ONLINE APPENDIX

Appendix A: Proofs

Proof or Proposition 2. For all $\kappa > \kappa_{1,}$, we can write total emissions (9) as:

$$X(\Delta) = \left(1 - \frac{\Delta \pi_M}{\kappa}\right)[1 - 2(1-\Delta)\pi_M] + \frac{\Delta}{\kappa\gamma}\pi_M(1-\pi_M). \tag{A.1}$$

Focusing first on the extremes of full competition and full collusion to get the main intuitions, the former is less polluting than the latter if $X(1) < X(1/2)$, or

$$\begin{aligned} 1 - \frac{\pi_M}{\kappa} + \frac{\pi_M(1-\pi_M)}{\kappa\gamma} &< \left(1 - \frac{\pi_M}{2\kappa}\right)(1-\pi_M) + \frac{\pi_M(1-\pi_M)}{2\kappa\gamma} \iff \\ \frac{\pi_M(1-\pi_M)}{2\kappa\gamma} &< \left(1 - \frac{\pi_M}{2\kappa}\right)(1-\pi_M) - \left(1 - \frac{\pi_M}{\kappa}\right) = \frac{\pi_M(1+\pi_M)}{2\kappa} - \pi_M, \end{aligned}$$

which simplifies to

$$\kappa < 1 - \frac{\gamma^{-\delta}}{2}\left(1 + \frac{1}{\gamma}\right) = \kappa_2, \tag{A.2}$$

where $\kappa_2 > 1 - \gamma^{-\delta} = \pi_M = \kappa_1$ was first defined in Proposition 2. Quite intuitively, for any given κ, (A.2) holds when γ or/and δ is large enough. Let us next determine where X achieves its maximum on $[1/2, 1]$:

$$\kappa\frac{\partial X}{\partial \Delta} = -4\pi_M^2\Delta + (2\kappa - 1 + 2\pi_M)\pi_M + \frac{1}{\gamma}\pi_M(1-\pi_M), \tag{A.3}$$

so $\partial X/\partial\Delta > 0$ if and only if

(A.4)

$$\Delta < \frac{1}{4\pi_M}\left(2\kappa - 1 + 2\pi_M + \frac{1-\pi_M}{\gamma}\right) = \frac{1}{2} + \frac{1}{4\pi_M}\left(2\kappa - 1 + \frac{\gamma^{-\delta}}{\gamma}\right) \equiv \hat{\Delta}_X(\kappa,\gamma,\delta).$$

Naturally, $\hat{\Delta}_X$ is increasing in κ and decreasing in both γ and δ. Moreover,

$$\hat{\Delta}_X(\gamma,\delta) < \frac{1}{2} \iff \kappa < \frac{1}{2}\left(1 - \frac{\gamma^{-\delta}}{\gamma}\right) = \kappa_2 - \frac{\pi_M}{2} \equiv \kappa_3, \tag{A.5}$$

$$\hat{\Delta}_X(\gamma,\delta) > 1 \iff \kappa > \kappa_3 + \pi_M = \kappa_2 + \frac{\pi_M}{2} \equiv \kappa_4 \tag{A.6}$$

where $\kappa_2 > \kappa_1 = \pi_M$ was first defined in Proposition 2, by equation (A.2). It then follows that (maintaining $\kappa > \kappa_1$, thus ensuring an interior optimum for z) :
(i) If $\kappa < \kappa_2 - \kappa_1/2$, Z is decreasing in Δ, and thus minimized at $\Delta = 1$.
(ii) If $\kappa > \kappa_2 + \kappa_1/2$, then Z is increasing in Δ, and thus minimized at $\Delta = 1/2$.
(iii) If $\kappa \in (\kappa_2 - \kappa_1/2, \kappa_2 + \kappa_1/2)$ then X is hump-shaped in Δ, with a maximum at $\hat{\Delta}_X(\gamma, \delta) \in (1/2, 1)$ and a minimum either at 1/2 or at 1, depending on $\kappa \gtreqless \kappa_2$ (recall that this is what defines κ_2).

Note, finally, that conditions $\kappa > \pi_M$ and $\kappa < \kappa_2 - \pi_M/2$ define a nonempty interval when $3\pi_M < 2\kappa_2$, that is, $\gamma^{-\delta}(2 - 1/\gamma) > 1$, or

$$\delta < \ln(2 - 1/\gamma) / \ln \gamma. \quad \blacksquare \tag{A.7}$$

Proof or Proposition 3. From (A.1), when $\kappa > \kappa_1$, we have

$$\kappa \frac{\partial X}{\partial \pi_M} = \frac{\Delta}{\kappa}\left[-1 + 2(1-\Delta)\pi_M + \frac{1-\pi_M}{\gamma}\right] - 2(1-\Delta)\left(1 - \frac{\Delta \pi_M}{\kappa}\right) - \frac{\Delta \pi_M}{\kappa \gamma}. \tag{A.8}$$

The last two terms are clearly negative, and so is the first, since $(1-\pi_M)/\gamma < 1 - \pi_M \leq 1 - 2(1-\Delta)\pi_M$ for all $\Delta \geq 1/2$. Recalling that $\pi_M = 1 - \gamma^{-\delta}$, it follows that $\partial X / \partial \delta < 0$. When $\kappa \leq \kappa_1$, R&D effort may be (depending on Δ) at a corner, $z = 1$, in which case $X = y_M/\gamma = 1/c\gamma^{-\delta-1}$, which decreases in δ. Finally, differentiating (A.3) in π_M,

$$\begin{aligned} \kappa \frac{\partial^2 X}{\partial \Delta \partial \pi_M} &= -4(1-2\pi_M)\Delta + \frac{1}{\gamma}(1-2\pi_M) - 2\kappa + 2(1-2\pi_M) - 1 \\ &= (1-2\pi_M)\left[\frac{1}{\gamma} + 2 - 4\Delta\right] - 1 - 2\kappa. \end{aligned}$$

If $1 - 2\pi_M \geq 0$, the right-hand side is bounded above by $(1-2\pi_M)/\gamma - 1 - 2\kappa < 1/\gamma - 1 - 2\kappa < 0$. If $1 - 2\pi_M < 0$, it is bounded above by $(2\pi_M - 1)(2 - 1/\gamma) - 1 - 2\kappa$, since $\Delta \leq 1$; but $\pi_M \leq 1$, so this expression is at most $1 - 1/\gamma - 2\kappa < 0$, since $\kappa > \kappa_1 = \pi_M = 1 - 1/\gamma^\delta > 1 - 1/\gamma$. Therefore, $\partial^2 X / \partial \Delta \partial \delta < 0$ for all Δ, as long as $\kappa > \kappa_1$.

Proof or Proposition 4. *Part (a).* This follows from the conjunction of $\partial X / \partial \Delta < 0$ for $\kappa \prec \kappa_2 - \kappa_1/2$, by Proposition 2, and

$$\frac{\partial U}{\partial \Delta} = \frac{\pi_M}{\kappa} \ln\left(\frac{1}{1 - 2(1-\Delta)\pi_M}\right) + \left(1 - \frac{\Delta \pi_M}{\kappa}\right)\frac{2\pi_M}{1 - 2(1-\Delta)\pi_M} > 0. \tag{A.9}$$

Part (b). Recalling (3), (5) and (10), we can rewrite

$$\text{(A.10)} \qquad U = \left(1 - \frac{\Delta\pi_M}{\kappa}\right)\ln\left(1 - 2(1-\Delta)\pi_M\right) + \ln\left(\frac{1}{c}\right),$$

$$\text{(A.11)} \qquad \frac{\partial U}{\partial \pi_M} = \frac{\Delta}{\kappa}\ln\left(\frac{1}{1-2(1-\Delta)\pi_M}\right) - \frac{2(1-\Delta)\pi_M}{1-2(1-\Delta)\pi_M}\left(1 - \frac{\Delta\pi_M}{\kappa}\right),$$

Thus, $\partial U/\partial \pi_M > 0$ if and only if

$$\text{(A.12)} \qquad \kappa < \Delta\left[\pi_M + f\left(\frac{2(1-\Delta)\pi_M}{1-2(1-\Delta)\pi_M}\right)\right],$$

where $f(t) \equiv \ln(1+t)/t$ for all $t > 0$ and $f(0) \equiv \lim_{t\to 0} f(t) = 1$. Note that f is a decreasing function, since $f'(t)$ has the sign of $g(t) \equiv t - (1+t)\ln(1+t)$, where clearly $g'(t) < 0 = g(0)$ for all $t > 0$. The right-hand side of (A.12) is thus increasing in Δ, so the inequality holds if and only if $\Delta > \underline{\Delta}(\pi_M, \kappa)$, with

$$\text{(A.13)} \qquad \underline{\Delta}(\pi_M, \kappa) < 1 \iff \kappa < 1 + \pi_M,$$

$$\text{(A.14)} \qquad \underline{\Delta}(\pi_M, \kappa) < 1/2 \iff \kappa < \frac{1}{2}\left[\pi_M + f\left(\frac{\pi_M}{1-\pi_M}\right)\right] \equiv \bar{\kappa}(\pi_M),$$

Condition (A.13) is always compatible with $\kappa > \pi_M$ and $\kappa < \kappa_2 - \pi_M/2$. Condition (A.14), which ensures that $\partial U/\partial \pi_M > 0$ for all values of $\Delta \in [1/2, 1]$, is more demanding since $\bar{\kappa}(\pi_M) < (1 + \pi_M)/2$ and compatible with $\kappa > \pi_M$, only if

$$\text{(A.15)} \qquad \pi_M < f\left(\frac{\pi_M}{1-\pi_M}\right) = \frac{\ln\left[1/(1-\pi_M)\right]}{\pi_M/(1-\pi_M)} \iff \pi_M^2 < (1-\pi_M)\ln\left(\frac{1}{1-\pi_M}\right),$$

which holds for instance when π_M is small enough, meaning that $\delta \ln \gamma$ is small enough. This finishes to establish (b).

Part (c). In (A.9), the first term is increasing in π_M, and while the second not always is, a sufficient condition is that $(\Delta\pi_M/\kappa)(1 - \Delta\pi_M/\kappa)$ be increasing, which occurs for $\Delta\pi_M/\kappa < 1/2$; conversely, $\pi_M/\kappa < 1/2$ is necessary the second term for that same term to be increasing in Δ up to $\Delta = 1$. Thus, when $\kappa > 2\pi_M = 2\kappa_1$, we have $\partial^2 U/\partial\Delta\partial\delta > 0$, hence the result since $\partial^2 X/\partial\Delta\partial\delta > 0$.

We check, finally, that this new lower bound on κ is compatible with key upper bounds previously defined, meaning that they jointly define a nonempty set of values for (κ, γ, δ). We have:

$$
\begin{aligned}
2\pi_M &< \kappa_2 - \pi_M/2 \iff 5(1-\gamma^{-\delta}) < 2\kappa_2 = 2 - \gamma^{-\delta}(1+1/\gamma) \iff \\
\text{(A.16)} \quad \delta &< \frac{\ln(4/3 - 1/3\gamma).}{\ln \gamma}. \\
2\pi_M &< \bar{\kappa}(\pi_M) \iff 3\pi_M < f\left(\frac{\pi_M}{1-\pi_M}\right) = \frac{\ln\left[1/(1-\pi_M)\right]}{\pi_M/(1-\pi_M)} \\
\text{(A.17)} \quad &\iff 3\pi_M^2 < (1-\pi_M)\ln\left(\frac{1}{1-\pi_M}\right).
\end{aligned}
$$

The first condition is naturally tighter than (A.7), so when it holds we have $\partial^2 U/\partial\Delta\partial\delta > 0$ for all Δ and $\partial U/\partial\delta > 0$ for Δ in some nonempty interval $(\underline{\Delta}, 1]$. If the second condition also holds (which is ensured by some additional upper bound on δ), then $\partial^2 U/\partial\Delta\partial\delta > 0 > \partial U/\partial\delta > 0$ for all $\Delta \in [1/2, 1]$. This, together with the fact that, from Proposition 2, $\partial^2 X/\partial\Delta\partial\delta \prec 0$ for all $\kappa \succ \kappa_1$, establishes Part (c). ■

Appendix B: Counterfactual Methodology

We can write our regression model in equation (12) as

$$
\text{(B.1)} \qquad Z_{j,t} \equiv \ln\left(PAT_{j,t} + 1\right) = \alpha V_{j,t} + \beta C_{j,t} + \gamma V_{j,t} \times C_{j,t} + \varphi F_{j,t} + \varepsilon_{j,t},
$$

where, for each firm j and time t, $PAT_{j,t}$ is the number of patents (families) of a given type (clean or dirty), $V_{j,t}$ and $C_{j,t}$ are its (average) degrees of exposure to prosocial values and competition respectively, and $F_{j,t}$ collects all other explanatory variables, such as oil prices, firm and period fixed effects, etc.

Denoting $\Delta X_{j,t} = X_{j,t} - X_{j,\tau}$ any historical or counterfactual change between dates τ and t, and given estimated coefficients $(\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\varphi})$, the implied patenting level at t is

$$
\text{(B.2)} \qquad \widehat{PAT}_{j,t} = (PAT_{j,\tau} + 1) \times \exp(\widehat{\Delta Z}_j) - 1,
$$

where (omitting time subscripts to lighten the notation):

$$
\text{(B.3)} \quad \widehat{\Delta Z}_j \equiv \hat{\alpha}\Delta V_j + \hat{\beta}\Delta C_j + \hat{\gamma}(\Delta V_j \times C_j) + \hat{\gamma}(V_j \times \Delta C_j) + \hat{\gamma}(\Delta V_j \times \Delta C_j) + \hat{\varphi}\Delta F_j.
$$

For small changes, $\widehat{\Delta PAT}_j$ is proportional to $\widehat{\Delta Z}_j$, and can thus be decomposed into the constituents of (B.3). Alternatively, one can use the fitted nonlinear model for counterfactual analysis, asking: "How much would the total patents of each type have increased or decreased between τ and t, if the *only* changing factor had been the variations in environmental values observed in the different countries, and thus firms' exposures $V_{j,t}$?" Or, replacing historical accounting by prospective simulations: "How much should we expect those patent numbers to increase between τ and (some future) t, if the only changing factor will be some assumed set of ΔV's (or/and ΔC's?"

The answer is obtained by setting, for each j, all terms in (B.3) to zero except for $\hat{\alpha}\Delta V_j + \hat{\gamma}(\Delta V_j \times C_j)$, then summing across firms the resulting $\widehat{\Delta PAT}_j$'s computed from (B.2). This total change can itself be attributed to the combination of a direct, "average" effect of the ΔV_j's (weighted by initial patenting activity), and one that reflects their *interaction,* and therefore their *correlation* pattern, with initial levels of competition, C_j. This is again clearest when understood as a first-order approximation,

$$\begin{aligned}
\widehat{\Delta PAT} &\equiv \sum_j \widehat{\Delta PAT}_j \approx \sum_j (PAT_j + 1)\,\widehat{\Delta Z}_j \\
&= \hat{\alpha} \sum_j (PAT_j + 1)\,\Delta V_j + \hat{\gamma} \sum_j (PAT_j + 1)\,C_j \times \Delta V_j \\
&= \left(\hat{\alpha} + \hat{\gamma}\bar{C}\right) \sum_j (PAT_j + 1)\,\Delta V_j + \hat{\gamma} \sum_j (PAT_j + 1)\left(C_j - \bar{C}\right)\Delta V_j, \qquad \text{(B.4)}
\end{aligned}$$

where $\bar{C} \equiv (1/N) \sum_j (PAT_j + 1)\,C_j$ is the average level of (firm exposure to) competition, with each firm weighted by its initial patenting activity. Alternatively, to get exact numbers we can simulate the nonlinear model, by:

(a) Setting, for all j, all changes in (B.3) except ΔV_j to zero, and equating all C_j's to $\bar{C}$; the results for clean, grey and dirty patents are given in Column 2 of Table 3. They correspond to what would have happened if every firm had faced the (patent-weighted) average attitudinal *change*, and the (patent-weighted) average *level* of market competition.

(b) Setting all terms but the $\hat{\gamma}(\Delta V_j C_j)$'s to zero, and subtracting $\hat{\gamma}\bar{C} \sum_j (PAT_j + 1) \times \Delta V_j$. This yields the results in Column 3, reflecting the (patent-weighted) extent to which firms that saw larger ΔV_j's in their markets were exposed there to higher or lower levels of competition.

Similarly, Columns 4 and 5 in Panel A compute the counterfactual changes in each number of patents (relative to total) corresponding to historical changes in competition *only,* doing so separately for the effect of the (patent-weighted) average change, evaluated at the mean level of environmental values, $\left(\hat{\beta} + \hat{\gamma}\bar{V}\right) \sum_j (PAT_j + 1) \times \Delta C_j$, and that reflecting the correlation pattern with initial attitudes, $\hat{\gamma} \sum_j (PAT_j + 1) \left(V_j - \bar{V}\right) \times \Delta C_j$, where $\bar{V} \equiv (1/N) \sum_j (PAT_j + 1) V_j$.

Column 7 incorporates all the above effects, plus those of the interaction in *changes,* $\hat{\gamma}(\Delta V_{jt} \times \Delta C_{j,t})$. Column 8 adds to the effects of column 7 those due to variations in oil prices.

The prospective exercise reported in Panel B of Table 3 is identical, except that the initial date is $\tau = 2012$ and the counterfactual $\Delta V_{j,t}$'s and $\Delta C_{j,t}$'s are taken to be *uniform* across firms, equal respectively to 0.78 and 0.08 standard deviations. As explained in Section 6 (see also Table B.1), these magnitudes are the historical ones observed in our sample, but with a sign reversal for the former –in line with the fact that, since 2012 (when our patent dataset ends), the previous general decline environmental values seems to have given way to an upswing.

Table B.1: Descriptive Statistics for Counterfactual Calculations

	Unweighted					Patent-Weighted	
	Mean	Std.Dev.	P1	P50	P99	Mean	Std.Dev.
$\Delta Values$	-0.779	0.761	-1.614	-1.059	2.454	0.142	1.630
$\Delta Comp$	0.089	0.371	-0.680	0.000	1.516	0.310	0.579
$(Comp - \overline{Comp}) \times \Delta Values$	-0.167	1.108	-1.963	-0.151	3.263	0.343	1.733
$\left(Values - \overline{Values}\right) \times \Delta Comp$	0.004	0.016	-0.030	0.000	0.066	0.013	0.025
$\Delta log(FuelPrice)$	0.556	0.046	0.522	0.537	0.754	0.576	0.057

Note: Patent weighting is defined in equation B.4, using firms' clean patent levels in 2002.

Table B.2: Correlations between key variables

	Clean	Dirty	Values	Competition	ΔValues	ΔCompetition
Clean	1.000					
Dirty	0.596	1.000				
Values	-0.050	-0.048	1.000			
Competition	0.078	0.050	-0.080	1.000		
ΔValues	0.104	0.123	-0.589	0.086	1.000	
ΔCompetition	0.046	0.049	-0.009	0.005	0.243	1

Note: Clean, Grey and Dirty correspond here to (one plus) each firms' number of patents in each category, in the 1997-2002 time period. The measures of Values and Competition also refer to the 1997-2002 sample period. The differenced variables refer to the difference between the 2008-2012 and the 1997-2002 time period.

Appendix C: Details on variable definition

C1. Classifying patents as clean, dirty or grey

Table C.1 reports the Cooperative Patent Classification (CPC) classification used to determine the different flavours of innovation.[20]

Table C.1: Patent CPC classification codes used

	Clean
Y02T10/60	Other road transportation technologies with climate change mitigation effect
Y02T10/70	Energy storage for electromobility
Y02T90/10	Technologies related to electric vehicle charging
Y02T90/34	Fuel cell powered electric vehicles
Y02T90/42	Hydrogen as fuel for road transportation
	Grey
Y02T10/10	Climate change mitigation technologies related to fuel injection
Y02T10/20	Climate change mitigation technologies related to exhaust after treatment
Y02T10/40	Climate change mitigation technologies related to engine Management Systems
Y02T10/50	Climate change mitigation technologies related to Intelligent Control Systems
	Dirty
F02	Combustion Engines
	Other Automotive
B60	Vehicles in General

C2. Values

The data on attitudes comes from the International Social Survey Program (ISSP) and the World Value Survey (WVS). Several questions could capture the values we are interested in, but they are often asked only in a limited set of countries during a single survey wave. Only one question is common to both surveys, allowing us to cover many countries for two time periods. In the ISSP, it is: *How willing would you be to pay much higher taxes in order to protect the environment?;*and in the WVS, *Can you tell me whether you strongly agree, agree, disagree or strongly disagree with the following statement: 'I would agree to an increase in taxes if the extra money were used to prevent environmental pollution'.* In both cases, answers are given on a 5-point scale. Answers to the ISSP question vary from 1 ('very willing') to 5 ('very unwilling')

[20]See `https://www.cooperativepatentclassification.org/index`, as well as also `https://www.wipo.int/classifications/ipc/en/` and `https://www.epo.org/news-issues/issues/classification/classification.html`.

and we reverse-code them, so that a higher value means a more pro-environmental attitude. In the WVS, answers to the corresponding question are 1 ('strongly agree'), 2 ('agree'), 4 ('disagree') and 5 ('strongly disagree'). We code as 3 the 'don't know' answers and reverse-code the others, as for the ISSP.

Because taxes pertain to public policy more directly than to consumer spending decisions, we also use one additional variable from each survey to create a synthetic index. For ISSP, the question is: *How willing would you be to pay much higher prices in order to protect the environment?* For the WVS, it is about (dis)agreement with the statement: *I would give part of my income if I were certain that the money would be used to prevent environmental pollution* . To ensure consistency, we code all answers so that higher values mean more pro-environmental attitudes. We then average all variables at the country-period level, transform them into z-scores, and average across all variables available for the country-period observation. We thus have data on willingness-to-pay for the environment for 25 countries for 2 periods, namely 2000 and 2010. Our data cover most major economies, and in particular most countries in which firms innovating in the automotive sector reside, with a few notable exceptions such as Italy and Spain.

C3. Computation of firm-level Lerner Index

We estimate firm-level measures of competition using a (revenue) production function framework. We assume a homothetic translog production function with materials $M_{i,t}$ and labor $L_{i,t}$ as flexible factors, and capital $K_{i,t}$ a quasi-fixed production factor. A firm's (log) revenue ($R_{i,t}$) growth can then be written as

$$\Delta r_{i,t} \approx \frac{\lambda}{\bar{\mu}_{i,t}} \Delta k_{i,t} + \bar{s}_{Mi,t} \left(\Delta m_{i,t} - \Delta k_{i,t}\right) + \bar{s}_{Li,t} \left(\Delta l_{i,t} - \Delta k_{i,t}\right) + \frac{1}{\bar{\mu}_{i,t}} \Delta \omega_{i,t}, \tag{C.1}$$

where $\Delta r_{i,t} = \ln(R_{i,t}/ \ln R_{i,t-1})$ (and equivalently for production factors), λ is a scale parameter, $\bar{s}_{Mi,t} = \left(s_{Mi,t} + s_{Mi,t-1}\right)/2$ the average share of materials expenditure in revenue between period t and $t-1$ (and equivalently for labor inputs), and $\omega_{i,t}$ a Hicks- neutral shifter of TFP or/and demand. $\bar{\mu}_{i,t}$ is the average markup of prices over marginal cost between period t and $t-1$, making $\bar{\mu}_{i,t}$ -1 a Lerner index specific to firm i at time t. Short run profit maximization implies

$$s_{Mi,t} = \frac{\alpha_{Mi,t}}{\mu_{i,t}}, \tag{C.2}$$

where $\alpha_{Mi,t}$ is the elasticity of output with respect to changes in production factor M (and analogously for labor). Note that in the translog case,

$$\alpha_{Mi,t} = \alpha_M + \alpha_{KM} k_{i,t} + \alpha_{LM} l_{i,t} + \alpha_{MM} m_{i,t}. \tag{C.3}$$

This specification is consistent with a wide variety of market structures. For further discussion see Martin (2012) and Forlani (2016). We can rewrite (C.1) as

$$\Xi_{i,t} \frac{\bar{\alpha}_{Mi,t}}{\lambda} - \Delta k_{i,t} = \frac{1}{\lambda} \Delta \omega_{i,t}, \tag{C.4}$$

where

$$\Xi_{i,t} \equiv \frac{\Delta r_{i,t} - \frac{\lambda}{\bar{\mu}_{i,t}} + \bar{s}_{Mi,t} \left(\Delta m_{i,t} - \Delta k_{i,t} \right) + \bar{s}_{Li,t} \left(\Delta l_{i,t} - \Delta k_{i,t} \right)}{\bar{s}_{Mi,t}}.$$

Given assumptions on the evolution of the $\Delta\omega_{it}$ shock, we can fit this to firm-level data using a GMM approach. Thus, if $\Delta\omega_{i,t}$ follows an AR(1) process, $\omega_{i,t} = \rho\omega_{i,t-1} + \eta_{i,t}$ where $\eta_{i,t}$ is iid, we can write

$$\hat{\eta}_{i,t} = \Xi_{i,t} \frac{\bar{\alpha}_{Mi,t}}{\lambda} - \Delta k_{i,t} - \frac{\rho}{\lambda} \left[\Xi_{i,t-1} \frac{\bar{\alpha}_{Mi,t-1}}{\lambda} - \Delta k_{i,t-1} \right],$$

and estimate the parameters $\delta = [\rho/\lambda, \alpha_M/\lambda, \alpha_{KM}/\lambda, \alpha_{LM}/\lambda, \alpha_{MM}/\lambda]$ using the moment conditions:

$$E \left[\hat{\eta}_{i,t} \times \left\{ \Xi_{i,t-1}, \frac{1}{\Delta k_{i,t}}, \frac{\bar{k}_{i,t}}{\Delta k_{i,t}}, \frac{\bar{l}_{i,t}}{\Delta k_{i,t}}, \frac{\bar{m}_{i,t}}{\Delta k_{i,t}} \right\} \right] = 0.$$

After identifying δ, we can compute $\widehat{\alpha_{Mi,t}/\lambda}$ using (C.3). Then, from (C.2) we can compute

$$\widehat{\frac{\lambda}{\mu_{i,t}}} = s_{Mi,t} \left(\widehat{\frac{\alpha_{Mi,t}}{\lambda}} \right)^{-1}, \tag{C.5}$$

which is an inverse Lerner Index, scaled by the returns to scale parameter λ; i.e. it tells us the excess of markups over returns to scale. While this is different from the markup over marginal costs, it is more relevant in terms of measuring market power, as revealed by excess earnings over what would be reasonable to compensate for increasing returns. We also implement a simpler version, assuming a Cobb Douglas production function, so that $\alpha_{Mi,t} = \alpha_M$. Both approaches lead to similar results.

Panel (a) of Figure 4 shows deciles of the distribution of markups over marginal costs – i.e., the inverse of the Lerner Index – across firms. It indicates that markups (and thus competition) have been flatlining over time, with the exception of the top decile, where we see an upward trend from 2003 onwards. Panel (b) shows changes in market power for continuing firms between 2002 and 2012: for the majority of automobile firms, the general picture is that of a *reduction in market power* during that time period.

C4. Exogenous competition indicators

We provide here more details on our construction of country and sector-specific competition indicators. Consider a simplified version of our main regression equation (12):

$$\Delta Innovation_j = \beta \Delta Competition_j + \varepsilon_j,$$

where competition is computed as the change in the inverse Lerner Index measured via markups $\mu_{j,t}$. Let us focus, for simplicity, on obtaining an unbiased estimate of the causal effect of competition β. A central concern is that shocks to innovation lead, almost by definition, to increases in market power, which could translate into lower competition measured as markups. Furthermore, an innovation shock to firm j could also affect other firms that operate in the same sector. Our identification assumption is that such effects only operate within 4-digit sectors.

Suppose that all firms in our sample produced for one country $c(j)$ only. Our strategy would then boil down to creating an indicator of exogenous shocks to competition, $\hat{comp}_{c(j),s(j),t}$ for each firm j, by averaging over the inverse markups for firms i in the same 2-digit sector as firm j but excluding those in the same 4-digit sector as firm j:

$$\text{(C.6)} \qquad \hat{comp}_{c(j),s(j),t} = Mean\left(\frac{1}{\mu_{i,t}} | s(i) = s(j) \text{ and } s4dig(i) \neq s4dig(j)\right).$$

Consequently, we can use as an index for changes in competition exposure for firm j

$$\text{(C.7)} \qquad \Delta \hat{comp}_j = \hat{comp}_{c(j),s(j),t} - \hat{comp}_{c(j),s(j),t-1},$$

so that $\Delta \hat{comp}_j \perp \varepsilon_j$.

In practice, the firms in our sample operate across several countries, which we measure by the share of patenting across various jurisdictions c via the weights $w_{c,j}$. To obtain exogenous shocks to competition for each firm j, we run regressions

$$\frac{1}{\mu_{i,t}} = \sum_{c} comp_{c,s(i),t} w_{c,i} + \varepsilon_{i,t} \tag{C.8}$$

over the sample of firms i such that $s(i) = s(j)$ and $s4dig(i) \neq s4dig(j)$. We then obtain a firm-level competition index for firm j as

$$\Delta c\hat{omp}_j = \sum_{c} \left(c\hat{omp}_{c,s(j),t} - c\hat{omp}_{c,s(j),t-1} \right) w_{c,j} \tag{C.9}$$

In the special case where a firm only operates in one country (C.8) and (C.9) are equivalent to (C.6) and (C.7).

Appendix D: Details on the Bartik research design

Suppose (clean) innovation by firm j at time t can be described as

$$I_{j,t} = J_j + \alpha S_{j,t} + \varepsilon_{j,t}$$

where, for simplicity, we consider only one Bartik-style variable

$$S_{j,t} = \sum_{c} w_{j,c} S_{c,t},$$

in which the $S_{c,t}$ are country-level shocks (e.g., pro-social attitudes) and the $w_{j,c}$ are firm-level weights measuring a firm's exposure to a particular country. J_j is a firm fixed effect. We assume that country-level shocks $S_{c,t}$ and firm-level shocks $\varepsilon_{j,t}$ can be decomposed as follows:

$$S_{c,t} = J_c + c_{c,t} + \eta_{c,t}, \tag{C.10}$$

$$\varepsilon_{j,t} = \gamma N_{j,t} + c_{c(j),t} + \nu_{j,t},$$

where

$$N_{j,t} = \sum_c w_{j,c} N_{c,t}$$

and the $N_{c,t}$ are additional country-level shocks that are affecting firm-level outcomes in accordance with firm-level exposure. We assume that the $N_{c,t}$ are not correlated with the $S_{c,t}$, i.e. $S_{c,t} \perp N_{c,t}$. $c_{c(j),t}$ is a time-varying country-level factor (where $c(j)$ denotes the country where firm j is based/headquartered). It is meant to capture specific capabilities that might emerge in a particular country (e.g., a strong supply-chain ecosystem favourable to clean technologies), which we also allow to feed into country level variables such as social attitudes (see equation C.10). J_c is a fixed component in the country-level shock, while $\nu_{j,t}$ and $\eta_{c,t}$ are iid.

In our baseline, we allow that weights could be determined by fixed country- and firm-level characteristics,

$$w_{j,c} = f(J_j, J_c) + \xi_{j,c}.$$

For instance, when firms established their patenting strategy during the pre-sample period, this may have been based on long-standing country characteristic known at the time, as well as on persistent firm capabilities. However, because ε_{jt} and ν_{jt} are iid, we can purge potential endogeneity by first differencing our regression equations:

$$\Delta I_j = \alpha \Delta S_{j,t} + \Delta \varepsilon_{j,t}, \tag{C.11}$$

with $\Delta \varepsilon_{j,t} \perp w_{j,c}$Note that this also ensures that

$$\sum_c w_{j,c} c_{c,t} \perp \Delta \varepsilon_{j,t}. \tag{C.12}$$

Next, since,

$$\Delta S_{j,t} = \sum_c w_{j,c} \left(\Delta c_{c,t} + \Delta \eta_{c,t} \right),$$

we also have $\Delta S_{j,t} \perp \Delta \varepsilon_{j,t}$, and thus (C.11) will provide an unbiased estimate of α. Similarly, $N_{j,t}$ will be independent between observations, so that no non-standard clustering is needed to estimate standard errors.

Alternatively, we can assume that persistent firm-level characteristic also influence innovation trends; our equation for firm level shocks would then be of the form

$$\varepsilon_{j,t} = \kappa J_{j,t} + \gamma N_{j,t} + c_{c(j),t} + \nu_{j,t},$$

where $\kappa J_{j,t}$ is a firm-specific trend. Note that the differenced Bartik explanatory variable $S_{j,t}$ remains orthogonal to most of the differenced firm level shocks

$$\Delta\varepsilon_{j,t} = \kappa J_j + \Delta N_{j,t} + \Delta c_{c(j,)t} + \Delta\nu_{j,t},$$

$$\Delta S_{j,t} \perp \kappa J_j, \Delta N_{jt}, \Delta\nu_{jt}.$$

However, because the weights $w_{j,c}$ are now correlated with the κJ_j part of the $\Delta\varepsilon_{j,t}$, we can no longer assume $\Delta S_{j,t} \perp \Delta\varepsilon_{j,t}$. This is easiliy rectified by including headquarter dummies $\alpha_{c(j)}$ in the regression equation, which becomes

$$\Delta I_j = \alpha\Delta S_{j,t} + \alpha_{c(j)} + \Delta\chi_{j,t},$$

where $\chi_{j,t} = \gamma\Delta N_{j,t} + \kappa J_j + \nu_{j,t}$. Because weights are no longer random, we can no longer assume $\Delta N_{j,t} \perp \Delta N_{i,t}$. To deal with this issue, we use the standard-error adjustment proposed by Adão et al. (2019). Finally, note that in addition to the $N_{c,t}$, we might have country-level shocks $O_{c,t}$ that are correlated with the $S_{c,t}$ and also affect firm innovation (e.g. fuel prices, R&D subsidies etc.). In that case, firm-level shocks would have the following structure:

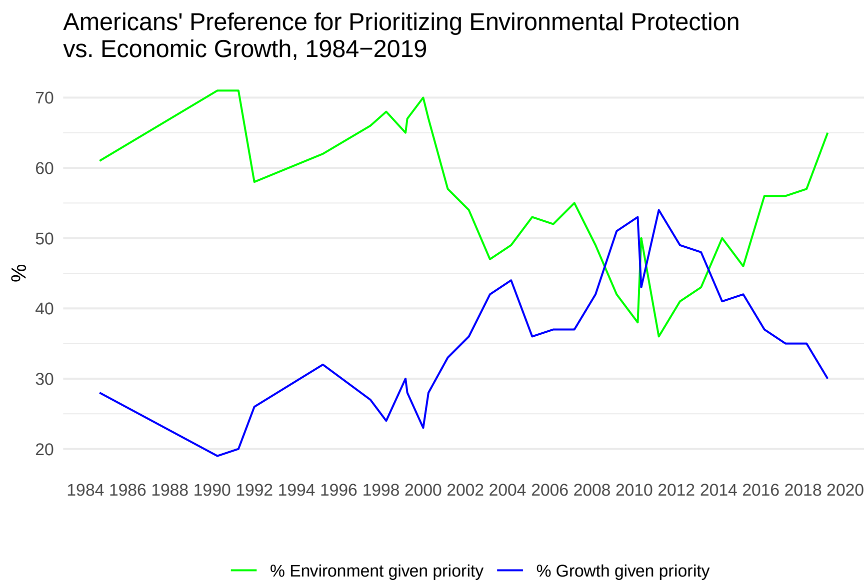
$$\varepsilon_{j,t} = \kappa J_{j,t} + \gamma N_{j,t} + \beta O_{j,t} + c_{c(j),t} + \nu_{j,t},$$

where $O_{j,t} = \sum_c w_{j,c} O_{c,t}$. To still obtain an unbiased estimate for α, we then need to include $O_{j,t}$ in the regression equation

$$\Delta I_j = \alpha\Delta S_{j,t} + \beta O_{j,t} + \alpha_{c(j)} + \Delta\chi_{j,t}.$$

Appendix Figures

Figure 3: Long run decline and recent reversal in pro-environmental concerns



Notes: Based on responses to the question "With which one of these statements about the environment and the economy do you most agree - protection of the environment should be given priority, even at the risk of curbing economic growth (or) economic growth should be given priority, even if the environment suffers to some extent" Source: Gallup (Saad, 2019)

Figure 4: Firm-level Markups

(a) Distribution over time (b) Change between the 2 periods

Notes: Panel (a) shows centiles (10th to 90th percentile) of firm-level markups (inverse of the Lerner index) over time. Panel (b) shows the distribution of changes in markups between 2002 and 2012. These markups are computed using ORBIS data.