Carbon taxes, Path Dependence and Directed Technical Change : Evidence from the Auto Industry*

Philippe Aghion[†], Antoine Dechezlepretre[‡]

David Hemous[§], Ralf Martin[¶] and John Van Reenen[‖]

August 30, 2012

Abstract

Can technical change be directed to combat climate change? We construct new firm-level panel data on auto industry innovation distinguishing between "dirty" (internal combustion engine) and "clean" (e.g. electric and hybrid) patents across 80 countries and 40 years. We show that firms tend to innovate relatively more in clean technologies when they face higher tax-inclusive fuel prices. Furthermore, there is path dependence in the type of innovation both from aggregate technology spillovers and from the firm's own innovation history. Using our model we simulate the increases in carbon taxes needed to allow clean to overtake dirty technologies.

JEL Classification: O3, O13, L62

Keywords: Climate Change, Innovation, Directed Technical Change, automobiles

---

*We would like to thank Daron Acemoglu, Michael Greenstone, Rebecca Henderson, Ken Judd, Joshua Gans, Ed Lazear and Nick Stern for comments. Participants at seminars in the AEA, Arizona, Cambridge, CIAR, Manchester, Paris, Stanford and Venice have all improved the paper. Financial support has come from the ESRC Centre for Economic Performance.

[†]Harvard University

[‡]Grantham Research Institute on Climate Change and the Environment and Centre for Economic Performance, London School of Economics

[§]Harvard University

[¶]Imperial College Business School and Centre for Economic Performance, London School of Economics

[‖]Centre for Economic Performance, London School of Economics and NBER

1 Introduction

With increasing concern over climate change there is much interest in how new technologies can help reduce greenhouse gas emissions like Carbon Dioxide (CO_2). Most climate change models assume exogenous technological change (e.g. Stern, 2006), but dealing with the challenge of global warming almost certainly requires more climate change related innovation (e.g. Henderson and Newell, 2011). But what policies can be used to achieve this?

Recent models of climate change with endogenous technical progress suggest that the market will generate insufficient innovation to reduce climate change and too much R&D investment directed at "dirty" technologies. For example, in Acemoglu et al. (2012) there is path-dependence in the *direction* of technical change. Firms in economies that have innovated a lot in dirty technologies in the past will find it more profitable to innovate in dirty technologies in the future. This path dependency feature when combined with the environmental externality (whereby firms do not factor in the loss in aggregate productivity or consumer utility induced by environmental degradation) will induce a laissez-faire economy to produce and innovate too much in dirty technologies compared to the social optimum. This in turn calls for government intervention to "redirect" technical change. For redirected technical change to be effective, the associated clean products need to be perceived as substitutes for dirty products by consumers. Cars are a good example of this as electric and hybrid vehicles are potential substitutes for vehicles based on internal combustion engines (another example would be renewable energy plants vs. fossil fuel plants).[1] Furthermore, transport accounts for around a quarter of worldwide CO_2 emissions, making it the second biggest greenhouse gas

[1] Innovations in fuel efficiency will also help combat climate change, but these will be offset if consumers respond by driving more. There are limits to how far this incremental innovation on an existing dirty technology can effectively reduce emissions. Our focus here is on more radical changes to vehicles depending on alternative cleaner fuel types.

emitting sector after energy production.[2] Consequently, the automobile sector is the focus of study in our paper. Note that even with such substitutable products as autos, however, there may be decreasing returns to dirty innovation and so the market would do part of the job of redirecting technical change towards clean technologies. Thus empirically evaluating the effects of past stocks of innovation on future trajectories has to be part of the aggregate dynamic evaluation of different policies.

In this paper, we construct new panel data on auto innovations to examine how firms redirect technical change in response to carbon prices in the context of path dependent innovation. Our main data are drawn from the European Patent Office's World Patent Statistical database (PATSTAT). These data cover close to the population of all worldwide patents (and their citations) since the beginning of the 20th century. To mitigate the well-known problem that many patents are of very low value, our innovation outcome measure focuses on "triadic" patents which are those that have been taken out in all three of the world's major patents offices: the European Patent Office (EPO), the Japanese Patent Office (JPO) and the United States Patents and Trademark Office (USPTO).[3] Our results are robust to alternative ways to controlling for heterogeneous patent values such as using wider categories of patents like biadics (US and EPO) or weighting by future citations. In automobiles since 1978[4] around 6,500 triadic patents in "clean" technologies (electric vehicles, hybrid vehicles, fuel cells for hydrogen vehicles, etc.) were filed versus about 18,500 triadic patents in "dirty"technologies which affect regular combustion engines.[5] Moreover, our database reports

[2] See http://www.wri.org/project/cait

[3] Triadic patents have been used extensively as a way to focus on high-value patents (Grupp et al., 1996; Grupp, 1998; Dernis, Guellec and van Pottelsberghe, 2001; Dernis and Khan, 2004; Guellec and van Pottelsberghe, 2004).

[4] Since the EPO was created in 1978 the triadic patent data only starts in that year.

[5] Overall, since the beginning of the 20th century, about 213,000 patents in "clean" technologies were filed worldwide, versus about 760,000 patents in "dirty".

the name of patent applicants which in turn allows us to match clean and dirty patents with distinct patent holders each of whom has her own history of clean versus dirty patenting.[6]

Our main results can be summarized as follows. First, higher (tax-inclusive) fuel prices induce firms to redirect technical change towards clean innovation and away from dirty innovation. Second, a firm's propensity to innovate in "clean" technologies appears to be stimulated by its exposure to past clean patents (and vice versa for dirty technologies) consistent with the path-dependency hypothesis. We measure this exposure by (i) an aggregate spillover index based on the (pre-sample) location of the firm's inventors across different countries combined with the changes in the country-specific stocks of clean and dirty auto patents since 1965; and (ii) the firm's own lagged stocks of clean and dirty innovations. In general econometric models including the carbon prices and both path dependency variables we find evidence for both firm-specific and aggregate path dependency.

Our paper relates to two main strands of literature. First, the literature on climate change, starting with Nordhaus (1994) who developed a dynamic model of climate change (the DICE model), which amounts to adding equations linking production to emissions into a Ramsey model. Subsequent contributions to this literature have looked at the implications of risk and discounting for the optimal design of environmental policy[7] or have looked at the choice between taxes and quotas, building on Weitzman (1974).

Second, our paper relates to the literature on directed technical change, in particular Acemoglu (1998, 2002; 2008) which itself was inspired by early contributions by Hicks (1932) and Habakkuk

[6] We do not consider radical innovations in upstream industries such as biofuels, for instance. To explore this is beyond the scope of the current paper which takes the more positive approach of exploring the determinants of clean innovation in vehicles.

[7] In particular see Stern (2006), Weitzman (2007, 2009), Dasgupta (2007, 2008), Nordhaus (2007), von Below and Persson (2008), Mendelsohn (2007), and Tol and Yohe (2006). More recent work by Golosov et al. (2009) characterizes the optimal policy in a model with exhaustible privately-owned resources.

(1962).[8] Also closely related to our paper is the empirical literature linking environmental policy and directed technical change. In particular, Popp (2002) uses aggregate U.S. patent data from 1970 to 1994 to study the effect of energy prices on energy-efficient innovations. He constructs two different measures of the knowledge stocks for the innovation regressions: (a) simple stock of previously U.S. granted patents and (b) future citation weighted stock of patents. In particular, he finds a significant impact from both, energy prices and the quality of the stock of knowledge available to the inventor, on directed innovations. This in turn reflects directed technical change as a response to change in energy prices. However since Popp uses aggregate data a concern is that his regressions also capture macro-economic shocks correlated with both innovation and the energy price. In this paper, we use international firm-level data which allows us to exploit differences in the extent to which firms in different countries are affected by policy-induced shocks to the energy price (e.g. fuel taxes). We can, for example, control for global macro shocks with a full set of time dummies. Further evidence of directed technical change applied to the context of saving energy can be found in Newell, Jaffe and Stavins (1999) which focuses on the air-conditioning industry, or in Crabb and Johnson (2010) who look at energy-efficient automotive technology. However, neither of these papers use multi-country data nor analyze whether there is path dependency in the direction of technical change. Hassler, Krussell and Olovsson (2011) find evidence for a trend increase in energy saving technologies following high oil prices. More ambitiously, papers such as Acemoglu, Akcigit, Hanley and Kerr (2012) and Golosov, Hassler, Krussell and Tsyvinski (2011) put these into

[8] The theoretical literature on directed technical change is well developed. See for example Messner (1997), Grubler and Messner, (1998), Goulder and Schneider (1999), Manne and Richels (2002), Nordhaus (2002), Van der Zwaan et al. (2002); Buonanno et al (2003), Nordhaus (2002), Sue Wing (2003) and Gerlagh (2008). in contrast, directed technical change has rarely been empirically tested. However, see Acemoglu and Linn (2004) and more recently Hanlon (2012) who shows how relative price changes induced by the Union blockade during the American Civil War induced British textile innovation in machines using Indian cotton.

a full blown quantitative dynamic general equilibrium model to examine optimal carbon taxes.

The paper is organized as follows. Section **??** develops a simple model to guide our empirical analysis. Section **??** presents the econometric methodology. The data is presented in Section **??** and some descriptives in Section **??**. Section **??** presents the results, discusses their robustness and some extensions. In Section **??** we explore the implications of our results for the evolution of future clean and dirty knowledge stocks and how this evolution would be affected by changes in the carbon price. Section **??** concludes.

2 Model

In this section we present a simple model to guide our empirical analysis. This model rationalizes path dependence in firms' own knowledge stock as well as the impact of a change in the price of fuel on clean and dirty innovation. We then show how one can add knowledge spillovers and energy efficiency innovations ("grey innovations") to our framework.

2.1 Basic framework

We consider a one-period model of an economy where consumers derive utility from an outside good and from motor vehicle services. Utility is quasi-linear with respect to the outside good C_0 (chosen as the numeraire) and β is the elasticity of consumption of motor vehicle services with respect to its index price. To consume motor vehicle services, consumers need to buy cars and fuel (call it a "dirty car bundle ") or cars and electricity (call it a "clean car bundle "). Utility is then given by.

$$U = C_0 + \frac{\beta}{\beta - 1}\left(\left(\int_0^1 \min\left(y_{ci}, \xi_c e_i\right)^{\frac{\sigma-1}{\sigma}} di\right)^{\frac{\sigma}{\sigma-1}\frac{\varepsilon-1}{\varepsilon}} + \left(\int_0^1 \min\left(y_{di}, \xi_d g_i\right)^{\frac{\sigma-1}{\sigma}} di\right)^{\frac{\sigma}{\sigma-1}\frac{\varepsilon-1}{\varepsilon}}\right)^{\frac{\varepsilon}{\varepsilon-1}\frac{\beta-1}{\beta}},$$

where e_i is the amount of electricity consumed for variety i of clean car, g_i is the amount of fuel consumed for variety i of dirty cars, ε is the elasticity of substitution between the clean and dirty cars, σ is the elasticity of substitution among varieties within each type of car and ξ_c (respectively ξ_d) is the energy efficiency of clean (respectively dirty) cars. We assume that $1 < \varepsilon \leq \sigma$, so that clean cars are more substitutable with each other than with dirty cars, and that $\varepsilon > \beta$: the elasticity of substitution between clean and dirty cars is larger than the price elasticity for cars as a whole. We denote by f_c the price of electricity and f_d the price of fuel. In the first part of the analysis innovation will be cost saving for producers, but later in this section we investigate energy saving innovation.

Cars are produced by local monopolists, and each monopolistic firm i produces one clean and one dirty variety with current productivities A_{zi}, $z = c, d$ in the respective sectors. That is, it takes $\frac{1}{A_{ci}}$ units of the outside good as input for firm i to produce one unit of clean car; similarly it takes $\frac{1}{A_{di}}$ units of the outside good as input for firm i to produce one unit of dirty car.

To complete the description of the model, we need to specify the innovation technology. Here, we assume that at the beginning of a period, by incurring total cost $\frac{1}{2}\psi x_{zi}^2$ entrepreneurs can increase their productivities in clean cars (A_{ci}) and dirty cars (A_{di}) according to:

$$A_{zi} = (1 + x_{zi}) A_{zi0} \text{ for } z \in \{c, d\}.$$

2.2 Solving the model

Define the price indexes for dirty and clean car bundles as:

$$P_z = \left(\int_0^1 \left(p_{zi} + \frac{f_z}{\xi_z} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}}, \text{ for } z \in \{c, d\}.$$

Inverse demand curves for clean and dirty cars are then given by

$$y_{zi} = \left(p_{zi} + \frac{f_z}{\xi_z}\right)^{-\sigma} P_z^{\sigma-\varepsilon}\left(P_c^{1-\varepsilon} + P_d^{1-\varepsilon}\right)^{\frac{\varepsilon-\beta}{1-\varepsilon}} \text{ for } z \in \{c, d\}. \tag{1}$$

Firm i's maximization problem can be expressed as:

$$\max_{y_{ci}, y_{di}} \pi_i = p_{ci} y_{ci} - \frac{1}{A_{ci}} y_{ci} + p_{di} y_{di} - \frac{1}{A_{di}} y_{di}.$$

where y_{zi} for $z \in \{c, d\}$ are given by (??).

This yields the following expressions for the equilibrium profits on clean and dirty cars sales by firm i:

$$\pi_{zi} = \frac{(\sigma-1)^{\sigma-1}}{\sigma^\sigma}\left(\frac{1}{A_{zi}} + \frac{f_z}{\xi_z}\right)^{1-\sigma} P_z^{\sigma-\varepsilon}\left(P_c^{1-\varepsilon} + P_d^{1-\varepsilon}\right)^{\frac{\varepsilon-\beta}{1-\varepsilon}} \text{ for } z \in \{c, d\}.$$

2.3 Firms' innovation efforts

We assume that ψ is sufficiently large that the equilibrium innovation intensities x_{ci} and x_{di} are uniquely defined by the first order conditions:.

$$x_{zi} = \frac{1}{\psi}\frac{\partial \pi_{ci}}{\partial x_{ci}} = \frac{(\sigma-1)^\sigma}{\psi\sigma^\sigma}\frac{1}{A_{zi0}(1+x_{zi})^2}\left(\frac{1}{A_{zi0}(1+x_{zi})} + \frac{f_z}{\xi_z}\right)^{-\sigma} P_z^{\sigma-\varepsilon}\left(P_c^{1-\varepsilon} + P_d^{1-\varepsilon}\right)^{\frac{\varepsilon-\beta}{1-\varepsilon}} \text{ for } z \in \{c, d\}. \tag{2}$$

Path dependence on firm's own history. The equilibrium innovation intensities x_{zi} increase in the firm's corresponding technology stocks A_{zi0} as long as the right hand side of (??) increases in A_{zi0}, which in turn is satisfied if the elasticity of substitution σ is sufficiently large.[9]**.** In this case, there is path dependence in clean and dirty innovation.[10]

[9]The precise condition is $\frac{(\sigma-1)}{A_{zi0}(1+z_{ci})} > \frac{f_z}{\xi_z}$, which is also satisfied if the price of fuel represents a sufficiently small share of the total costs of a car.

[10]Alternatively, one could have assumed that the cost function takes the following form $\Psi(x_c, x_d) = \frac{1}{2}\left(\psi x_c^2 + \psi x_d^2 + \chi(x_c + x_d)^2\right)$, for $\chi \neq 0$, there would be strategic substitutability in innovation efforts on clean versus dirty technologies within firms. A large stock of dirty knowledge would discourage clean innovation. However, we did not find empirical support for this hypothesis.

Redirecting innovation through changes in the fuel price. We now investigate the impact of a change in the fuel price on clean and dirty innovation. Totally differentiating (??) for $z = c$ with respect to the fuel price, and then using the notation $\widehat{X} = \frac{dX}{X}$,we get:

$$(1-\omega)\,\widehat{x_{ci}} = \left(\sigma - \varepsilon + (\varepsilon - \beta)\,\frac{P_c^{1-\varepsilon}}{P_c^{1-\varepsilon} + P_d^{1-\varepsilon}}\right)\widehat{P}_c + (\varepsilon - \beta)\,\frac{P_d^{1-\varepsilon}}{P_c^{1-\varepsilon} + P_d^{1-\varepsilon}}\widehat{P_d}, \tag{3}$$

where $\omega \equiv \frac{d}{dx_{zi}} \ln\left(\frac{1}{(1+x_{zi})^2}\left(\frac{1}{A_{ci0}(1+x_{ci})} + \frac{f_z}{\xi_z}\right)^{-\sigma}\right) < 1$.[11]

For sufficiently small innovation intensities, one can neglect the indirect impact of an increase in fuel price via the innovation response of other firms, so that $\widehat{P}_c \approx 0$, and $\widehat{P_d}$ has the sign of $\widehat{f}_d$.[12] Then:

$$(1-\omega)\,\widehat{x_{ci}} \approx (\varepsilon - \beta)\,\frac{P_d^{1-\varepsilon}}{P_c^{1-\varepsilon} + P_d^{1-\varepsilon}}\widehat{P_d}.$$

This in turn implies that the equilibrium intensity of clean innovation increases with fuel price since we assumed $\varepsilon \geq \beta$.

Similarly, we get:

$$(1-\omega)\,\widehat{x_{di}} \approx \sigma\left(\widehat{P_d} - \frac{\frac{f_d}{\xi_d}}{\frac{1}{A_{di}} + \frac{f_d}{\xi_d}}\widehat{f}_d\right) - \left(\varepsilon\frac{P_c^{1-\varepsilon}}{P_c^{1-\varepsilon} + P_d^{1-\varepsilon}} + \beta\frac{P_d^{1-\varepsilon}}{P_c^{1-\varepsilon} + P_d^{1-\varepsilon}}\right)\widehat{P_d} \tag{4}$$

once we neglect the indirect impact of a change in fuel price on the price indexes working through the innovation response.

The second term has the opposite sign from that of $\widehat{f}_d$: namely, an increase in fuel price reduces the benefit of dirty innovation both because it induces substitution towards clean cars and because it reduces the overall consumption of cars.

[11] That ω be less than 1 follows from the fact that at the equilibrium the left-hand side of (??) crosses the right-hand side from below.

[12] Another reason to neglect this indirect impact is that firms typically operate in several markets, with different exposures to each market for each firm. Therefore the allocation of innovation of the competitors does not depend only on the fuel price in a given country but also on the fuel price in other countries.

The first term would be zero if all firms had the same dirty technologies, otherwise it has the sign of $\widehat{f}_d$ for the least productive dirty firms and the opposite sign otherwise. This term captures a reallocation effect of an increase in fuel price from most to least productive dirty firms. But, for "average firms" for which $\frac{\frac{f_d}{\xi_d}}{\frac{1}{A_{di}}+\frac{f_d}{\xi_d}}\widehat{f}_d \approx \widehat{P_d}$, dirty innovation decreases with the fuel price.

Overall, our analysis suggests that an increase in fuel price should decrease dirty innovation intensity (as long as the firm's initial productivity in dirty technologies is not too low compared to that of other firms) and it should increase clean innovation intensity.

2.4 Extensions

Knowledge spillovers. In the empirical part of the paper we investigate not only the effects of firms' own past knowledge but also the effects of aggregate knowledge spillovers across firms in the country where innovation occurs. To introduce the possibility of such aggregate spillovers in our model, one can simply assume that each producer i benefits from positive knowledge spillovers from a set Ω_i of varieties with average technologies $\overline{A_{ci}}$ and $\overline{A_{di}}$, for example if the R&D cost function is of the form $\frac{1}{2}\left(\psi\left(\overline{A_{ci}}\right)x_{ci}^2+\psi\left(\overline{A_{di}}\right)x_{di}^2\right)$, where ψ is a decreasing function of its argument. This extension directly yields that clean innovation increases with the aggregate stock of clean knowledge and dirty innovation with the aggregate stock of dirty knowledge.

"Grey" innovation. Innovations in the car industry may also involve improvements in energy efficiency. Innovations that increase energy efficiency for clean cars would not react differently to an increase in fuel price than cost reducing innovations

Now, let us analyze innovations that improve dirty energy efficiency. More specifically, suppose that at the beginning of every period, a firm can increase its energy efficiency from ξ_{di0} to $\xi_{di} = (1+x_{\xi i})\,\xi_{di0}$ if it spends $\frac{1}{2}\psi x_{\xi i}^2$ units of the outside good in R&D. We refer to this type of innovations as “grey” innovations. The equilibrium intensity in grey innovation is given by

$$x_{\xi i} = \frac{(\sigma-1)^{\sigma}}{\psi\sigma^{\sigma}} \frac{f_d}{\xi_{di0}\left(1+x_{\xi i}\right)^2} \left(\frac{1}{A_{di0}\left(1+x_{ci}\right)} + \frac{f_d}{\xi_{di0}\left(1+x_{\xi i}\right)}\right)^{-\sigma} P_d^{\sigma-\varepsilon}\left(P_c^{1-\varepsilon}+P_d^{1-\varepsilon}\right)^{\frac{\varepsilon-\beta}{1-\varepsilon}}.$$

For sufficiently high ψ, i.e for small innovation intensity, totally differentiating with respect to the fuel price leads to:

$$(1-\omega)\widehat{x_{\xi i}} \approx \sigma\left(\widehat{P_d} - \frac{\frac{f_d}{\xi_{di}}}{\frac{1}{A_{di}}+\frac{f_d}{\xi_{di}}}\widehat{f}_d\right) - \left(\varepsilon\frac{P_c^{1-\varepsilon}}{P_c^{1-\varepsilon}+P_d^{1-\varepsilon}} + \beta\frac{P_d^{1-\varepsilon}}{P_c^{1-\varepsilon}+P_d^{1-\varepsilon}}\right)\widehat{P_d} + \widehat{f}_d.$$

This expression is similar to (**??**) except for the last term which captures a direct positive effect of an increase in the fuel price on energy efficiency innovation. The overall impact of an increase in the fuel price on grey innovation is therefore theoretically ambiguous. In the case where all dirty firms have similar productivities (both in grey and in dirty technologies), and where the clean bundle is much more competitive than the dirty one this expression simplifies into $(1-\omega)\widehat{x_{\xi i}} \approx \left(1-\beta\frac{\frac{f_d}{\xi_d}}{\frac{1}{A_d}+\frac{f_d}{\xi_d}}\right)\widehat{f}_d$, so that the impact of an increase in fuel price on grey innovation is negative when the price of fuel is a large share of the total cost of a dirty car bundle and the price elasticity of the dirty car bundle is large. What happens then is that as the fuel price increases, the negative impact on innovation of the reduction in the demand for dirty cars is larger than the positive impact of a bigger marginal benefit for grey innovations.

We use the term ”grey innovations” as the impact of these innovations on the environment is also ambiguous: on the one hand these innovations increase energy efficiency and therefore reduce the amount of fuel consumption per car; on the other hand these innovations make fossil fuel cars cheaper, thereby increasing total consumption of these cars. Formally, one gets

$$g_i = \frac{y_{di}}{\xi_{di}} = \frac{\sigma}{\sigma-1}\frac{1}{\xi_{di}}\left(\frac{1}{A_{di}}+\frac{f_d}{\xi_{di}}\right)^{-\sigma} P_d^{\sigma-\varepsilon}\left(P_c^{1-\varepsilon}+P_d^{1-\varepsilon}\right)^{\frac{\varepsilon-\beta}{1-\varepsilon}},$$

so that

$$\frac{dg_i}{d\xi_{di}} = \left((\sigma - 1) \frac{f_d}{\xi_{di}} - \frac{1}{A_{di}} \right) \frac{1}{\xi_{di}^2} \left(\frac{1}{A_{di}} + \frac{f_d}{\xi_{di}} \right)^{-\sigma-1} \frac{\sigma}{\sigma - 1} P_d^{\sigma-\varepsilon} \left(P_c^{1-\varepsilon} + P_d^{1-\varepsilon} \right)^{\frac{\varepsilon-\beta}{1-\varepsilon}},$$

which is ambiguously signed. The expression is negative if the price of fuel is sufficiently low relative to other costs, but it is positive if the elasticity of substitution across cars is sufficiently large.

Empirically, our definition of dirty innovation is a mix of these grey innovations that improve energy efficiency for internal combustion engines and purely dirty innovations. We will therefore show the robustness of the results to the exclusion and inclusion of grey innovation.

2.5 Summarizing our main predictions

We have presented a model of clean versus dirty innovation, in which we predict:

1. producers have a higher propensity to innovate in clean than dirty the larger the fuel price, f_d.

2. for σ sufficiently high, producers have a higher propensity to innovate in clean technologies the higher A_c^0 and the lower A_d^0, i.e. the higher their initial stock of clean versus dirty technologies

3. producers have a higher propensity to innovate in clean (respectively, dirty) technologies the higher the aggregate level of clean (respectively, dirty) technology in neighboring varieties or in the aggregate economy.

We now confront these three predictions with our panel data on clean and dirty innovation in the auto industry.

3 Econometrics

3.1 General approach

Consider the following Poisson (or more generally any log-link) specification for the determination of firm innovation in "clean" technologies[13], as measured by patents:

$$PAT_{CLEAN,it} = \exp(\beta_{C,P} \ln FP_{it} + A_{C,it}) + e_{C,it} \quad (5)$$

where $PAT_{Clean,it}$ is the number of patents applied for in clean technologies by firm i in year t; A_{Cit} is the firm's knowledge stock relevant for clean innovation at the beginning of year t, which depends both upon its own stocks of past clean and dirty innovation and upon potential aggregate spillovers from other firms (discussed below), $e_{C,it}$ is an error term and FP_{it} is our primary price policy variable. Since no country has established meaningful carbon pricing yet, we use tax-inclusive fuel prices for the various countries in our sample, exploiting cross-country variations in taxation and differential firm-specific market exposure. We also test the robustness of our results to explicitly using fuel taxes as the policy variable. Because different firms operate differentially in geographical markets (for example GM has some "home bias" towards the US market whereas Toyota has a "home bias" towards the Japanese market) they are differently exposed to tax changes in different countries. This home bias could be due to product differentiation and heterogeneous tastes or it may result from government policies to promote domestic firms (e.g. trade barriers). To take this heterogeneity into account we construct a firm-specific policy variable. More specifically, we use the firm's history of patent filing across countries to assess the relative importance of the various markets the firm is operating in and then construct firm-specific weights on fuel prices - as well as

[13] In our regressions we use an equivalent equation for dirty technologies. We initially discuss only one of these equations to simplify the notation.

other policy and control variables - from the corresponding markets. We discuss this in more detail in the Data Section, but the basic idea is that firms will seek greater IP protection in those markets that they deem will seem more important in the future.[14]

We consider several measures of a firm's knowledge stock, as our theory implies that the history and exposure of the firm to different types of innovation will matter. In particular we parameterize the knowledge stock as:

$$A_{Cit} = \beta_{C,1} \ln SPILL_{C,it} + \beta_{C,2} \ln SPILL_{D,it} + \beta_{D,3} \ln K_{C,it} + \beta_{D,4} \ln K_{D,it} \tag{6}$$

Since we expect that firms "stand on the shoulders of giants", their knowledge stock will depend on knowledge spillovers from other firms which in turn could be both in clean technologies ($SPILL_{C,it}$) and in dirty technologies ($SPILL_{D,it}$). We construct the spillover pools in different ways, but our baseline measures use the past stocks of total clean or dirty patents since 1978.

Drawing on the evidence that knowledge has a local component (e.g. Jaffe, Trajtenberg, and Henderson, 1993) we use the firm's pre-sample distribution of inventors across countries to weight the country spillover stocks. In other words, if the firm has many inventors in the US (regardless of whether the HQ of the firm is in Tokyo or Detroit) then the knowledge stock in the US is given a higher weight (see Data Section). The firm's knowledge may also depend on its own past history of innovation and we denote this as $K_{C,it}$ (past own stock of clean innovation) and $K_{D,it}$ (past own stock of dirty innovation).[15]

[14] The measure is far from perfect of course as the costs and benefits of IP vary between countries (e.g. China is a fast growing market but patents are weakly protected. Since these factors are reasonably common across firms, however, how filing rates differ across countries at the firm level should still reveal something about beliefs over relative importance.

[15] We construct stocks using the perpetual inventory method, but then show robustness to using non-parametric distributions of patent flows and to considering alternative assumptions over knowledge depreciation rates. See Data Section.

There are of course other factors that will influence innovation in addition to the carbon price and the past history of innovation. These include policy measures such as R&D incentives for clean innovation and controls over emissions; the size and wealth of the country (proxied by GDP and GDP per capita) which we denote by the vector $w_{C,it}$ and other unobservable factors that we control for through introducing a firm fixed effect ($\eta_{C,i}$), a full set of time dummies (T_t) and an error term that is uncorrelated with the right hand side variables ($u_{C,it}$). Adding these extra terms and substituting equation (**??**) into (**??**) gives us our main estimating equation for clean innovation:

$$\begin{aligned} PAT_{CLEAN,it} = \exp(&\beta_{C,P} \ln FP_{it} + \beta_{C,1} \ln SPILL_{C,it} + \beta_{C,2} \ln SPILL_{D,it} \\ &+ \beta_{C,3} \ln K_{C,it} + \beta_{C,4} \ln K_{D,it} \\ &+ \beta_{C,w} w_{it} + \ln \eta_{C,i} + T_{C,t}) + u_{C,it} \end{aligned} \tag{7}$$

Similarly, we can derive an estimating equation for dirty innovation:

$$\begin{aligned} PAT_{DIRTY,it} = \exp(&\beta_{D,P} \ln FP_{it} + \beta_{D,1} \ln SPILL_{C,it} + \beta_{D,2} \ln SPILL_{D,it} \\ &+ \beta_{D,3} \ln K_{C,it} + \beta_{D,4} \ln K_{D,it} \\ &+ \beta_{D,w} w_{it} + \ln \eta_{D,i} + T_{D,t}) + u_{D,it} \end{aligned} \tag{8}$$

The theory yields predictions on the coefficients in these two equations. If the carbon tax induces more clean than dirty innovation then the marginal effect of the fuel price must be larger on clean innovation than on dirty innovation: $\beta_{C,P} > \beta_{D,P}$. Typically we would expect that $\beta_{C,P} > 0$ and $\beta_{D,P} < 0$, however this is not a necessary condition for redirecting technical change relatively towards clean (e.g. the fuel price could depress clean innovation, but depress dirty innovation by far more - cf Gans, 2011).

Next, for there to be path dependence in the direction of innovation it must be that (ceteris paribus) firms that are exposed to stronger dirty spillovers become more prone to conduct dirty innovation rather than clean innovation: $\beta_{D,2} > 0$ and $\beta_{D,2} > \beta_{C,2}$. And path dependence may involve similar effects working through a firm's own accumulated knowledge: $\beta_{D,4} > 0$ and $\beta_{D,4} > \beta_{C,4}$. Also, we would expect that the effect of dirty spillovers and dirty knowledge stocks on dirty innovation be larger than the effects of clean spillovers and clean knowledge stocks: $\beta_{D,2} > \beta_{D,1}$ and $\beta_{D,4} > \beta_{D,3}$. The reverse predictions should apply for the clean equation: $\beta_{C,2} < \beta_{C,1}$ and $\beta_{C,4} > \beta_{D,1}$

3.2 Dynamic count data models with fixed effects

Our main regression equations are (**??**) and (**??**) which we can re-write succinctly as

$$PAT_{zit} = \exp\left(x_{it}\beta_z\right)\eta_{zi} + u_{zit} \tag{9}$$

where $z \in \{DIRTY, CLEAN\}$, x_{it} is the vector of regressors and $u_{zit} = \exp(v_{zit})$.

In the results section we present a number of econometric specifications that seek to control for firm fixed effects η_{zi} in these dynamic non-linear count data models. Further details are provided in Appendix A (see also Blundell, Griffith and Windmeijer, 2002 and Cameron and Trivedi, 2005). The traditional estimator is the the Hausman, Hall and Griliches (1984, HHG) approach which is analogous to the standard within groups method in linear panel data (least squares dummy variables). The disadvantage of HHG is that it requires strict exogeneity just like the within groups estimator and this assumption will be violated by the lagged dependent variables in our path dependent models. Quasi-differenced GMM procedures such as Chamberlain (1992) that do not require strict exogeneity tend to have poor finite sample properties due to the problem of weak instruments.

In response to these problems, Blundell, Griffith and Van Reenen (1999, 1995, BGVR) propose a "mean scaling" estimator using information on the long history of patents in the pre-estimating sample period to control for unobserved heterogeneity. Essentially this is a control function approach where we use a proxy (functions of the pre-sample average levels of firm innovation) to control for the bias associated with unobserved fixed effects.

A potential problem with BGVR in our context is that the number of clean innovations only takes off in the mid 1980s (see Figure 4) so the pre-sample estimator may not do a good job at controlling for the fixed effect in the clean equations. Consequently, we develop a third econometric model which we label CFX, the Conditional Fixed Effects estimator. Rather than using information from the pre-sample period to calibrate the control function like BGVR, we use *future* data. Using the analogy with linear dynamic panel data whereas BGVR uses long backward differencing, CFX uses forward orthogonal deviations (see Arellano and Honore, 2005; Arellano, 2003). Essentially we estimate the main regression equation simultaneously with a second equation for the control function. Full details are provided in Appendix A.

Below, we provide results using CFX, BGVR and HHG. We also contrast these results to simple log-linear OLS models where we add an arbitrary constant (unity) to the patents variables in order to take logarithms. Although all these methods deliver similar qualitative results we focus on the CFX as our baseline specification as it should be more robust to the econometric problems discussed above.

4 Data

4.1 Main dataset

Our main data are drawn from the World Patent Statistical Database (PATSTAT) maintained by the European Patent Office.[16] Patent documents are categorized using the international patent

[16]PATSTAT is available from the EPO at http://www.epo.org/searching/subscription/raw/product-14-24.html

classification (IPC) and national classification systems. We have extracted all the patents pertaining to "clean" and "dirty" technologies in the automotive industry. "Dirty" includes patents related to the internal combustion engine. "Clean" includes patents specifically related to clean car technologies such as electric, hybrid and hydrogen vehicles. Our selection of relevant IPC codes for clean technologies follows the same procedures as the OECD.[17] The precise description of the IPC codes used to identify relevant patents can be found in Table 1.

To measure innovation, we use counts of patent applications. The advantages and limitations of patenting as a measure of innovation have been discussed at length in the literature.[18] For our purposes, the main advantage of using patent data is that they are available at a highly disaggregated level. In particular, we can distinguish innovations in the auto industry according to specific technologies, such as control systems specially designed for hybrid vehicles. R&D investment cannot be disaggregated by type of innovation in this way. Further, R&D is not reported for many small and medium sized firms, especially in Europe and in the US privately listed firms are also exempt from the accounting requirement to report R&D. Moreover, autos are a large, R&D intensive sector where patents are perceived as a reasonably efficient means of protection against imitation (Cohen et al., 2000).[19] These considerations make patents a reasonably good indicator of innovative activity in the sector.

However, patent-based indicators also suffer from a number of limitations. The first is that patents are not the only way to protect innovations, although a large fraction of the most economi-

[17]See OECD (2011) www.oecd.org/environment/innovation, Vollebergh (2010) and Hačič et al (2008).

[18]See Griliches (1990) and for a recent overview, OECD (2009).

[19]Cohen et al. (2000) conducted a survey questionnaire administered to 1478 R&D labs in the U.S. manufacturing sector. They rank sectors according to how effective patents are considered as a means of protection against imitation, and find that the top 3 industries according to this criterion are medical equipment and drugs, special purpose machinery and automobile.

cally significant innovations appear to have been patented (Dernis et al., 2001). Another problem is that patent values are highly heterogeneous with most patents having a very low valuation. Finally, the number of patents that are granted for a given innovation varies significantly across patent offices with concerns over increasing laxity in recent years particularly in the US (e.g. Jaffe and Lerner, 2004).

To mitigate these problems, we focus on "triadic" patents[20] which are those patents that have been taken out in all three of the world's major patents offices: the European Patent Office (EPO), the Japanese Patent Office (JPO) and the United States Patents and Trademark Office (USPTO).[21] Focusing on triadic patents has a number of advantages. First, triadic patents provide us with a common measure of innovation worldwide, which is robust to administrative idiosyncracities of the various patent offices. For example, if the same invention is covered by one patent in the US and by two patents in Japan, all of which are part of the same triadic patent family, we will count it as one single invention. Secondly, triadic patents cover only the most valuable inventions,[22] which explains why they have been used so extensively to capture high-quality patents (Grupp et al., 1996; Grupp, 1998; Dernis, Guellec and van Pottelsberghe, 2001; Dernis and Khan, 2004; Guellec and van Pottelsberghe, 2004). Thirdly, triadic patents typically protect inventions that have a potential worldwide application. These patents are thus relatively independent of the countries in which they are filed. This is important for us as we will regress innovation activity on a weighted average of

[20] To identify triadic patents we use the INPADOC dataset in PATSTAT. For details on the construction of patent families see Martinez, 2010.

[21] In fact, following standard practice we use all patents filed at the EPO and JPO and granted by the USPTO. This is because the USPTO only revealed granted patents until 2001, when they changed policy. For consistency we thus consider only triadic patents granted by the USPTO both before and after 2001. For the official definition of triadic patents and how triadic patent families are constructed, see Dernis and Kahn, 2004, and Martinez, 2010.

[22] It has been empirically demonstrated that the number of countries in which a patent is filed is correlated with other indicators of patent value (see, for example, Lanjouw et al, 1998, Harhoff et al, 2003).

fuel prices across various countries (see below).

Our data set includes 6,419 "clean" and 18,652 "dirty" triadic patents.[23] Since the EPO was created in 1978 our triadic patent data only starts in that year. The last year of fully comprehensive triadic data is 2005, so this is our end year.[24] Our basic dataset consists of all those applicants (both firms and individuals) who applied for at least one of these clean or dirty auto patents. We identify 3,423 distinct patent holders, which breaks down into 2,427 companies and 996 individuals. Our results are robust to dropping individuals, but it seems more natural to include "garage" inventors who may be more radical than firms. For every patent holder we subsequently identify the number of clean, dirty and "other" (i.e. neither clean nor dirty) triadic patents. We also extract other pieces of information based on this sample. For example, we identify all the other patents filed by holders of at least one clean or dirty triadic patent, which represents a total of 4,467,362 patent applications.

4.2 Tax-inclusive fuel prices

To estimate the impact of a carbon tax on innovation in clean and dirty technologies, we use information on fuel prices and fuel taxes. Data on tax-inclusive fuel prices are available from the International Energy Agency (IEA) for 25 countries (including some non-OECD countries), from 1978 onwards.[25] Since data are available for both diesel and gasoline fuels, we construct a time-varying country-level fuel price defined as the average of diesel and gasoline prices. The average

[23]In total, the PATSTAT data set includes 213,668 "clean" and 762,708 "dirty" patent applications. We thus focus on the high end of the distribution in terms of patent quality.

[24]The number of triadic patent families in all technologies (i.e., not clean/dirty related) drops starting in 2006. This is because of time lags between application and grant date at the USPTO.

[25]The IEA reports some incomplete data for an additional 13 countries.

fuel price across these 25 countries for our regression period 1986-2005[26] are shown on Figure 1A. Following a peak in the early 1980s, prices have remained broadly flat until the late 1990s when they rose. Average taxes have increased steadily during the 1990s (see Figure 1B).

What is more striking, however, is the high degree of heterogeneity across countries in the annual change in fuel prices, much of it being driven by cross-country differences in tax policies (see Figures 2 and 3). This heterogeneity will help us identify the effect of variations in fuel prices on innovation in a way which is firm-specific. An important issue noted above is that data on fuel prices are available only at the country level, whereas we would like to also exploit across firm variation. A related issue is that the car market is global and government policies abroad might be at least as important for firms' innovation decisions as domestic policies in the country where the firm is headquartered, especially for smaller countries. It is likely, however, that some markets matter more for some auto firms than others. This is for two reasons. First, auto manufacturers have different styles of vehicles reflecting their heterogenous capabilities and branding. Consequently, demand for a firms' products will be more popular in some nations than others depending on local tastes (e.g. Berry et al, 1995; Goldberg, 1995). Second, there is typically some home bias towards car makers that are considered to be "national champions" in national tastes and government policies (for example, the recent auto bailouts in Detroit and elsewhere). The upshot of this is that auto firms display heterogeneous current and expected future market shares across nations and their R&D decisions will be more influenced by prices and policies in some countries than others.

To operationalize this idea we construct a fuel price variable for each firm as a weighted average of fuel prices across countries based on a proxy of where the firm expects to sell cars in the future.

[26] In our baseline regressions we use data up until 1985 as a pre-sample period required for the BGVR estimator and to construct patent portfolio and inventor weights as discussed further below.

Our price index is defined as:

$$\ln FP_{it} = \sum_c w_{ic}^P \ln FP_{ct} \tag{10}$$

where P_{ct} is the tax-inclusive fuel price in country c at time t and w_{ic}^P is a firm-specific weight. The weight is determined by the importance of that country as a market outlet for that particular firm. We measure this as the proportion of the firm's total auto-related patents taken out in county c. The rationale for doing this is that a firm will seek intellectual property (IP) protection in jurisdictions where it believes it will need to sell in the future (even if it licenses the technology, the value of license will depend on whether it has obtained IP protection in relevant growth markets). For every patent applied for, we know that the patenting firm has paid the cost of legal protection in a discrete number of countries. For example, a firm may choose to enforce its rights in all EU countries or only in a subset of EU countries, say Germany and the UK. Similarly, the firm may decide to apply for patent protection in the US but not in smaller markets. In order to reflect the greater importance of larger countries, we also weight by each country's average GDP over 1965-1985. Finally, in order to make sure that the computed exposures are a (weakly) exogenous source of variation across firms, the weights are calculated using the patent portfolio of each company over the 1965-1985 "pre-sample" period, whereas we run regressions over the period 1986-2005.[27] We then perform robustness tests using different pre-sample periods to check that nothing important is driven by the precise year of cut-off (e.g. 1965-1990 and estimate from 1990-2005).

Why not use an alternative weighting scheme which simply reflects where firms currently sell their products (e.g. Bloom, Schankerman and Van Reenen, 2010)? First, we believe that the information on where firms choose to take patent protection is a potentially better measure because it reflects their *expectations* of where their future markets will be (which is relevant for R&D

[27] For further details see the Data Appendix.

decisions). Second, there is a data constraint: although sales distributions by geographic area are available for larger firms they are not available for smaller firms - and there are many patents from these smaller firms. We show our weights compared to sales weights for some of the largest car firms in Table 2 - Toyota, Volkswagen, Ford, Honda and Peugeot. The correlation is generally high between the two (0.95) suggesting that the weights we choose do a reasonable job at reflecting market shares.[28]

4.3 Patent and spillover stocks

Firm patent stocks are calculated in a straightforward manner using the patent flows ($PAT_{z,it}$) described above. Following Cockburn and Griliches (1988) and Peri (2005), the patent stock is calculated using the perpetual inventory method:

$$K_{zit} = PAT_{zit} + (1 - \delta)K_{zit-1} \tag{11}$$

where $z \in \{DIRTY, CLEAN\}$. We take δ, the depreciation of R&D capital, to be equal to 20%, as is commonly assumed in the literature, but we check the robustness of our results to other plausible values of this parameter.

To construct aggregate spillovers for a firm, we use information on the geographical location of the various inventors in that firm. These are geographically located regardless of nationality of the firm's headquarters or the location of the office where the patent was filed (e.g. the patents of Toyota's scientists working in US labs contribute to the US spillover pool). Implicit in our approach is the view that the geographical location of an inventor is likely to be a key determinant of knowledge spillovers rather than the jurisdiction over which the patent is taken out (which matters

[28]One exception is that Volkswagen appears to have a much higher patent share in Germany (it's home country) than its sales would suggest.

more as a signal of where the market for sales is likely to be). Many papers have documented the importance of the geographical component of knowledge spillovers in patents and other indicators (e.g. Henderson, Jaffe and Trajtenberg, 1993, 2005 and Griffith, Lee and Van Reenen, 2011).

Importantly, the distribution of the patent portfolio across countries and the distribution of inventors vary considerably across firms. This is illustrated in Figure A1. To construct a *firm*-specific spillover pool we use an analogous empirical strategy to that for the fuel price. The spillover weight w_{ic}^{S} is the share of all firm i's inventors (i.e. where the inventors lived and worked) in country c between 1965 and 1985. The spillover pool for firm i can then be calculated by:

$$SPILL_{zit} = \sum_{c} w_{ic}^{S} SPILL_{zct} \tag{12}$$

where $z \in \{DIRTY, CLEAN\}$ and $SPILL_{zct}$ is the spillover pool in country c at time t. The spillover pool in country c at time t, $SPILL_{zct}$ is defined as:

$$SPILL_{zct} = \sum_{j \neq i} w_{jc}^{S} K_{zjt} \tag{13}$$

i.e. the spillover pool of a country for firm i is the sum of all other firms patent stocks that have inventors in the country (weighted by the number of inventors). The aggregate stocks in (**??**) are thus entirely based on firm level stocks. This allows us to make out of sample simulations of aggregate stocks using firm level equations only. As an alternative strategy we could simply construct country level spillover stocks by aggregating over all patents of inventors based in that country:

$$SPILL_{zct} = \sum_{j \in \text{Inventors based in c}} K_{z,jt} \tag{14}$$

where $K_{zict} = PAT_{zit} + (1-\delta) K_{zjct-1}$ and PAT_{zjt} are the patents filed that associated with inventor j in year t. In our baseline specification both methods give very similar results. For consistency with our simulation results we use the first method (equation (**??**)) throughout the paper.

As noted above, a common problem is that patents values are highly heterogeneous. We mitigate this problem by conditioning on triadic patents, which screen out the very low value patents. But we also do two other checks. First, we weight patents by the number of future citations. Second, we use "biadic" patents filed at the EPO and at the USPTO, following Henderson and Cockburn (1993) who argued that patents were important if they had been applied in at least two of the three major economic regions. Our results are robust to these two variants.

5 Descriptive statistics

5.1 Aggregate statistics

Aggregate patenting in clean and dirty technologies has been rising over time (see Figure 4): the number of triadic patents in dirty technologies rose steadily between 1978 and 1988 and then again between 1992 and 2000. It has been decreasing during the last five years of our dataset. The number of clean patents remained very low for a decade and a half after 1980 before rising sharply between 1995 and 2002, reaching 724 in 2002. The rate of innovation has been stable subsequently. As a consequence, while the number of clean patents represented only 10% of the number of dirty patents during the 1980s, this ratio has grown to around 60% by 2005.

Clean and dirty patents are identified using a number of relevant International Patent Classification (IPC) codes. Table 1 shows the definitions of the IPC classes that we use to identify clean and dirty auto patent which follows OECD definitions. Patents relevant to the internal combustion engine are relatively straightforward to define (IPC Code F02 excluding sub-classes C/G/K). Clean innovations include those pertaining to electric, hybrid and hydrogen vehicles and the relevant sub-classes as shown in Table 1. The PATSTAT database holds information about the set of countries where the same invention is patented.

For every patent in our data set we know whether the invention has also been filed (prior to or following the filing of the patent at USPTO, EPO and JPO) at any other patent office included in PATSTAT (over 80 offices). When a patent is filed, it must include citations to earlier patents that are related to the new invention. Citations to earlier patents - or backward citations - are indicative of the accumulated knowledge used by the inventor to develop the new invention (e.g. Jaffe and Trajtenberg, 2002). We collect this information from the PATSTAT database. This represents 181,151 citations for all clean and dirty triadic patents included in our data set, which amount to 13.1 citations for the average patent. Table 3 reports the distribution of citations between clean and dirty categories. We see that among the patents cited by clean patents, 47% are clean, whereas 5% are dirty. The remaining 48% refer to other – i.e. neither clean or dirty – patents. To get a sense of what these figures imply, suppose that spillovers between clean and dirty patent categories were uniform; i.e. a clean patent – on average – facilitates subsequent clean innovation no more than it would facilitate subsequent dirty innovation and vice versa. Considering that even at the end of our sampling period there are about two times as many dirty patents as there are clean patents (see Figure 4), we would expect that the likelihood of a clean or dirty patent citing a dirty patent be at least three times higher than that of the clean patent citing a clean patent. Interestingly, we find that the likelihood of a clean on clean citations (47%) is almost as high as the likelihood of dirty on dirty citations (59%), suggesting that within category spillovers are vastly higher than between category spillovers. This suggests path-dependence in the direction of innovation as the theory suggests. We will formally test this in the next section.

5.2 Firm patent portfolios

Table 4 displays the top 10 patenters in clean technologies between 1978 and 2005 and Table 5 displays the top 10 patenters in dirty technologies between 1978 and 2005. They show the predominance of Japanese and German companies. An interesting finding from Tables 4 and 5 is that the vast majority of top innovators in clean technologies are not strictly specialized in this field. Most top companies' patent portfolios include both clean and dirty patents. The only exception is Samsung SDI, a battery specialist. Recall that this is based on triadic patents and US companies tend to file disproportionally more patents in just the US than in Europe and Japan. This explains why companies such as General Motors is not among the top 10 patenters in terms of triadic patents.

In order to illustrate this further we present in Tables 6 to 9 the top 10 clean and dirty patenters at the EPO and at the USPTO (Table 8). We find for example that General Motors is the third largest patenter of clean technologies at the USPTO. While it is clear that there a number of big companies active in both clean and dirty automotive patenting, computing a Herfindahl Index (HHI) for patenting over 1978 to 2005 for clean innovation we find a HHI of 0.023 and for dirty we find a HHI of 0.038, both of which reflect a low degree of concentration. The top 10 patent holders in clean account for 35.6% of patents over 1978 to 2005 whereas the corresponding figure is 46.6% for dirty, suggesting that innovation in dirty is more concentrated than innovation in clean.

Descriptive statistics for our dataset used in the regressions are shown in Table 10. We condition on firms who have produced at least one clean or dirty triadic patent since 1978. In any given year, the average number of dirty patents is 0.22 and the average number of clean patents is 0.08.

6 Results

6.1 Main Results

Our main results are shown in Table 11. The first three columns use the flow of a firm's clean patents as the dependent variable and the last three columns use the flow of dirty patents as the dependent variable. All estimates include firm fixed effects using the CFX approach (described in detail in Appendix **??**), year dummies and country GDP per capita (using the same weights as we do for fuel prices). We experimented with a wide range of other country specific variables and report some of the important ones below.[29] In column (1), we observe that higher tax inclusive fuel prices have a significantly positive correlation with clean patents with an elasticity of around unity. A 10% higher fuel price is associated with about 10% more clean patents. The spillovers and lagged patent stocks take signs consistent with the path dependency hypothesis. Firms who were more exposed to clean innovation by other firms ("clean spillovers") are significantly more likely to produce clean patents, whereas those benefiting more from dirty spillovers are significantly less likely to innovate in clean technologies. A 10% increases in clean spillovers is associated with a 4% increase in clean patenting whereas a 10% increases in dirty spillovers is associated with a 3% fall in clean patenting.

In addition to path dependence at the economy level through spillovers, path dependence also appears at the firm level. Firms which have innovated in clean technologies in the past are much more likely to continue to do so in the future. According to column (1) in Table 11 a 10% increase in lagged clean innovation stocks is associated with an increase of 5% in current clean innovation. A history of dirty innovation is also associated with more clean innovation, but this coefficient of

[29]Other country variables (GDP, population, inflation, etc.) did not appear to be robustly significant in our more general models.

0.25 is much smaller than both the coefficient on the lagged stock of clean innovation (0.51) and the corresponding coefficient in the dirty equation of column (4) of 0.64. In other words, a history of dirty innovation helps both types of innovation in the future but is more important for for future dirty innovation, leading to path dependence in the type of innovation.

Column (2) of Table 11 includes R&D subsidies for clean technologies and column (3) includes a control for emission regulations. Neither of these additional policy variables is significant. Moreover, the point estimate on the tax inclusive fuel price hardly change, which suggests that our results are not driven by a correlation between fuel prices and other country-level policies.

Column (4) of Table 11 repeats the specification of column (1) but uses dirty patents as the dependent variable. The coefficient on fuel prices is negative and significant as expected indicating that a 10% increase in prices is associated with a 5.3% decrease in dirty innovation. The signs on on the spillovers and own firm lagged innovation are broadly symmetric to those in the clean equation. Exposure to clean spillovers reduces dirty patenting, exposure to dirty patenting fosters it. A history of either dirty or clean patenting has a positive effect on further clean patenting, however the effect of dirty patenting is stronger. Importantly, both for spillovers and own knowledge stocks we have that own effects (e.g. clean on clean) are stronger than the corresponding cross effects (e.g. dirty on clean) thus meeting the necessary condition for path dependence to occur.

In summary, Table 11 appears to offer considerable support for our model. First, higher carbon prices significantly encourage clean innovation and have a negative effect on dirty innovation; and second there is path dependence in the direction of technical change: countries and firms that have a history of relatively more clean (dirty) innovation are more likely to innovate in clean (dirty) technologies in the future.

6.2 Alternative dynamic count data models

Table 12 considers alternative econometric approaches to deal with count data models and unobserved fixed heterogeneity. First, we control for fixed effects in our Poisson model following the Hausman, Hall et al approach ("HHG") in column (1) for clean patents and column (3) for dirty patents. The signs of coefficients are generally the same as in our baseline models, but the marginal effect of fuel price is much greater in absolute magnitude for dirty innovation and smaller (and insignificant) for clean. Indeed, the magnitude of the estimated elasticity for dirty patents seems unreasonably large, a 10% increase in fuel prices being associated with a 24.5% decrease in dirty innovation. We suspect that the assumption of strict exogeneity underlying HHG is problematic in our context, as we have a highly dynamic specification.

Columns (2) and (4) of Table 12 implement the Blundell et al (1999) estimator ("BGVR") . The pattern of the spillover effects and dynamics remain similar to the baseline regression, and we still obtain a positive and significant effect of fuel prices on clean innovation and an negative and significant effect on dirty innovation. The value of the fuel price coefficients are comparable to the baseline case. The elasticity of fuel prices on clean patenting is 0.672 which a bit smaller than Table 11 (0.992), whereas in the dirty equation it is -0.614 slightly larger in absolute magnitude than Table 11 (-0.539). BGVR produces somewhat larger values for path dependency (e.g. the effects of lagged clean knowledge stocks on current clean patenting) than in the baseline results which may be because the BGVR approach is not fully controlling for fixed effects.

The final two columns of Table 12 present an even simpler model where we use relative patenting in clean $\ln\left(\frac{1+PAT_{CLEAN,it}}{1+PAT_{DIRTY,it}}\right)$ as the dependent variable in an OLS regression. Column (5) shows that there is a significant and positive effect of fuel prices on relative clean innovation. Column (6) shows that this result is robust to including a full set of country by year fixed effects to absorb any potential country specific time varying policy variables.

6.3 Extensions and robustness

Table 13 repeats regressions for our three main count data approaches (BGVR, HHG and CFX) but restricts the sample to firms which had at least one patent filed prior to 1985. This leads to small changes in the point estimates but no changes in the overall qualitative patterns. Table 14 repeats our regressions using just fuel taxes (rather than the tax-inclusive fuel prices used until now). The results are comparable to the baseline results, although one difference is that the absolute magnitude of the coefficients on taxes are smaller for both types of innovation. This is to be expected if firms are guided by fuel prices and fuel suppliers bear some of the cost burden of a tax.

In Table 15 use a longer pre-sample period to construct weights (1965-1990) and run the regressions only on the 1991 to 2005 to period. The coefficients are again comparable, although some of the fuel price variables are no longer significant presumably due to smaller sample size. Importantly, we always find a significantly different point estimate for fuel price on clean and on dirty, consistent with there being directed technological change. In Table 16 we report regressions based on biadic rather than triadic patents; i.e. we include all innovations into the construction of the innovation and knowledge stock variables that are registered with the EPO and the USPO but not necessarily the Japanese patent office. As a consequence the size of our sample increases slightly, but the results do not change much. In Table 17 we construct the knowledge stock variables - including the spillover variables - using citation weighted counts for all worldwide patents. This leads to results which are similar in qualitative terms; i.e. it remains true that the price response is larger for clean patents than for dirty patents ensuring a clean inducing price effect. Equally, the relative size of the knowledge and spillover stocks continue to be consistent with path dependence.

However, the price effects on clean with both CFX and BGVR are now implausibly high with an implied elasticity of more than 2. Also, the spillover effects tend to be much larger whereas the own stock effects are diminished. An explanation for this could be that using citation weights is prone to larger measurement errors than using multi patent office patenting. However, these measurement errors tend to be smoothed out when computing the aggregative spillover stocks so that they affect the own stocks more.

7 Simulation results

Our regression results imply that there is path dependence in the type of innovation pursued, both through internal knowledge stock effects and through external spillovers. Here we explore how strong this path dependence is in quantitative terms by studying the simulated future evolution of both clean and dirty knowledge stock implied by our fitted models. In particular we are interested to see under which conditions *the clean knowledge stock for the aggregate economy exceeds the dirty knowledge stock*. In line with Acemoglu et at. (2012) this would be a requirement for clean technologies to be able to compete with dirty ones, even without policy intervention. Our projections should be considered as an exploration into the strength of path dependence rather than necessarily realistic forecasts of future patenting.

We recursively compute values of expected patenting under different policy scenarios, use those to update the knowledge stock variables (including the spillover variables) and feed them into the next iteration.[30] Hence if we split the right hand side variables x_{it} in equation (**??**) into variables that are functions of the knowledge stock (k_{it}) and other variables such as the fuel price(p_{it}), we

[30]Below we also provide simulation runs providing recursively generated knowledge stocks over the sample period (1986-2005). We can compare those to the actual values to further examine the quality of our empirical models.

can write $x_{it} = [k_{it}, p_{it}]$ and a particular iteration is defined by:

$$\widehat{PAT}_{zit+T} = \exp\left(k_{it+T-1}\beta_{kz} + p^{CF}_{it+T-1}\beta_{pz}\right)\eta_{zi} \tag{15}$$

where $k_{it+T} = f\left(k_{it+T-1}, \widehat{\text{PAT}}_{CLEAN,t+T}, \widehat{\text{PAT}}_{DIRTY,t+T}\right)$ and where $\widehat{\text{PAT}}_{CLEAN,t+T}$, $\widehat{\text{PAT}}_{DIRTY,t+T}$ are predicted future patent flows for firms in the sample and p^{CF}_{it+T} are potentially counterfactual values of the policy and other control variables. We focus our exploration on periods up to 2030 with 2020 as a focal point. This is somewhat arbitrary but in line with scenarios by the International Energy Agency (IEA)[31] suggesting that globally fossil fuel use must peak by that year to avoid highly risky climate change. It is also consistent with the European Commission's 2020 targets.[32]

Figure 5 reports simulations based on the regressions from Table 13 columns (1) and (4). The knowledge stocks in clean and dirty technologies are reported on the y-axes.[33] In Figure 5A we report the baseline case without any changes to policy variables compared to their 2005 values.[34] We see that our regressions imply that path dependence is strong enough to lead to the economy becoming ever more locked into dirty technologies; i.e. the aggregate knowledge stock grows much faster over time than the clean knowledge stock so that the initial advantage of dirty widens over time. In panel B we see that this remains true even if we increase fuel prices globally by 10%. In this scenario, however, the clean stock grows faster and the dirty stock more slowly than in Panel A. The same pattern is repeated in Panel C when we increase prices by 20%, but in neither case will clean overtake dirty before 2003. Panel C shows that for a 30% price increase clean overtakes

[31] http://blogs.ft.com/energy-source/2009/11/10/fossil-fuel-use-must-peak-by-2020-warns-iea/#axzz1tQmZyLoy

[32] see http://ec.europa.eu/news/economy/100303_en.htm

[33] For the simulations we restrict the sample to the firms where we have pre sample information. In this way we do not have to make further assumptions as to how changes in the spillover and policy variables would affect firms where these variables are essentially missing.

[34] We equally keep other exogenous variables such as GDP per capita at their 2005 levels, as well as the year fixed effect.

dirty by 2030 and Panel D shows that for clean and dirty parity by 2020 we need an increase in price in the order of 40%.

In Figure 6 we explore the importance of own lagged innovation vs. spillovers in leading to path dependence. In Panels A and B we report simulations where we fix the spillover variables at 2005 levels and only update own innovation stocks when simulating future changes. In Figure 6A we find that the gap between dirty and clean increases more slowly over time compared to Figure 5A. This is not surprising given that both own knowledge stock effects and spillovers imply path dependence and Figure 6A switches off spillovers. While this might look like better news for clean innovation, it is now also the case that any efforts to help clean through fuel price increases are less effective, as we have switched off one channel through which they feed through the economy. Hence in Figure 6B we see that a fuel price increase of 40% is no longer enough to tip the balance between clean and dirty by 2020. In other words, path dependence acts as a double-edge sword: it increases the gap between dirty and clean technologies in a status quo situation, but facilitates a catch-up of clean technologies when the fuel price increases. The same happens when in Figures 6C and 6D where we fix own lagged knowledge capital stocks at 2005 levels and only update spillovers. Note that quantitatively, spillover effects on current innovation are smaller than effects from own knowledge stocks. Consequently, the series grow logarithmically in Figures 6C and 6D rather than exponentially in Figures 6A and 6B.

Finally in Figure 7 we fix both own stocks and spillovers. We then explore how aggregate flows respond to successively increasing the fuel price. Hence we can ask: "By how much would we have to increase carbon prices to have the flow of clean exceed the flow of dirty patents?" We can think of this as a kind of upper bound estimate to tip the balance in favour of clean by means of a fuel price intervention. From the graph we see that this requires a very large fuel price increase on the order of 120%. Figure A3 in appendix shows simulation results for HHG and BGVR models. In both cases we find as in the baseline CFX case that without price increases clean will not catch up with dirty innovation stocks. In quantitative terms the simulated knowledge stocks for dirty knowledge are comparable to CFX in the HHG case but are growing much faster in the BGVR case. This is a result of the larger estimates for lagged own knowledge stocks found in the regressions above. Nevertheless, in terms of fuel price increases that are required for clean knowledge stocks to get ahead of dirty within a 15 year horizon we find a comparable although somewhat smaller order of magnitude. For both BGVR and HHG the turning point is at approximately a 30% increase. In Figure A4 we examine the required price increase to bring about higher flows of clean from the base year onwards. Not surprisingly[35] for HHG this is rather low at around 75% whereas for BGVR it is around 120%, similar to the value found in the baseline estimates. Finally, we compute simulated knowledge stocks for our sample period (1986-2005) which we can compare to actual knowledge stocks. The results are reported in Figure A4. Both CFX and HHG track the actual dynamics rather well. BGVH tends to overestimate the growth of dirty knowledge stocks and underestimate the observed growth of clean knowledge stocks.

8 Conclusion

In this paper we have combined several patent datasets to analyze directed technical change in autos, which is a key industry of concern for climate change innovations. Consistent with a simple model, we find that "clean" innovation does positively respond to increases in the tax-inclusive fuel prices, exploiting the fact that prices evolve differentially across countries and firms are differentially exposed to these price changes because of their different expected sales profiles across geographical

[35]Considering the point estimates of the fuel price coefficients in Table 14.

markets. Our second key result is that there is strong evidence of "path dependence" in the sense that firms with more inventors exposed to clean innovation are more likely to direct their research energies to clean innovation in the future (a directed knowledge spillover effect). Similarly, firms with a history of clean innovation in the past are more likely to focus on clean innovation in the future.

The fact that such path dependence holds for clean (as well as for dirty) innovation highlights the desirability of acting sooner to shift incentives for climate change innovation. Since the stock of dirty innovation is greater than that of clean innovation, the path dependence effect ("building on the shoulders of giants") will tend to lock economies into high carbon emissions, even after the introduction of a mild carbon tax or R&D subsidies for clean. This in turn reinforces the case for stronger action now, which could be relaxed in the future as the economy's stock of knowledge shifts in more of a clean direction. While there is some difference in quantitative terms across the different econometric models, there is a remarkable consistency across specifications to the effect that without further policy intervention clean technologies will not catch up with the more advanced dirty technologies. Increases in carbon price instead can bring about a change in direction. Our baseline results suggest an increase of 40% of fuel prices with respect to the 2005 price will allow clean innovation stocks to overtake dirty stocks after fifteen years. These estimates depend upon multiplier effects of a given price increase working through knowledge stocks and aggregate spillovers. Our analysis could be extended in several directions. One extension would be to use micro data to estimate the relative efficiency of R&D investments in clean versus dirty innovation, and also the elasticity of substitution between the two types of production technologies. As argued in Acemoglu et al (2012), these parameters play as important a role as the discount rate in characterizing the optimal environmental policy. These and other equally important extensions are left for further research.

Web Appendix for "Carbon Taxes, Path Dependence and Directed Technical Change: evidence from the Auto Industry"

A Econometric Models

We separately examine clean and dirty patent counts using a standard Poisson model

$$PAT_{zit} = \exp\left(x_{it}\beta_z\right)\eta_{zi} + u_{zit} \tag{16}$$

where z$\in$ $\{DIRTY, CLEAN\}$ and x_{it} is a vector of regressors including functions of the lagged dependent variable. For identification we assume $E(u_{zit}|x_{it}) = 0$.[36] We consider four alternative estimation techniques that allow for the possibility of firm level fixed effects η_{zit} in the propensity to patent. The standard approach is Hausman, Hall and Griliches (1984, HHG) who suggest a transformation akin to the within groups estimator in the linear panel data context. In GMM terms, their estimator can be expressed as relying on the following moment condition for identification (e.g. Blundell, Griffith, Windmeijer, 2002):

$$E\left\{\left(PAT_{zit} - \mu_{zit}\frac{P\bar{A}T_{zi}}{\bar{\mu}_{zi}}\right)x_{kit}\right\} = 0$$

for all variable in x_{it} where $\mu_{zit} = \exp\left(x_{it}\beta_z\right)$ and a bar represents the average of a variable over time for a specific firm. Note that

$$PAT_{zit} - \mu_{zit}\frac{P\bar{A}T_{zi}}{\bar{\mu}_{zi}} = u_{it} - \frac{\mu_{zit}}{\bar{\mu}_{zi}}\bar{u}_{zit}$$

implying that we require strict exogeneity, i.e. the shock u_{zit} must be uncorrelated with x_{it} not only contemporaneously, but in all periods; i.e. $E\left\{u_{zit}|x_{i\tau}\right\} = 0$ for all t and τ. When using regressors that depend on past realizations of the dependent variable such as the knowledge capital stocks, this assumption is violated.

[36]Note that we can equivalently represent the model in terms of a multiplicative shock ν_{zit} with $E(\nu_{zit}|x_{it}) = 1$. We would have

$$\nu_{zit} = 1 + \frac{u_{zit}}{\exp\left(x_{it}\beta_z\right)\eta_{zi}}$$

and our assumptions concerning u_{zit} imply $E(\nu_{zit}|x_{it}) = 1$.

Blundell, Griffith and Van Reenen (1999, BGVR) proposed an alternative estimator which is robust to relaxing the strict exogeneity assumption. It relies on introducing a control function term for the fixed effects, which is identified from realizations of the dependent variable in a pre-sample period. Hence, the idea is to think of the fixed effect as the combination of a control term $\phi(\cdot)$ and an error, ω_i.

$$\eta_{zi} = \phi\left(\ln P\bar{A}T_{zi0}, I\left\{P\bar{A}T_{zi0} = 0\right\}\right) + \omega_i$$

where $P\bar{A}T_{zi0}$ is the average amount of patenting by firm i in the pre-sample period. BGVR show that with $\phi(\cdot) = \exp\left(\phi_{zl} \ln P\bar{A}T_{zi0} + \phi_{z2} I\left\{P\bar{A}T_{zi0} = 0\right\}\right)$, pre-determined x_{it}[37] and stationarity in the dynamic system implied by equation (**??**) estimates of β_z are unbiased as the duration of the pre-sample period becomes large. Thus, effectively we estimate the following model:

$$PAT_{zit} = \exp\left(x_{it}\beta + \phi_{zl} \ln P\bar{A}T_{zi0} + \phi_{z2} I\left\{P\bar{A}T_{zi0} = 0\right\}\right) + u_{zit}$$

The BGVR approach requires the realizations of the dependent variable in the pre-sample period to be representative of a firm's behavior over the sample period. Formally, the series must be mean stationary (conditionally on the time dummies). It is easy to see why this might be violated in particular for clean patents, whose realizations are concentrated towards the end of our sample period. Consequently, for many firms we do not observe any clean patenting in the pre-sample period which could inform us about variations in their fixed propensity to patent in clean.

To address this problem we propose a new estimator in the same spirit of using a control function as in BGVR. However, rather than using information from the pre-sample period to calibrate the control function, we simultaneously exploit future data. We estimate the main regression equation as well as a second equation allowing us to identify the control function from future data. The key idea is the following. In general a control term $\breve{\phi}_{zit}(\cdot)$ will lead to consistent estimates, if the resulting error term $\breve{\omega}_{zit} = \eta_{zit} - \breve{\phi}_{zit}(\cdot)$ is orthogonal to x_{it}; i.e. $E\{\breve{\omega}_{zit} | x_{it}\} = 0$. Note, that given a parameter vector β we can obtain such an estimate by regressing[38]

$$\frac{PAT_{ziT}}{\mu_{ziT}} = \eta_{zi} + \frac{u_{ziT}}{\mu_{ziT}} = \breve{\phi}_z(x_{it}) + \breve{\omega}_{zit} \tag{17}$$

with $T > t$, provided that the variables in x_{it} are pre-determined because then

$$E\left\{\frac{u_{ziT}}{\mu_{ziT}} | x_{it}\right\} = 0 \tag{18}$$

and we can intepret $\breve{\phi}_{zit}(x_{it})$ as the expectation of the fixed effects given x_{it}:

[37] i.e. $E\{u_{i\tau} | x_{it}\} = 0 \; for \; \tau \geq t$.

[38] For notational simplicity we write the following equation with just one future term. In practice we can improve efficiency by regressing on an average of future values $\frac{1}{T-t+1}\sum_{\tau=t}^{T} \frac{PAT_{zi\tau}}{\mu_{\_zi\tau}}$.

$$\breve{\phi}_z(x_{it}) = E\{\eta_i | x_{it}\}$$

As in the standard case we parametrise $\breve{\phi}_z(x_{it})$ as an exponential function[39] ,

$$\breve{\phi}_z(x_{it}) = \exp(x_{zit}\gamma)$$

Notice, that given this control function we can transform our main regression equation as

$$\frac{PAT_{zit}}{\breve{\phi}_z(x_{it})} = \exp(x_{it}\beta_z) + \exp(x_{it}\beta_z)\frac{\breve{\omega}_{zit}}{\breve{\phi}_z(x_{it})} + \frac{u_{zit}}{\breve{\phi}_z(x_{it})} \tag{19}$$

where we replaced η_i by $\breve{\phi}_{zit}(x_{it}) + \breve{\omega}_{zit}$ and divided by $\breve{\phi}_{zit}(x_{it})$. Because the x_{it} are are predetermined, given the definition of $\breve{\omega}_{zit}$ and recalling the definition $\mu_{zit} = \exp(x_{it}\beta_z)$ we have that

$$E\left\{\left(\mu_{zit}\frac{\breve{\omega}_{zit}}{\breve{\phi}_z(x_{it})} + \frac{u_{zit}}{\breve{\phi}_z(x_{it})}\right) | x_{it}\right\} = 0 \tag{20}$$

Hence, we have two equations that depend on each other as well as two sets of moment conditions. We can consequently estimate equations (**??**) and (**??**) as a system of two simultaneous equations using the sample analog of the following moments

$$E\left\{\left(\begin{array}{c} \frac{PAT_{zit}}{\breve{\phi}_z(x_{it})} - \mu_{zit} \\ \frac{PAT_{ziT}}{\mu_{ziT}} - \breve{\phi}_z(x_{it}) \end{array}\right) | x_{it}\right\} = 0$$

We refer to this approach below as the control function fixed effects estimator (CFX).

In addition to these three dynamic count data approaches we also explore the common practice of implementing equation (**??**) as a linear panel data estimator by taking logs of the dependent variable after simply adding the value of unity (an arbitrary constant); i.e. the regression equation becomes:

$$\ln(1 + PAT_{zit}) = x_{it}\beta_z + \alpha_{zi} + \varepsilon_{zit}$$

Although this model has undesirable features like generating negative predicted values of patenting it is attractive because it is straightforward to estimate a relative clean vs. dirty regression; i.e.

$$\ln(1 + PAT_{CLEANit}) - \ln(1 + PAT_{DIRTYit}) = x_{it}(\beta_{CLEAN} - \beta_{DIRTY}) \tag{21}$$

$$+ (\alpha_{CLEANi} - \alpha_{DIRTYi}) + (\varepsilon_{CLEANit} - \varepsilon_{DIRTYit}) \tag{22}$$

[39]In theory we can even allow a more flexible specification where the conditional expectation varies over time; i.e. $\breve{\phi}_{zt}(x_{it}) = E_t\{\eta_i | x_{it}\}$. This could reflect firms learning more about their fixed effect over time for instance. In practice this increases the number of parameters to be estimated greatly and becomes computationally very burdensome. In our baseline results we therefore fix $\breve{\phi}_z(\cdot)$ over time.

We show in the results section that the results are qualitatively similar no matter which precise estimation technique we use.

B Data Appendix

B.1 Identifying unique patent holders

Our patent data is drawn from the World Patent Statistical Database (PATSTAT) maintained by the European Patent Office. We use the September 2009 version of PATSTAT. The PATSTAT database reports the name of patent applicants, but a common problem with patent data is that the name of patentees often varies, because of spelling mistakes, typographical errors and name variants. To identify unique patent holders we use the ECOOM-EUROSTAT-EPO PATSTAT Person Augmented Table (EEE-PPAT) database, available at http://www.ecoom.be/nl/eee-ppat, which provides a dictionary of harmonized patent applicants' names produced through a computer algorithm followed by visual inspection. We then manually check the name match, which allows us to put together companies that a typical computer algorithm would consider distinct. For example we match Ford Motor Company with Ford Werke, its German subsidiary. As a result, we are able to reduce the number of distinct patent holders from 20,916 to 3,423, 2,427 of which are companies and 996 are individuals.

B.2 Firm-level weights

B.2.1 Weights based on patent portfolios

As explained above in the main text, the firm-specific fuel price is computed as the weighted geometric mean of the fuel prices across countries with weights reflecting the shares of the corresponding countries in the firm's patent portfolio. Our price variable is thus defined as:

$$\ln FP_{it} = \sum_c w_{ic}^{FP} \ln FP_{ct} \tag{23}$$

where FP_{ct} is the tax-inclusive fuel price in country c at time t and w_{ic}^{P} is the firm-specific weight for country c. In order to make sure that the computed exposures are an exogenous source of variation across firms, the weights are calculated using the patent portfolio of each company over the 1965-1985 "pre-sample" period (with the regressions performed on the 1986-2005 period).

To make matters concrete consider the example of Hitachi, a large Japanese car parts manufacturer, who filed 90,381 patents between 1965 and 1985. 63,175 of these filings were in Japan, 8,315 in the US and 3,498 in Germany. The rest were in a large number of other patent offices. Note that there are a larger number of filings than there are patents, as one invention can be filed in multiple patent offices. For example, Hitachi's patent 11464997 (this is the DOCDB family number) was developed by a Japanese inventor and filed in 1980 both in Japan and in the US. This patent enters twice in the patent-portfolio weight: once for Japan and once for US, since it indicates that both the US and Japan matter for Hitachi. Hitachi's 90,381 patents filed between 1965 and 1985 correspond to only 70,526 distinct inventions, some of which were patented in several countries even though almost all of Hitachi's R&D activities are conducted in Japan (we use inventor location below for spillovers). In order to reflect the greater importance of larger countries when constructing fuel price weights, we take each country's average GDP over 1965-1985 into account. The firm-specific weight for country c is thus equal to:

$$w_{ic}^P = \frac{s_{ic}^P gdpc}{\sum_c s_{ic}^P gdpc} \tag{24}$$

where s_{ic}^P is the share of country c in Hitachi's patent portfolio between 1965 and 1985 and $gdpc$ is the share of country c in the world's GDP over 1965-1985. The weights used for Hitachi are 68.8% for Japan, 23.9% for US and 2.7% for Germany. The weights summed across all other countries was 4.6% so the total weights sum to 100

We use the patent-portfolio weights to construct the price, tax, and emission regulations variables:

$$\ln P_{it} = \sum_c w_{ic}^P \ln P_{ct} \tag{25}$$

where P_{ct} is the tax-inclusive fuel price per liter in country c at time t. We use exactly the same weights to calculate the the fuel tax per liter and level of automobile pollution regulation in country c at time t. Note that in constructing the weights we use all patent filings from applicant firms who have filed at least one auto-related patent. These are all applicants who have filed a dirt or clean patent as defined by Table 1 from the OECD or in an IPC class defined as autos according to the OECD's cross walk. We could have also included patent filings by applicants who were part of the auto-related firms who had never filed for a clean or dirty auto patent according to our definitions. This would have increased our sample of patent filings from 4.5m to about 16m. We chose not to do this as many of these patents are only distantly related to autos and so would not be relevant for tracking the demand for cars. Going in the other direction, we could narrow our definition to include only patents in IPC classes we deem as clean or dirty and exclude all other patents by the same applicants. Building weights from this narrower pool led to similar results to those presented here.

B.2.2 Weights based on location of inventors

To construct the firm-specific spillover pools in clean and dirty knowledge we use an analogous empirical strategy to that for the fuel price. The firm-specific spillover pool is computed as the weighted geometric mean of the knowledge pools across countries with weights reflecting the shares of the corresponding countries in the firm's pool of inventors. The spillover pool for firm i is calculated as:

$$SPILL_{zit} = \sum_{c} w_{ic}^{S} SPILL_{zict} \tag{26}$$

where $z \in \{DIRTY, CLEAN\}$ and $SPILL_{z,ct}$ is the spillover pool in country c at time t, which can be firm specific (more on that below). The spillover weight w_{ic}^{S} is the share of all firm i's inventors (i.e. where the lead patent holder lived and worked) in country c between 1965 and 1985.

This weight differs from the patent-portfolio weight w_{ic}^{P} described above in two ways. First, instead of using information on where each patent was filed (for example, the US patent office) we use the location of the patent inventors (who are more likely to benefit from other research conducted locally). Inventor countries are counted fractionally, so if a patent is filed by two inventors, one from Germany and one from the US, each country will receive one half.[40] Note that we use information on the *country of residence* of the inventor, not on his nationality. This seems natural because the geographical location of the inventor is likely to be the critical issue for knowledge spillovers. The second difference with respect to patent-portfolio weight w_{ic}^{P} is that each invention is only counted once, no matter in how many patent offices it has been patented. This is to avoid double counting. Returning to Hitachi's patent 11464997 filed in 1980 both in Japan and in the US, this patent enters twice in the patent-portfolio weight but only once in the inventor location weight, as a Japan-developed invention. So although $w_{Hitachi,Japan}^{P} = 0.688$ as above, $w_{Hitachi,Japan}^{S} = 0.99$. This indicates that although almost all Hitachi's R&D is based in Japan it sells car parts to a much wider geographical market.

The spillover pool in country c at time t of firm i, $SPILL_{zict}$ is defined as:

$$SPILL_{z,ct} = \sum_{j \neq i} w_{jc}^{S} K_{z,jt} \tag{27}$$

i.e. the spillover pool of a country for firm i is the sum of all other firms patent stocks that have inventors in the country (weighted by the number of inventors). The aggregate stocks in (**??**) are thus entirely based on firm level stocks. This allows us to make out of sample simulations of aggregate stocks below using firm level equations only. As an alternative strategy we could simply construct country level spillover stocks by aggregating over all patents of inventors based in that country:

$$SPILL_{zct} = \sum_{j \in \text{Inventors based in c}} K_{z,jt} \tag{28}$$

where $K_{zict} = PAT_{zit} + (1-\delta)K_{zjct-1}$ and PAT_{zjt} are the patents filed that associated with inventor j in year t. In our baseline specification both methods give very similar results. For consistency with our simulation results we use the first method (Equation **??**) throughout the paper.

[40] We do this in order to avoid giving an artificially higher weight to a patent with multiple inventors compared to one with just a single named inventor

Note that we also use the inventor weights to construct the amount of R&D in energy-efficient transportation in country c at time t.

B.2.3 Other data

Fuel price and fuel tax come from the International Energy Agency's Energy Prices and Taxes database, available online at http://data.iea.org. We use Households End-Use Prices in USD PPP/unit. Since data are available for both diesel and gasoline fuels, we define fuel price as the average of diesel and gasoline prices.

Data on public R&D expenditures comes from the IEA's Energy Technology Research and Development database, available online at http://data.iea.org. We use Total R&D in Million USD (2010 prices and exchange rates). We use data on Energy efficiency in transportation (Flow 13).

Data for environmental standards governing maximum permissible levels of tailpipe emissions for pollutants from new automobiles were sourced from a dataset originally constructed by Perkins and Neumayer (2012). Countries' regulatory stringency is coded on a scale of 0 to 5. The basis of the classification scheme is the European Union's (EU) Euro emission standards which were originally implemented across member states in 1992 and have subsequently been tightened in a series of incremental steps. Countries are coded 0 if they had no national emissions standards in place for new vehicles, or if standards were less stringent than the equivalent of Euro 1, during the year in question. Countries where Euro 1 or its equivalent was legally enforceable are coded 1, and so on, with 5 for countries having implemented the equivalent of the Euro 5 standard.

Data on GDP, GDP per capita and population are taken from the World Bank's World Development Indicators, available at http://data.worldbank.org/. GDP and GDP per capita are PPP and constant 2005 USD.

Sales data used to compare the patent weights with sales distribution are from the following sources:

HONDA
http://world.honda.com/investors/library/annual_report/2006/ar2006.pdf

TOYOTA
http://www.toyota-global.com/investors/ir_library/annual/pdf/2005/pdf/04.pdf
http://www.toyota-global.com/company/profile/figures/vehicle_production_sales_and_exports_by_region.html

VW
http://www.volkswagen.co.uk/assets/common/content/volkswagen-world/annual-report-2003.pdf
http://www.volkswagen.co.uk/assets/common/content/volkswagen-world/annual-report-2004.pdf
http://www.volkswagen.co.uk/assets/common/content/volkswagen-world/annual-report-2005.pdf

FORD
http://corporate.ford.com/doc/2005_AR_full.pdf

http://corporate.ford.com/doc/2004annualReport.pdf
http://corporate.ford.com/doc/2003_annual_report.pdf

PEUGEOT
http://www.psa-peugeot-citroen.com/document/publication/annual_report_20051151075591.pdf
http://www.psa-peugeot-citroen.com/document/publication/2004_Annual_Report1117556773.pdf
http://www.psa-peugeot-citroen.com/document/amf/2003_Annual_report1168420831.pdf