technology (clean or dirty) will be dominant. In addition, we set large upper (100) and lower (-100) bounds on the step gap distribution space. The algorithm proceeds as follows:

1. Posit an initial guess for the value function at time zero of the form

$$v^j_{n,t}(0) = \frac{\pi^j_n}{\rho + \bar{z}} \quad \forall t$$

   where $\bar{z}$ represents an estimated rate of creative destruction (we use $\bar{z} = 0.15$). Instantiate the technology gap distribution using the patent data with

$$\mu_{n,t}(0) = \mu_{n,0} \quad \forall t.$$

2. Iterate forward in time from $t = 0$ to $t = T$ by finding innovation rates x^j_t and z^j_t given value function and product distributions guesses at time $t+1$, $v^j_{n,t+1}(k)$ and $\mu_{n,t+1}(k)$. Using these innovation rates, update the time $t+1$ product distribution $\mu_{n,t+1}(k+1)$ using discrete time versions of the flow equations in (21).

3. Find the implied dominant technology at the terminal period by determining which technology type has a higher aggregate innovation rate as some late stage period $T - T_P$ (we use $T_P = 200$). Use the corresponding terminal value functions to update $v^j_{n,T}(k+1) = v^j_{n,\infty}$.

4. Iterate backward in time from $t = T$ to $t = 0$ by updating value function $v^j_{n,t-1}(k+1)$ using $v^j_{n,t}(k)$ and $\mu_{n,t}(k)$ according to a discretized version of the HJB equation in Lemma 1, re-solving for innovation rates x^j_t and z^j_t in the process.

5. Repeat steps 2-4 until the convergence criterion

$$\max_{n,t} \left| v^j_{n,t}(k+1) - v^j_{n,t}(k) \right| < \varepsilon$$

   is met. We use $\varepsilon = 10^{-6}$.

In order to avoid any instability, particularly when one is close a threshold where the asymptotically dominant technology switches over, we also introduce heterogeneity in incumbent fixed costs as explained above.[15]

[15] Our algorithm introduced a similar heterogeneity in entrant fixed costs, but at the end, entrants are never in the region where this heterogeneity matters.

Using up-to-date computer hardware, the equilibrium solver takes anywhere from five seconds to two minutes, depending on the speed of convergence. The code is written mostly in Python, with core routines written in C/C++.

Estimation To find the moments used in the SMM estimation, we simulate a panel of 16384 firms using equilibrium variables from the above model solving routine. Each firm has a portfolio of product lines with various technology gap values. We cap the maximum number of product lines a firm can have at 200. In order to determine the sales and R&D activity of the firm, we must keep track of both the number of product lines it is currently operating in, as well as the knowledge stock of the firm, which can in general differ. We simulate this panel of firms for 5 years to replicate the data generating process, and discretize time to have 100 subperiods per year, so that the simulations have 500 periods.

Optimal Policy We compute optimal policies for both the constant and time-varying cases. In the constant case, we use a straightforward grid search to find the optimum. In the time-varying case, we parameterize policies using a three stage carbon tax and a three stage research subsidy. Within each stage, whose boundaries are also parameterized and optimized over, we have constant policy levels. We then search over this space of functions using a combination of a simple simulated annealing algorithm (Kirkpatrick et al, 1983) and a Nelder-Mead (simplex) algorithm (Nelder and Mead, 1965).

4 Estimation Results

In this section, we provide the results of the simulated method of moments estimation and discuss the fit of our model to non-targeted moments. Finally, we show how atmospheric carbon concentration, temperatures and aggregate output evolve given these parameters in a laissez-faire equilibrium (with no policy intervention) starting from the observed distribution of technology gaps.

4.1 Parameter Estimates

Table 3 summarizes our parameter estimates.

Table 3. Parameter Estimates

Parameter	Description	Value
θ	Innovation productivity	0.500
λ	Innovation step size	1.075
F_I	Fixed cost of incumbent R&D	0.002
F_E	Fixed cost of entry	0.035

Our innovation productivity estimate implies that one unit of labor with a single product line generates an innovation with probability of about 8% a year. We estimate the innovation step size as 1.075 which implies a gross profit margin 7%. Finally our model predicts a sizable fixed cost advantage for incumbent firms. Their fixed cost of operation is equivalent to 6% of the entrants' fixed cost.

4.2 Goodness of Fit

Here we describe the goodness of fit of our model, first focusing on the four targeted moments, and then a range of diverse non-targeted moments.

4.2.1 Targeted Moments

Table 4 shows the values of the four moments used for estimation in the data and the model implications.

Table 4. Moment Matching

Moments	**Model**	**Data**
Entry Share	0.013	0.013
Exit Rate	0.018	0.018
Average R&D/Sales	0.066	0.066
Aggregate Sales per Employee Growth	0.007	0.012

On the whole, there is a very good match between the model and the data. The entry share, exit rate and R&D intensity are identical between the data and the model. Moreover the aggregate sales per worker growth is also very close.

4.2.2 Non-targeted Moments

Our main method of evaluating the quantitative fit of our model is to look at a range of non-targeted moments, which are presented (together with the model implications) in Tables 5A-5C.

We choose the non-targeted moments to represent aspects of the firm size distribution and its growth properties, which are quite different from the moments we targeted in our estimation. Our first non-targeted moment compares the size ratio of the median entrant to the median incumbent firm. Our targeted moments on entry/exit rates, overall sector growth, and R&D intensity do not directly impose strong constraints on this size distribution. Table 5A contrasts the size ratios in the model and data with respect to employment, sales, and sales per employee, and shows that our model implications match the data very closely with respect to the latter two metrics, though not as well for employment.[16]

Table 5A. Entrant Size Ratio to Incumbents

	Ratio of Median Sizes	
Size Measure:	**Model**	**Data**
Employment	0.17	0.03
Sales	0.18	0.20
Sales per Employee	1.12	1.05

Notes: Table compares non-targeted moments in model and data.

Our next point of comparison is for the structure of the growth distribution. We first calculate the unconditional growth rate of employment for each firm in the model and data defined as $(Emp_t - Emp_{t-1})/((Emp_t + Emp_{t-1})/2)$. This formula divides the employment change across the period by the average of the start and end values. As argued by Davis, Haltiwanger and Schuh (1996), this approach has attractive properties like a symmetric treatment of positive and negative growth and bounded values that minimize outliers. We calculate growth over five-year intervals. We then calculate the probability that firms experience substantial movements in either positive or negative directions. Comparing these movements in the model to the data is meaningful as it provides insights into how well the innovation step sizes and associated firm dynamics mirror the sector's true performance. Table 5B summarizes this comparison:

Table 5B. Comparison of Growth Distribution

	Employment Growth Probability	
Change over 5-Years:	**Model**	**Data**
Decrease 75% or more	0.17	0.11
Decrease 50% or more	0.20	0.15
Decrease 25% or more	0.27	0.26
Increase 25% or more	0.24	0.31
Increase 50% or more	0.17	0.20
Increase 75% or more	0.15	0.14
Increase 100% or more	0.08	0.11

Notes: Table compares non-targeted moments in model and data.

On this dimension, the model matches the data quite well. Compared to the data, the model somewhat over-predicts major downward employment declines of 50% or more. Recall that we matched the exit rate itself almost exactly, so this indicates an over-prediction of large employment declines conditional on survival. On the other hand, the model matches the data well for predicting employment decreases or increases of 25% or more. Likewise, the model and data are in very close agreement on the relative probabilities and sizes of large employment increases for firms.

[16] To pass Census Bureau disclosure restrictions, the empirical medians are "fuzzy" median estimates that use the average values over the 45th to 55th percentiles.

Our final comparison is the variation in conditional growth rates for employment across the firm size distribution, again over a five-year period. We divide firms into quantiles based upon their initial size in each period. We then compute the growth rates using the above formula, and the following table provides the comparison:

Table 5C. Comparison of Growth over Size Distribution

	5-Year Conditional Growth Rates	
Quantile of Sizes:	**Model**	**Data**
Smallest	18%	31%
2nd	25%	14%
3rd	18%	11%
4th	-5%	-1%
Largest	-0%	-10%

Notes: Table compares non-targeted moments in model and data.

The comparison is again reasonable. We match the general feature in the data that conditional growth rates are highest for small firms. The model's employment distribution is a little less fine-grained than the data, as about 50% of our firms have one product and employment is partially proportional to product counts. In consequence, there is limited variation across the smaller quantiles in the model compared to the more regular distribution in the data. The model and data then match quite well in identifying lack of growth for the top two quantiles compared to the bottom three (though the model does not predict employment declines in the largest firms that are present in the data).

Overall, our reading of the evidence is that for this range of diverse moments, which we did not target in our estimation, the model performs reasonably well, and this bolsters our confidence that our quantitative model is able to capture production and innovation dynamics in the energy sector.

4.3 Climate Dynamics in the Laissez-faire Economy

We next describe the implied future equilibrium and atmospheric carbon paths of the model given our estimates with the case of no carbon taxes and research subsidies. Given the initial distribution of technology gaps, dirty innovation is more productive and with no policy intervention, most R&D is initially targeted to the dirty technology as shown in Figure 4. Moreover, at these innovation rates, technology gaps and the profitability of dirty technologies increases relative to those of clean technologies, and clean R&D converges to zero. Consequently, in the long-run clean technologies disappear completely and dirty technologies take over the whole economy.

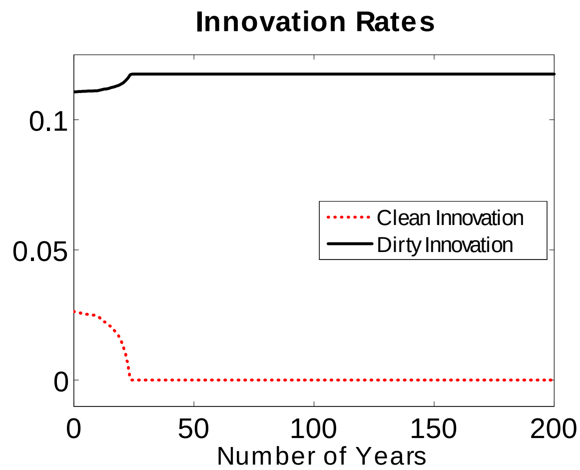

FIGURE 4: INNOVATION RATES, x_t^c AND x_t^d

The obvious implication of this time path of innovations is a steady increase in dirty energy production and carbon emissions. There are two ways of ascertaining the implications of these growing carbon emissions from our economy estimated and calibrated to US data. The first is to ignore emission growth from the rest of the world (i.e., keeping their emissions at a constant level). This is done in Figure 5A, which shows an increase in temperature of an additional 2.5°C in the next 200 years.[17] The alternative is shown in Figure 5B and assumes that emissions from the rest of the world will grow at the same rate as the US.

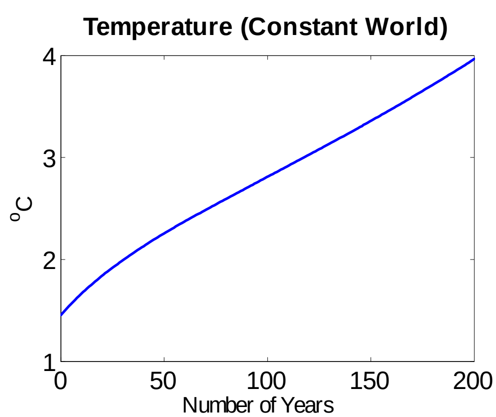

FIGURE 5A

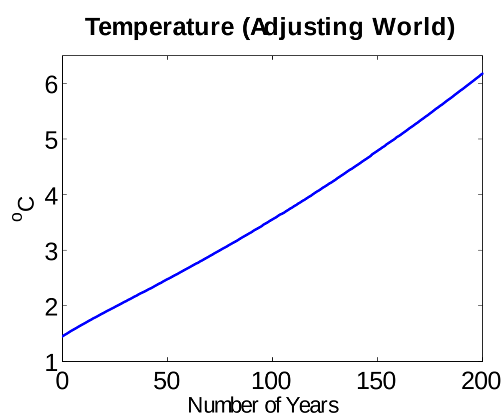

FIGURE 5B

The impact on global temperature is considerably exacerbated in view of the fact that we are now showing the increase in global temperature resulting from growth in global emissions, not

[17] We use the following formula to compute the temperature changes:

$$\Delta temperature = \frac{\lambda \left(\ln S_t - \ln \bar{S}\right)}{\ln 2}.$$

just US emissions. It is important to recall that, as noted above, given the functional form in (2), the exact path of emissions from the rest of the world has no impact on optimal policy (since the marginal damage created by US emissions are independent of the level of global emissions). This is the reason why we did not have to take a position on the time path of emissions in the rest of the world until this point (and now we are only taking a position in order to translate its implied path of emissions into changes in global temperatures).

5 Policy Analysis

In this section, we derive the policy sequence that maximizes discounted welfare. Throughout, we do not allow the social planner to correct for monopoly distortions, thus limiting ourselves to the policy instruments discussed above—a carbon tax and subsidy to clean research.[18] In addition, our theoretical analysis makes it clear that what is relevant is the differential tax and subsidy rates for clean vs. dirty energy, thus we just look at taxes on dirty production, which we refer to as "carbon taxes," and subsidies to clean innovation. Finally, we restrict the subsidies to entrants and incumbents to be the same, i.e., $s_{Et} = s_{It}$ for all t (based on early results which suggest that when both instruments are allowed to vary they are often equal or very close to each other). Throughout, we consider a private discount rate of $\rho = 1\%$ and present results for two different social discount rates: $\rho_{sp} = 1\%$ and $\rho_{sp} = 0.1\%$.[19] In both cases, paths that involve no switch to clean technology will lead to unbounded atmospheric carbon and temperature increases and consumption limiting to zero because of the economic distortions created by the unbounded increase in atmospheric carbon (see equation (2)), and we assign minus infinite social welfare to such paths, so that when feasible, a switch to clean technology is preferred.[20] Finally, we first focus on optimal constant policies (where carbon taxes and research subsidies are constant over time), which have several advantages: they are simpler and more transparent and the optimal time-varying policies, which we also characterize below, are time-inconsistent, raising some caveats about interpretation.

[18] As mentioned above, in the one-sector version of our model (either with only dirty or only clean technology), taxes or subsidies to research would only affect relative wages of skilled workers (employed in the research sector), and crucially not the aggregate rate of innovation. For this reason, subsidies to clean research or taxes on dirty research are identical in our model.

[19] The reasoning here is that, following Stern (2007), the social planner—society—may have a lower discount rate than that implied by market interest rates. Thus 0.1% is what should be applied to the welfare analysis of the social planner, while still keeping the discount rate that firms use in their decisions at 1%.

[20] This is implied whenever the growth rate g is greater than ρ_{sp}, which is always satisfied when $\rho_{sp} = 0.1\%$, but may or may not be satisfied when $\rho_{sp} = 1\%$ depending on the growth rate.

5.1 Optimal Constant Policies

Table 6 shows optimal constant policies for the two values of social discounting, $\rho_{sp} = 1\%$ (high), $\rho_{sp} = 0.1\%$ (low).

TABLE 6. OPTIMAL CONSTANT POLICY

	$\rho_{sp} = 1\%$	$\rho_{sp} = 0.1\%$
τ	16%	44%
s	61%	95%

In both cases, there is a very aggressive research subsidy for clean technology. With $\rho_{sp} = 1\%$, the carbon tax is fairly low, at 16%, while research directed at clean technologies receives a 61% government subsidy (meaning that for every dollar of R&D spending, there is a 61 cents subsidy). With the social discount rate of $\rho_{sp} = 0.1\%$, carbon taxes are raised to 44%, but now clean research subsidies are even more aggressive, at 95%.

The intuition for why optimal policy relies so much on subsidies to clean research is instructive. The social planner would like to induce a switch from R&D directed at carbon intensive dirty technologies towards clean technologies. She can do so by choosing a sufficiently high carbon tax rate today and in the future, because this would reduce the profitability of production using dirty technologies and secure both a switch to clean production and, on the basis of this, to research directed at clean technologies. However, this is socially costly because given the current state of technology, switching most production to clean technology has a high consumption cost (because the marginal costs of production of clean technologies are initially significantly higher than those of dirty technologies). Hence it is a better strategy for the social planner to choose the carbon tax to only deal with the carbon emission externality and rely on the research subsidy to induce the switch to clean technologies in the long run. Figure 6 in fact shows that the social planner is able to do this, particularly with $\rho_{sp} = 0.1\%$, where the optimal policy involves a very rapid ramp up of clean innovation rates and the disappearance of all research directed to dirty technologies in about 130 years. Interestingly, however, with $\rho_{sp} = 1\%$, the social planner chooses not to completely replace dirty research with clean research in the first several hundred years. Instead, she subsidizes clean research just enough to make sure that clean research and thus clean technologies also survive for a long time, but not so much that they overtake the dirty sector completely. As a result, dirty innovation survives for several hundreds years (clean innovation exceeds dirty innovation only around year 500), but throughout, innovation rates in the clean technology are significantly higher than in the laissez-faire equilibrium shown in Figure 4 where they converged to zero in about 25 years.

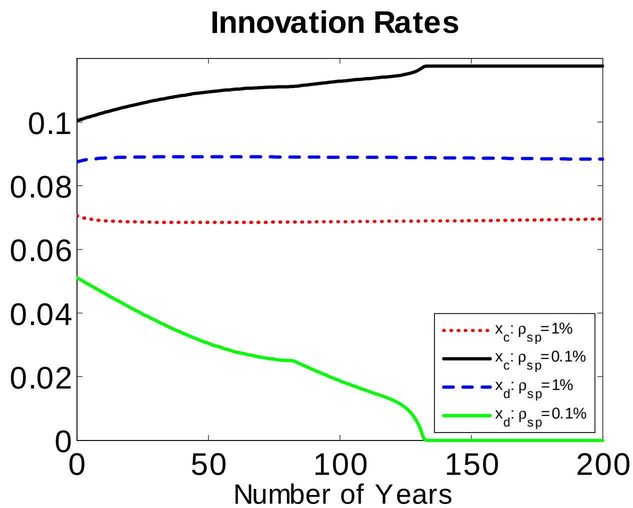

FIGURE 6: INNOVATION RATES UNDER OPTIMAL POLICIES

Figure 7 depicts the implied path of temperature increases under the optimal policies, in the same two ways as we have done in Figure 5—assuming either that emissions from the rest of the world are constant or that they grow at the same rate as US emissions. In both cases, global temperature increases less, and in fact significantly less with $\rho_{sp} = 0.1\%$, than in Figure 5.

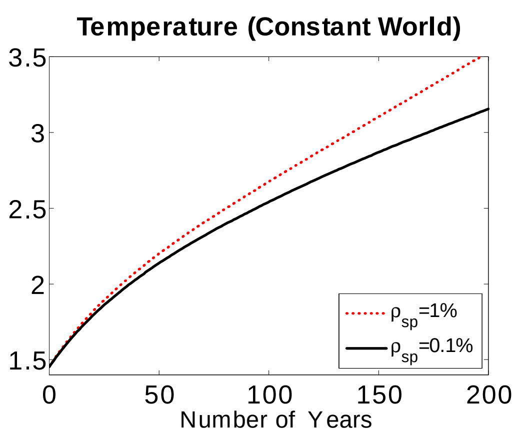

FIGURE 7A

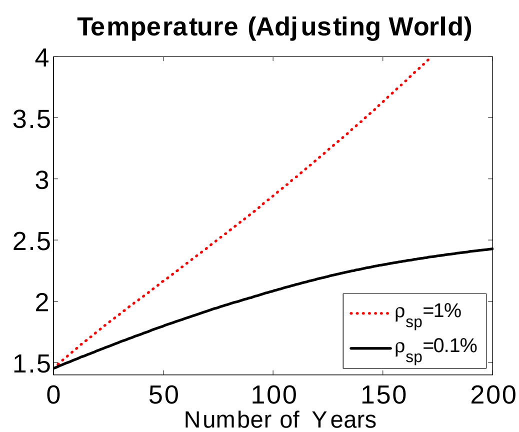

FIGURE 7B

5.2 Optimal Time-Varying Policies

We now return to optimal time-varying policies and characterize the welfare gains from using time-varying rather than constant policies. For computational reasons, we look for policies that take a simple "step function form" with three endogenously determined switch points.[21]

The resulting optimal policies are shown in Figure 8. A couple of features are worth noting. First, the subsidy rate for clean research is very similar to the constant policies in both cases. With a social discount rate of $\rho_{sp} = 0.1\%$, this subsidy rate is roughly constant (starting at 95% and declining to 93% at year 54).[22] Moreover, in this case, the carbon tax starts somewhat higher than with the optimal constant policy but then declines from 54% to 34% at year 92). With $\rho_{sp} = 1\%$, the pattern is different: the subsidy rate starts and remains at 25%, again very close to its constant optimal policy value, but the carbon tax starts at zero and stay there for quite a while, and then increases dramatically to 650% in year 328.

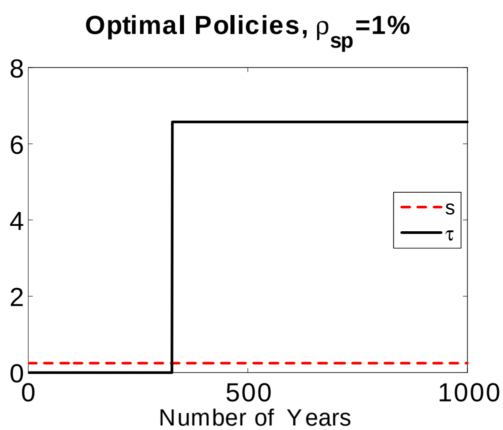

FIGURE 8A

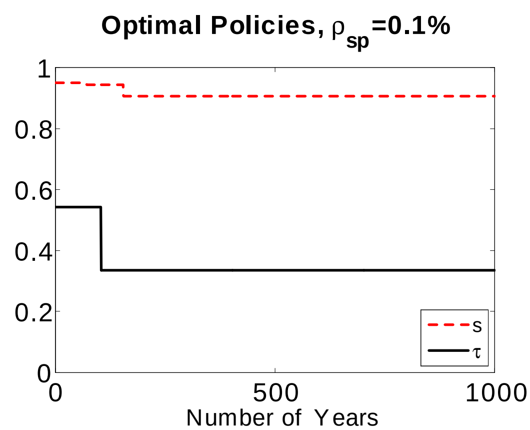

FIGURE 8B

The patterns shown in Figure 8 result from the interplay of two counteracting forces. First, all else equal, the social planner would like to delay as much as possible the consumption loss from switching to clean technologies.[23] Second, carbon taxes early on are more effective in both switching production and reducing emissions (given the long half life of carbon in the atmosphere imposed in our model of the carbon cycle). With $\rho_{sp} = 1\%$, the first effect is dominant because with this relatively high social discount rate, high consumption during the early years is highly valued, encouraging the planner to delay the start of high carbon taxes for quite a while. In consequence, in this case carbon taxes are sharply backloaded, and in fact, as in the constant policy case, dirty innovation disappears only very slowly—over several hundreds of years. With $\rho_{sp} = 0.1\%$, the future is less heavily discounted, strengthening the

[21] For our baseline results, we experimented with increasing the number of switch points and with alternative formulations of time variation, with broadly similar results.

[22] Note that research subsidies in the far future may not matter very much because early research subsidies may have already induced a large reallocation of research from dirty to clean. Nevertheless, research subsidies in the future are not undetermined because there is always some positive fraction of firms with a dirty portfolio of product lines which will then have incentives to undertake research in the dirty technology (this fraction declines to zero asymptotically in the optimal allocation with $\rho_{sp} = 0.1\%$).

[23] This effect in part reflects the fact that the social planner is committing to a policy path and is not "time consistent".

second effect and making carbon taxes frontloaded and the complete switch to clean innovation much more rapid. In both cases, however, the average values of the carbon tax in the first 200 years is in the ballpark of the constant optimal policies (16% with $\rho_{sp} = 1\%$ and 44% with $\rho_{sp} = 0.1\%$).

Table 7 shows that the welfare loss from using constant policies is quite small, 0.3% with $\rho_{sp} = 0.1\%$, but sizable, about 16%, with $\rho_{sp} = 1\%$, which reflects the benefits for social planner's utility resulting from high consumption growth at the expense of high emissions in the first 300 years. This pattern suggests that the results with the low social discount rate, $\rho_{sp} = 0.1\%$, are more plausible in a range of dimensions.[24]

Table 7. Welfare Costs

$\rho_{sp} = 1\%$	$\rho_{sp} = 0.1\%$
16%	0.3%

5.3 Counterfactual Policy Analysis

Our model enables us to investigate the welfare and climatic implications of a range of counterfactual policies. Here we focus on two counterfactuals. The first is relying just on a carbon tax (i.e., no research subsidy) as the policy tool, and the second is delaying intervention for 50 years and then choosing the optimal policy from that point onwards. We focus on time-varying optimal policies, which are shown for these two counterfactual policies in Figure 9.

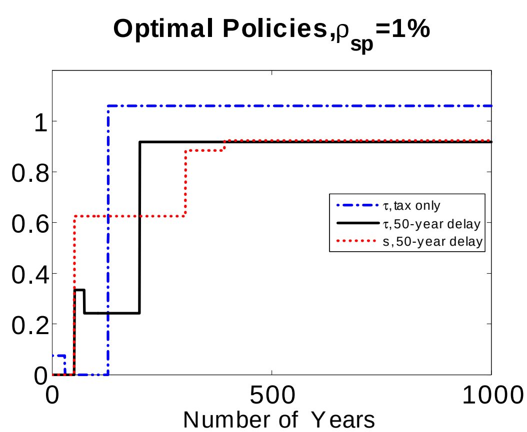

Figure 9a

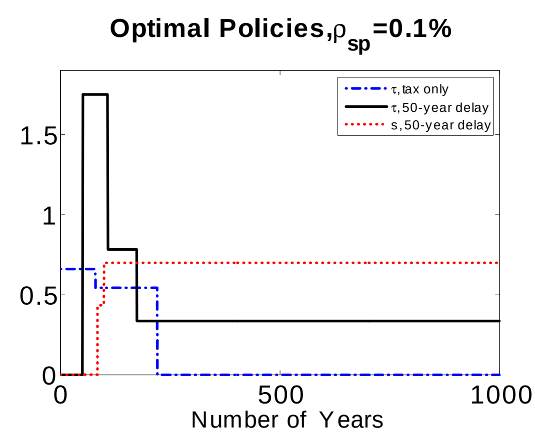

Figure 9b

Optimal policies following the 50-year delay are more aggressive than the baseline optimal policies, and when the only policy tool is the carbon tax, this tax also is typically higher.[25]

[24] We have also verified that our main results with an intermediate social discount rate of $\rho_{sp} = 0.5\%$ are very similar to those with $\rho_{sp} = 0.1\%$, again making us trust these results more than the ones based on $\rho_{sp} = 1\%$.

[25] When just relying on the carbon tax and with $\rho_{sp} = 0.1\%$, the carbon tax reaches zero earlier than in the baseline shown in Figure 8. Nevertheless, it is effectively more aggressive than the baseline since it starts at a

For example, in the carbon tax only counterfactual, this tax is higher because it is also being used to redirect innovation towards clean technologies. As a result, with $\rho_{sp} = 1\%$, aggregate temperature increases less at long horizons under this constrained optimal policy than our actual optimal policy, but this is at the expense of slower output growth, especially early on. As a result, the cost of just relying on carbon tax for optimal policy, shown in Table 8, is 4.2% with $\rho_{sp} = 1\%$ and 3.4% with $\rho_{sp} = 0.1\%$. Delaying the start of optimal policies by 50 years leads to greater losses—a consumption equivalent welfare cost of 8% with $\rho_{sp} = 1\%$, and 16.6% with $\rho_{sp} = 0.1\%$. These numbers indicate that delaying policy interventions to combat carbon emissions could have very significant welfare costs, especially when the social discount rate is low. Moreover, just relying on carbon taxes—eschewing research subsidies—could also have sizable welfare costs.

Table 8. Welfare Costs

Carbon Tax Only		50-year Delay	
$\rho_{sp} = 1\%$	$\rho_{sp} = 0.1\%$	$\rho_{sp} = 1\%$	$\rho_{sp} = 0.1\%$
4.2%	3.4%	8.0%	16.6 %

Finally, we also evaluate what the climatic and welfare implications of maintaining current US policies (here interpreted for the whole world) would be relative to adopting an optimal policy moving forward. For this purpose, we have tried to estimate the carbon taxes implied by US policies and the current subsidies to clean innovation (relative to dirty R&D) in our sample of firms. There is much uncertainty about what the carbon tax in the United States will be moving forward. A cap-and-trade program is likely to be implemented, but it is unclear what the implied carbon tax rate will be. On the other hand, Greenstone et al. (2011) estimate a social cost of carbon equal to about \$21 in 2010, expressed in 2007 dollars, and this number is currently being used for cost-benefit analysis by US agencies. This social cost estimate is the central tendency across a number of models and scenarios considered. The social cost increases in real 2007 terms to \$45 in 2050 as a consequence of future marginal emissions becoming ever more harmful. We therefore use two values for the "business-as-usual" carbon tax, 0% consistent with the current situation, and 24% (approximately implied by \$45 social cost of carbon in 2050, a relatively early point in the transition path).[26] We estimate the current clean research subsidy from our sample as follows: over our full 30 year period, 49% of all R&D expenditures by our clean firms are federally funded, while the same number is 11%

higher level (66% instead of 54%) and remains at a higher level (54% instead of 34%) for the first 220 years, and this induces both a much more rapid switch to clean production and also encourages a switch to clean innovation despite the absence of research subsidies in this case.

[26] In particular, US carbon emissions are 1.58 billion tons in 2002. One metric ton of carbon is equivalent to 3.667 units of carbon dioxide. Our dirty firms have sales of approximately one trillion dollars in this year. The \$45 social cost is \$39 in 2002 terms. These numbers imply a real tax rate in 2050 of about 23% ($(39 \times 3.667 \times 1.58 \times 10^9)/10^{12} \simeq 0.23$). We approximate this with an 24% tax rate (since our taxes have to be multiples of λ). This carbon tax rate is much less than currently used in Sweden (see Golosov et al. 2011) and also less than the numbers suggest that by Nordhaus (2008).

for our dirty firms. This implies a 43% ($(1 - 0.49) / (1 - 0.11) \simeq 1 - 0.43$) subsidy for clean R&D relative to dirty R&D.

The scenario with a zero carbon tax, regardless of the discount rate, involves 100% welfare costs because, in this case, temperature increases rapidly and continues to grow unboundedly. Essentially, 43% R&D subsidy for clean is insufficient to redirect technological change towards clean with no carbon tax. The resulting significant damage to the environment leads to a disastrous welfare result. Interestingly, however, even with this less than optimal subsidy to clean research, it turns out that the temperature increase can be contained if there is a moderate carbon tax at 24%. As a result, with this moderate carbon tax, the welfare costs are still sizable but limited as shown in Table 9.

Table 9. Welfare Costs

$\tau = 24\%$, $s = 43\%$		$\tau = 0$, $s = 43\%$	
$\rho_{sp} = 1\%$	$\rho_{sp} = 0.1\%$	$\rho_{sp} = 1\%$	$\rho_{sp} = 0.1\%$
18%	8%	100%	100%

6 Robustness and Extensions

In this section, we investigate how our estimation, optimal policy and counterfactual results are affected by a range of different modeling assumptions or variations on parameter estimates. Throughout, to economize on space we only report the implied optimal policies (even when the variation in question involves reestimating the parameters of the underlying model).

6.1 Alternative Damage Elasticity γ

As noted above, actual damages from atmospheric carbon may be greater than the estimates commonly used in the economics literature. We now show the sensitivity of our results to higher values of these damages, captured by the parameter γ, in our model. Table 10 depicts constant optimal policies for two cases, when γ is twice and 10 times as large as our baseline value, $\gamma = 5.3 \times 10^{-5}$, and Figure 10 shows optimal time-varying policies for the same two cases.

Table 10. Optimal Constant Policies

	$\gamma = 2\times$			$\gamma = 10\times$	
	$\rho_{sp} = 1\%$	$\rho_{sp} = 0.1\%$		$\rho_{sp} = 1\%$	$\rho_{sp} = 0.1\%$
τ	16%	44%	τ	24%	54%
s	61%	95%	s	95%	95%

Overall, the results are remarkably similar to those in our baseline. Interestingly, with $\rho_{sp} = 1\%$, optimal constant policies are identical when γ is twice as large as the baseline. This result, which at first appears counterintuitive, is because the optimal policy in this case does not eliminate but chooses to contain carbon emissions (and does not even eliminate the dirty sector). When γ is doubled, the social planner still prefers to maintain this containment strategy, making optimal policies very similar to the baseline. When γ is taken to be much larger (10 times as large as the baseline), this is no longer optimal, and we now see a more aggressive carbon tax and a much more aggressive research subsidy, utilized for eliminating the dirty sector.

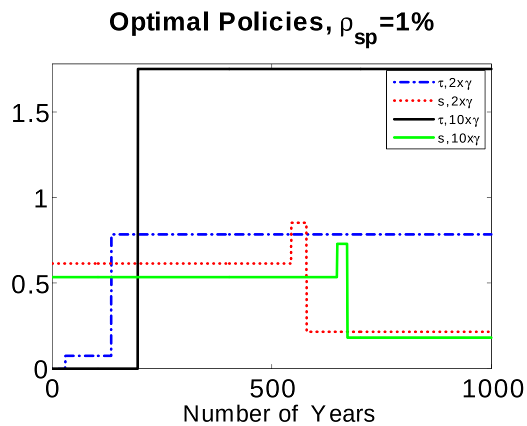

FIGURE 10A

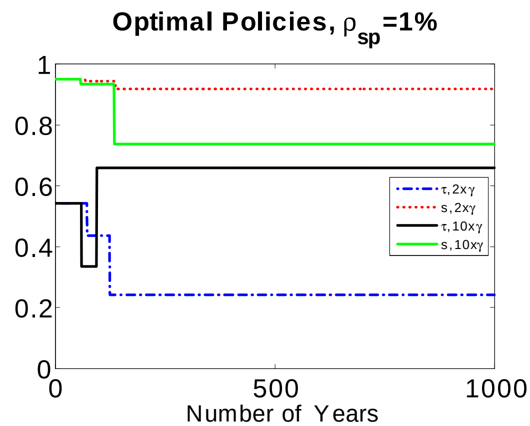

FIGURE 10B

Optimal time-varying policies, which are shown in Figure 10, are also quite similar—but not identical—to the baseline.

Overall, with the exception of the last case mentioned, the results suggest that the qualitative, and even quantitative, messages from our analysis are fairly robust to different economic damages from atmospheric concentration.

6.2 Costly Research Subsidy

We next investigate the robustness of our results to assuming that R&D subsidies create direct distortions. In particular, we assume that for every dollar of subsidy, $1+\chi$ dollars need to be spent, so that χ is a waste, which we subtract from consumption. We consider two values of χ, 50% and 100%, both of which are very aggressive choices on the distortion or implications of research subsidies. The results for constant policies are shown in Table 11.

TABLE 11. OPTIMAL CONSTANT POLICIES

50% Consumption Cost			100% Consumption Cost		
	$\rho_{sp}=1\%$	$\rho_{sp}=0.1\%$		$\rho_{sp}=1\%$	$\rho_{sp}=0.1\%$
τ	16%	54%	τ	16%	66%
s	61%	53%	s	61%	0%

We find that except in one case, optimal policy still makes use of aggressive research subsidies despite the significant waste that these create. The reason for this is that, as implied by our discussion above, the carbon tax is a poor substitute for research subsidies; it also encourages clean research but does so at the cost of creating more intra-temporal distortions. In consequence, there is ample room for research subsidies even when they are distortionary. In fact, Table 11 shows that with a discount rate of $\rho_{sp} = 1\%$, optimal constant policies are identical to the case without any distortions. Intuitively, the social planner finds it optimal to leave the carbon tax unchanged (recall that the carbon tax can only change in steps), and with unchanged carbon tax, the research subsidy also remains constant. With the lower social discount rate, $\rho_{sp} = 0.1\%$, the carbon tax becomes more aggressive; in fact, with 100% distortions from research subsidies and this lower social discount rate, the optimal constant policy increases the carbon tax significantly and ceases to use research subsidies. However, Figure 11 shows that optimal time-varying policies in this case still involve heavy use of positive research subsidies. Moreover, the qualitative pattern of optimal time-varying policies is quite similar to the baseline, shown in Figure 8, and again involve backloading of carbon taxes for $\rho_{sp} = 1\%$, frontloading of carbon taxes for $\rho_{sp} = 0.1\%$, and fairly aggressive use of research subsidies, especially in the first few hundred years. The fact that research subsidies are now phased out entirely with $\rho_{sp} = 0.1\%$ is also very intuitive: research subsidies early on are sufficient to switch most innovation to clean, and start influencing only a few firms after a while (as most leading-edge technologies are now clean); because they are costly and become largely unnecessary, it is natural for the social planner to rely less on them. With $\rho_{sp} = 1\%$, this does not happen because the transition to clean technology is slower and research subsidies are still useful for several hundreds of years.

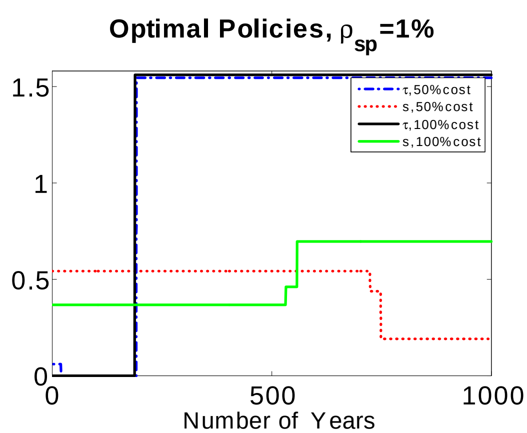

Figure 11a

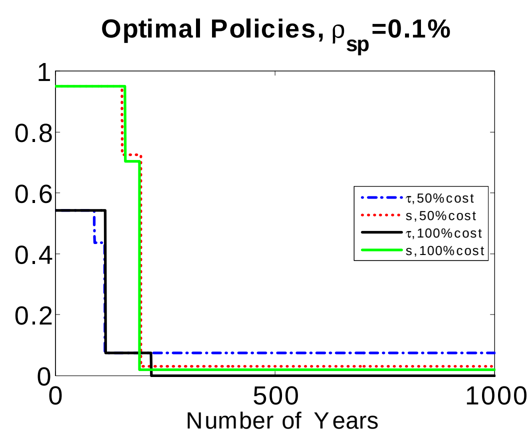

Figure 11b

Overall, we conclude that with reasonable values of distortions, and even with certain extreme values of distortions, the broad pattern of optimal policies is quite similar to the baseline case, and research subsidies are still an essential part of the portfolio of optimal policies, even if they may be significantly distortionary.

6.3 Alternative R&D Elasticities η

Our baseline results are for $\eta = 0.45$ which averaged across cross-sectional and first-difference estimates. We now reestimate the model using first a value of η in the ballpark of the cross-sectional estimates ($\eta = 0.65$) and then for a value corresponding to the first-difference estimates ($\eta = 0.35$), and investigate the implications of this for the fit of the model and for optimal policy. Overall, the fit of the model is not affected much by the change in η, and the implications for optimal constant policies are shown in Table 12 and optimal time-varying policies are shown in Figure 12.

Table 12. Optimal Constant Policy Rates

	$\eta = 0.35$			$\eta = 0.65$	
	$\rho_{sp} = 1\%$	$\rho_{sp} = 0.1\%$		$\rho_{sp} = 1\%$	$\rho_{sp} = 0.1\%$
τ	16%	34%	τ	24%	66%
s	0%	95%	s	84%	84%

When the elasticity of innovation to R&D effort is higher than in our baseline, at $\eta = 0.65$, the results are also remarkably similar to the baseline both with constant and time-varying policies. With the lower elasticity, $\eta = 0.35$ and $\rho_{sp} = 0.1\%$, they are also fairly similar. However, with $\eta = 0.35$ and $\rho_{sp} = 1\%$, the optimal constant policy is quite different. To understand the reason why, recall that in our baseline with $\rho_{sp} = 1\%$, the optimal policy involves positive research effort directed both towards dirty and clean technologies. When the elasticity of innovation to R&D declines, the social planner, restricted to a constant policy and with a reasonably high discount rate, finds it optimal to induce a slow switch to clean technology, which can be achieved with just a carbon tax. This is partly an artifact of constant policies; Figure 12 shows that optimal time-varying policies still heavily rely on research subsidies in this case. We thus conclude that the main message from our baseline results continue to apply with a fairly wideband of elasticities.

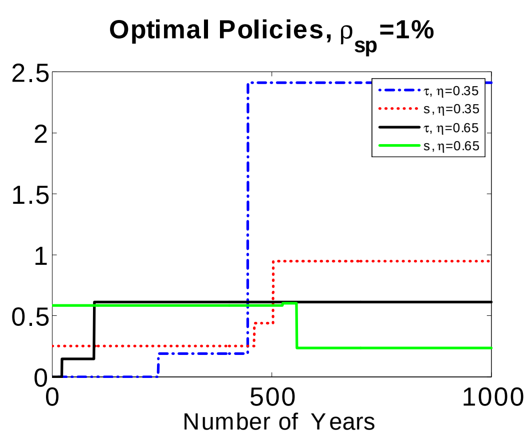

Figure 12a

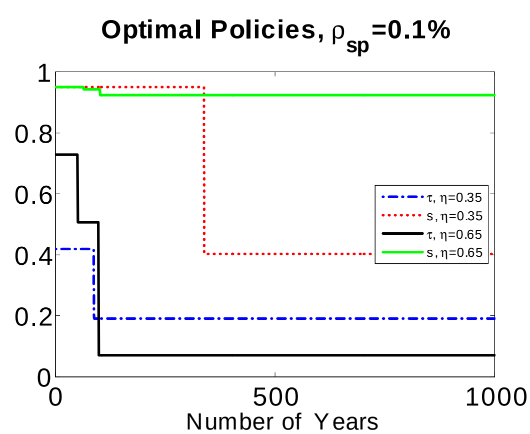

Figure 12b

6.4 Alternative Leapfrogging Probabilities α

We also investigate the implications of different values of α, in particular, focusing on a lower and higher estimate of α ($\alpha = 0.03$ and $\alpha = 0.05$). The results reported in Table 13 and Figure 13 are quite similar to the baseline results both quantitatively and qualitatively. In particular, optimal constant policies are in the ballpark of our baseline with $\alpha = 0.04$, and optimal time-varying policies have the same backloading and frontloading properties and similar values, though the exact switch points do differ.

Table 13. Optimal Constant Policy Rates

	$\alpha = 0.03$			$\alpha = 0.05$	
	$\rho_{sp} = 1\%$	$\rho_{sp} = 0.1\%$		$\rho_{sp} = 1\%$	$\rho_{sp} = 0.1\%$
τ	24%	54%	τ	16%	44%
s	95%	95%	s	32%	95%

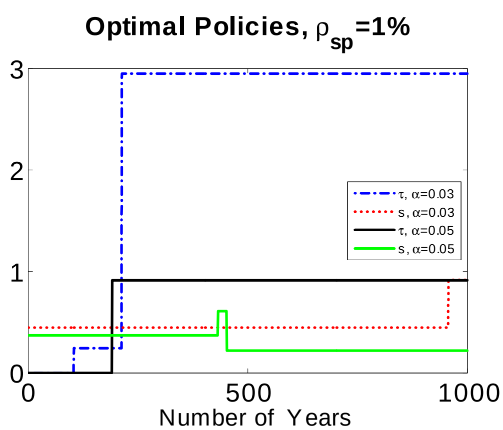

Figure 13a

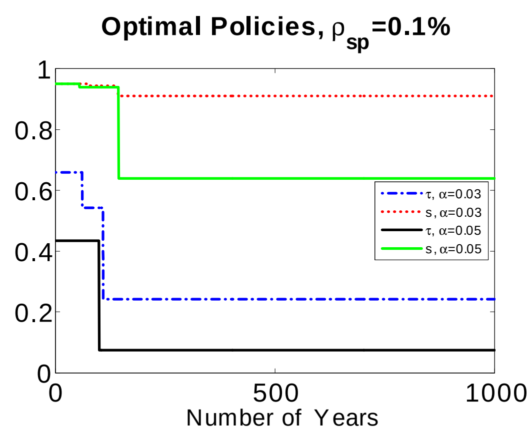

Figure 13b

6.5 Alternative Initial Technology Distribution

Finally, we also considered an initial technology gap distribution defined with several modifications from our baseline. First, rather than just sum patent counts, we weight patents by the normalized citation counts the patent receives. Second, we consider four-digit industries rather than three-digit industries. And third, we only consider industries within the manufacturing and energy sectors. Using these criteria, there are 332 SIC4 industries that are of sufficient size in terms of innovative firm counts to pass Census Bureau disclosure restrictions. Among these industries, 9.4% are led by the clean-energy stock. Table 14 summarizes some moments of this distribution:

TABLE 14. INITIAL CONDITION DISTRIBUTIONS SIC4

Metric:	**Clean Energy**	**Dirty Energy**
Mean Patent Total	140	663
Standard Deviation	401	1242
Share: [0,20]	53%	2%
Share: [21,100]	23%	18%
Share: [101,500]	17%	48%
Share: [500+]	6%	33%

The average gap to the frontier for dirty-patents stocks in the 9% of cases where clean patents have the lead is 463 patents, or in relative terms, 33% of the total patenting in that line to date. The average gap to the frontier for clean-patent stocks in the 91% of cases where dirty patents have the lead is 624 patents and 82% in relative terms. The conversion factor in this case is 12.6/0.161. The distribution graph has a broadly similar shape as Figure 3 and we omit it to save space. The fraction of product lines with a non-zero gap in terms of step sizes is 82%. Clean energy leads by one or more step sizes in 7% of cases. Dirty energy has a lead of 20 and 50 steps sizes or more in 8% and 2% of technologies, respectively.

TABLE 15. OPTIMAL CONSTANT POLICY RATES

	$\rho_{sp} = 1\%$	$\rho_{sp} = 0.1\%$
τ	16%	54%
s	74%	95%

Using this alternative distribution of initial technology gaps has fairly limited impact on optimal constant and time-varying policies, which are shown in Table 15 and Figure 14. Both optimal constant and time-varying policies are remarkably similar to the baseline, making us conclude that our qualitative and even quantitative results are fairly robust to reasonable variations in the initial distribution of technology gaps.

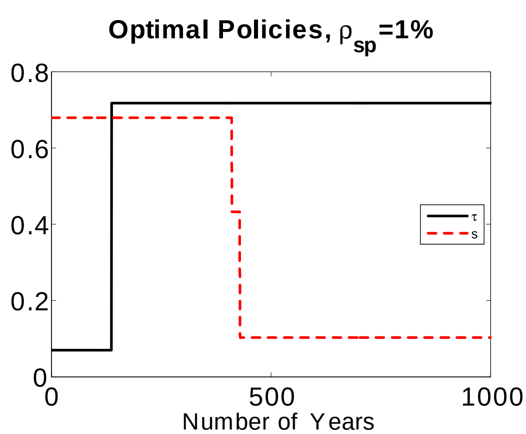

FIGURE 14A

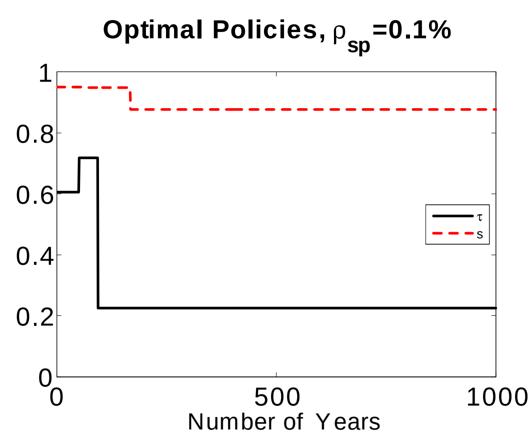

FIGURE 14B

7 Conclusion

One of the central challenges facing the world economy is reducing carbon emissions, which appears to be feasible only if a successful transition to clean technology can be induced. This paper has investigated the nature of a transition to clean technology theoretically and empirically. We developed a microeconomic model where clean and dirty technologies compete in production and innovation. If dirty technologies are more advanced to start with, the potential transition to clean technology can be difficult both because clean research must climb several steps to catch up with dirty technology and because this gap discourages research effort directed towards clean technologies. We characterized several properties of the equilibrium in this model and then estimated its key parameters from microdata on production, employment, R&D, patents and entry and exit of firms in the US energy sector, using regression analysis and simulated method of moments. Our model performs fairly well in matching a range of patterns in the data that were not directly targeted in the estimation, giving us confidence that it is potentially useful for the analysis of the transition to clean technology in the US energy sector.

Theoretically, carbon taxes and research subsidies encourage production and innovation in clean technologies. The key question we investigate using our estimated quantitative model is whether optimal policy will indeed secure a transition to clean technology, and if so how rapidly, and whether it will do so using carbon taxes or a combination of carbon taxes and research subsidies. A naïve intuition would be that only carbon taxes should be used because externalities are created by carbon (in the absence of these carbon externalities, the social planner would have no reason to interfere with or subsidize research).

In contrast to this intuition, we find that optimal policy heavily relies on research subsidies, and this result is fairly robust across a range of variations and for different damages and social discount rates. We also use the model to evaluate the welfare consequences of a range of alternative policy structures. For example, just relying on carbon taxes or delaying intervention both have significant welfare costs.

Though, to the best of our knowledge, it is the first attempt to develop a microeconomic model of the transition to clean technology and to quantitatively characterize optimal policy in such a setup, our paper has inevitably left several questions unanswered and taken a number of shortcuts, all of which constitute interesting areas for future research and investigation. We list some of these we view as particularly important here:

1. Our damage function enabled us to abstract from emissions in the rest of the world. Though very convenient, this approach left out several interesting considerations. The first is the interaction between US and global emissions.

2. The second is the potential impact of US transition to clean technology on technology

choices in the rest of the world. In particular, to the extent that there is such an impact, optimal policy may be more aggressive (for reasons discussed in Acemoglu et al., 2012a) because it might have the power to also induce a switch to clean technology in the rest of the world also.

3. These concerns naturally fit into another important topic: game-theoretic interactions in emissions and technology choice across several countries in the global economy (Harstad, 2012, Dutta and Radner, 2006).

4. For reasons we have explained, we did not allow for nonlinear threshold effects in the impact of atmospheric carbon on economic efficiency. Such nonlinearities are likely to be important and their exact position might be uncertain. Incorporating such nonlinearities, together with an explicit approach to uncertainty along the lines of Weitzman (2009), would be an important area for future research. This would also necessitate a much more detailed investigation of the interactions between US emissions and the rest of the world.

5. Our optimal policies are characterized under the assumptions of commitment to the policy sequence by the social planner. In the absence of such commitment, there will be a time inconsistency problem. An obvious important next step is to characterize time-consistent optimal policy.

6. Another interesting area is to investigate the interactions between international trade, technology and emissions (see Hemous, 2012).

References

Acemoglu, Daron. 1998. "Why do New Technologies Complement Skills? Directed Technical Change And Wage Inequality," *Quarterly Journal of Economics* 113: 1055-1090.

Acemoglu, Daron. 2002. "Directed Technical Change." *Review of Economic Studies* 69: 781-810.

Acemoglu, Daron and Ufuk Akcigit. 2012. "Intellectual Property Rights Policy, Competition and Innovation." *Journal of the European Economic Association* 10: 1-42.

Acemoglu, Daron, Philippe Aghion, Leonardo Bursztyn, and David Hemous. 2012a. "The Environment and Directed Technical Change." *American Economic Review* 102: 131-166.

Acemoglu, Daron, Ufuk Akcigit, Nick Bloom, and William Kerr. 2012b. "Innovation, Reallocation, and Growth." Working Paper.

Adda, Jerome, and Russell Cooper. 2003. *Dynamic Economics: Quantitative Methods and Applications*, MIT Press.

Aghion, Philippe, Antoin Dechezlepretre, David Hemous, Ralf Martin, and John Van Reenen. 2012. "Carbon Taxes, Path Dependency and Directed Technical Change: Evidence from the Auto Industry." Harvard University and LSE mimeo.

Aghion, Philippe, Christopher Harris, Peter Howitt, and John Vickers. 2001. "Competition, Imitation and Growth with Step-by-Step Innovation." *Review of Economic Studies* 68: 467-492.

Aghion, Philippe, and Peter Howitt. 1992. "A Model of Growth Through Creative Destruction." *Econometrica* 110: 323-351.

Akcigit, Ufuk, and William Kerr. 2010. "Growth through Heterogeneous Innovations." NBER Working Paper 16443.

Archer, David. 2005. "Fate of Fossil Fuel CO2 in Geologic Time." *Journal of Geophysical Research* 110: C09S05.

Balasubramanian, Natarajan, and Jagadeesh Sivadasan. 2011. "What Happens When Firms Patent? New Evidence from US Economic Census Data." *Review of Economics and Statistics* 93: 126-146.

Bovenberg, Lans, and Sjak Smulders. 1995. "Environmental Quality and Pollution-Augmenting Technological Change in a Two-Sector Endogenous Growth Model." *Journal of Public Economics* 57: 369-391.

Bovenberg, Lans, and Sjak Smulders. 1996. "Transitional Impacts of Environmental Policy in an Endogenous Growth Model." *International Economic Review* 37: 861-893.

Conesa, Juan Carlos and Dirk Krueger. 1998. "Social Security with Heterogeneous Agents", *Review of Economic Dynamics*, 2: 757-795.

Davis, Steven, John Haltiwanger, and Scott Schuh. 1996. *Job Creation and Destruction*, MIT Press.

Dutta, Prajit K. and Radner, Roy. 2006. "A Game-Theoretic Approach to GlobalWarming," *Advances in Mathematical Economics* 8: 135-53.

Emanuel, Kerry. 2005. "Increased Destructiveness of Tropical Cyclones Over the Past 30 Years." *Nature* 436: 686-688.

Energy Modeling Forum Study 19. 2004. *Energy Economics* 26.

Foster, Lucia, and Cheryl Grim. 2010. "Characteristics of the Top R&D Performing Firms in the US: Evidence from the Survey of Industrial R&D." CES Working Paper 10-33.

Ghosh, Kaushik, and William Kerr. 2010. "Patent-SSEL Matching." Census Bureau Technical Paper.

Golosov, Mikhail, John Hassler, Per Krusell, and Aleh Tsyvinski. 2011. "Optimal Taxes on Fossil Fuel in General Equilibrium." NBER Working Paper 17348.

Goulder, Lawrence, and Koshy Mathai. 2000. "Optimal CO2 Abatement in the Presence of Induced Technological Change." *Journal of Environmental Economics and Management* 39: 1-38.

Goulder, Lawrence, and Stephen Schneider. 1999. "Induced Technical Change and the

Attractiveness of CO2 Abatement Policies." *Resource and Energy Economics* 21: 211-253.

Gouriéroux, Christian, and Alain Monfort. 1996. *Simulation-Based Econometric Methods*, Oxford University Press.

Greenstone, Michael, Elizabeth Kopits, and Anne Wolverton. 2011. "Estimating the Social Cost of Carbon for Use in US Federal Rulemakings: A Summary and Interpretation." MIT Working Paper 11-04.

Griliches, Zvi. 1990. "Patent Statistics as Economic Indicators: A Survey." *Journal of Economic Literature* 28: 1661-1707.

Grimaud, Andre, Gilles Lafforgue, and Bertrand Magné. 2011. "Climate Change, Mitigation Options and Directed Technical Change: A Decentralized Equilibrium Analysis." *Resource and Energy Economics* 33: 938-962.

Grossman, Gene, and Elhanan Helpman. 1991. *Innovation and Growth in the Global Economy,* MIT Press.

Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg. 2001. "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools." NBER Working Paper 8498.

Harris, Christopher, and John Vickers. 1995. "Innovation and Natural Resources: A Dynamic Game with Uncertainty." *RAND Journal of Economics* 26: 418-430.

Harstad, Bard. 2012. "Climate Contracts: A Game of Emissions, Investments, Negotiations, and Renegotiations." *Review of Economic Studies*, forthcoming.

Hartley, Peter, Kenneth Medlock, Ted Temzelides, and Xinya Zhang. 2011. "Energy Sector Innovation and Growth: An Optimal Energy Crisis." Rice University Working Paper.

Hassler, John, and Per Krusell. 2012. "Economics and Climate Change: Integrated Assessment in a Multi-Region World." *Journal of the European Economic Association*, forthcoming.

Hassler, John, Per Krusell, and Conny Olovsson. 2011. "Energy-Saving Technical Change." NBER Working Paper 18456.

Hemous, David. 2012. "Environmental Policy and Directed Technical Change in a Global Economy: The Dynamic Impact of Unilateral Environmental Policies?", INSEAD Mimeo.

IPCC (Intergovernmental Panel on Climate Change). 2007. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change,* Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller (eds.). Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jaffe, Adam, David Popp, and Richard Newell. 2010. "Energy, the Environment, and Technological Change." in *Handbook of the Economics of Innovation: Vol. 2*, Bronwyn Hall and Nathan Rosenberg, eds., Academic Press/Elsevier, 873-937.

Kerr, William, and Shihe Fu. 2008. "The Survey of Industrial R&D—Patent Database Link Project." *Journal of Technology Transfer* 33: 173-186.

Kerr, William, and William Lincoln. 2010. "The Supply Side of Innovation: H-1B Visa Reforms and US Ethnic Invention." *Journal of Labor Economics* 28: 473-508.

Kirkpatrick, S., C.D. Gelatt, and M.P. Vecchi. 1983. "Optimization by Simulated Annealing." *Science* 220: 671-680.

Klette, Tor Jakob, and Samuel Kortum. 2004. "Innovating Firms and Aggregate Innovation." *Journal of Political Economy* 112: 986-1018.

Krusell, Per, and Anthony Smith. 2009. "Macroeconomics and Global Climate Change: Transition for a Many-Region Economy." Working Paper.

Landsea, Christopher. 2005. "Atlantic Hurricanes and Global Warming." *Nature* 438: E11-E12.

Lentz, Rasmus, and Dale Mortensen. 2008. "An Empirical Model of Growth through Product Innovation." *Econometrica* 76: 1317-1373.

MacCracken, Christopher, James Edmonds, Son Kim, and Ronald Sands. 1999. "The Economics of the Kyoto Protocol." *The Energy Journal* 20: 25–72.

McFadden, Daniel. 1989. "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration." *Econometrica* 57: 995-1026.

Mendelsohn, Robert, William Nordhaus, and Daigee Shaw. 1994. "The Impact of Global Warming on Agriculture: A Ricardian Analysis." *American Economic Review* 84: 753-771.

Messner, Sabine. 1997. "Endogenized Technological Learning in an Energy Systems Model." *Journal of Evolutionary Economics* 7: 291-313.

Mueller, Nicholas, Robert Mendelsohn, and William Nordhaus. 2011. "Environmental Accounting for Pollution in the United States Economy." *American Economic Review* 101: 1649-1675.

Nelder, John, and Roger Mead. 1965. "A Simplex Method for Function Minimization." *The Computer Journal* 7: 308-313.

Newell, Richard, Adam Jaffe, and Robert Stavins. 1999. "The Induced Innovation Hypothesis and Energy-Saving Technological Change." *Quarterly Journal of Economics* 114: 941-975.

Nicholls, Robert, and Jason Lowe. 2006. "Climate Stabilisation and Impacts of Sea-level Rise." in *Avoiding Dangerous Climate Change*, H.J. Schellnhuber, ed., 195-201.

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

Nordhaus, William. 2007. "To Tax or Not to Tax: The Case for a Carbon Tax." *Review of Environmental Economics and Policy*, 26-44.

Nordhaus, William. 2008. *A Question of Balance: Weighing the Options on Global Warming Policies*, Yale University Press.

Nordhaus, William, and Joseph Boyer. 2000. *Warming the World: Economic Modeling of Global Warming,* MIT Press.

OECD Environment Directorate. 2011. "Patent Search Strategies for the Identification of Selected Environment-Related Technologies." www.oecd.org/environment/innovation.

Pakes, Ariel, and David Pollard. 1989. "Simulation and the Asymptotics of Optimization Estimators." *Econometrica* 57: 1027-1057.

Pizer, William. 1999. "The Optimal Choice of Climate Change Policy in the Presence of Uncertainty." *Resource and Energy Economics* 21: 255-287.

Popp, David. 2002. "Induced Innovation and Energy Prices." *American Economic Review* 92: 160-180.

Popp, David. 2004. "ENTICE: Endogenous Technological Change in the DICE Model of Global Warming." *Journal of Environmental Economics and Management* 48: 742-768.

Popp, David. 2006. "Innovation in Climate Policy Models: Implementing Lessons from the Economics of R&D." *Energy Economics* 28: 596-609.

Popp, David, and Richard Newell. 2012. "Where Does Energy R&D Come From? Examining Crowding Out From Energy R&D." *Energy Economics* 3: 980-991.

Stern, Nicholas. 2007. *The Economics of Climate Change: The Stern Review*, Cambridge University Press.

Stott, Peter, D.A. Stone, and M.R. Allen. 2004. "Human Contribution to the European Heatwave of 2003." *Nature* 432: 610-614.

Van der Zwaan, Robert, Reyer Gerlagh, G. Klaassen, and L. Schrattenholzer. 2002. "Endogenous Technical Change in Climate Change Modelling." *Energy Economics* 24: 1-19.

Weitzman, Martin. 2009. "On Modeling and Interpreting the Economics of Catastrophic Climate Change." *Review of Economics and Statistics* 91: 1-19.

Zangwill, Willard, and C.B. Garcia. 1981. *Pathways to Solutions, Fixed Points, and Equilibria*, Prentice-Hall.