Online Appendix:
Not for Publication Unless Requested

A-1 Proofs and Additional Theoretical Results

A-1.1 Proof of Lemma 1

Proof of Lemma 1. Substituting (15) into (14), we obtain

$$
\begin{aligned}
& r_t \left(\sum\nolimits_{i=1}^{u} v_{n_i,t}^{j}\right) Y_t - \left(\sum\nolimits_{i=1}^{u} \dot{v}_{n_i,t}^{j}\right) Y_t - \left(\sum\nolimits_{i=1}^{u} v_{n_i,t}^{j}\right) \dot{Y}_t \\
= \; & \sum\nolimits_{i=1}^{u} \left[\begin{array}{c} \pi_{n_i}^{j} + z^{j}\left(-v_{n_i,t}^{j} Y_t\right) + \\ z^{-j}\left(1-\alpha+\mathbb{I}_{(n_i^j \le 0)}\alpha\right)\left(-v_{n_i,t}^{j} Y_t + v_{n-1,t}^{j} Y_t\right) + \mathbb{I}_{(n_i^j > 0)} z^{-j}\alpha\left(-v_{n_i,t}^{j} Y_t + v_{-1,t}^{j} Y_t\right) \end{array} \right] \\
& + \int \max_{x^j \ge 0} \left[u^j x^j \mathbb{E}_n v_{n_i,t}^{j} Y_t - \left(1-s_I^j\right) u^j w^s \left(\left(x^j\right)^{\frac{1}{\eta}} \theta^{-\frac{1}{\eta}} + \mathbb{I}_{(x^j>0)} F_I\right)\right] dF_I
\end{aligned}
$$

This can further be simplified to

$$
\begin{aligned}
& r_t \sum\nolimits_{i=1}^{u} v_{n_i,t}^{j} - \sum\nolimits_{i=1}^{u} \dot{v}_{n_i,t}^{j} - \sum\nolimits_{i=1}^{u} v_{n_i,t}^{j} g_{Y,t} \\
= \; & \sum\nolimits_{i=1}^{u} \left[\begin{array}{c} \tilde{\pi}_{n_i}^{j} - z^{j} v_{n_i,t}^{j} + \\ z^{-j}\left(1-\alpha+\mathbb{I}_{(n_i^j \le 0)}\alpha\right)\left(v_{n-1,t}^{j} - v_{n_i,t}^{j}\right) + \mathbb{I}_{(n_i^j > 0)} z^{-j}\alpha\left(v_{-1,t}^{j} - v_{n_i,t}^{j}\right) \end{array} \right] \\
& + \int \max_{x^j \ge 0} \left[u^j x^j \bar{v}_t^{j} - \left(1-s_I^j\right) u^j \tilde{w}^s \left(\left(x^j\right)^{\frac{1}{\eta}} \theta^{-\frac{1}{\eta}} + \mathbb{I}_{(x^j>0)} F_I\right)\right] dF_I.
\end{aligned}
$$

where

$$\tilde{w}^s \equiv \frac{w^s}{Y} \text{ and } \tilde{\pi} = \frac{\pi}{Y}.$$

Defining $\tilde{r}_t \equiv r_t - g_{Y,t}$ delivers the desired result. ∎

A-1.2 The Evolution of Quality Distributions

The quality indices in dirty sectors evolve according to following flow equations:

$$\begin{aligned}
\dot{Q}^d_{n>1,t} &= z^d_t \lambda Q^d_{n-1,t} + (1-\alpha) z^c_t Q^d_{n+1,t} - z_t Q^d_{n,t} \\
\dot{Q}^d_{1,t} &= z^d_t \lambda Q^d_{0,t} + (1-\alpha) z^c_t Q^d_{2,t} + \alpha z^d_t \lambda \sum_{n<0} Q^c_{n,t} - z_t Q^d_{1,t} \\
\dot{Q}^d_{0,t} &= (1-\alpha) z^d_t \lambda Q^d_{-1,t} + (1-\alpha) z^c_t Q^d_{1,t} - z_t Q^d_{0,t} \\
\dot{Q}^d_{-1,t} &= (1-\alpha) z^d_t \lambda Q^d_{-2,t} + z^c_t Q^d_{0,t} + \alpha z^c_t \sum_{n>0} Q^d_{n,t} - z_t Q^d_{-1,t} \\
\dot{Q}^d_{n<-1,t} &= (1-\alpha) z^d_t \lambda Q^d_{n-1,t} + z^c_t Q^d_{n+1,t} - z_t Q^d_{n,t}
\end{aligned}$$

The first line expresses the change in the quality index of sectors with $n > 1$. When there is a new dirty innovation in sector with $n-1$ at the rate z^d_t, on average its quality is improved by a multiplicative factor λ. Likewise, when there is a new incremental clean innovation at the rate $(1-\alpha)\, z^c_t$ in sectors with $n+1$, they enter into state n. Finally, there is an outflow from state n through clean or dirty innovations that occur at the rate z_t. The other lines represent the flow equations for state $n-1$, $n = 0$ and $n < 0$ respectively. Similar flow equations can be expressed for clean technologies as follows:

$$\begin{aligned}
\dot{Q}^c_{n>1,t} &= z^d_t Q^c_{n-1,t} + (1-\alpha) z^c_t \lambda Q^c_{n+1,t} - z_t Q^c_{n,t} \\
\dot{Q}^c_{1,t} &= z^d_t Q^c_{0,t} + (1-\alpha) z^c_t \lambda Q^c_{2,t} + \alpha z^d_t \sum_{n<0} Q^c_{n,t} - z_t Q^c_{1,t} \\
\dot{Q}^c_{0,t} &= (1-\alpha) z^d_t Q^c_{-1,t} + (1-\alpha) z^c_t \lambda Q^c_{1,t} - z_t Q^c_{0,t} \\
\dot{Q}^c_{-1,t} &= (1-\alpha) z^d_t Q^c_{-2,t} + z^c_t \lambda Q^c_{0,t} + \alpha z^c_t \lambda \sum_{n>0} Q^d_{n,t} - z_t Q^c_{-1,t} \\
\dot{Q}^c_{n<-1,t} &= (1-\alpha) z^d_t Q^c_{n-1,t} + z^c_t \lambda Q^c_{n+1,t} - z_t Q^c_{n,t}.
\end{aligned}$$

A-1.3 Optimal Constant Policy

To develop an intuition about the nature of optimal policy, we now consider a simplified optimal policy problem. In particular, we abstract from dynamics and from the exhaustible resource, and assume that the dirty output enters linearly into the production function, so that

$$\ln Y = \int_0^1 \ln y_i di - \gamma \int_0^1 y_i^d di. \tag{A-1}$$

This implies that the objective of the social planners to maximize net output. In the absence of the exhaustible resource, the unskilled labor market clearing condition becomes

$$1 = \int_0^1 l_i^c di + \int_0^1 l_i^d di.$$

Define

$$x_i \equiv \frac{q_i^d}{q_i^c}.$$

Let us rank sectors from 0 to 1 with ascending x_i, and also for ease of exposition, let us make the following assumptions:

1. x_i's distribution F is continuous and differentiable with density f,

2. all sectors charge a markup of λ.

Production is linear and satisfies $y_i = q_i^j l_i^j$, depending on which $j \in \{d, c\}$ is used. Furthermore, let us also normalize $q_i^d = 1$. We can order the product lines by x_i and define the cutoff to be μ, which will be a function of the tax τ, so that we can rewrite (A-1) as

$$\ln Y = -\int_\mu^1 \ln x_i di + \underbrace{\int_0^\mu \ln l_i^d di + \int_\mu^1 \ln l_i^c di}_{\equiv L} - \gamma Y^d. \tag{A-2}$$

Note that the marginal product line can be defined as

$$x^* = 1 + \tau \text{ and } \mu = 1 - F(1+\tau). \tag{A-3}$$

In addition, noting that all dirty sectors will use the same labor and all clean sectors will use the same labor, and taking into account the impact of the carbon tax on dirty labor demand, we have

$$l^d = \frac{1}{\mu + (1+\tau)(1-\mu)} \text{ and } l^c = \frac{1+\tau}{\mu + (1+\tau)(1-\mu)}.$$

Now we can show that

$$\begin{aligned} L &= \mu \ln \frac{1}{\mu + (1+\tau)(1-\mu)} + (1-\mu) \ln \frac{1+\tau}{\mu + (1+\tau)(1-\mu)} \\ &= (1-\mu) \ln(1+\tau) - \ln[\mu + (1+\tau)(1-\mu)] \end{aligned}$$

Likewise

$$Y^d = \frac{\mu}{\mu + (1+\tau)(1-\mu)}.$$

Let us assume that

$$F(x) = 1 - \frac{1}{x}.$$

Then from (A-3) we get the fraction of dirty sectors as

$$\mu = \frac{1}{1+\tau}$$

and the total sum of the dirty output in (23)

$$Y^d = \frac{1}{1 + (1+\tau)\tau}$$

Now the objective function can be expressed as

$$\max_{\tau}\left\{-\int_{\frac{1}{1+\tau}}^{1}\ln x_i di+\frac{\tau}{1+\tau}\ln\left(1+\tau\right)-\ln\left[\frac{1}{1+\tau}+\tau\right]-\frac{\gamma}{1+\left(1+\tau\right)\tau}\right\}.$$

The first-order condition of the maximization problem is

$$\frac{\tau}{\left(1+\tau\right)^{2}}-\frac{\left(1+\tau\right)^{2}-1}{\left(1+\tau\right)\left[1+\left(1+\tau\right)\tau\right]}+\gamma\frac{1+2\tau}{\left[1+\left(1+\tau\right)\tau\right]^{2}}=0$$

This can be expressed as

$$\gamma=\frac{\tau\left[1+\left(1+\tau\right)\tau\right]}{\left(1+\tau\right)^{2}}\equiv\Psi\left(\tau\right)$$

Note that $\Psi\left(0\right)=0$ and $\lim_{\tau\to\infty}\Psi\left(\tau\right)=\infty$. Note also that $\Psi\left(\tau\right)$ is monotonically increasing in τ. Hence from intermediate value theorem the optimal tax rate $\tau^{*}\left(\gamma\right)$ exists and it is unique. Moreover, the optimal tax rate is increasing in the damage parameter γ.

A-2 Computational and Estimation Algorithms

A-2.1 Computational Algorithm

Our theoretical analysis shows that the key microeconomic decisions are 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). To compute this fixed point, we first update the product line distribution (μ) in the forward direction and the value functions (V^{j}) in the backward direction, using the long-term (clean) steady state as the terminal condition.

For the estimation, we simulate a large panel of 16,384 firms, each with an evolving portfolio of product lines with various technological leads (n). After a lengthy burn-in time, we then sample the targeted statistics in order to find the best fit parameters. We use standard optimization routines to compute optimal policies. When focusing on constant policies, we use a straightforward grid search to find the optimum. In the time-varying case, we parameterize policies using a quartic carbon tax and a quartic research subsidy. 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).

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 (27).

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 to a threshold where the asymptotically dominant technology switches over, we also introduce heterogeneity in incumbent fixed costs as explained in the text.[1]

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++.

Finally, in our of policy analysis, we approximate optimal research subsidies and carbon taxes as quartic functions of time. In Section 6, we verify that the qualitative patterns are similar when we approximate these policies by step functions.

[1] A similar heterogeneity can be introduced in entrant fixed costs, but because entrants are never in the region where this heterogeneity matters, this makes no difference to our numerical procedures.

A-2.2 Estimation Algorithm

To find the moments used in the SMM estimation,[2] we simulate a panel of $16,384$ 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.

A-3 Further Details on Data and Sample Construction

Our patent database was 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. Our focus is on industrial assignees with inventors in United States. In a representative year, 1997, this group comprised about 77,000 patents (about 40% of the total USPTO patent count in 1997). We focus on technology codes, that number of 150,000, given the broad nature of patent classes (about 450 groups). This is important as energy-sector patents are spread out over multiple patent classes—two examples related to solar energy are "Power Plants/utilizing natural heat/Solar" and "Stoves and Furnaces/Solar heat collector". Moreover, we use patenting technologies 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). As a representative year, 1997, our energy-related

[2] McFadden (1989), Pakes and Pollard (1989) and Gourieroux and Monfort (1996) characterize the statistical properties of the SMM estimator.

patents comprised 7.6% of the total US patent count.

Our classifications follow closely Popp (2002) and Popp and Newell (2012), and we report results for our key parameters that just use their classification system. When seeking to extend their work, we use three steps. We first utilize 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. We next use information on energy-related R&D data from the NSF R&D Survey. 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 final 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.[3]

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.[4] 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 merg-

[3]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.

[4]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).

ers 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; 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).

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 (it is biannual for most of our sample period). 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.

When calculating our innovation production function for the sector (e.g., the η and α parameters), 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.[5] 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.

The main text outlines how we focus on firms that are initially innovative, including those that transition out of innovation, but do not allow non-innovative incumbents to enter. 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.

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.[6] In addition to the statistics reported in the main text, 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 56,000 patents during 1995-1999.

[5]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.

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

Several additional points are worth noting about the sample and our data approach. 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.

Tables A1 and A2 provide additional results mentioned in the main text:

Table A1. 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 Table 1.

Table A2. 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%

Table $A3$ summarizes the alternative initial technology distribution considered in the robustness analysis. The average gap to the frontier for dirty-patent 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 distribution graph has a broadly similar shape as Figure 3. 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 step sizes or more in 8% and 2% of technologies, respectively.

Table A3. 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%

Additional References

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.

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

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.

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

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

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

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

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

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

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

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.

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

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.