Market Power, Innovation, and the Green Transition

Rik Rozendaal*

JOB MARKET PAPER

This version: April 23, 2024 Latest version

Abstract

This paper studies the relationship between climate policy, market power and innovation. Using linked data on patent applications and firms' balance sheets, I document five empirical facts, some of them novel to the literature. Most importantly, I find that firms with a higher degree of market power are, on average, more invested in dirty technologies than their direct competitors. These findings motivate me to develop a model of directed technical change and the environment with strategic innovation incentives, incorporating all five facts. Firms compete for market power by innovating in clean or dirty technologies. A carbon tax can decrease the effective technological distance between two competitors, and thus affects both the intensity and the direction of innovation. In the model, the increase of a carbon tax can sharply increase clean innovation while also increasing dirty innovation by some firms. Calibration results show that while the transition to a green economy may temporarily decrease aggregate market power and permanently increase innovation, the market power effect is harmful from a welfare perspective.

*Tilburg University, Department of Economics. Email: r.l.rozendaal@tilburguniversity.edu. Firstly, I thank Sjak Smulders for his valuable feedback throughout this project. I also thank Sina Ates, Jan Boone, Dallas Burtraw, Greg Casey, Malik Curuk, Reyer Gerlagh, Josh Linn, Gerard van der Meijden, Herman Vollebergh and participants at the Danish Central Bank, Utrecht University, UC3M, Resources for the Future, Tilburg University, SERE (ETH Zurich), Dutch Economists Day (The Hague), EAERE (Limassol), KVS New Paper Sessions (The Hague) and DEARE Day (Amsterdam) for helpful comments and discussions, and Berk Aktas for helping me with some of the data.

1 Introduction

Decarbonizing the economy requires firms to transition from fossil fuel-based production to clean production, which requires a wide range of innovations (IPCC, 2023). Theories of directed technical change (DTC) and the environment predict that climate policy shifts innovation efforts from dirty technologies to clean ones, and are supported by empirical evidence.[1] Yet, despite the fact that a large literature shows the importance of competition, market power and strategic behavior for innovation, growth and business dynamism,[2] these topics receive little attention in the DTC-environment literature. Little is known about the role of market power and strategic behavior in the context of the green transition. This paper aims to fill that gap in the literature.

Market power, innovation and the green transition are, in fact, closely connected. Competing firms may have different productivity levels for both clean and dirty production processes. Hence, the effects of a climate policy, which makes dirty production (relatively) more expensive, may differ between firms within an industry. That is, climate policy can affect the technological lead of market leaders over their competitors, and thus the degree of market power in an industry. Market power, in turn, affects firms' incentives to innovate. Firms that have more to gain from a successful innovation invest more than firms that have little to gain. Climate policy can thus affect not only the direction but also the intensity of innovation, and make markets more (or less) dynamic. Finally, total investments in clean and dirty innovation determine the speed and success of the green transition.

This paper studies the interaction of climate policy, innovation and market power. I first introduce five motivating facts. Using historical data on patents and sales, I explore the relationship between market power and the direction of innovation. I find that (i) both clean and dirty patents are filed by firms that are active in a wide range of sectors; (ii) the direction of innovation is path dependent; (iii) market leaders are, on average, more invested in dirty technologies than their direct competitors; (iv) there is some evidence that differences in technology gaps between leaders and laggards are correlated with market competition and

[1] See, e.g., Acemoglu et al. (2012, 2016) for theory and Popp (2002) and Aghion et al. (2016) for empirics. This literature is discussed in more detail below.

[2] See, for instance, Aghion et al. (2005) and Akcigit and Ates (2023). This literature is also discussed more elaborately below.

the direction of innovation; (v) both clean and dirty patenting correlate positively with firm size. Facts (iii) and (iv) are novel to the literature, the others are confirming earlier findings using a larger and more recent data set. Together, they suggest that market leaders may respond differently to climate policies than smaller firms in the same sectors. Several of these facts cannot be explained by existing models of DTC and the environment.

The facts above motivate me to develop a new DTC-environment model with strategic innovation incentives that incorporates all of the empirical findings. The model, which builds on state-of-the-art work in the endogenous growth literature (Akcigit and Ates, 2023), features a continuum of intermediate input sectors in which two firms, which each use either a clean or a dirty production technology, compete on prices and innovate strategically. That is, firms invest more in R&D if the technological distance with their competitor is small because the expected gains from innovation are greater. Climate policy can change the effective technological distance between two competitors, and thus affects both the intensity and the direction of innovation. As a result, a main theoretical finding is that, under some conditions, a carbon tax can increase both clean and dirty innovation in an industry.

I then use micro data to calibrate the model to match recent macro moments. Next, I simulate the effects of the introduction of a carbon tax in 2024. I find, among other things, that the policy change leads to a decrease in aggregate market power as measured by markups and an increase in the intensity of innovation along the green transition. Quantifying the role of the market power channel that this paper adds to the literature, I find that, despite the decrease in markups and the increase in innovation, the optimal carbon tax when this channel is switched off is about 44% higher than the optimal carbon tax in the baseline model. This is due to a costly reallocation of labor from production to research, which amplifies the negative effect on output at the moment when the tax is introduced.

Competition and strategic behavior are important for innovation but understudied in the context of the green transition. To see how strategic innovation matters, consider the following two examples. First is Tesla, which disrupted the car market with its electric vehicles. In doing so, it essentially forced incumbents to start taking a clean technology seriously, and to change the direction of their innovation efforts towards clean. Second is the case of steel manufacturing. Incumbent firms in this industry tend to be very carbon intensive. An alternative way of producing is available in the form of green hydrogen, but this technology is far from competitive. Climate policy, in the form of carbon pricing or subsidies for clean production, may bring the clean technology closer to the dirty one, and presents incumbent firms with a dilemma: keep improving their dirty product to stay competitive, or transition to clean production. My model captures these types of interactions between firms and policy.

The literature on innovation and the environment is centered around the concept of directed technical change (DTC), which was applied to the environment by Acemoglu et al. (2012, 2016), among others. Earlier theoretical work on the topic includes Smulders and de Nooij (2003). The idea is that the direction of innovation is endogenous and determined by relative prices, market sizes and the accumulated stocks of knowledge of clean and dirty inputs. Recent contributions to this literature include Hassler et al. (2021), who focus on the scarcity of natural resources, Casey (2023), who focuses on energy efficiency and Acemoglu et al. (2023), who study the effects of the shale gas revolution. Aghion et al. (2023) model the effect of consumers' green preferences and competition on clean innovation, but do not model dirty technologies separately. Aghion et al. (2024) focus on the role of financial frictions for the green transition. Similarly to my model, they have heterogeneous degrees of path dependence in the direction of innovation across firms. In their model younger firms are cleaner than older firms, whereas in my model laggards tend to be cleaner than leaders (within sectors). Contrary to their model, I introduce strategic incentives for innovation. This paper adds to the theoretical literature on DTC and the environment by showing how the interaction of climate policy and market power affects both the direction and the intensity of innovation.

Policy-induced innovation is also at the center of the Porter Hypothesis, which states that environmental policy is not necessarily bad (and possibly even good) for competitiveness, as it creates a market for new−clean−technologies (Porter, 1990; Porter and van der Linde, 1995). Using a different mechanism I show how climate policy may increase market competition within sectors and in the aggregate economy.

To study the effect of within-industry competition on the direction of innovation, I build a model of strategic innovation that follows the structure of Akcigit and Ates (2023). The literature on this topic started with Aghion and Howitt (1992) and Grossman and Helpman (1991) and is centered around growth through creative destruction, in the tradition of Schumpeter's (1942) idea that economic growth is driven by new, productive firms that replace old, unproductive ones. These papers model technology as a ladder of quality, where each successful innovation leads a firm to improve its technology by one step. Firms compete to climb the ladder and gain market power. A key finding in this literature is that the free market may produce a growth rate that is too high from a welfare perspective because of the destructive nature of innovation. Instead of one technology ladder per sector, I have two−one clean, one dirty−which means that firms now also face a choice on the direction of their R&D. As long as carbon emissions are not priced, the model is similar to, e.g., the one by Akcigit and Ates (2023). A carbon tax changes the effective "height" of the dirty ladder by increasing the marginal costs of using the dirty technology, which may create different innovation incentives for leaders and laggards.

This paper also builds on the empirical literature on innovation, competition and the environment. Empirical work has shown that the direction of innovation (clean or dirty) is affected by the DTC mechanisms and by specific climate policies.[3] The influential work by Aghion et al. (2005) shows both theoretically and empirically the relationship between competition and innovation within industries, but does not consider the direction of innovation. Earlier work on this topic includes Blundell et al. (1995). Few papers assess the effect of competition on the direction of innovation. Bremer (2020) finds that the effect of electricity prices on clean innovation is positive and strengthens with market competition, whereas the effect of natural gas prices is positive but weakens with competition. Aghion et al. (2023) show that green preferences affect green innovation, especially when competition is strong. Aghion et al. (2024) show that credit constraints disproportionately affect young firms and green innovation, which is in line with earlier findings by Noailly and Smeets (2021).[4]

[3]For the effect of the DTC mechanisms on innovation see Newell et al. (1999), Popp (2002), Linn (2008), Noailly and Smeets (2015) and Aghion et al. (2016). For the effect of policies see Jaffe and Palmer (1997), Johnstone et al. (2010), Calel and Dechezleprêtre (2016) and Rozendaal and Vollebergh (2024). The empirical literature is summarized in Grubb et al. (2021).

[4]This paper relates to two other strands of literature that deal with strategic interaction, innovation and the environment. The first is mostly focused on the interaction between a fossil fuel cartel and a fuel importing country that may invest in renewables to replace fossil fuels (Dasgupta et al., 1983; Harris and Vickers, 1995;

The rest of this paper is organized as follows. Section 2 describes the data on patents and financials and then introduces the motivating facts. Section 3 builds the theoretical model. Section 4 shows the calibration and the quantitative exercises. Section 5 concludes.

2 Motivating facts

This section empirically analyses firms' innovation behavior using a large data set that links data on their finances to data on their patent applications. I use patent counts as a measure of innovation, as is common in the literature. I first briefly introduce the data. I then show descriptive evidence on innovation in technologies aimed at mitigating climate change, which are referred to as clean, and technologies relating to fossil fuels, which I refer to as dirty. I highlight five stylized facts which motivate a model of directed technical change and the environment with heterogeneous degrees of path dependence within industries and a direct link between climate policy, market power and innovation.

2.1 Data

This section briefly introduces the data. I make use of the Orbis Intellectual Property and Orbis Historical databases, both of which are managed by Bureau van Dijk. Orbis IP contains millions of patent applications, which can be linked using a firm identifier to financial data about the patents' applicants from Orbis Historical. I use patent data from 1978 until 2018 and financial data from 2010 until 2018. I use patent counts as a measure of innovation. Specifically, I count "triadic patent families" at the applicant level.[5] Critically, patents can be classified as clean, i.e., contributing to climate change mitigation, or dirty, i.e., related to fossil fuels, using their technology codes. I follow the most recent literature on classifying

Gerlagh and Liski, 2011; Jaakkola, 2019). The second strand uses real options theory to study the strategic timing of technology adoption and strategic investment in R&D for new product lines (Huisman and Kort, 2003, 2004; Compernolle et al., 2022; Dawid et al., 2023). I depart from these papers by studying strategic innovation behavior between direct competitors in a general equilibrium setting using an endogenous growth model.

[5] A patent family consists of all patent applications that protect a single invention. A family is classified as triadic if it consists of at least one patent application at the patent offices of the EU, Japan and the US. Focusing on triadic patent families eliminates low value patents from the sample.

clean and dirty patents (Jee and Srivastav, 2023).

Using the firm identifier I can link patent applications to balance sheet data and other financial variables about patent applicants. This allows me to compare the distribution of clean and dirty innovators across countries and sectors. It also provides me with a direct link between the direction of innovation and variables that reflect market power at the firm level, such as firm size, age and profitability.

Appendix A elaborates on the data in much more detail and shows trends in clean and dirty patenting, the composition of clean and dirty technology groups and patenting by applicant country. It also elaborates on the matching between the databases.

To measure past inventions, which are relevant for current inventions if innovation is path dependent, I follow the literature and define firm i's knowledge stocks as

$$K_{it}^{T} = \sum_{s=1978}^{t} P_{is}^{T}, \tag{1}$$

where P^T is the count of patents in technology group $T \in \{clean, dirty, all\}$.[6,7] To measure the direction of innovation let us define the innovation gap and the technology gap as

$$\text{Innovation gap}_{it} = \sinh^{-1}(P_{it}^{C}) - \sinh^{-1}(P_{it}^{D}), \tag{2}$$

$$\text{Technology gap}_{it} = \sinh^{-1}(K_{it}^{C}) - \sinh^{-1}(K_{it}^{D}). \tag{3}$$

I follow Acemoglu et al. (2023) and use the inverse hyperbolic sine transformation to accommodate zeroes in the patent counts and knowledge stocks.[8] The gaps above are approximately equal to log ratios of clean over dirty for positive values of both counts or both stocks.

To study the role of market power I use financial variables from Orbis Historical on revenue, employment, profit margin and firm age. I also define market leaders and laggards

[6] One can also discount past patents to account for the fact that knowledge becomes obsolete. Common rates of discounting are 10 and 20% per year. I do not use discounting to be consistent with my theoretical model but all the results shown in this section are qualitatively robust to including a 20% discount rate.

[7] To accommodate zero values when taking the log of the knowledge stock I follow Aghion et al. (2016) and set zeros to ones, and include a dummy variable that equals 1 in case the knowledge stock equals 0.

[8] The use of this transformation has recently been criticized by Mullahy and Norton (2022), especially when zeros are common (which is the case in my data). They recommend to use untransformed variables instead. The results in this section are qualitatively similar when I use untransformed variables. I choose to follow Acemoglu et al. (2023) as a baseline because the regressions with untransformed variables are more sensitive to outliers.

to draw comparisons between the two groups. Leaders are defined as the 10 largest firms in terms of revenue in a 2 digit industry (NACE Rev. 2) in a country in a given year, and laggards are those firms ranked 11 until 20.[9] Finally, to measure market competition at the sector level I define a concentration index as the total revenue of the top 10 firms divided by the total revenue of the top 20 firms (i.e., the revenue of market leaders over the combined revenue of leaders and laggards).

2.2 Stylized facts

This section documents five facts about clean and dirty patenting and market power, some of them novel to the literature. Together, these facts motivate the development of a new model of directed technical change and the environment, which allows for (i) a direct effect of climate policy on market power, (ii) a relationship between market power and innovation and (iii) heterogeneous degrees of path dependence across firms and sectors.

Fact 1. *Both clean and dirty patents are filed by firms that are active across a wide range of sectors.*

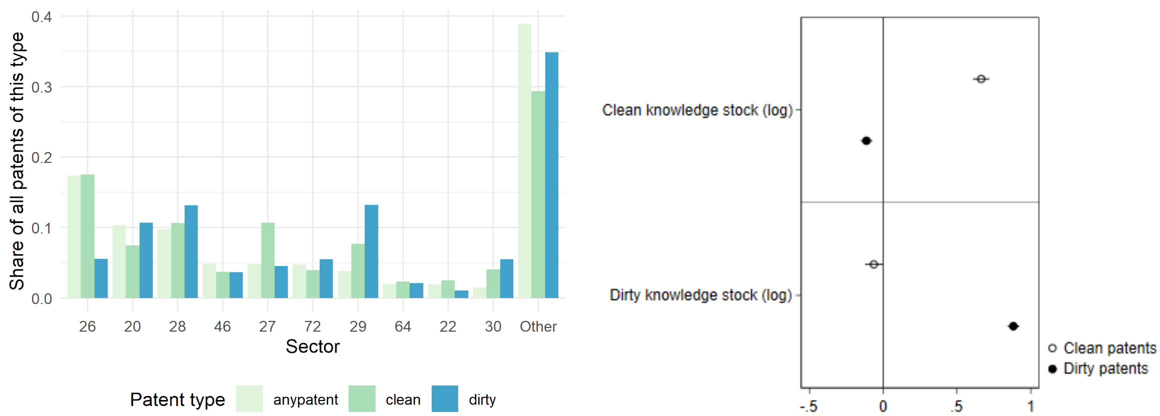
Perhaps unsurprisingly, given the diversity of technologies that can be classified as clean or dirty (see Figures A3 and A4 in the Appendix), both clean and dirty patent applicants are active in a diverse set of sectors. Figure 1a shows that no single sector dominates in clean or dirty patenting. While energy technologies constitute a substantial share of both clean and dirty patents, firms in the energy sector are not among the most active patent applicants, which suggests that these firms largely depend on others to supply them with new technologies. This finding also suggests that papers that focus only on the energy sector miss a large part of clean innovation which may be a limitation, depending on the precise goal of the paper.

Fact 2. *The direction of innovation is path dependent.*

[9] I use 2 digit rather than 4 digit industries because these are large firms that may be active in multiple (narrow) industries. I only use the top 20 firms because coverage of large firms in Orbis Historical is better and more consistent across countries and years than coverage of small firms. The results in this section are robust to various changes in the definition of leaders and laggards.

Figure 1: Motivating facts (I)

(a) Patents by applicant sector **(b)** Path dependence in innovation



Notes: Data sources: Orbis IP and Historical. Panel 1a shows the distribution of patents by applicant sector. Anypatent refers to the entire sample (including clean, dirty and neutral). Sectors are classified using the NACE Rev. 2 classification. The sectors in the figure are the following. 26: Manufacture of computer, electronic and optical products; 20: Manufacture of chemicals and chemical products; 28: Manufacture of machinery and equipment n.e.c.; 46: Wholesale trade, except of motor vehicles and motorcycles; 27: Manufacture of electrical equipment; 72: Scientific research and development; 29: Manufacture of motor vehicles, trailers and semi-trailers; 64: Financial service activities, except insurance and pension funding; 22: Manufacture of rubber and plastic products; 30: Manufacture of other transport equipment. Panel 1b shows the results of Poisson regressions of clean and dirty patents on the first lag of the clean and dirty patent stock. Table E1 in the Appendix reports the full results of these regressions. Standard errors are clustered at the firm level. The sample covers the years 1978-2018.

Path dependence in innovation is well established. For instance, Blundell et al. (1995) find strong evidence that stocks of accumulated past patents predict patenting behavior by firms.[10] In the context of clean and dirty innovation, Aghion et al. (2016) find a strong effect of clean (dirty) knowledge stocks on clean (dirty) innovation in a sample of firms innovating in car technologies. Aghion et al. (2019) discuss the topic in more detail and conclude that it is difficult to shift the innovation system from dirty to clean technologies due to inertia caused by past investments.

Figure 1b shows the results of Poisson regressions of firms' clean and dirty patenting on

[10]In fact, they develop a fixed effects estimator that uses past patenting as a proxy for the firm fixed effect and show that this method controls for unobserved differences between firms.

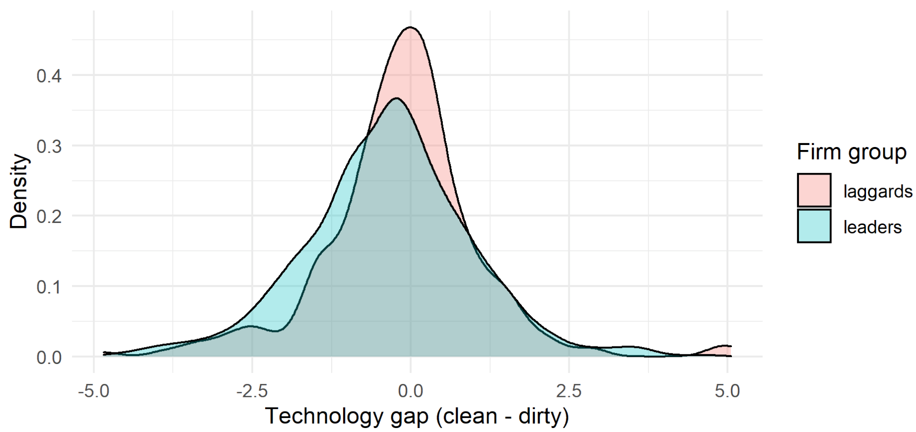
the first lag of their clean and dirty knowledge stocks. Consistent with the literature, but using a much larger sample, I find that having applied for many clean (dirty) patents in the past means that a firm is more likely to apply for a clean (dirty) patent today. Variation in knowledge stocks explains over 50% of the variation in clean and dirty patenting. Moreover, I find that a firm with a large stock of dirty (clean) patents is less likely to apply for a clean (dirty) patent, conditional on its stock of clean patents. Table E1 in the Appendix shows that, conditional on country-sector-year fixed effects, the innovation gap is positively correlated with the clean knowledge stock and the technology gap, and negatively with the dirty knowledge stock.[11] These findings indicate strong path dependence in the direction of innovation, suggesting that firms with large dirty knowledge stocks (relative to clean) need a larger incentive to switch from dirty to clean innovation.

Fact 3. *Market leaders are, on average, more invested in dirty technologies than their direct competitors.*

Perhaps the most important and novel finding in this section is that leaders tend to be dirtier than the laggards in the same sector. Figure 2 plots the distribution of the technology gap (clean - dirty) for leaders and laggards in 2018. Technology gaps close to zero are much more common for laggards than for leaders. Large negative gaps, particularly those between -2.5 and -1 are much more common for leaders than for laggards, whereas this is not the case for large positive gaps. Figure 2 suggests that dirty leaders tend to be more invested in dirty technologies than their direct competitors, whereas clean leaders are not more invested in clean technologies than clean laggards. Columns 1 and 2 in Table E2 in the Appendix confirm this finding. This table shows the results of a set of regressions of the technology gap on indicators of market power, controlling for country-sector-year fixed effects and thus comparing direct competitors. Column 1 shows that larger firms tend to have a more negative technology gap, meaning that they are more invested in dirty technologies. While employment, profits and age are not statistically significant when conditioning on revenue, Table E3 shows that each of these variables is negatively correlated with the technology gap

[11] Since the innovation gap is approximately normally distributed, I use the OLS estimator, which allows me to control for country-sector-year fixed effects. The Poisson fixed effects estimator requires strict exogeneity of the independent variables. Since knowledge stocks are based on past values of the dependent variable, they do not satisfy this requirement. Hence, I do not include fixed effects in the Poisson regressions.

Figure 2: Technology gap distribution for leaders and laggards



Notes: Data sources: Orbis IP and Historical. The technology gap is defined in (3). Leaders are the top 10 firms in terms of revenue in their 2 digit NACE Rev. 2 industry and country, laggards are the firms ranked 11 until 20 in those same industries. Graph is for the year 2018 and includes only firms that applied for at least one patent in that year.

when included separately.[12] Very much in line with this finding, column 2 in Table E2 shows that leaders tend to have a lower technology gap than the firms that are classified as laggards, and the remaining firms (which are not in the top 20 in their industry).

Fact 4. *There is some evidence that differences in technology gaps between leaders and laggards are correlated with market competition and the direction of innovation.*

The findings above suggest that leaders and laggards may respond differently to climate policies. That is, the results suggest that leaders need a larger incentive to switch from dirty to clean innovation because they are more invested in dirty technologies. There is some evidence that these differences between direct competitors are heterogeneous across two dimensions, which is relevant for modeling the market power and innovation effects of climate policy as accurately as possible. Figure 3a shows that, pooling all leader and laggard firms with negative technology gaps in 2018, technology gaps tend to get more negative with market concentration for laggards but not for leaders.[13]

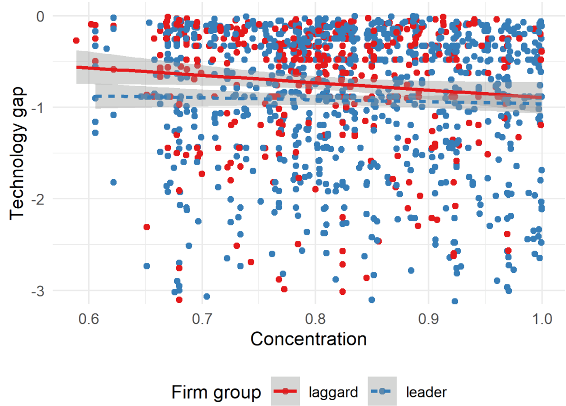
Table E4 in the Appendix shows more formally that, within sectors, the difference in

[12]Figure E1 shows that leaders, who have higher revenues than laggards by definition, are, on average, also larger in terms of employees, more profitable and older than laggards.

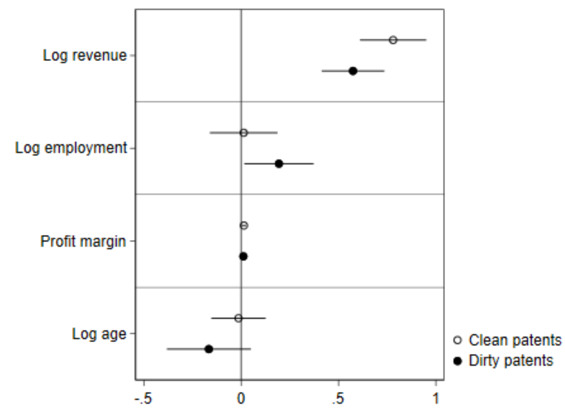
[13]The difference between the trend lines is small but statistically significant for the range of concentration

Figure 3: Motivating facts (II)

(a) Competition and tech. gaps for dirty firms



(b) Firm size and innovation



Notes: Data sources: Orbis IP and Historical. Panel 3a plots technology gaps as defined in (3) against industry concentration for leaders and laggards with a negative technology gap. The sample includes all firms that are classified as leader or laggard in the year 2018 and that applied for at least one patent in the past. Panel 3b plots the coefficients from two Poisson regressions of clean and dirty patents on various indicators of firm size and market power. The regressions include country-sector-year fixed effects. Standard errors are clustered at the firm level. The sample covers the years 2010-2018. The sample consists of all firms that have applied for at least one clean or dirty patent between 1978 and 2018 and for which the financial variables were available. Table E2 in the Appendix also shows these results.

technology gaps between leaders and laggards is different between high and low concentration sectors, and between clean and dirty sectors. In particular, leaders in highly concentrated industries are more invested in dirty technologies than leaders in less concentrated industries. Moreover, leaders in sectors in which the median firm has a negative technology gap are more invested in dirty technologies than their competitors, whereas the technology gaps of leaders and laggards are not statistically significantly different in clean sectors. While the results in Table E4 indicate clear heterogeneity across sectors, potential measurement error in both concentration and technology gaps warrants caution in interpreting these findings. Still, heterogeneity across the dimensions of market competition and innovation direction means that climate policy may affect innovation differently across industries. Calibrating the joint

between roughly 0.63 and 0.86. Figure E2 in the Appendix shows no such difference for firms with a positive technology gap.

distribution of market competition and the degree to which leaders and laggards are invested in dirty technologies takes such differences into account.

Fact 5. *Both clean and dirty patenting correlate positively with firm size.*

The findings above may suggest that clean innovation is mostly driven by small firms. That is not the case at all, according to Figure 3b. Here, I use a Poisson regression to show that firm size and profitability are positively correlated with clean patenting on the intensive margin, conditional on country-sector-year fixed effects. The same is true for dirty patenting. Table E2 in the Appendix also shows these results. These results are in line with findings that large firms tend to patent more in general (Akcigit and Kerr, 2018) and that clean innovation in Sweden is driven mostly by large firms (Ustyuzhanina et al., 2022). They suggest that theoretical models of directed technical change and the environment should allow for a relationship between firm size and innovation intensity for both clean and dirty innovation.

The empirical findings above highlight several patterns in the data that existing models of directed technical change and the environment cannot explain or incorporate. First, while a large literature shows that the competitive environment in which firms operate affects their innovation behavior (e.g., Aghion et al., 2005), market power and strategic incentives are absent in most of the existing DTC-environment models.[14] The finding that, within sectors, leaders tend to be more invested in dirty technologies than laggards makes this limitation especially relevant. It suggests that climate policy may affect directly competing firms differently, and may thus affect market power and strategic innovation incentives. Incorporating strategic incentives into a DTC-environment model allows me to capture effects that other models cannot capture. Second, the empirical findings indicate a rather complex heterogeneity across sectors in three variables: the degree of competition, the technology gap of leaders and the technology gap of laggards. I capture this heterogeneity in my theoretical model by calibrating the joint distribution across those three variables using micro data.

[14] Acemoglu et al. (2016) and Aghion et al. (2024) do have have models in which firms innovate to gain market power. However, since they rely on a Klette and Kortum (2004) type of structure in which any firm can improve over any product line, firms in these models do not have direct competitors. Hence, there is no strategic incentive for incumbents to deter a competitor's innovation by innovating more itself.

3 Model

This section develops a model of directed technical change and the environment with strategic innovation incentives. The model incorporates all of the five stylized facts documented above. It is in continuous time and features a continuum of intermediate input sectors, all of which produce their product using either a clean or a dirty technology. In the model, the direction of innovation is path dependent at the firm level, and the degree to which firms are invested in clean or dirty technologies is heterogeneous across firms (within sectors) and across sectors. The model allows for this heterogeneity across sectors to be correlated with the degree of market competition. Finally, the intensity of innovation depends on strategic innovation incentives and is correlated, within sector, with firm size. The starting point of the analysis is the model in Akcigit and Ates (2023), to which I add climate change and directed technical change in each intermediate input sector. That means that firms decide not only on the intensity of their innovation efforts but also on the direction.[15]

A main difference between this paper and existing models on DTC and the environment is that I integrate climate change and directed technological change in a model of strategic innovation (Akcigit and Ates, 2023), rather than relying on the Klette and Kortum (2004) structure of innovation. To my knowledge, I am the first to take this approach. The main difference between the two approaches is that Klette and Kortum (2004) models assume that firms are active in multiple intermediate input markets,[16] while the strategic innovation models assume that firms are active in a single market. The former type of models then assume that a successful innovation leads a firm to take over a random product line and add it to its portfolio. In the latter type of models a successful innovation leads to a stepwise improvement to a firm's own product. This distinction has several important implications. First,

[15] Various versions exist of Akcigit and Ates (2023) model. Earlier versions are Akcigit and Ates (2019) and Akcigit and Ates (2021). The 2023 paper has the most general version of the model. The 2021 paper uses a simplified version of the model and focuses more on 10 stylized facts about business dynamism in the US. The 2019 paper is the working paper version of the 2023 one. I take elements from all three papers. Like Akcigit and Ates (2019), I have constant innovation step sizes. Like Akcigit and Ates (2019, 2021), I model the goods by the two producers within a sector as perfect substitutes. Like Akcigit and Ates (2019, 2023), I do not restrict the technology gap between firms within a sector to 1 (in most of this section) and I have R&D costs in terms of labor. Like Akcigit and Ates (2021), I shut down entry and exit and radical innovations.

[16] Intermediate inputs are also referred to as product lines in these models.

strategic innovation models allow for a more detailed modeling of competition and innovation incentives within sectors, implementing mechanisms from the industrial organization literature. Second, strategic innovation models do not rely on the assumption that a successful innovation leads to an improvement to a random intermediate input, which arguably makes them more intuitive because innovation efforts are targeted.[17,18]

Most of this section follows Akcigit and Ates (2023) quite closely. It first describes consumers and producers, then static competition and outcomes, and finally dynamic decisions to arrive at the seemingly paradoxical result that the introduction or increase of a carbon tax may increase dirty innovation by some firms. I then aggregate across sectors to solve for the general equilibrium in closed form.

3.1 Model setup

3.1.1 Consumers

The representative consumer's utility is

$$U_t = \int_{s=t}^{\infty} \exp\big(-\rho(s-t)\big)\ln(C_s)ds, \tag{4}$$

where C is consumption and $\rho > 0$ is the rate of time preference. The budget constraint is

$$P_t C_t + \dot{A}_t = w_t L_t + r_t A_t + G_t, \tag{5}$$

where labor L is supplied inelastically, A are asset holdings and G denotes lump-sum taxes or transfers from the government. I normalize $L_t = 1$. The price of the final good, P, is normalized to 1 and w and r denote wage and interest rate, respectively. Households own the firms, so $A_t = \int_{\mathcal{F}} V_{ft} df$, where $\mathcal{F}$ denotes the set of firms and V_f is the value of firm f.

[17] That is, if a steel manufacturer invests in innovation, it can improve a clean steelmaking process but not a random clean technology like solar panels or electric vehicles.

[18] Another implication of the assumption that innovation happens within sectors is that it allows for the persistent existence of sectors that are hard (costly) to abate. Clean technologies can only become competitive in these sectors as a result of multiple successful clean innovations or if climate policy is made very stringent. Acemoglu et al. (2016) do allow for sectors to vary in the technological distance between clean and dirty (they estimate these distances for a set of industries in their Section 3D), but in their model a successful (radical) clean innovation lets a clean technology overtake a dirty technology in a random product line, irrespective of the initial technological distance between the two or the degree of competition in that market.

3.1.2 Final good production

Final good Y is produced from a continuum of intermediate goods y, and global warming causes losses in output:

$$\ln Y_t = -\frac{\gamma}{2}T_t^2 + \int_0^1 \ln y_{jt} dj, \tag{6}$$

where j indicates an intermediate good and T_t is the temperature increase since pre-industrial times. While there is much uncertainty about the correct damage function in economic models and the empirical evidence is heavily criticized (Pindyck, 2013, 2021), I follow Dietz and Venmans (2019) in modeling damages as exponential-quadratic. In an extensive survey of the literature, Nordhaus and Moffat (2017) show that this function fits the available data well.

Final good production is perfectly competitive and the final good is used for consumption only. That is, the resource constraint is $Y_t = C_t$. The Euler equation that results from the household problem and the resource constraint is

$$r_t = g_t + \rho, \tag{7}$$

where $g_t \equiv \frac{\dot{Y}_t}{Y_t}$ is the growth rate of aggregate output.

3.1.3 Global warming

For the climate part of the model I follow the most recent economic literature on the topic, which exploits recent advances in climate science. Specifically, I follow Dietz and Venmans (2019) and model temperature as linear proportional to cumulative emissions of CO_2. This approach has the benefit of being simple and intuitive while not suffering from several inconsistencies with climate science models that are present in many Integrated Assessment

Models (IAMs).[19] Cumulative emissions are denoted by S_t:

$$S_t = \int_0^t E_s ds, \tag{8}$$

where E_s denotes the flow of emissions at time s.

Global warming since pre-industrial times T_t follows the following law of motion,

$$\dot{T}_t = \varepsilon(\zeta S_t - T_t), \tag{9}$$

where ζ is the Transient Climate Response to Cumulative Carbon Emissions (TCRE), which is essentially the slope of the linear relationship between cumulative emissions and warming, and ε is the "initial pulse-adjustment timescale of the climate system" (Dietz and Venmans, 2019). That is, it parameterizes the delay between emissions and warming.

3.1.4 Intermediate good production

Most of the action in this model happens within intermediate good sectors, which are duopolies. Firms compete on prices and use either a clean or a dirty technology to produce. As a result, the interaction between climate policy and the technology gaps between the firms in the two different technologies determines whether firms innovate in clean or dirty technologies.

Each intermediate good sector j consists of two firms, i and $-i$. Total production of intermediate good j is simply the sum of the two firms' outputs:

$$y_{jt} = y_{ijt} + y_{-ijt}. \tag{10}$$

A firm produces its intermediate good j using either a clean or a dirty technology:

$$y_{ijt} = y^C_{ijt} + y^D_{ijt} = q^C_{ijt} l^C_{ijt} + q^D_{ijt} \min\left\{ l^D_{ijt}, \frac{e_{ijt}}{\kappa} \right\}, \tag{11}$$

[19] Most IAMs model the carbon cycle with various degrees of detail (e.g., Nordhaus 2014; Golosov et al. 2014). Dietz et al. (2021) show that most of these models overestimate the delay between emissions and warming and abstract away from positive feedback effects. In addition, temperatures start declining after peaking too late, which is inconsistent with climate models. Empirical evidence from climate science has shown that the relationship between temperature and cumulative emissions is approximately linear and that the delay between emissions and warming is short (about 10 years between emission and full effect on temperature) (e.g., Matthews et al. 2009; Ricke and Caldeira 2014). Hence, to be consistent with climate science, Dietz et al. (2021) recommend taking the approach of Dietz and Venmans (2019) and simply model temperature as a linear function of cumulative emissions.

where C indicates clean and D indicates dirty, q is the technology, l is labor and e is emissions. κ is a scaling parameter. Essentially, each intermediate good can be produced using a clean technology, which uses only labor, or using a dirty technology, which uses labor and emits carbon. The two types of the good (clean and dirty) are perfect substitutes. Note that the dirty good producer optimally chooses $e_{ijt} = \kappa l_{ijt}^D$. I omit the the subscript j until aggregate variables are discussed.

Total costs are

$$TC_{it} = TC_{it}^C + TC_{it}^D = w_t l_{it}^C + w_t l_{it}^D + \tau_t^E e_{it} = w_t l_{it}^C + w_t(1 + \kappa\tau_t) l_{it}^D, \tag{12}$$

where w is the wage and $\tau_t^E = \tau_t w_t$ is the emission tax. Since dirty production is Leontieff in labor and emissions, the carbon tax is essentially a tax on the wage bill of firms that use their dirty technology. Note that by picking τ, the policymaker fixes the relative factor prices of labor and emissions. I will refer to τ as a carbon tax, though a more precise label would be the price of emissions relative to labor. Specifying the tax in this way, rather than as a fixed charge per unit of emissions, drastically simplifies the dynamic problem, as well as the general equilibrium outcomes.

Marginal costs are as follows,

$$MC_{it} = \min\{MC_{it}^C, MC_{it}^D\} = \min\Big\{\frac{w_t}{q_{it}^C}, \frac{w_t(1 + \kappa\tau_t)}{q_{it}^D}\Big\}. \tag{13}$$

Firms can innovate to increase their total factor productivity q. A successful innovation increases a firm's technology by factor $\lambda > 1$:

$$q_{i(t+\Delta t)}^F = \lambda q_{it}^F, \tag{14}$$

with $F \in \{C, D\}$. Assuming firms start from $q_{i0}^F = 1$, we can write $q_{it}^F = \lambda^{n_{it}^F}$, where n_{it}^F is the number of innovation steps that firm i has taken for technology F at time t.

Let us define the technology gaps as follows. $m_{it}^T = n_{it}^C - n_{it}^D$ is the firm's own technology gap, i.e. the difference between its own clean and dirty technologies. $m_{it}^C = n_{it}^C - n_{-it}^C$ and $m_{it}^D = n_{it}^D - n_{-it}^D$ are the clean and dirty technology gaps, respectively, and show how many steps firm i is ahead of its competitor in the given technology.

Maximizing profits, a firm chooses to use the technology that allows it to produce at the lowest marginal costs. That is, it uses the clean technology if and only if

$$\begin{aligned}
\frac{w_t}{q_{it}^C} &\leq \frac{w_t(1+\kappa\tau_t)}{q_{it}^D} \\
\lambda^{n_{it}^D - n_{it}^C} = \lambda^{-m_{it}^T} &\leq 1+\kappa\tau_t \\
-m_{it}^T \ln(\lambda) &\leq \ln(1+\kappa\tau_t) \\
-m_{it}^T &\leq \frac{\ln(1+\kappa\tau_t)}{\ln(\lambda)} \equiv \tilde{\tau}_t \\
m_{it}^T + \tilde{\tau}_t &\geq 0. \qquad (15)
\end{aligned}$$

That is,

$$y_{it} = \begin{cases} y_{it}^C & \text{if} \quad m_{it}^T + \tilde{\tau}_t \geq 0 \\ y_{it}^D & \text{if} \quad m_{it}^T + \tilde{\tau}_t < 0 \end{cases}. \qquad (16)$$

Note that $\tilde{\tau}$ essentially expresses the carbon tax in terms of the minimum number of innovation steps a firm's dirty technology needs to be ahead of its clean technology for the firm to still use dirty. Since marginal costs do not depend on quantity and clean and dirty are perfect substitutes, firms use either clean or dirty (whichever has lower marginal costs), but never a combination of both technologies.

Next, let us define the effective technology gap, which will be central to the rest of the analysis. Given that firms only use the technology that has lower marginal costs the definition of the effective gap is

$$m^E(m_{it}^C, m_{it}^D, m_{it}^T, \tau_t) = \begin{cases} m_{it}^C & \text{if} \quad m_{it}^T + \tilde{\tau}_t \geq 0, \quad m_{-it}^T + \tilde{\tau}_t \geq 0 \\ m_{it}^D + m_{it}^T + \tilde{\tau}_t & \text{if} \quad m_{it}^T + \tilde{\tau}_t \geq 0, \quad m_{-it}^T + \tilde{\tau}_t < 0 \\ m_{it}^C - m_{it}^T - \tilde{\tau}_t & \text{if} \quad m_{it}^T + \tilde{\tau}_t < 0, \quad m_{-it}^T + \tilde{\tau}_t \geq 0 \\ m_{it}^D & \text{if} \quad m_{it}^T + \tilde{\tau}_t < 0, \quad m_{-it}^T + \tilde{\tau}_t < 0 \end{cases}. \qquad (17)$$

So, while technology gaps m^C, m^D, m^T are always integer values (there are no half innovations), a positive carbon tax can create an effective technology gap that is not an integer if the two firms use different technologies (one clean, the other dirty).

At any point in time, firms can invest in either clean or dirty innovation. Firms hire workers to work in R&D. The production function of ideas is

$$x_{it} = \left(\beta \frac{h_{it}}{\alpha}\right)^{\frac{1}{\beta}}, \qquad (18)$$

where x is the arrival rate of a new invention (moving the firm a step up the clean or dirty technology ladder), h is the number of R&D workers employed, β determines the shape of the function, and α is used for scaling. Research workers earn the same wage as production workers, meaning that a firm's total R&D investments are

$$R_{it} = \alpha \frac{x_{it}^{\beta}}{\beta} w_t. \tag{19}$$

Like Akcigit and Ates (2023), I model knowledge diffusion from a market leader to a laggard as an exogenous arrival rate δ, which brings the effective technology gap down to 0. Note that when knowledge diffusion takes place, the laggard firm becomes equally good at the technology that the market leader is using. Hence, if the two firms were using different technologies before diffusion took place, then diffusion causes the (former) laggard to switch to the (former) leader's technology.

3.1.5 Graphical representation

Figure 4 presents a graphical representation of the main (static) elements of the model. The most important difference with the bulk of papers in the stepwise innovation literature is that firms compete on two innovation ladders, rather than one, and the two ladders interact in the sense that an innovation step on one ladder can lead a firm to take over the market from a firm that leads on the other ladder. Panel 4a shows firm 1's clean, dirty and own technology gaps ($m_{1t}^C, m_{1t}^D, m_{1t}^T$, respectively). In this case, firm 2 leads in the clean technology (left ladder) and firm 1 leads in the dirty technology (right ladder). Firm 1's own technology gap is -2 (its clean technology is 2 steps behind its dirty one) and firm 2's own technology gap is 0. Firm 1 leads this market and uses the dirty technology to produce. The effective technology gap is equal to 1.

Panel 4b shows the effects of a low and a high carbon tax ($\tilde{\tau}_L$ and $\tilde{\tau}_H$, respectively). Firm 1's dirty technology is the most productive, but the carbon tax increases firm 1's marginal costs. With the low tax, firm 1 remains market leader, as its dirty technology is still cheaper to operate than firm 2's clean technology. The effective technology gap is reduced from 1 to 0.5 by the introduction of the tax. With the high tax, firm 2 becomes the market leader, as its clean technology is now cheaper to operate than firm 1's dirty technology. The effective gap is -0.5 from the perspective of firm 1.

Finally, panel 4c shows the effects of a successful innovation. In this case, firm 1 innovates in the dirty technology, while firm 2 innovates in clean.

3.2 Equilibrium

3.2.1 Static competition and sector level outcomes

Intermediate good demand follows from the final good producer's problem:

$$y_{jt} = \frac{Y_t}{p_{jt}}, \tag{20}$$

which implies that the final goods producer spends an equal amount Y_t on each intermediate input j.

Each market consists of a duopoly with Bertrand competition. There can be two types of markets, namely neck-and-neck (or leveled) markets, where the two firms' marginal costs are equal, and unleveled markets, where one firm has lower marginal costs than its competitor. I will refer to the firm that has lower marginal costs than its competitor as the market leader and to the other firm as the laggard. Note that a firm's effective technology gap m_{it}^E is positive if it is the market leader, zero if it is a neck-and-neck firm and negative if it is a laggard:

$$MC_{it} < MC_{-it} \iff m_{it}^E > 0. \tag{21}$$

Intuitively, if both firms use the same technology, the more productive firm in the relevant technology is the market leader. If the firms use different technologies, then the tax and the gap between one firm's clean and the other firm's dirty technology determine who leads.

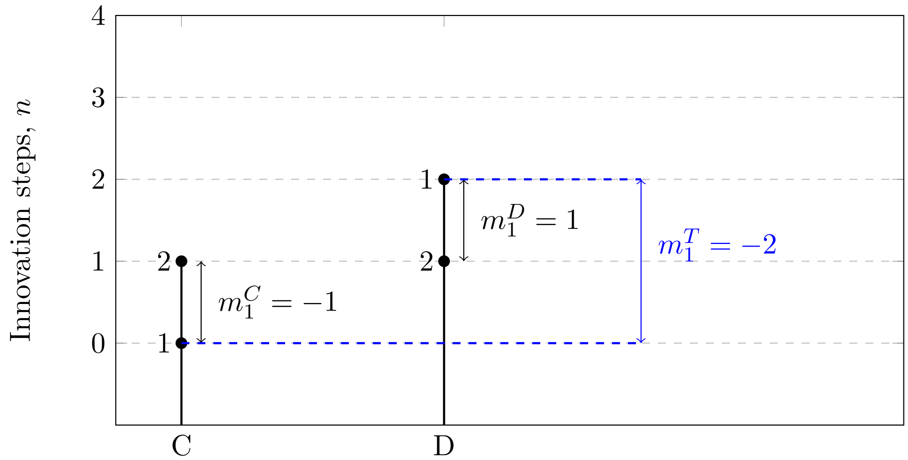
Since we are in a setting of Bertrand competition, firms engage in limit pricing. That is, the market leader sets its price equal to its competitor's marginal costs and supplies the entire market.

$$p_{jt} = \begin{cases} MC_{-it} & \text{if} \quad m_{it}^E \geq 0 \\ MC_{it} & \text{if} \quad m_{it}^E \leq 0 \end{cases}. \tag{22}$$

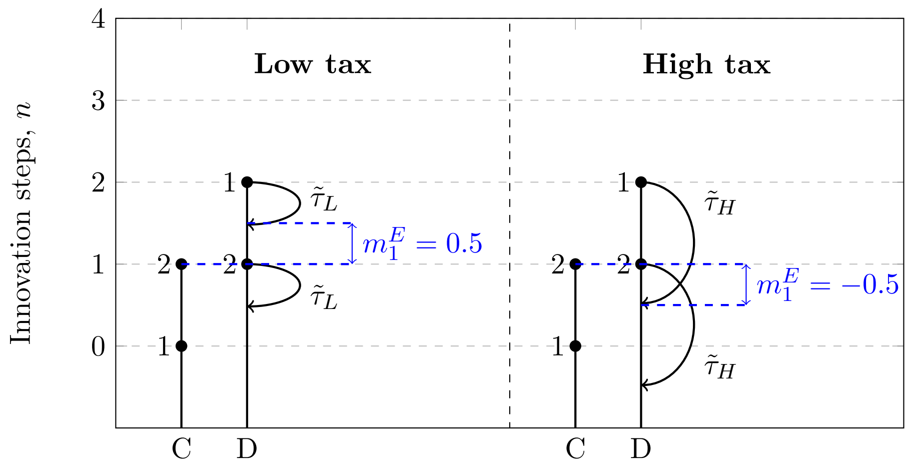
The laggard makes zero profits, while the leader can charge a markup that depends on its effective technology gap with respect to its competitor. Neck-and-neck firms each supply

Figure 4: Graphical representation of the model

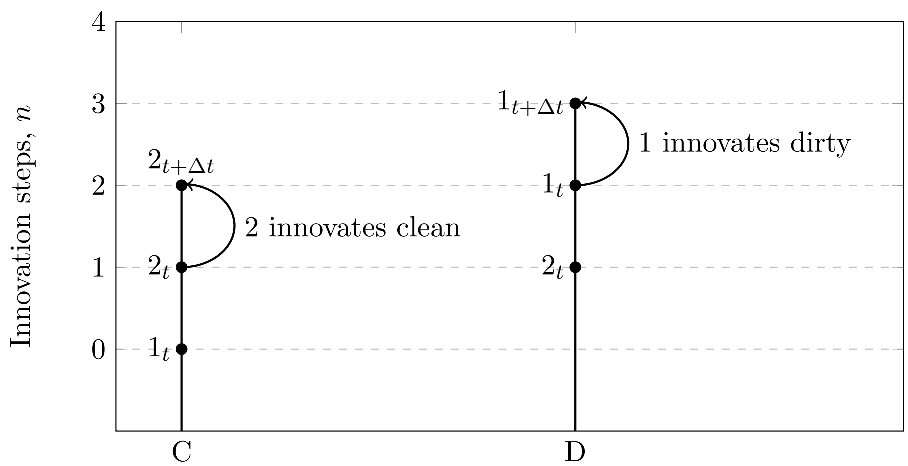
(a) Own, clean and dirty technology gaps



(b) A carbon tax affects the effective technology gap



(c) Clean and dirty innovation



half of the market at a price equal to marginal costs and thus make zero profits. Let π denote profits excluding expenditures on R&D, i.e. revenues minus production costs. Using (20) and (22) gives

$$\pi(m_{it}^E) = \begin{cases} (p_{jt} - MC_{it})y_{it} = \left(1 - \frac{MC_{it}}{MC_{-it}}\right)Y_t = \left(1 - \frac{1}{\lambda^{m_{it}^E}}\right)Y_t & \text{if} \quad m_{it}^E > 0 \\ 0 & \text{if} \quad m_{it}^E \leq 0 \end{cases}. \tag{23}$$

That is, a firm with an effective technology lead of m_{it}^E can charge a markup of price over marginal costs equal to $\frac{p_{it}}{MC_{it}} = \lambda^{m_{it}^E}$.

In this model the degree of competition and the degree of market power in an intermediate input sector are exact opposites and fully determined by the effective technology gap. An increase in competition or a decrease in market power in this context means a decrease in the effective technology gap and thus in the market leader's markup. This interpretation of competition and market power follows Akcigit and Ates (2023) and differs from Aghion et al. (2005, 2023), who interpret the degree of market competition as the extent to which neck-and-neck firms can collude. In my model an effective gap of 0, which means that a market is neck-and-neck, is essentially equivalent to perfect competition in the sense that firms set their price equal to marginal costs and make no profits.

A firm's total output also follows directly from (20) and (22):

$$y_{it} = \begin{cases} \frac{Y_t}{MC_{-it}} & \text{if} \quad m_{it}^E > 0 \\ \frac{Y_t}{2MC_{it}} & \text{if} \quad m_{it}^E = 0 \\ 0 & \text{if} \quad m_{it}^E < 0 \end{cases}. \tag{24}$$

Firms' labor demand and emissions follow from the production function (11), together with (24) and the firm's decision to use clean or dirty (15):

$$l_{it} = \begin{cases} \frac{y_{it}}{q_{it}^C} & \text{if} \quad m_{it}^T + \tilde{\tau}_t \geq 0 \\ \frac{y_{it}}{q_{it}^D} & \text{if} \quad m_{it}^T + \tilde{\tau}_t < 0 \end{cases}, \tag{25}$$

$$e_{it} = \begin{cases} 0 & \text{if} \quad m_{it}^T + \tilde{\tau}_t \geq 0 \\ \frac{\kappa y_{it}}{q_{it}^D} & \text{if} \quad m_{it}^T + \tilde{\tau}_t < 0 \end{cases}. \tag{26}$$

Appendix C.1 shows a full characterization of prices, profits, output, labor and emissions, plugging in the different combinations of clean and dirty production by the two firms.

3.2.2 Firm values and innovation

For the remainder of this paper, let us focus on the case where τ_t is fixed over time and any changes to it are immediate, permanent and unanticipated. Firms maximize the net present value of lifetime profits by choosing innovation efforts x, which are either clean or dirty. Note that, at a given point in time, a firm can innovate only in one of the two technologies. Hence, if τ is fixed over time, firms invest only in the technology for which their marginal costs are lower given τ. That is, they invest in clean R&D if $m^T_{it} + \tilde{\tau} \geq 0$ and in dirty R&D otherwise.[20] The firm's maximization problem is

$$\max_{x^{\infty}_{i,s=t}} \int_{s=t}^{\infty} \exp\big(-r_s(s-t)\big)\big(\pi(m^E_{is}, Y_s) - \alpha \frac{x^{\beta}_{ms}}{\beta} w_s\big) ds.$$

To solve the dynamic problem of the firm, let us write down the value function for a firm that has an effective technology gap of $m^E_{it} \in [-\overline{m}, \overline{m}]$, where $\overline{m}$ is the maximum number of steps a firm can be ahead in a particular technology. Having a maximum technology gap is necessary to have a finite number of states (for a given τ). In most of this paper I consider $\overline{m}$ to be large. In some applications I set it to 1 to derive analytical results. Dropping the subscripts and superscripts on m^E_{it} for convenience gives the following value function,

$$\begin{aligned} r_t V_{mt} - \dot{V}_{mt} = \max_{x_{mt}} \Big\{ & \pi(m) - \alpha \frac{x^{\beta}_{mt}}{\beta} w_t + x_{mt}[V_{m+1,t} - V_{mt}] \\ & + x_{-mt}[V_{m-1,t} - V_{mt}] + \delta[V_{0,t} - V_{mt}] \Big\}, \end{aligned} \tag{27}$$

where the first term on the right hand side captures operating profits (23) and the second term is R&D expenditures (19). The third term captures that by investing in innovation, the firm increases its technology gap by 1 step with arrival rate x_{mt}. The fourth term captures that the firm's competitor is also investing in R&D and moves a step up with arrival rate x_{-mt}. The final term captures the exogenous arrival of knowlegde diffusion, which brings the effective technology gap down to 0.

Next, let us follow Akcigit and Ates (2023) and normalize firm values by aggregate output. Define the normalized value of a firm that is m steps ahead as $v_{mt} \equiv \frac{V_{mt}}{Y_t}$. Then, substituting

[20] Profits depend only on Y_t, which is outside the firm's control, and m^E_{it}, which it wants to maximize. Innovating in the technology at which it is already better increases m^E_{it} by 1, whereas innovating in the other technology increases m^E_{it} by less than one (possibly by 0). Hence, the firm innovates in the technology at which it is already better.

for V in (27), dividing by Y_t, and using the Euler equation (7) gives

$$\rho v_{mt} - \dot{v}_{mt} = \max_{x_{mt}} \Big\{ 1 - \frac{1}{\lambda^m} - \alpha \frac{x_{mt}^{\beta}}{\beta}\omega_t + x_{mt}[v_{m+1,t} - v_{mt}] \\ + x_{-mt}[v_{m-1,t} - v_{mt}] + \delta[v_{0,t} - v_{mt}] \Big\}, \tag{28}$$

for market leaders ($m > 0$) and

$$\rho v_{mt} - \dot{v}_{mt} = \max_{x_{mt}} \Big\{ - \alpha \frac{x_{mt}^{\beta}}{\beta}\omega_t + x_{mt}[v_{m+1,t} - v_{mt}] \\ + x_{-mt}[v_{m-1,t} - v_{mt}] + \delta[v_{0,t} - v_{mt}] \Big\} \tag{29}$$

for laggards and neck-and-neck firms ($m \leq 0$). Note that $\omega_t \equiv \frac{w_t}{Y_t}$ denotes the normalized wage.

Taking the first order conditions yields the innovation efforts of a firm that has an effective technology gap $m_{it}^E = m$:

$$x_{mt} = \Big(\frac{1}{\alpha \omega_t} [v_{m+1,t} - v_{mt}] \Big)^{\frac{1}{\beta - 1}}, \tag{30}$$

which gives firms' R&D worker demand

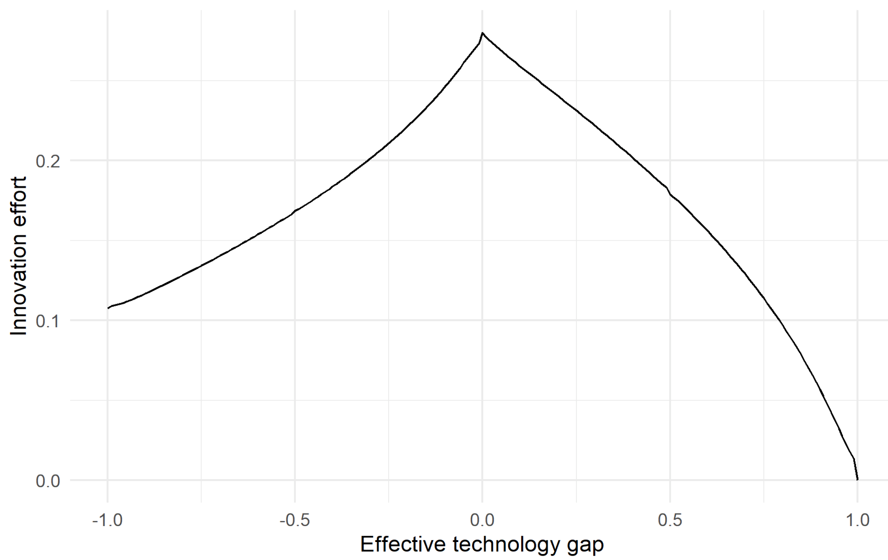
$$h_{mt} = \frac{\alpha x_{mt}^{\beta}}{\beta}. \tag{31}$$

3.2.3 Partial equilibrium results

Before turning to the general equilibrium outcomes of the model, let us consider some partial equilibrium results. In this part of the paper, I will assume that the aggregate economy is on a balanced growth path (BGP) with $\omega_t = \omega$, and hence $\dot{v}_{mt} = 0$. I will assume that the carbon tax is implemented or changed in only a single sector. Since there is a continuum of sectors, the tax change will not affect the aggregate economy, which thus remains on the BGP.

Furthermore, I set $\overline{m} = 1$ and $\beta = 2$ in this subsection to derive analytical solutions for firms' research intensities. Akcigit and Ates (2021) also make these two simplifications to their more elaborate model (Akcigit and Ates, 2019) to show analytical results. Notice that $\overline{m} = 1$ implies $m_{it}^C, m_{it}^D \in \{-1, 0, 1\}$ and $m_{it}^E \in [-1, 1]$. I further assume that if a firm with

Figure 5: The relationship between the technology gap and innovation efforts



$m^E_{it} \in (0, 1)$ innovates successfully, its gap does not increase *by* 1 but *to* 1.[21] Under these assumptions, the relationship between the effective technology gap and an individual firm's innovation efforts has an "inverted V" shape with a peak at $m^E_{it} = 0$. Figure 5 shows how a firm's innovation efforts depend on its effective technology gap.[22]

Proposition 1. *If $\overline{m} = 1$, $\beta = 2$ and the aggregate economy is on a BGP, so $\omega_t = \omega$, then innovation efforts x_{it} are increasing in m^E_{it} if $m^E_{it} < 0$, and decreasing in m^E_{it} if $m^E_{it} > 0$. Innovation efforts are highest when $m^E_{it} = 0$ and zero when $m^E_{it} = 1$.*

Proof. See appendix B.

Note that the shape of this relationship between the technology gap and innovation efforts

[21] If the leader in the market has a positive non-integer lead, then the two firms must be using different technologies. If the leader innovates, it improves not only its own technology but also its competitor's because the effective gap cannot exceed 1. Hence, the effective gap becomes 1 because the laggard switches to the same technology that the leader is using. Note that this assumption essentially makes "knowledge diffusion" dependent on the tax. This is not a good property, and hence I do not make this assumption when $\overline{m}$ is set higher in the general equilibrium part of the paper. In that case, if a firm has the largest non-integer lead, it cannot gain anything from innovating.

[22] To obtain the figure, I solve the system of seven equations for different values of $\tilde{\tau}$ (see Appendix B). I fix parameters at the following values, $\lambda = 1.05, \rho = 0.01, \alpha = 1, \beta = 2, \delta = 0.02, \omega = 1$. These parameters affect the height and steepness of the inverted V, but not the shape itself.

implies a positive effect of the degree of competition on innovation. Competition, measured by (the lack of) markups, is strongest when firms are neck-and-neck. As the proposition shows, this is also when innovation efforts are highest. In other words, the model features a strong "escape competition effect" (Aghion et al., 2005).

Another implication of Proposition 1 is that the increase or introduction of a carbon tax may increase innovation in dirty technologies at the firm level.

Proposition 2. *If $\overline{m} = 1$, $\beta = 2$ and the aggregate economy is on a BGP, so $\omega_t = \omega$, then the increase or introduction of a carbon tax in a single sector can increase a firm's dirty innovation efforts.*

Proof. See appendix B.

The key here is that the level of the carbon tax at which firms switch from dirty to clean may differ between firms. If one firm switches from dirty to clean while the other remains dirty, then the effective technology gap between them decreases.[23] This implies that the the degree of competition is increased by the tax, meaning that both firms increase their innovation efforts. Thus, the firm that remains dirty is innovating more heavily in dirty than it did before the tax change.

Note that this effect is only present if firms have different own technology gaps m^T, and that the effect is only there for values of $\tilde{\tau}_t$ that are high enough to cause one firm to switch but not so high that both firms switch. If both firms switch, the effective technology gap may still change because m^C (the new gap) need not be equal to m^D (the old gap), but there is no more dirty innovation. Note further that if a market has two dirty firms that are not neck-and-neck but that are equally good at the clean technology (which they are not using), the introduction of a carbon tax that is high enough to make both firms switch, leads to a large increase in innovation as it reduces the effective gap to 0. Finally, though the above propositions prove the presence of the escape competition effect and the seemingly contradictory finding that a carbon tax may increase dirty innovation in a simplified setting, i.e. $\overline{m} = 1$, $\beta = 2$ and balanced growth, these effects are also present in the more complete version of the model.[24]

[23] It has to decrease because firms switch only if it is optimal for them to switch.

[24] When $\overline{m} > 1$, the maximum innovation effort shifts somewhat to the right but never further right than

3.2.4 General equilibrium

This subsection solves for the general equilibrium outcomes of the model in closed form. To do so, it aggregates the continuum of intermediate input sectors, which vary along three relevant dimensions, namely the effective technology gap, which determines the profits and innovation incentives for the firms, and the own technology gap for both the leader and the laggard in each market. It is important to keep track of these own technology gaps because they determine what share of sectors switch technologies at a given change in the carbon tax. I first introduce some notation which is used to aggregate sectors and to keep track of the joint distribution of the effective and own technology gaps. I then solve sequentially for the normalized wage, aggregate emissions, total output, wages and R&D expenditures. Next, I characterize the evolution of aggregate productivity and of the technology gap distribution across sectors. Finally, I define the dynamic general equilibrium and the balanced growth path of the model.

Let us start by introducing some notation. Denote the market leader in industry j with L and the laggard (or follower) with F. That is, $MC_{Ljt} \leq MC_{Fjt}$ and hence, $m^E_{Ljt} \geq 0 \geq m^E_{Fjt}$. Let q_{ijt} denote the technology that firm i uses,

$$q_{ijt} = \begin{cases} q^C_{ijt} & \text{if} \quad m^T_{ijt} + \tilde{\tau}_t \geq 0 \\ q^D_{ijt} & \text{if} \quad m^T_{ijt} + \tilde{\tau}_t < 0 \end{cases}, \tag{32}$$

and let Q_t denote the aggregate productivity index as in Akcigit and Ates (2019):

$$Q_t = \exp\Big(\int_0^1 \ln(q_{Ljt})dj\Big). \tag{33}$$

Improvements to market leaders' productivity and firms taking over markets from competitors (and thus increases to Q_t) drive aggregate growth in this model.

Next, let us define a set of $(2\overline{m}^T + 1) \times (2\overline{m}^T + 1)$ matrices, where $\overline{m}^T$ is the maximum gap between a firm's own clean and dirty technologies. The point of these matrices is to track combinations of firms' own technology gaps and the effective technology gaps in their

$m = 1$.

markets.

$$\Psi_{mt} = \begin{bmatrix} \psi_{-\overline{m}^T,-\overline{m}^T,m,t} & \cdots & \psi_{-\overline{m}^T,\overline{m}^T,m,t} \\ \vdots & \ddots & \vdots \\ \psi_{\overline{m}^T,-\overline{m}^T,m,t} & \cdots & \psi_{\overline{m}^T,\overline{m}^T,m,t} \end{bmatrix}, \tag{34}$$

where

$$\psi_{klmt} = \int_0^1 \mathbb{1}\Big\{m^T_{Ljt} = k \wedge m^T_{Fjt} = l \wedge m^E_{Ljt} = m\Big\} dj. \tag{35}$$

That is, ψ represents the share of all sectors with a particular m^T_L, m^T_F, m^E_L combination.[25]

From this set of matrices follows the effective technology gap distribution, as in the Akcigit and Ates papers:

$$\mu_{mt} = \sum_{k=-\overline{m}^T}^{\overline{m}^T} \sum_{l=-\overline{m}^T}^{\overline{m}^T} \psi_{klmt}, \tag{36}$$

where μ is the share of all sectors with a given effective technology gap $m = m^E$.

For the general equilibrium outcomes, we need such a distribution for all 4 combinations of technologies used by market leaders and laggards. For this purpose, define $\mathcal{M}^D_t$ and $\mathcal{M}^C_t$ as the set of integers in the intervals $[-\overline{m}^T, -\tilde{\tau}_t)$ and $[-\tilde{\tau}_t, \overline{m}^T]$, respectively. Then,

$$\mu^{DD}_{mt} = \sum_{k\in\mathcal{M}^D_t} \sum_{l\in\mathcal{M}^D_t} \psi_{klmt}, \tag{37}$$

$$\mu^{CD}_{mt} = \sum_{k\in\mathcal{M}^C_t} \sum_{l\in\mathcal{M}^D_t} \psi_{klmt}, \tag{38}$$

$$\mu^{DC}_{mt} = \sum_{k\in\mathcal{M}^D_t} \sum_{l\in\mathcal{M}^C_t} \psi_{klmt}, \tag{39}$$

$$\mu^{CC}_{mt} = \sum_{k\in\mathcal{M}^C_t} \sum_{l\in\mathcal{M}^C_t} \psi_{klmt}, \tag{40}$$

where, for instance, μ^{DC}_{mt} is the share of sectors in which the leader uses the dirty technology, the laggard uses the clean technology, and the effective gap is m.

[25] Note that many cells in the Ψ matrices can only be equal to 0. For instance, all sectors in which both m^T_L and m^T_F are greater (smaller) than $-\tilde{\tau}$ have two firms that use the clean (dirty) technology, meaning that the effective technology gap is an integer. So ψ_{klmt} for both k and l greater (smaller) than $-\tilde{\tau}$ and m not equal to an integer must be equal to 0.

Next, for convenience, let us order the continuum of sectors as follows. Sectors in which both firms use the dirty technology are indexed $j \in [0, \theta_{1t}]$ and are referred to as DD sectors. Sectors in which the leader uses the clean technology and the laggard uses the dirty one are indexed $j \in (\theta_{1t}, \theta_{2t}]$ and are referred to as CD. Their counterparts DC are indexed $j \in (\theta_{2t}, \theta_{3t}]$. Finally, sectors where both firms use clean are CC and indexed $j \in (\theta_{3t}, 1]$. Thresholds θ are time dependent as laggard firms can become market leaders (and vice versa) and firms can switch technologies. The thresholds are

$$\theta_{1t} = \int_0^1 \mathbb{1}\Big\{m_{Ljt}^T + \tilde{\tau}_t < 0 \wedge m_{Fjt}^T + \tilde{\tau}_t < 0\Big\} dj = \sum_{m \in \mathcal{M}_t} \mu_{mt}^{DD}, \tag{41}$$

$$\theta_{2t} = \theta_{1t} + \int_0^1 \mathbb{1}\Big\{m_{Ljt}^T + \tilde{\tau}_t \geq 0 \wedge m_{Fjt}^T + \tilde{\tau}_t < 0\Big\} dj = \theta_1 + \sum_{m \in \mathcal{M}_t} \mu_{mt}^{CD}, \tag{42}$$

$$\theta_{3t} = 1 - \int_0^1 \mathbb{1}\Big\{m_{Ljt}^T + \tilde{\tau}_t \geq 0 \wedge m_{Fjt}^T + \tilde{\tau}_t \geq 0\Big\} dj = 1 - \sum_{m \in \mathcal{M}_t} \mu_{mt}^{CC}. \tag{43}$$

Now, let us solve, in turn, for the normalized wage ω, total emissions, wages, total output, and R&D expenditures. Using the labor market clearing condition $1 = \int_0^1 l_{ijt} + l_{-ijt} + h_{ijt} + h_{-ijt} dj$ together with labor demand in production (25) and research $h_{ijt} = \frac{\alpha}{\beta} x_{ijt}^{\beta}$, yields the normalized wage ω_t:

$$1 = \int_0^1 \mathbb{1}\Big\{m_{Fjt}^T + \tilde{\tau}_t < 0\Big\} \frac{Y_t}{w_t(1+\kappa\tau_t)\lambda^{m_{Ljt}^E}} + \mathbb{1}\Big\{m_{Fjt}^T + \tilde{\tau}_t \geq 0\Big\} \frac{Y_t}{w_t \lambda^{m_{Ljt}^E}} + \frac{\alpha}{\beta}(x_{Ljt}^{\beta} + x_{Fjt}^{\beta}) dj$$

$$\iff$$

$$\omega_t = \left(\sum_{k \in \mathcal{M}_t} \frac{\mu_{kt}^{DD} + \mu_{kt}^{CD}}{(1+\kappa\tau_t)\lambda^k} + \frac{\mu_{kt}^{DC} + \mu_{kt}^{CC}}{\lambda^k} \right) \left(1 - \frac{\alpha}{\beta} \sum_{k \in \mathcal{M}_t} \mu_{kt} (x_{Ljt}^{\beta} + x_{Fjt}^{\beta}) \right)^{-1}, \tag{44}$$

where $\mathcal{M}_t$ is the set of all possible values in the interval $[0, \overline{m}]$ that m_{it}^E can take at time t.[26]

Total emissions $E_t = \int_0^1 e_{ijt} + e_{-ijt} dj$ are just the sum of each firm's emissions (26), and

[26] It is time dependent because it depends on τ_t. Note that if $\tau_t = 0$, then $\mathcal{M}_t$ consists of only integers (as in Akcigit and Ates 2019).

can be solved in closed form by plugging in ω_t:

$$
\begin{aligned}
E_t = \frac{\kappa}{\omega_t} \int_0^1 & \mathbb{1}\Big\{ m_{Ljt}^T + \tilde{\tau}_t < 0 \wedge m_{Fjt}^T + \tilde{\tau}_t < 0 \Big\} \frac{1}{(1+\kappa\tau_t)\lambda^{m_{Ljt}^E}} \\
& + \mathbb{1}\Big\{ m_{Ljt}^T + \tilde{\tau}_t < 0 \wedge m_{Fjt}^T + \tilde{\tau}_t \geq 0 \Big\} \frac{1}{\lambda^{m_{Ljt}^E}} dj \\
= \frac{\kappa}{\omega_t} & \left[\int_0^{\theta_{1t}} \frac{1}{(1+\kappa\tau_t)\lambda^{m_{Ljt}^E}} dj + \int_{\theta_{2t}}^{\theta_{3t}} \frac{1}{\lambda^{m_{Ljt}^E}} dj \right] = \frac{\kappa}{\omega_t} \sum_{k \in \mathcal{M}_t} \frac{\mu_{kt}^{DD}}{(1+\kappa\tau_t)\lambda^k} + \frac{\mu_{kt}^{DC}}{\lambda^k}. \qquad (45)
\end{aligned}
$$

Aggregate emissions depend on several things, and climate policy can thus affect emissions through several channels. First, since only leaders that use the dirty technology emit carbon, they depend on the proportion of sectors that have a dirty leader. Second, within those sectors, emissions depend negatively on the degree of market power. Powerful firms charge high prices and thus produce and emit little. In other words, "the monopolist is the conservationist's friend" (Hotelling 1931; Solow 1974). Third, emissions depend negatively on the innovation intensities x. As demand for R&D increases, labor is reallocated from production to research, and hence emissions decrease. While the effect of climate policy on emissions through the first channel is always (weakly) negative (as the tax increases, fewer sectors have a dirty leader), the effect through the second channel may be positive and may thus partially offset the first channel. That is, climate policy may decrease the effective technology gap in those sectors that have a dirty leader, increasing those leaders' emissions. The third channel can in principle go either way, though if leaders are dirtier than laggards, the third channel reduces emissions as climate policy decreases markups and thus increases demand for R&D.

Total emissions are then used to update cumulative emissions (8) and global warming (9). Plugging sector level output $y_{jt} = y_{Ljt} + y_{Fjt}$, which follows from (24), and the updated temperature into final good production (6) gives

$$
\begin{aligned}
\ln(Y_t) = \int_0^1 & \mathbb{1}\Big\{ m_{Fjt}^T + \tilde{\tau}_t \geq 0 \Big\} \ln\left(\frac{Y_t}{w_t} q_{Fjt}^C \right) + \mathbb{1}\Big\{ m_{Fjt}^T + \tilde{\tau}_t < 0 \Big\} \ln\left(\frac{Y_t}{w_t(1+\kappa\tau_t)} q_{Fjt}^D \right) dj \\
& - \frac{\gamma}{2} T_t^2.
\end{aligned}
$$

Cancelling out $\ln Y_t$ on both sides and taking the wage to the left hand side yields a closed form solution for the wage w_t:

$$w_t = \exp\left[\int_0^1 \mathbb{1}\Big\{m^T_{Fjt} + \tilde{\tau}_t \geq 0\Big\} \ln(q^C_{Fjt}) + \mathbb{1}\Big\{m^T_{Fjt} + \tilde{\tau}_t < 0\Big\}\Big(\ln(q^D_{Fjt}) - \ln(1+\kappa\tau_t)\Big)dj \right. \\ \left. - \frac{\gamma}{2}T_t^2\right],$$

which can be further simplified to

$$w_t = \frac{\exp\big(\int_0^1 \ln(q_{Fjt})dj\big)\exp\big(-\frac{\gamma}{2}T_t^2\big)}{\exp\big(\int_0^1 \mathbb{1}\{m^T_{Fjt} + \tilde{\tau}_t < 0\}\ln(1+\kappa\tau_t)dj\big)} = \frac{Q_t\lambda^{-\sum_{k\in\mathcal{M}_t}\mu_{kt}k}\exp\big(-\frac{\gamma}{2}T_t^2\big)}{(1+\kappa\tau_t)^{\theta_{2t}}}, \tag{46}$$

The first part of the numerator in (46) is identical to the wage in Akcigit and Ates (2019). The second part adds climate damages, and the denominator corrects for the fact that Q and q do not just reflect labor productivity (like in the Akcigit and Ates papers, where labor is the only factor), but total factor productivity. Note that the correction applies to sectors in which the laggard (which does not produce unless the firms are neck-and-neck) uses the dirty technology. This comes from the limit pricing structure, where the laggard's marginal costs, which include the carbon tax in these sectors, determine the leader's production.

Combining (46) and (44) yields total final good production:

$$Y_t = \frac{w_t}{\omega_t} = \frac{Q_t\lambda^{-\sum_{k\in\mathcal{M}_t}\mu_{kt}k}\exp\big(-\frac{\gamma}{2}T_t^2\big)\Big(1-\sum_{k\in\mathcal{M}_t}\mu_{kt}(x^\beta_{Ljt}+x^\beta_{Fjt})\Big)}{(1+\kappa\tau_t)^{\theta_{2t}}\Big(\sum_{k\in\mathcal{M}_t}\frac{\mu^{DD}_{kt}+\mu^{CD}_{kt}}{(1+\kappa\tau_t)\lambda^k} + \frac{\mu^{DC}_{kt}+\mu^{CC}_{kt}}{\lambda^k}\Big)} \tag{47}$$

Tax revenues are $G_t = \tau^E_t E_t = \tau_t w_t E_t$ and are distributed lump sum to the household.

Total R&D expenditures on clean innovation are

$$\begin{aligned} R^C_t &= \int_{\theta_{1t}}^{\theta_{2t}} R_{Ljt}dj + \int_{\theta_{2t}}^{\theta_{3t}} R_{Fjt}dj + \int_{\theta_{3t}}^{1} R_{Ljt} + R_{Fjt}dj \\ &= \frac{\alpha w_t}{\beta}\sum_{k\in\mathcal{M}_t}\mu^{CD}_{kt}x^\beta_{kt} + \mu^{DC}_{kt}x^\beta_{-kt} + \mu^{CC}_{kt}(x^\beta_{kt}+x^\beta_{-kt}), \end{aligned} \tag{48}$$

and total R&D expenditures on dirty innovation are

$$\begin{aligned} R^D_t &= \int_0^{\theta_{1t}} R_{Ljt} + R_{Fjt}dj + \int_{\theta_{1t}}^{\theta_{2t}} R_{Fjt}dj + \int_{\theta_{2t}}^{\theta_{3t}} R_{Ljt}dj \\ &= \frac{\alpha w_t}{\beta}\sum_{k\in\mathcal{M}_t}\mu^{DD}_{kt}(x^\beta_{kt}+x^\beta_{-kt}) + \mu^{CD}_{kt}x^\beta_{-kt} + \mu^{DC}_{kt}x^\beta_{kt}. \end{aligned} \tag{49}$$

Finally, like the Akcigit and Ates papers, let us determine the evolution of aggregate productivity and the technology gap distribution. In equilibrium, Q_t will determine the growth rates of wages and output. Along the balanced growth path, it grows at a constant rate. Q_t changes when a market leader or a neck-and-neck firm innovates, or when a firm that is less than 1 step behind innovates. For $\tilde{\tau} > 0$ but not equal to an integer or half of an integer, it looks as follows,[27]

$$\begin{aligned}\ln(Q_{t+\Delta t}) - \ln(Q_t) = &\Big[2\mu_{0t}x_{0t} + \sum_{k\in\mathcal{M}_{\geq 1t}} \mu_{kt}x_{kt} + \mu_{pt}(x_{pt} + (1-p)x_{-pt}) \\ &+ \mu_{1-pt}(x_{1-pt} + px_{p-1t})\Big]\ln(\lambda)\Delta t + o(\Delta t),\end{aligned} \tag{50}$$

where $\mathcal{M}_{\geq 1t}$ is the set of possible effective technology gaps in the interval $[1, \overline{m}]$ at time t, p is one of the technology gaps in the interval $(0, 1)$,[28] and $o(\Delta t)$ contains second order terms that account for simultaneous innovations by the two firms in a sector. It disappears as the time increment Δt goes to 0 (that is, $\lim_{\Delta t\to 0} o(\Delta t) = 0$).

The distribution of own and effective technology gaps Ψ changes over time as firms innovate and knowledge diffuses. Note that the own gap distributions are not constant along the balanced growth path, as clean (dirty) firms innovate only in the clean (dirty) technology, meaning that their m_{it}^T keeps increasing (decreasing). For $k \in [-\overline{m}^T + 1, \overline{m}^T - 1]$, $l \in [-\overline{m}^T + 1, \overline{m}^T - 1]$, $m \in (1, \overline{m} - 1]$ and $\tilde{\tau}_t$ not equal to an integer, the law of motion of a cell ψ_{klmt} in matrix Ψ_{mt} is as follows,

$$\begin{aligned}\frac{\psi_{k,l,m,t+\Delta t} - \psi_{k,l,m,t}}{\Delta t} = &\,\mathbb{1}\Big\{k + 1 + \tilde{\tau}_t < 0\Big\}\psi_{k+1,l,m-1,t}x_{m-1,t} \\ &+ \mathbb{1}\Big\{k - 1 + \tilde{\tau}_t > 0\Big\}\psi_{k-1,l,m-1,t}x_{m-1,t} \\ &+ \mathbb{1}\Big\{l + 1 + \tilde{\tau}_t < 0\Big\}\psi_{k,l+1,m+1,t}x_{-m-1,t} \\ &+ \mathbb{1}\Big\{l - 1 + \tilde{\tau}_t > 0\Big\}\psi_{k,l-1,m+1,t}x_{-m-1,t} \\ &- \psi_{k,l,m,t}(x_{m,t} + x_{-m,t} + \delta) + \frac{o(\Delta t)}{\Delta t}.\end{aligned} \tag{51}$$

The intuition here is that a clean (dirty) innovation increases (decreases) a firm's m^T by 1, and an innovation by a market leader (laggard) increases (decreases) the effective technology

[27]If $\tau = 0$ or a positive integer, all terms that have a p disappear because non-integer gaps are not possible. Further, if $2\tilde{\tau}$ is equal to an integer, then the last term inside the brackets in (50) disappears as $p = 1 - p$.

[28]For instance, if $\tilde{\tau} = 3.2$, then p is either 0.2 or 0.8. This term enters because a firm that is less than a full step behind improves the sector's technology level if it makes an innovation step (it increases q by less than factor λ, though).

gap by 1. The law of motion for ψ_{klmt} outside the intervals for k, l, m mentioned above are shown in Appendix C.2.

Contrary to the own technology gap distribution, the effective gap distributions μ are constant along the balanced growth path (which implies that the thresholds θ are constant, too). The share of sectors with a given effective gap changes as follows during the transition for sectors with $m^E_{Ljt} \in (1, \overline{m}-1]$ and $FF \in \{DD, CD, DC, CC\}$,

$$\frac{\mu^{FF}_{m,t+\Delta t} - \mu^{FF}_{m,t}}{\Delta t} = \mu^{FF}_{m-1,t} x_{m-1,t} + \mu^{FF}_{m+1,t} x_{-m-1,t} - \mu^{FF}_{m,t}(x_{m,t} + x_{-m,t} + \delta) + \frac{o(\Delta t)}{\Delta t}. \quad (52)$$

That is, the share of sectors with $m^E_{Ljt} = m$ increases when a leader with $m-1$ innovates or when the competitor of a leader with $m+1$ innovates (bringing the gap to m in both cases). It decreases when a leader with gap m or its competitor innovates or when knowledge diffusion takes place. Appendix C.2 shows the evolution of the share of sectors with gaps in the intervals $[0, 1]$ and $(\overline{m}-1, \overline{m}]$.

Finally, let us define the dynamic equilibrium and the balanced growth path.

Definition 1. *A dynamic general equilibrium in this economy is an allocation*

$$\left\{r_t, w_t, p_{jt}, y_{jt}, l_{jt}, h_{jt}, x_{jt}, e_{jt}, R^C_t, R^D_t, L_t, Y_t, E_t, T_t, G_t, Q_t, \{\Psi_{mt}\}_{m\in\{0,\dots,\overline{m}\}}\right\}^{t\in[0,\infty)}_{j\in[0,1]}$$

such that (i) intermediate input producers choose their technology (clean or dirty) according to (16); (ii) the sequence of prices and quantities $\{p_{jt}, y_{jt}\}$ satisfies equations (22) and (24) and maximizes the operating profits of the intermediate input producers in market j; (iii) the R&D decisions x_{jt} satisfy (30), and aggregate clean and dirty R&D satisfy (48) and (49); (iv) labor supply $L = 1$ is equal to the sum of intermediate input producers' optimal production worker demand (25) plus their optimal R&D worker demand (31); (v) Y_t satisfies (47) and $C_t = Y_t$; (vi) w_t clears the labor markets at every t; (vii) interest rate r_t satisfies the Euler equation (7); (viii) total emissions E_t satisfy (45); (ix) all carbon taxes are transferred lump-sum to the household such that the government's budget is balanced at all times; (x) T_t, Q_t and Ψ_{klmt} evolve as in (9), (50) and (51), respectively.

Definition 2. *A balanced growth path in this economy is a general equilibrium allocation such that (i) output and wages grow at the same rate g_t, meaning that ω_t is constant and $\dot{v}_{mt} = 0 \ \forall \ m \in \{-\overline{m}, \dots, \overline{m}\}$; (ii) innovation intensities x_{mt} are such that the effective gap distribution $\{\mu_{mt}\}_{m\in\{0,\dots,\overline{m}\}}$ is constant, i.e. $\dot{\mu}_{mt} = 0 \ \forall \ m \in \{0, \dots, \overline{m}\}$.*

Note the following features of this model's balanced growth path (BGP). First, output need not grow at a constant rate along the BGP. From the value functions (28) and (29), together with the equilibrium solution for ω_t in (44), it is clear that the dynamic problem of the intermediate input firms does not depend on the climate. Hence, the fact that damages from global warming may not grow at a constant rate, which means that output and wages do not grow at a constant rate, does not affect the BGP as the effective gap distribution is unaffected. Total factor productivity does grow at a constant rate along the BGP. Second, because of knowledge diffusion, there are no "mixed" sectors, i.e. sectors with one clean and one dirty firm, on the BGP.[29] Third, the own technology gap distribution is not fixed but evolves endogenously along the BGP. That is, the distance between clean and dirty within sectors keeps increasing as clean firms innovate in clean technologies and dirty firms innovate in dirty technologies. Hence, the longer the policymaker waits with implementing a higher carbon tax, the higher is the required tax to make some fixed share of firms switch from dirty to clean. In other words, the model features strong path dependence.

4 Calibration and quantitative exercises

Having established the dynamic general equilibrium and the balanced growth path, I calibrate the model to be able to use it to draw policy and welfare conclusions. Since the goal of this paper is to understand the effects of stringent climate policy in the future, I use recent data for the calibration. I assume that the economy starts from a BGP in the 2010s, shock the carbon tax, and simulate the transition to the new BGP. This section first discusses calibration procedure, which includes using micro data to set the initial technology gap distribution and matching macro level moments. It then discusses the solution algorithm used to solve for the transition from one BGP to a new one. Finally, it presents a set of quantitative exercises that illustrate the economic effects of a carbon tax in the current setting and that investigate welfare.

[29] Remember that when knowledge diffusion takes place, the laggard learns about the leader's technology. It automatically switches to that technology because it must be better than its own technology (otherwise that firm would not be the laggard).

4.1 Calibration

Calibrating the model involves a series of steps. First, I set a number of parameters externally based on existing literature. Second, I set some of the initial conditions of the model based on real data, namely the full history of global CO_2 emissions and the share of firms that use a clean technology. I also set the initial carbon tax at this stage. Third, I calibrate the remaining parameters internally to match a set of macro level moments. Fourth, I use micro level data to set the remaining initial conditions, namely the joint distribution across sectors of the effective technology gap (m^E) and the own technology gaps of leaders and followers (m_L^T and m_F^T, respectively).

4.1.1 External calibration

The parameters set externally are shown in Table 1. The first is ρ, the rate of time preference, which is the subject of a large debate in environmental economics. Whereas Nordhaus (2014) sets ρ at 1.5%, others such as Stern (2007) argue that is should be much lower at 0.1%. In my model the rate of time preference is important not only because climate change is a long run problem but also because innovation is forward looking. I follow Acemoglu et al. (2016), who also have a model with both climate change and endogenous innovation, and set ρ to 1%. I set β, the curvature of the R&D production function, equal to 1/0.35 as in Akcigit and Ates (2023), who follow the earlier literature on the topic (and who call this parameter γ).

For the climate damages, as well as the climate science parameters, I follow Dietz and Venmans (2019). The value of γ at 0.01 implies that at 2° C warming, damages are 2% of output (and 8% at 4° C warming). For the TCRE I also follow Dietz and Venmans (2019) and multiply the rate from the literature by 1.1 to account for other greenhouse gases than CO_2.[30] For ε the value of 0.5 means that it takes about 10 years for emissions to reach their full effect on the global temperature.

[30] Dietz and Venmans (2019) argue that this characterization (TCRE times 1.1) is supported by evidence for the past 150 years, while it is unclear whether it will be higher or lower in the future.

Table 1: Externally calibrated parameters

Parameter	Value	Description	Source
ρ	1%	Rate of time preference	Acemoglu et al. (2016)
β	1/0.35	R&D cost curvature	Akcigit and Ates (2023)
γ	0.01	Climate damage elasticity	Dietz and Venmans (2019); Nordhaus and Moffat (2017)
ζ	0.00048×1.1	Transient Climate Response to Cum. Carbon Emissions	Dietz and Venmans (2019); Matthews et al. (2009)
ε	0.5	Initial pulse-adjustment time-scale of the climate system	Dietz and Venmans (2019); Ricke and Caldeira (2014)

4.1.2 Initial conditions

The model introduced in this paper features a degree of hysteresis, meaning that the choice of initial conditions affects the outcomes of the model. Firstly, carbon emissions have lasting effects on output. Secondly, the share of sectors that use a dirty technology determines both how many firms are responsible for total global emissions, which will be one of the internal calibration targets, and how many firms are potentially directly affected by a change in the carbon tax. Thirdly, the gap between clean and dirty technologies within firms determines the size of the tax that is needed to induce a given share of firms to switch from dirty to clean production at a given point in time. I address these issues as follows.

First, I import cumulative emissions between the industrial revolution and 2019, the year I use for the calibration using data from Friedlingstein et al. (2022). I then compute the temperature in 2019 using equation (9). This yields the climate damages to GDP to be used in the internal calibration.

Second, I jointly set the initial carbon tax and the share of sectors that use the clean technology on the initial BGP, given the initial tax. I choose to set the initial tax to 0,[31] and I obtain the share of firms that use the clean technology from the same patent data

[31] In fact, I set it close to zero ($\tau_t = 0.00001$) for numerical convenience in solving the model. The choice of the initial τ in itself does not meaningfully affect the results in this section. Rather, the share of firms that are clean given the initial tax matters.

that was used for Section 2. In the model, a firm uses the clean technology if the condition $m_{it}^T + \tilde{\tau}_t \geq 0$ is satisfied. Thus, I select the same sample that is used for Figure 2,[32] and find the proportion of firms that have a weakly positive technology gap (as defined in (3)). This is the case for 40.1% of firms.

Having set the history of emissions and the proportion of firms that are dirty given the initial tax, I proceed with the internal calibration, which is discussed below. A fully specified technology gap distribution Ψ_{mt} is not required to solve for the balanced growth path. From the internal calibration I obtain the effective gap distribution μ_{mt}, which I use to set Ψ_{mt} based on the same micro data discussed above.

To set the initial joint distribution of effective and own technology gaps I first convert patent counts into innovation steps to obtain a measure of m^T. Next, I compute a measure of concentration at the sector level (top 10 revenue divided by top 20 revenue) to sort sectors by their degree of competition. I then use the BGP effective gap distribution, obtained from the internal calibration, to find the distribution of own technology gaps of leaders and laggards for each possible effective gap. For instance, 12.3% of sectors have a technology gap of 0 along the initial BGP, so I use the m^T distribution of the firms that are active in the 12.3% of sectors with the lowest concentration index to fill in $\Psi_{m=0,t=0}$. Appendix D.1 elaborates on the procedure for setting the initial Ψ_{mt} matrices in more detail.

4.1.3 Internal calibration

The parameters to be calibrated internally are innovation step size λ, diffusion arrival rate δ, R&D production function scaling parameter α and emission scaling parameter κ. The moments to be targeted in the calibration are the average markup, the profit share, TFP growth and total emissions. The first two are computed as follows in the model.

$$\text{Average markup} = \int_0^1 \lambda^{m_{Ljt}^E} = \sum_{k \in \mathcal{M}_t} \mu_{kt} \lambda^k, \tag{53}$$

$$\text{Profit share} = \frac{\int_0^1 \pi_{jt} dj}{Y_t} = 1 - \sum_{k \in \mathcal{M}_t} \mu_{kt} \frac{1}{\lambda^k}. \tag{54}$$

[32]These are all firms that have applied for at least one clean or dirty triadic patent family and that are in the top 20 firms by revenue in a country-sector for which I have at least 20 firms in my data set for the year 2018.

The other two follow directly from (50) and (45).

The data for these variables comes from a variety of sources. A challenge here is that data is typically available at the country level, while global emissions are relevant for the climate. Hence, I use data that is as close as possible to being representative of the world economy. For the average markup I take the most recent estimate by Díez et al. (2021), who use data on both rich and developing countries and both private and public firms to estimate markups. Their estimate for 2015 is 1.29.[33] For the profit share I use the value of 19% found for the US in 2018 by Eggertsson et al. (2021). I compute average total factor productivity (TFP) growth using data from the OECD and find that it was 0.44% per year over the period 2011-2019.[34] For carbon emissions I take the year 2019 from the same data that is used for all historical emissions (Friedlingstein et al., 2022).

The calibration procedure is as follows. For a given combination of $\{\lambda, \delta, \alpha, \kappa\}$, I compute the effective tax $\tilde{\tau}_t$. I then guess a value for the normalized wage ω_t, and solve for the innovation intensities from the value functions combined with the first order conditions (28)-(30).[35] Using the innovation intensities I solve for the BGP effective technology gap distribution μ_{mt} by setting its law of motion (52) equal to 0. I then compute the implied ω_t and compare it to the initial guess. I repeat the process with the implied ω_t as the new guess until the process converges. With the resulting μ_{mt} and x_{mt} I compute the moments mentioned above and compare them to their data counterparts. I evaluate them using the same objective function as Akcigit and Ates (2023):

$$\text{Objective} = \sum_{k=1}^{4} \frac{|\text{model}(k) - \text{data}(k)|}{\frac{1}{2}|\text{model}(k)| + \frac{1}{2}|\text{data}(k)|}, \tag{55}$$

[33]De Loecker and Eeckhout (2018) also estimate global markups and find an average of 1.6. I choose to use the estimate by Díez et al. (2021) for several reasons. First, their paper is published and peer-reviewed, while De Loecker and Eeckhout (2018) is a working paper. Second, their analysis includes both public and private companies, while the analysis by De Loecker and Eeckhout (2018) includes only public companies.

[34]I use MFP from the public OECD database to compute this variable. Growth rates are weighted by GDP and the included countries are Australia, Austria, Belgium, Canada, Switzerland, Germany, Denmark, Spain, Finland, France, Great Britain, Greece, Ireland, Israel, Italy, Japan, Korea, Luxembourg, the Netherlands, Norway, New Zealand, Portugal, Sweden and the United States.

[35]Note that $\dot{v}_{mt} = 0$ on the BGP, so plugging in the first order conditions yields a system of K equations in K unknowns where K is the number of possible values for the effective technology gap.

Table 2: Internally calibrated parameters

Parameter	Value	Description
λ	1.0496	Innovation step size
δ	0.0156	Diffusion arrival rate
α	186.1686	R&D scaling parameter
κ	66.2997	Emission scaling parameter

Table 3: Model fit

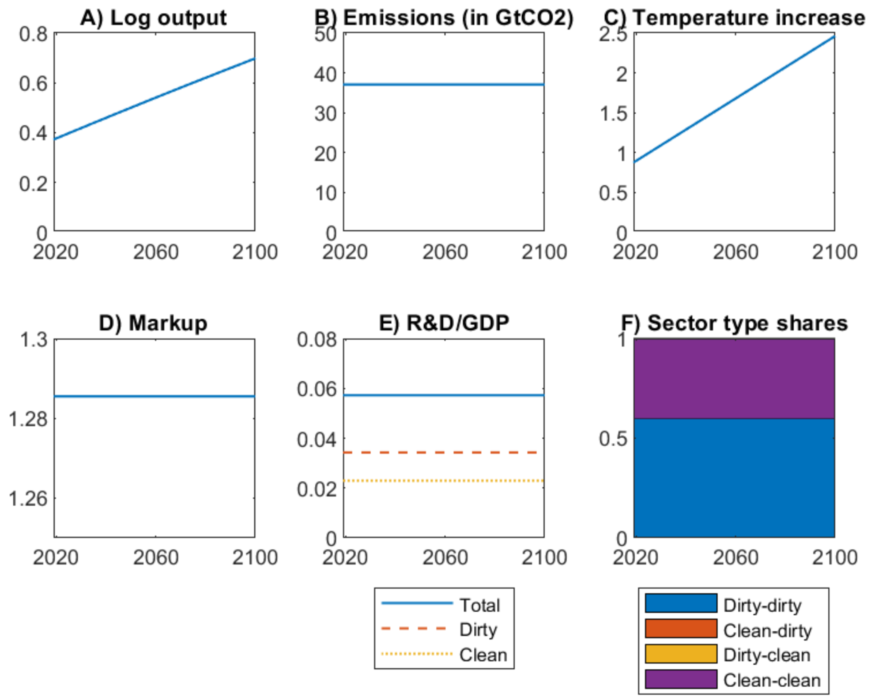
Moment	Model	Data	Source
Average markup (2015)	1.2855	1.29	Díez et al. (2021)
Profit share (2018)	19%	19%	Eggertsson et al. (2021)
TFP growth (average 2011-2019)	0.4356%	0.4356%	OECD
Emissions (2019, in $GtCO_2$)	37.0856	37.0829	Friedlingstein et al. (2022)

where k denotes each moment. I then find the combination of parameters that minimizes the above objective.

Table 2 shows the resulting parameter values and Table 3 shows the model fit. The average markup is slightly below the value in the data, while the other three moments are matched almost perfectly. Compared to Akcigit and Ates (2019), which is the closest paper to mine in terms of the model but targets moments from the US economy in the 1980s, the value of λ is highly similar while δ is smaller in my calibration, which is consistent with the findings of Akcigit and Ates (2019, 2023) that this parameter has decreased since the 1980s. α is quite a bit higher in my calibration because innovation needs to be very costly to match low TFP growth while firms care very much about the future.[36]

[36] Akcigit and Ates (2019) set ρ at 5% while I use 1% as is common in the environmental economics literature. In addition, TFP growth was almost a full percentage point higher in the 1980s, which are targeted by Akcigit and Ates (2019).

Figure 6: Balanced growth path (no additional climate policy)

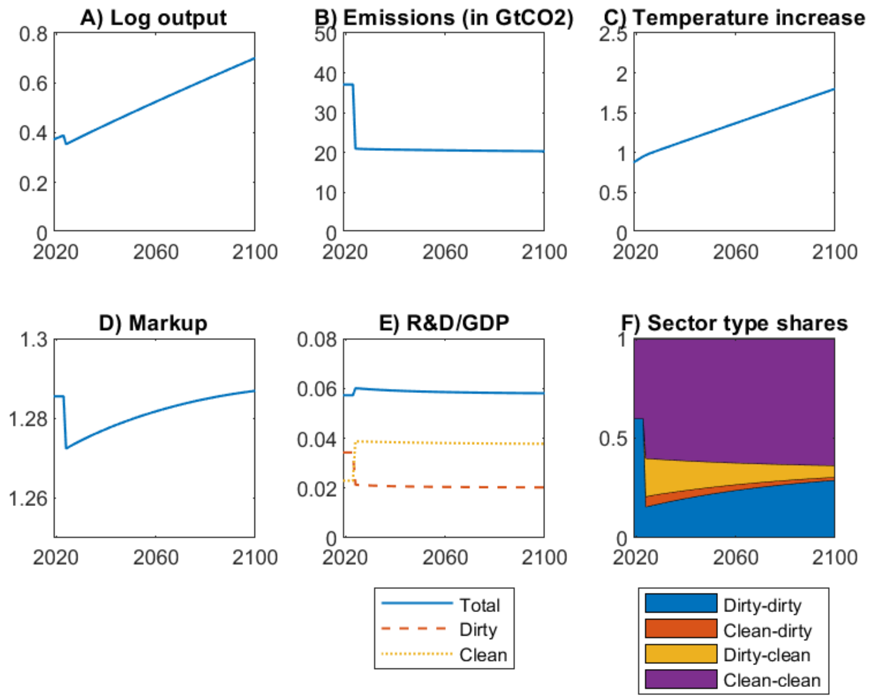


4.2 Quantitative exercises

Having calibrated the model's parameters and set the initial conditions, I turn to several quantitative exercises in which I simulate the transition to a green(er) economy. A shock to the carbon tax directly affects the effective gap size distribution, which implies that it affects the normalized wage ω_t. When ω_t is not constant, solving the value functions in (28) and (29) becomes a non-trivial problem. Appendix D.2 elaborates on the solution algorithm used to solve for the transition, which is largely based on the procedure in Akcigit and Ates (2023).

In the quantitative exercises I first simulate forward the BGP, which can be interpreted as a business as usual scenario in which no additional climate policy is implemented. Next, I numerically find the optimal one-time permanent tax change and study its effects. Then, I simulate the effects of the tax change that, according to the model, would be sufficient to reach 1.5° C temperature increase by the year 2100. Finally, I investigate the relevance of the market power channel that my model adds to the literature by manipulating the initial conditions such that this channel is essentially switched off. I then compare outcomes to the simulation in which the market power channel is present.

Figure 7: Transition after the optimal increase in τ



4.2.1 No climate action

If no further climate policy is implemented, the economy remains on the calibrated BGP. One can think of this as a business as usual scenario. Figure 6 shows what happens to the model's main variables in this scenario. Panel A shows that output growth, which consists of TFP growth minus growth in climate damages, gradually decreases as TFP growth is constant while damage growth increases. The model predictions on both emissions and the temperature path are consistent with intermediate scenarios modeled by the IPCC (2023).[37] On the BGP, the aggregate markup and R&D expenditures as a share of GDP are constant. In addition, dirty firms remain dirty and clean firms remain clean, meaning that there are no changes in the shares of clean and dirty sectors.

[37]The prediction that emissions are flat under currently implemented policies is close to the median scenario in the top figure on page 22 of the Summary for Policymakers in IPCC (2023). The temperature increase to 2.5 degrees by 2100 is close to the intermediate scenario in the top figure on page 17 of the Summary for Policymakers in IPCC (2023).

4.2.2 The "optimal" carbon tax

To find the optimal unanticipated, one-time and permanent change to the carbon tax in 2024, I solve for the transition and new balanced growth path for a grid of tax changes and compute welfare.[38] It turns out that the optimal $\tilde{\tau}$ is equal to 2.07 innovation steps, which corresponds to 138 dollars per tonne of CO_2 at the moment of implementation.[39,40] Keeping the price of emissions relative to labor fixed, the carbon tax τ_t^E grows with the wage rate w_t, which grows at the rate of the economy along the BGP but at a different rate along the transition. The optimal tax change improves welfare by a consumption equivalent of 1.88% compared to business as usual (no policy changes).

Figure 7 shows what happens to various variables as the optimal tax change is implemented and economy goes through the green transition. At the moment of implementation output drops by 3.5% as a substantial share of dirty firms start using a clean technology at which they are less productive (that is, their $q^D > q^C$). After the implementation of the tax all firms with m_i^T equal to -1 or -2 switch from dirty to clean production. This reduces emissions almost by half, and would lead to a temperature increase in the year 2100 of about 1.8° C.[41] Panel F shows that the share of sectors with two dirty firms (dirty-dirty) decreases dramatically. These sectors become either completely clean (clean-clean) or mixed, meaning that they have a clean leader and a dirty laggard (clean-dirty) or the other way around (dirty-clean). There are no mixed sectors along the BGP.

Interestingly, as the tax change is implemented, markups jump down by about 1.4 percentage points (from 1.286 to 1.272). This is because in all sectors in which at least one firm switches from dirty to clean production, the effective technology gap changes. If the leader has a more negative m^T than the laggard, which is the case on average according to

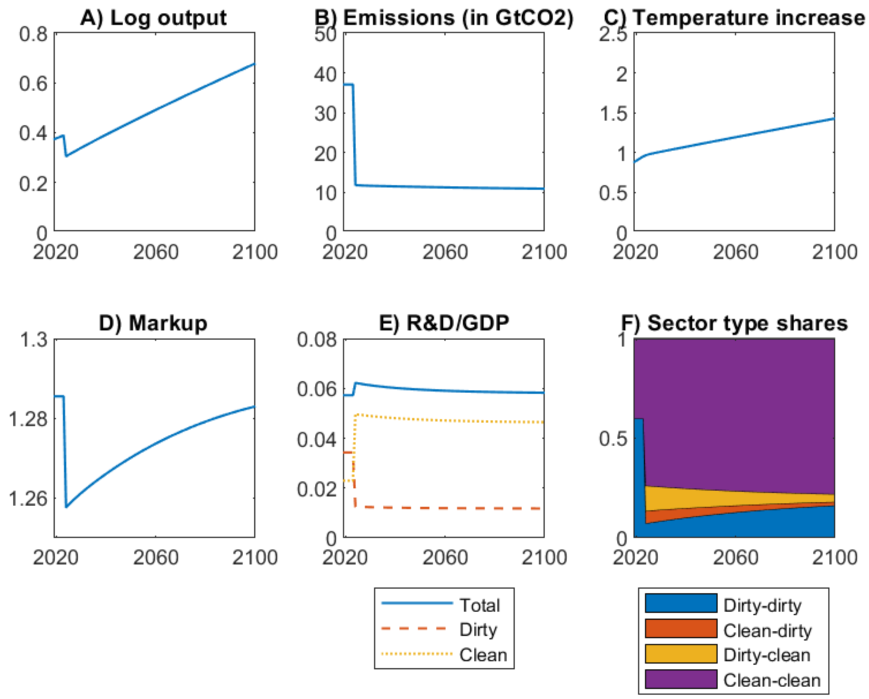
[38] I use the discrete version of the utility function (4) to find welfare: $W = \sum_{s=2024}^{3024} \frac{1}{(1+\rho)^{s-2024}} \ln(C_s)$. Figure E3 in the Appendix shows welfare levels for different values of τ.

[39] The carbon tax is expressed in units of final good. To convert to dollars per tonne of CO_2, I divide $\tau_{2024}^E = \tau_{2024} w_{2024}$ by the model's Y_{2024}, multiply by 100.88 trillion (world GDP in current dollars in 2022 according to the World Bank: `https://data.worldbank.org/indicator/NY.GDP.MKTP.CD`), and convert from gigatonnes to tonnes.

[40] Please note that the precise value of the optimal tax is quite sensitive to the calibration of the initial conditions, specifically to the conversion from patents to innovation steps.

[41] Note that, since emissions do not go down to 0, temperature keeps increasing after the year 2100.

Figure 8: Transition after increase in τ that limits warming to 1.5° C



the evidence shown in Section 2, then the effective gap goes down and markups decrease. Over the course of the green transition, markups increase to their new BGP value which is slightly above the initial level (1.290).[42] At the same time, R&D spending as a percentage of total output jumps up slightly. This happens because the share of sectors with a low effective technology gap increases, and firms in more competitive sectors invest more in R&D because the relative gains are larger (see Figure 5).[43] The increase in R&D, together with the decrease in emissions (and thus in the growth of climate damages), leads to a persistent increase in the growth rate of the economy after the policy is implemented.

[42] Markups are (slightly) different across BGPs because ω_t is different across BGPs (it depends on τ_t), meaning that the costs of R&D and thus the effective gap distribution are also different.

[43] At the same time, wages decrease more than total output, meaning ω_t jumps down at the time of the shock, putting downward pressure on R&D/GDP. The positive effect dominates in this case, leading to an increase.

4.2.3 Limiting warming to 1.5° C

The optimal tax change, according to the model and given the calibration as described above, is not high enough to reach the Paris goal of 1.5° C. One could argue that the optimal tax is in fact the one that reaches the goal that governments of the world have set for themselves at lowest cost. So, this section implements the lowest tax that reaches the Paris goal in 2100. This tax is 4.01 innovation steps or 283 dollars per tonne. Figure 8 shows the results of this exercise. All effects that are visible for the optimal tax are amplified for the tax that reaches 1.5° C. Output drops by 8.1% at the moment of implementation, and a larger share of sectors switches to clean production. Markups decrease by almost 3 percentage points to 1.258. R&D investments scaled by GDP increase from 5.7% to 6.2%.

4.2.4 Quantifying the role of the market power channel

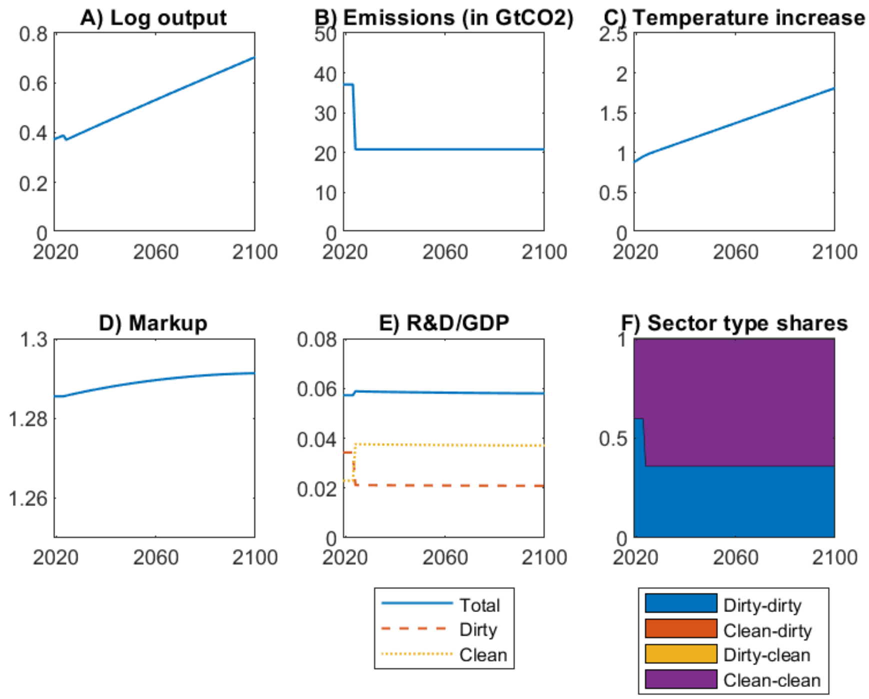
To quantify the role of the market power channel I manipulate the joint distribution of m_L^E, m_L^T and m_F^T in such a manner that there is no immediate effect of the carbon tax on effective technology gaps.[44] Next, I simulate the transition after a change in the $\tilde{\tau}$ to 2.07, which is the optimal tax change according to the model when the market power effect is in place (as in Figure 7). Figure 9 reports the results. Unsurprisingly, I find that fewer firms switch from dirty to clean when the market power channel is switched off because all laggards now have their leader's technology gap, which is more negative on average. Yet, emissions at the moment of the shock decrease slightly more than in Figure 7 (to 20.83 compared to 21.02). That is because a dirty firm with a higher markup produces and thus emits less than an otherwise similar dirty firm with a lower markup (see the discussion under (45); the monopolist is the conservationist's friend), so, all else equal, a decrease in markups puts upward pressure on emissions.[45]

What is surprising, however, is that the drop in output at the moment when the tax is introduced is substantially lower when the market power effect is switched off (1.8% compared

[44] I do this by setting m_{Fj}^T equal to m_{Lj}^T, which means that either both leader and laggard switch or neither firm switches. In both cases the effective gap between them stays the same.

[45] A counteracting force here is that ω_t decreases more at the moment of the shock when the market power effect is switched off, putting upward pressure on emissions (see (45)). In this case the effect through markups dominates.

Figure 9: Transition with no immediate effect on market power ($\tilde{\tau} = 2.07$)



to 3.5%). In fact, this leads to an improvement in welfare. That is, despite lower markups under the market power channel, consumers are better off when the market power channel is switched off. The reason for this surprising result is that less labor is reallocated from production to research, mitigating the initial drop in output caused by the introduction of the tax. When the market power channel is in place the carbon tax increases the mass of sectors that have a small effective technology gap, thus decreasing average markups. These sectors have high research intensities because the escape competition effect is strongest when the effective gap is small. Hence, the carbon tax creates an increase in demand for researchers, leading to a reallocation of labor from production to research and thus an amplified drop in total output (not only because of a switch in technologies but also because of a drop in labor). This reallocation is switched off in the current exercise, meaning that output falls less. Under the current calibration the additional growth due to increased innovation efforts does not fully compensate the initial drop in output, leading to a welfare loss. This is a case of the classical result by Aghion and Howitt (1992) that models of creative destruction can produce growth that is higher than optimal because of conflicting distortionary effects.

Note further that markups gradually increase and that R&D slightly increases as the tax is introduced. This is because the tax change leads to a decrease in the normalized wage ω_t, making R&D cheaper and thus increasing innovation efforts.

As a complementary exercise I numerically computed the optimal tax when the market power channel is switched off. This turns out to be 2.99 innovation steps or 197 dollars at the moment of implementation, which is 44% higher than when the market power effect is present. Figure E4 in the Appendix shows the simulation results for this exercise.

5 Conclusion

This paper makes several contributions to the literature on technological change and the environment. First, I document five empirical facts using a large set of patents linked to financial data. Most importantly, I find that, within sectors, market leaders tend to be more invested in dirty technologies than laggards. Because the direction of innovation is path dependent, this suggests that leaders require a stronger incentive to switch to clean production than their direct competitors. This is a novel finding that cannot be explained using current DTC models.

The second contribution of this paper is to develop an endogenous growth model with directed technical change in clean and dirty technologies and strategic innovation incentives, incorporating the empirical findings discussed above. The model produces a number of interesting insights. First, this paper is the first to show that some incumbent firms have an incentive to increase their dirty innovation investments in response to a carbon tax increase because of strategic incentives. Second, this paper shows how climate policy can directly affect markups, both at the firm level and in aggregate. As a result, climate policy affects not only the direction of innovation but also its intensity. Quantitative exercises using the calibrated model show that for reasonable levels of the carbon tax, climate policy has a modest negative effect on aggregate market power and a modest positive effect on R&D investments. Counter-intuitively, the decrease in markups harms consumers compared to a scenario in which climate policy does not have an immediate effect on markups, because it reallocates labor from production to R&D. This amplifies the initial fall in output and is not fully compensated by faster growth during the transition.

The results of this paper have potentially important policy implications. The finding that climate policy can decrease aggregate market power and increase innovation suggests that climate policy may be less costly than initially anticipated. While the current calibration shows that the reallocation of labor that is caused by the carbon tax change actually harms welfare, a carbon tax combined with another policy intervention, such as a subsidy to (clean) production labor or a tax on (dirty) R&D, could potentially capture the benefits of the market power effect of climate policy. The paper thus highlights that a combination of carbon taxes and clean innovation subsidies, as is proposed in the canonical work by Acemoglu et al. (2012), may not be enough, and should be combined with policies that mitigate the adverse effects of the reallocation of labor. A related lesson that can be drawn from this paper is that climate policies can have complex and counter-intuitive general equilibrium effects. Those who advocate technological solutions to climate change should realize that innovation can be costly if it takes the nature of creative destruction. A final lesson for policymakers is that strategic interactions drive firms' innovation decisions, and are thus critical for the green transition.

This paper also highlights some promising areas for future research. Firms' strategic responses to climate policies and the role of market power are understudied and potentially highly relevant for the green transition. Interesting research areas are the ranking of various policy instruments and complementarities between instruments in a setting with climate policy, market power and strategic innovation incentives. Another promising area is to empirically test some of this paper's predictions, such as the conditions under which incumbents may increase their dirty innovation efforts and the potential for climate policy to strengthen market competition.

References

Acemoglu, D., Aghion, P., Barrage, L., and Hémous, D. (2023). Climate change, directed innovation, and energy transition: The long-run consequences of the shale gas revolution. Working Paper 31657, National Bureau of Economic Research.

Acemoglu, D., Aghion, P., Bursztyn, L., and Hemous, D. (2012). The environment and directed technical change. *American Economic Review*, 102(1):131–66.

Acemoglu, D., Akcigit, U., Hanley, D., and Kerr, W. (2016). Transition to clean technology. *Journal of Political Economy*, 124(1):52–104.

Aghion, P., Bergeaud, A., De Ridder, M., and Van Reenen, J. (2024). Lost in transition: Financial barriers to green growth.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., and Howitt, P. (2005). Competition and innovation: An inverted-U relationship. *The Quarterly Journal of Economics*, 120(2):701–728.

Aghion, P., Bénabou, R., Martin, R., and Roulet, A. (2023). Environmental preferences and technological choices: Is market competition clean or dirty? *American Economic Review: Insights*, 5(1):1–20.

Aghion, P., Dechezleprêtre, A., Hémous, D., Martin, R., and Van Reenen, J. (2016). Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry. *Journal of Political Economy*, 124(1):1–51.

Aghion, P., Hepburn, C., Teytelboym, A., and Zenghelis, D. (2019). Path dependence, innovation and the economics of climate change. *Handbook on green growth*, pages 67–83.

Aghion, P. and Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, 60(2):323–351.

Akcigit, U. and Ates, S. T. (2019). What Happened to U.S. Business Dynamism? NBER Working Paper 25756, National Bureau of Economic Research.

Akcigit, U. and Ates, S. T. (2021). Ten facts on declining business dynamism and lessons from endogenous growth theory. *American Economic Journal: Macroeconomics*, 13(1):257–98.

Akcigit, U. and Ates, S. T. (2023). What happened to US business dynamism? *Journal of Political Economy*, 131(8):2059–2124.

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

Bajgar, M., Berlingieri, G., Calligaris, S., Criscuolo, C., and Timmis, J. (2020). Coverage and representativeness of Orbis data. Science, Technology and Industry Working Paper 2020/06, OECD.

Blundell, R., Griffith, R., and Reenen, J. V. (1995). Dynamic Count Data Models of Technological Innovation. *The Economic Journal*, 105(429):333–344.

Bremer, L. (2020). Energy prices, competition, and environment-related innovation. Working paper.

Calel, R. and Dechezleprêtre, A. (2016). Environmental Policy and Directed Technological Change: Evidence from the European Carbon Market. *The Review of Economics and Statistics*, 98(1):173–191.

Casey, G. (2023). Energy Efficiency and Directed Technical Change: Implications for Climate Change Mitigation. *The Review of Economic Studies*. rdad001.

Compernolle, T., Kort, P. M., and Thijssen, J. J. (2022). The effectiveness of carbon pricing: The role of diversification in a firm's investment decision. *Energy Economics*, 112:106115.

Dasgupta, P., Gilbert, R., and Stiglitz, J. (1983). Strategic considerations in invention and innovation: The case of natural resources. *Econometrica*, 51(5):1439–1448.

Dawid, H., Keoula, M. Y., Kopel, M., and Kort, P. M. (2023). Dynamic investment strategies and leadership in product innovation. *European Journal of Operational Research*, 306(1):431–447.

De Loecker, J. and Eeckhout, J. (2018). Global market power. Working Paper 24768, National Bureau of Economic Research.

Dechezleprêtre, A., Muckley, C. B., and Neelakantan, P. (2021). Is firm-level clean or dirty innovation valued more? *The European Journal of Finance*, 27(1-2):31–61.

Dietz, S., van der Ploeg, F., Rezai, A., and Venmans, F. (2021). Are economists getting climate dynamics right and does it matter? *Journal of the Association of Environmental and Resource Economists*, 8(5):895–921.

Dietz, S. and Venmans, F. (2019). Cumulative carbon emissions and economic policy: In search of general principles. *Journal of Environmental Economics and Management*, 96:108–129.

Díez, F. J., Fan, J., and Villegas-Sánchez, C. (2021). Global declining competition? *Journal of International Economics*, 132:103492.

Eggertsson, G. B., Robbins, J. A., and Wold, E. G. (2021). Kaldor and Piketty's facts: The rise of monopoly power in the United States. *Journal of Monetary Economics*, 124:S19–S38. The Real Interest Rate and the MarginalProduct of Capital in the XXIst CenturyOctober 15-16, 2020.

Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck, J., Le Quéré, C., Luijkx, I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Alkama, R., Arneth, A., Arora, V. K., Bates, N. R., Becker, M., Bellouin, N., Bittig, H. C., Bopp, L., Chevallier, F., Chini, L. P., Cronin, M., Evans, W., Falk, S., Feely, R. A., Gasser, T., Gehlen, M., Gkritzalis, T., Gloege, L., Grassi, G., Gruber, N., Gürses, O., Harris, I., Hefner, M., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Jain, A. K., Jersild, A., Kadono, K., Kato, E., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Landschützer, P., Lefèvre, N., Lindsay, K., Liu, J., Liu, Z., Marland, G., Mayot, N., McGrath, M. J., Metzl, N., Monacci, N. M., Munro, D. R., Nakaoka, S.-I., Niwa, Y., O'Brien, K., Ono, T., Palmer, P. I., Pan, N., Pierrot, D., Pocock, K., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Rodriguez, C., Rosan, T. M., Schwinger, J., Séférian, R., Shutler, J. D., Skjelvan, I., Steinhoff, T., Sun, Q., Sutton, A. J., Sweeney, C., Takao, S., Tanhua, T., Tans, P. P., Tian, X., Tian, H., Tilbrook, B., Tsujino, H., Tubiello, F., van der Werf, G. R., Walker, A. P., Wanninkhof, R., Whitehead, C., Willstrand Wranne, A., Wright, R., Yuan, W., Yue, C., Yue, X., Zaehle, S., Zeng, J., and Zheng, B. (2022). Global carbon budget 2022. *Earth System Science Data*, 14(11):4811–4900.

Gerlagh, R. and Liski, M. (2011). Strategic resource dependence. *Journal of Economic Theory*, 146(2):699–727.

Golosov, M., Hassler, J., Krusell, P., and Tsyvinski, A. (2014). Optimal taxes on fossil fuel in general equilibrium. *Econometrica*, 82(1):41–88.

Gopinath, G., Kalemli-Özcan, , Karabarbounis, L., and Villegas-Sanchez, C. (2017). Capital Allocation and Productivity in South Europe*. *The Quarterly Journal of Economics*, 132(4):1915–1967.

Grossman, G. M. and Helpman, E. (1991). Quality Ladders in the Theory of Growth. *The Review of Economic Studies*, 58(1):43–61.

Grubb, M., Drummond, P., Poncia, A., McDowall, W., Popp, D., Samadi, S., Penasco, C., Gillingham, K. T., Smulders, S., Glachant, M., Hassall, G., Mizuno, E., Rubin, E. S., Dechezleprêtre, A., and Pavan, G. (2021). Induced innovation in energy technologies and systems: a review of evidence and potential implications for CO_2 mitigation. *Environmental Research Letters*, 16(4):043007.

Harris, C. and Vickers, J. (1995). Innovation and natural resources: A dynamic game with uncertainty. *The RAND Journal of Economics*, 26(3):418–430.

Hassler, J., Krusell, P., and Olovsson, C. (2021). Directed technical change as a response to natural resource scarcity. *Journal of Political Economy*, 129(11):3039–3072.

Haščič, I. and Migotto, M. (2015). Measuring environmental innovation using patent data. OECD Environment Working Papers 89, OECD.

Hotelling, H. (1931). The economics of exhaustible resources. *Journal of Political Economy*, 39(2):137–175.

Huber, K. (2018). Disentangling the effects of a banking crisis: Evidence from german firms and counties. *American Economic Review*, 108(3):868–98.

Huisman, K. J. and Kort, P. M. (2003). Strategic investment in technological innovations. *European Journal of Operational Research*, 144(1):209–223.

Huisman, K. J. and Kort, P. M. (2004). Strategic technology adoption taking into account future technological improvements: A real options approach. *European Journal of Operational Research*, 159(3):705–728.

IEA (2021). Methodology for identifying fossil fuel supply related technologies in patent data. Annex to report Patents and the energy transition, EPO and IEA/OECD.

IPCC (2023). Climate change 2023: Synthesis report. Synthesis report of the sixth assessment report, Intergovernmental Panel on Climate Change, Geneva, Switzerland.

Jaakkola, N. (2019). Carbon taxation, opec and the end of oil. *Journal of Environmental Economics and Management*, 94:101–117.

Jaffe, A. B. and Palmer, K. (1997). Environmental Regulation and Innovation: A Panel Data Study. *The Review of Economics and Statistics*, 79(4):610–619.

Jee, S. J. and Srivastav, S. (2023). Knowledge spillovers between clean and dirty technologies. *Available at SSRN 4233536*.

Johnstone, N., Haščič, I., and Popp, D. (2010). Renewable energy policies and technological innovation: evidence based on patent counts. *Environmental and Resource Economics*, 45(1):133–155.

Kalemli-Ozcan, S., Sorensen, B., Villegas-Sanchez, C., Volosovych, V., and Yesiltas, S. (2023). How to construct nationally representative firm level data from the orbis global database: New facts and aggregate implications. *American Economic Journal: Macroeconomics*.

Klette, T. J. and Kortum, S. (2004). Innovating firms and aggregate innovation. *Journal of Political Economy*, 112(5):986–1018.

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

Linn, J. (2008). Energy Prices and the Adoption of Energy-Saving Technology. *The Economic Journal*, 118(533):1986–2012.

Matthews, H. D., Gillett, N. P., Stott, P. A., and Zickfeld, K. (2009). The proportionality of global warming to cumulative carbon emissions. *Nature*, 459(7248):829–832.

Mullahy, J. and Norton, E. C. (2022). Why transform y? a critical assessment of dependent-variable transformations in regression models for skewed and sometimes-zero outcomes. Working Paper 30735, National Bureau of Economic Research.

Newell, R. G., Jaffe, A. B., and Stavins, R. N. (1999). The Induced Innovation Hypothesis and Energy-Saving Technological Change*. *The Quarterly Journal of Economics*, 114(3):941–975.

Noailly, J. and Smeets, R. (2015). Directing technical change from fossil-fuel to renewable energy innovation: An application using firm-level patent data. *Journal of Environmental Economics and Management*, 72:15–37.

Noailly, J. and Smeets, R. (2021). Financing energy innovation: Internal finance and the direction of technical change. *Environmental and Resource Economics*, pages 1–25.

Nordhaus, W. (2014). *A question of balance: Weighing the options on global warming policies*. Yale University Press.

Nordhaus, W. D. and Moffat, A. (2017). A survey of global impacts of climate change: Replication, survey methods, and a statistical analysis. Working Paper 23646, National Bureau of Economic Research.

Pindyck, R. S. (2013). Climate change policy: What do the models tell us? *Journal of Economic Literature*, 51(3):860–72.

Pindyck, R. S. (2021). What we know and don't know about climate change, and implications for policy. *Environmental and Energy Policy and the Economy*, 2:4–43.

Popp, D. (2002). Induced innovation and energy prices. *American Economic Review*, 92(1):160–180.

Popp, D., Pless, J., Haščič, I., and Johnstone, N. (2022). 4. innovation and entrepreneurship in the energy sector. In Andrews, M. J., Chatterji, A., Lerner, J., and Stern, S., editors, *The Role of Innovation and Entrepreneurship in Economic Growth*, pages 175–248. University of Chicago Press, Chicago.

Porter, M. E. (1990). *The Competitive Advantage of Nations*. Free Press, New York.

Porter, M. E. and van der Linde, C. (1995). Toward a new conception of the environment-competitiveness relationship. *Journal of Economic Perspectives*, 9(4):97–118.

Probst, B., Touboul, S., Glachant, M., and Dechezleprêtre, A. (2021). Global trends in the invention and diffusion of climate change mitigation technologies. *Nature Energy*, 6(11):1077–1086.

Ricke, K. L. and Caldeira, K. (2014). Maximum warming occurs about one decade after a carbon dioxide emission. *Environmental Research Letters*, 9(12):124002.

Rozendaal, R. and Vollebergh, H. (2024). Policy-induced innovation in clean technologies: Evidence from the car market. Working paper.

Schumpeter, J. A. (1942). *Capitalism, Socialism and Democracy*. Harper and Brothers, New York.

Smulders, S. and de Nooij, M. (2003). The impact of energy conservation on technology and economic growth. *Resource and Energy Economics*, 25(1):59–79.

Solow, R. M. (1974). The economics of resources or the resources of economics. *American Economic Review*, 64(2):1–14.

Stern, N. (2007). *The Economics of Climate Change: The Stern Review*. Cambridge University Press.

Ustyuzhanina, P., Ferguson, S., Norbäck, P.-J., and Persson, L. (2022). Who is doing clena tech innovation? Evidence from Swedish firms. Working paper.

Appendix

A Data

I make use of the Orbis Intellectual Property (IP) and Orbis Historical databases, both of which are managed by Bureau van Dijk. Orbis IP consists of about 130 million separate patent applications at patent offices all over the world that go back to the 1800s. The novel feature of the Orbis IP database is that patent applicants are assigned an identifier that can be linked to the Orbis Historical database, which has balance sheets and other financial data on millions of firms in many countries. Whereas the Orbis financial data is used quite widely in economics, the the patent data set is rarely used.[1] The link between financial and patent data is critical for this paper, as it allows me to distinguish market leaders and laggards. The limitations of the Orbis Historical database are well documented (Bajgar et al., 2020).[2] I follow Kalemli-Ozcan et al. (2023) in constructing a representative sample and focus on large firms and recent years, for which coverage is good. My sample runs from 1978, the year in which the European Patent Office was established, until 2018.[3] When using the financial data, I only cover 2010 until 2018 because the coverage of the financial variables is limited in earlier years.

I use patent counts as a measure of innovation, as is common in the literature. Patent applications are an outcome of the innovation process and are thus a measure of *successful* innovation efforts. While not all innovations are protected by patents, aggregate patenting can account for substantial variation in economic growth and productivity (Kogan et al., 2017). I use the number of "triadic patent families" as my innovation measure, which results

[1]Orbis Historical is used by Gopinath et al. (2017) and Díez et al. (2021), for instance. Kalemli-Ozcan et al. (2023) and Bajgar et al. (2020) investigate the coverage and representativeness of Orbis Historical and provide guidance in constructing a representative sample. Huber (2018) and Noailly and Smeets (2021) are among the few papers that use the link between the patenting and financial data sets.

[2]The coverage of firms in the Orbis Historical database varies widely over time and between countries, as does the coverage of variables. Bajgar et al. (2020) document that large firms are overrepresented in the database, even within commonly used size bins.

[3]Due to lags in the grant and publication process of patents, databases are only complete after about 4 years (Aghion et al., 2016). Indeed, I see a sharp drop in total patents in 2019 (I collected my data in 2022), so I end my sample in 2018.

in a sample of 1.4 million innovations between 1978 and 2018.[4] An important advantage of using patents as a measure of innovation is that they provide detailed information about the technology that they protect. This allows researchers to classify patents into different categories, such as those aimed at mitigating climate change, or those related to extracting fossil fuels.

I follow the most recent literature on clean and dirty patent classifications (Jee and Srivastav, 2023) to classify all triadic patent families as either “clean”, “dirty” or “neutral” (neither clean nor dirty).[5] Both clean and dirty technologies can be placed broadly in the categories energy, manufacturing and transportation, though clean includes a few additional categories like buildings and carbon capture.[6] Dirty technologies can be disaggregated into those that improve efficiency (making a dirty technology less dirty), which are referred to as “gray”, and those that do not.

Figure A1 shows the share of all triadic patents that are classified as clean or dirty over time. Perhaps not surprisingly, dirty technologies have historically accounted for a larger share of innovations than their clean counterparts. Over 5% of all innovations have been

[4]Firms can apply for multiple patents to protect a single invention. All patents that cover one invention are grouped together as a patent family. I count patents at the family level to avoid double counting. INPADOC families are used for this study, which include all patents with a common priority patent. The priority year (filing year of the priority patent) is used to assign a year to an invention. Patent values are highly heterogeneous. I therefore restrict my sample to “triadic” patent families, which are inventions for which a patent has been filed at the patent offices of the EU, Japan and the US (Aghion et al., 2016). Focusing on triadic families eliminates low value patents from the sample. The idea is that, since a patent application is costly, firms only apply at the three main patent offices if their invention is sufficiently valuable and has the potential for international adoption. Orbis IP contains 130 million patent applications that belong to 69 million patent families, of which 1.4 million are triadic.

[5]Jee and Srivastav (2023) collect all the relevant CPC and IPC technology classes from the literature in Supplementary Table 1 in their Appendix. Their table includes technology classes from Haščič and Migotto (2015), Aghion et al. (2016), Dechezleprêtre et al. (2021), IEA (2021) and Popp et al. (2022).

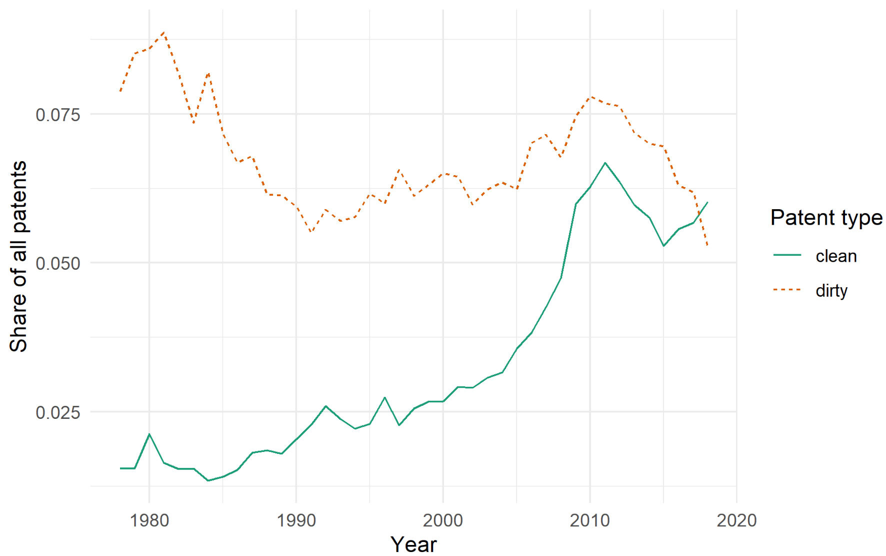
[6]The included clean technologies are buildings, carbon capture, storage and sequestration, clean energy (geothermal, hydro, nuclear, photovoltaic, solar thermal, thermal-PV hybrids, wind and others), clean ICT, manufacturing (agriculture, chemical, consumer products, enabling technologies, metal), smart grids, transportation (electric, hybrid and hydrogen vehicles and fuel cells), and waste management. Dirty technologies are energy from fossil fuels (upstream, processing and downstream, transmission and distribution), manufacturing (relating to several industries like steel and chemicals), and transport (internal combustion engine).

related to fossil fuels in every year between 1978 and 2018. Dirty innovation is thus persistent, although its share has been falling quite sharply since 2010. The clean innovation share was low until a rapid increase that started around 2005. It then fell after peaking in 2011 but has since slightly overtaken dirty innovation in terms of yearly counts.[7] Figure A2 further shows that while the shares of clean and dirty patenting have been quite stable over time, their absolute numbers have drastically increased over the sample period. Figure A3 breaks down clean innovations into different technology groups. Transport constitutes the largest share, especially since 2000. Figure A4 shows that the majority of dirty patents are related to manufacturing, and figure A5 shows that the share of dirty innovations that are classified as gray has increased from less than half in the 1980s to almost three quarters in 2018.

The firm identifier in the IP data set allows me to match firms to the Historical data set. Of the 1.4 million triadic patents, 1.33 million can be matched, meaning that I observe the country where the firm is based and the main sector in which it operates. Figures A6 and 1a in the Appendix show the distribution of patenting across countries and sectors, respectively, for the entire sample and for clean and dirty patents separately. A few things stand out. First, Japanese firms are by far the most actively patenting. Over 40% of clean triadic patents families are filed by Japanese firms. Second, the largest European countries (Germany, France, UK) have a larger share of dirty than of clean patents, whereas the largest Asian countries (in terms of patenting, Japan, Korea and China) have a larger share for clean than for dirty technologies. The US has a similar share for both types. Third, firms from the top 10 patenting countries have applied for almost 90% of all triadic patents (for the entire sample, as well as for clean and dirty specifically). Turning to the sector distribution, the most striking finding is perhaps that no sector is particularly dominant and that both clean and dirty patenting occur in a wide variety of sectors.

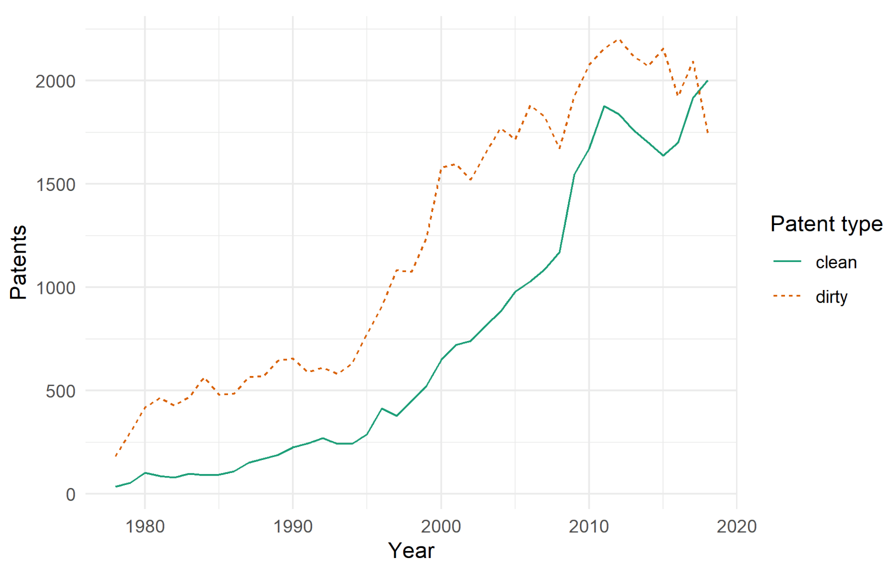
[7]This fall in clean innovation is attributed to low energy prices, clean technologies reaching maturity, and the effects of the great recession which hit financing for clean R&D especially hard (Probst et al., 2021; Acemoglu et al., 2023; Aghion et al., 2024).

Figure A1: Share of clean and dirty patents over time



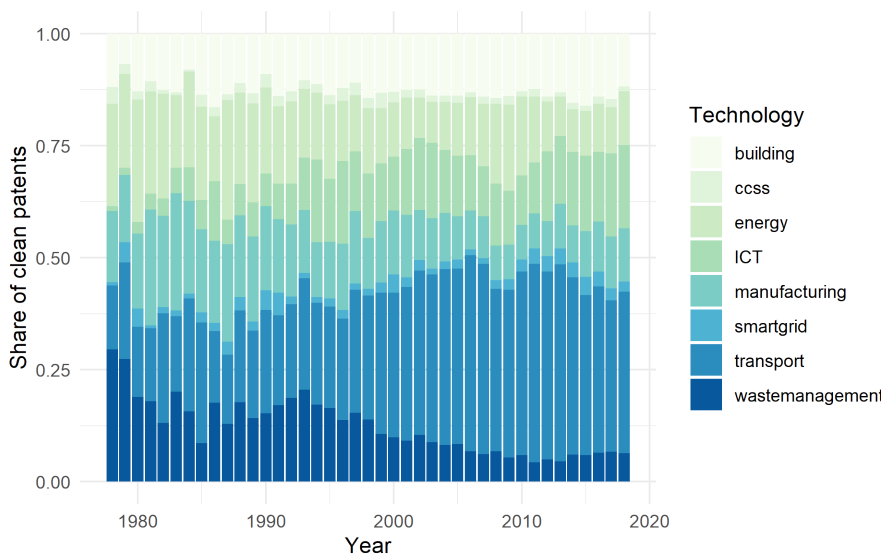
Notes: Data source: Orbis IP. Share of triadic patent families that are classified as clean or dirty. See text for details about patent classifications.

Figure A2: Total clean and dirty patents over time



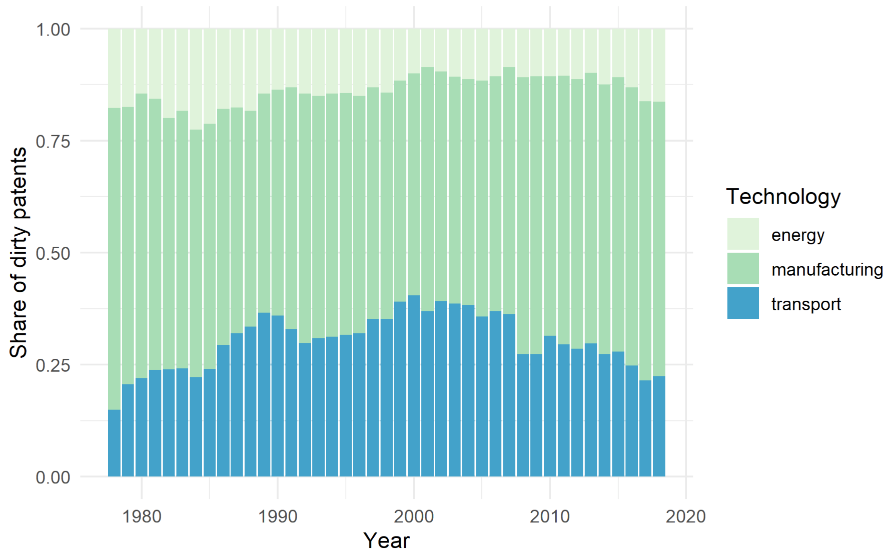
Notes: Data source: Orbis IP. Total number of triadic patent families that are classified as clean or dirty. See text for details about patent classifications.

Figure A3: Shares of different clean technology types over time



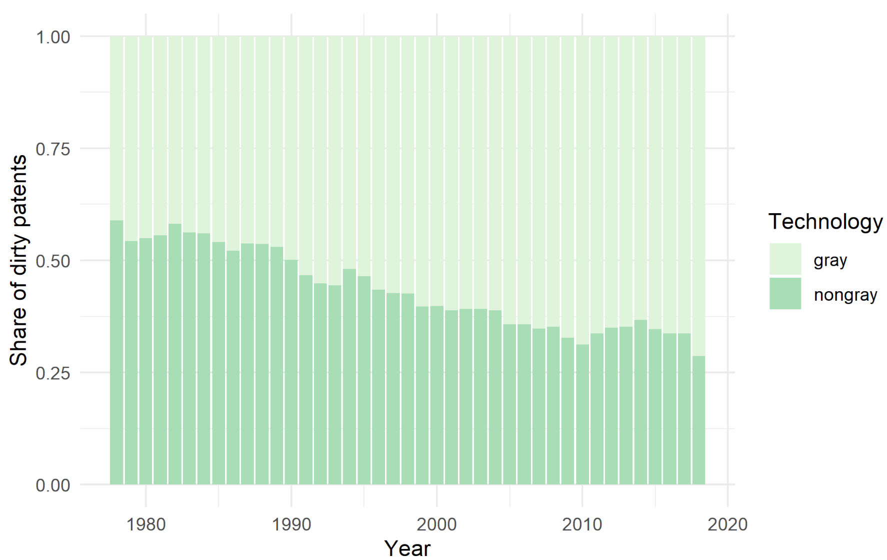
Notes: Data source: Orbis IP. Share of clean triadic patent families that belong to different technology groups. The included clean technologies are buildings, carbon capture, storage and sequestration, clean energy (geothermal, hydro, nuclear, photovoltaic, solar thermal, thermal-PV hybrids, wind and others), clean ICT, manufacturing (agriculture, chemical, consumer products, enabling technologies, metal), smart grids, transportation (electric, hybrid and hydrogen vehicles and fuel cells), and waste management.

Figure A4: Shares of different dirty technology types over time



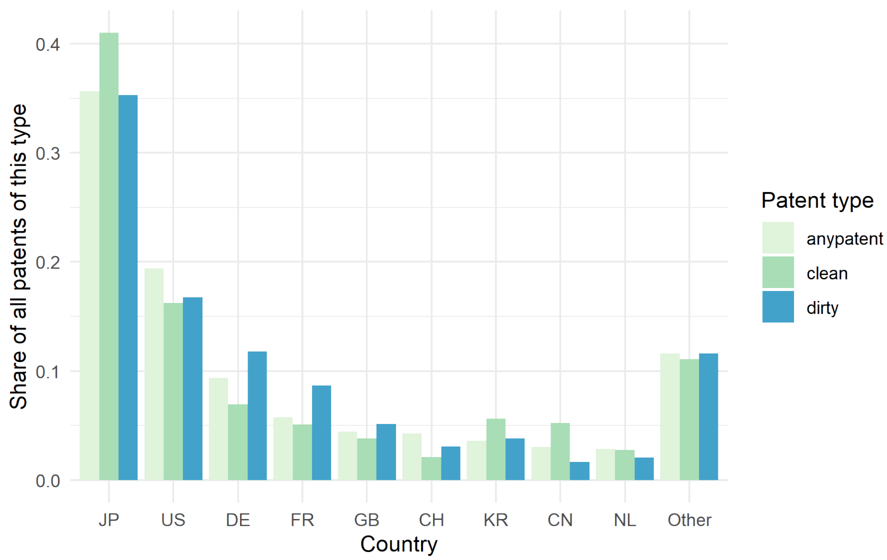
Notes: Data source: Orbis IP. Share of dirty triadic patent families that belong to different technology groups. The included dirty technologies are energy from fossil fuels (upstream, processing and downstream, transmission and distribution), manufacturing (relating to several industries like steel and chemicals), and transport (internal combustion engine).

Figure A5: Share of dirty patents that are gray



Notes: Data source: Orbis IP. Share of dirty triadic patent families that are classified as gray, meaning that they improve the (energy) efficiency of the particular technology.

Figure A6: Patents by applicant country



Notes: Data sources: Orbis IP and Historical. Anypatent refers to the distribution across countries of the entire sample (including clean, dirty and neutral).

B Proofs

Proof of Proposition 1.

If $\overline{m} = 1$, and the aggregate economy is on a BGP, so $\omega_t = \omega$, then the dynamic problem of the firm consists of 7 value functions. Define a positive number in the interval between 0 and 1, $p \in (0,1)$. The value functions are then

$$\begin{aligned}
\rho v_1 &= \max_{x_1} \Big\{ 1 - \frac{1}{\lambda} - \alpha \frac{x_1^2}{2} \omega + x_{-1}[v_0 - v_1] + \delta[v_0 - v_1] \Big\}, \\
\rho v_0 &= \max_{x_0} \Big\{ -\alpha \frac{x_0^2}{2} \omega + x_0[v_1 - v_0] + \overline{x}_0[v_{-1} - v_0] \Big\}, \\
\rho v_{-1} &= \max_{x_{-1}} \Big\{ -\alpha \frac{x_{-1}^2}{2} \omega + x_{-1}[v_0 - v_{-1}] + \delta[v_0 - v_{-1}] \Big\}, \\
\rho v_p &= \max_{x_p} \Big\{ 1 - \frac{1}{\lambda^p} - \alpha \frac{x_p^2}{2} \omega + x_p[v_1 - v_p] + x_{-p}[v_{p-1} - v_p] + \delta[v_0 - v_p] \Big\}, \\
\rho v_{-p} &= \max_{x_{-p}} \Big\{ -\alpha \frac{x_{-p}^2}{2} \omega + x_1[v_{1-p} - v_{-p}] + x_p[v_{-1} - v_{-p}] + \delta[v_0 - v_{-p}] \Big\}, \\
\rho v_{1-p} &= \max_{x_{1-p}} \Big\{ 1 - \frac{1}{\lambda^{1-p}} - \alpha \frac{x_{1-p}^2}{2} \omega + x_{1-p}[v_1 - v_{1-p}] + x_{p-1}[v_{-p} - v_{1-p}] + \delta[v_0 - v_{1-p}] \Big\}, \\
\rho v_{p-1} &= \max_{x_{p-1}} \Big\{ -\alpha \frac{x_{p-1}^2}{2} \omega + x_{p-1}[v_p - v_{p-1}] + x_{1-p}[v_{-1} - v_{p-1}] + \delta[v_0 - v_{p-1}] \Big\}.
\end{aligned}$$

So, given τ, the effective technology gap can take 7 values, of which 3 are integers $(-1, 0, 1)$ and 4 are non-integers $(p, -p, 1-p, p-1)$ (only 2 non-integers if $p = 0.5$). Note that the first three value functions (the integer ones) do not depend on the non-integer states. The reason for this is that once the gap is an integer, it can never become a non-integer again as long as τ is fixed (because firms innovate only in their better technology and a successful innovation always leads to a step of 1. Note further that if firm i has a gap of $m_{it}^E = -p$, i.e. it is a laggard and uses a different technology than its competitor, and its competitor innovates successfully, then m_{it}^E becomes -1 as firm i switches to the technology that its competitor is using. This is the case because $\overline{m} = 1$ implies that a successful innovation by a leader in a particular technology (the firm already has $m_{it}^C = 1$ or $m_{it}^D = 1$ but still innovates in that technology) also leads its competitor to step up by 1, so the gap stays 1.

The first order conditions associated with the above value functions are

$$\begin{aligned}
x_1 &= 0, \\
x_0 &= \frac{1}{\alpha\omega}[v_1 - v_0], \\
x_{-1} &= \frac{1}{\alpha\omega}[v_0 - v_{-1}], \\
x_p &= \frac{1}{\alpha\omega}[v_1 - v_p], \\
x_{-p} &= \frac{1}{\alpha\omega}[v_{1-p} - v_{-p}], \\
x_{1-p} &= \frac{1}{\alpha\omega}[v_1 - v_{1-p}], \\
x_{p-1} &= \frac{1}{\alpha\omega}[v_p - v_{p-1}].
\end{aligned}$$

Profits are strictly increasing in m_{it}^E for $m_{it}^E > 0$ and thus in p. Hence, we can order 5 of the above value functions as follows:

$$v_1 > v_p > v_0 > v_{-p} > v_{-1} > 0.$$

From this, together with the first order conditions, it directly follows that $x_1 < x_p < x_0$ and that $\frac{dx_p}{dp} < 0$. So, as the positive technology gap gets closer to 1, innovation efforts get closer to 0. This proves the downward sloping part of the innovation efforts for positive m_{it}^E.

To prove the upward sloping efforts for negative m_{it}^E, let us first take the following difference.

$$\begin{aligned}
\rho v_0 - \rho v_{-1} &= \alpha\omega\frac{x_0^2}{2} - x_0(v_0 - v_{-1}) - \alpha\omega\frac{x_{-1}^2}{2} - \delta(v_0 - v_{-1}) \\
x_0^2 &= x_{-1}^2 + \frac{2}{\alpha\omega}(\rho + \delta + x_0)(v_0 - v_{-1})
\end{aligned}$$

Now, since $v_0 > v_{-1}$ and $\alpha, \omega, \rho, \delta, x_0 > 0$, the second term on the right hand side is positive, and hence, $x_0 > x_{-1}$.

Let us now take another difference to compare x_{-p} to x_0 and x_{-1}.

$$\begin{aligned}
\rho v_{-p} - \rho v_{-1} &= \alpha\omega\frac{x_{-p}^2}{2} - x_p(v_{-p} - v_{-1}) - \delta(v_{-p} - v_0) - \alpha\omega\frac{x_{-1}^2}{2} - \delta(v_0 - v_{-1}) \\
x_{-p}^2 &= x_{-1}^2 + \frac{2}{\alpha\omega}(\rho + \delta + x_p)(v_{-p} - v_{-1})
\end{aligned}$$

Now, clearly, by the same argument as above, $x_{-p} > x_{-1}$. Furthermore, since $x_p < x_0$ and $v_{-p} < v_0$, comparing x_0^2 to x_{-p}^2 yields that $x_{-p} < x_0$. In fact, since $\frac{dx_p}{dp} < 0$ and $\frac{dv_{-p}}{dp} < 0$, it turns out that x_{-p} is decreasing in p. So, as p gets closer to 0, both x_p and x_{-p} approach x_0. As p gets closer to 1, x_{-p} gets closer to x_{-1}.

This establishes proposition 1.

Proof of Proposition 2.

Denote the market leader by L, the laggard or follower by F, and the time period just before the (unanticipated) tax change as 0. Now suppose that $\tilde{\tau}_0 = 0$, $m^T_{L0} = -1$, $m^T_{F0} = 0$, $m^D_{L0} = 1$, $m^C_{L0} = 0$, and hence $m^E_{L0} = 1$. That is, the leader uses the dirty technology, in which it has a lead of 1, while both firms are equally good at the clean technology. It follows from Proposition 1 that since the leader has the maximum lead, it does not invest in innovation.

Now suppose the carbon tax is increased and denote this period by 1: $\tilde{\tau}_1 = \frac{1}{2}$. The leader still uses the dirty technology (since $m^T_{L1} + \tilde{\tau}_1 = -\frac{1}{2} < 0$) but its lead is decreased from 1 to $m^E_{L1} = \frac{1}{2}$. Following Proposition 1, $x_{\frac{1}{2}} > x_1$, so the leader innovates more than it did before. Since it still uses the dirty technology ($m^T_{L1} + \tilde{\tau}_1 < 0$), the tax causes dirty innovation by firm L to increase.

This establishes Proposition 2.

C Full characterization of results

C.1 Firm/sector level

Below are the full characterizations of prices, profits, output, labor and emissions. Note that $\tilde{\tau}_t \equiv \frac{\ln(1+\kappa\tau_t)}{\ln(\lambda)}$.

$$p_{jt} = \begin{cases} \frac{w_t}{q^C_{-it}} & \text{if} \quad m^T_{it} + \tilde{\tau}_t \geq 0, \quad m^T_{-it} + \tilde{\tau}_t \geq 0, \quad m^C_{it} \geq 0 \\ \frac{w_t}{q^C_{it}} & \text{if} \quad m^T_{it} + \tilde{\tau}_t \geq 0, \quad m^T_{-it} + \tilde{\tau}_t \geq 0, \quad m^C_{it} \leq 0 \\ \frac{w_t(1+\kappa\tau_t)}{q^D_{-it}} & \text{if} \quad m^T_{it} + \tilde{\tau}_t \geq 0, \quad m^T_{-it} + \tilde{\tau}_t < 0, \quad m^D_{it} + m^T_{it} + \tilde{\tau}_t \geq 0 \\ \frac{w_t}{q^C_{it}} & \text{if} \quad m^T_{it} + \tilde{\tau}_t \geq 0, \quad m^T_{-it} + \tilde{\tau}_t < 0, \quad m^D_{it} + m^T_{it} + \tilde{\tau}_t \leq 0 \\ \frac{w_t}{q^C_{-it}} & \text{if} \quad m^T_{it} + \tilde{\tau}_t < 0, \quad m^T_{-it} + \tilde{\tau}_t \geq 0, \quad m^C_{it} - m^T_{it} - \tilde{\tau}_t \geq 0 \\ \frac{w_t(1+\kappa\tau_t)}{q^D_{it}} & \text{if} \quad m^T_{it} + \tilde{\tau}_t < 0, \quad m^T_{-it} + \tilde{\tau}_t \geq 0, \quad m^C_{it} - m^T_{it} - \tilde{\tau}_t \leq 0 \\ \frac{w_t(1+\kappa\tau_t)}{q^D_{-it}} & \text{if} \quad m^T_{it} + \tilde{\tau}_t < 0, \quad m^T_{-it} + \tilde{\tau}_t < 0, \quad m^D_{it} \geq 0 \\ \frac{w_t(1+\kappa\tau_t)}{q^D_{it}} & \text{if} \quad m^T_{it} + \tilde{\tau}_t < 0, \quad m^T_{-it} + \tilde{\tau}_t < 0, \quad m^D_{it} \leq 0 \end{cases}$$

$$\pi_{it} = \begin{cases} \left(1 - \frac{1}{\lambda^{m^C_{it}}}\right) Y_t & \text{if} \quad m^T_{it} + \tilde{\tau}_t \geq 0, \quad m^T_{-it} + \tilde{\tau}_t \geq 0, \quad m^C_{it} > 0 \\ 0 & \text{if} \quad m^T_{it} + \tilde{\tau}_t \geq 0, \quad m^T_{-it} + \tilde{\tau}_t \geq 0, \quad m^C_{it} \leq 0 \\ \left(1 - \frac{1}{\lambda^{m^D_{it} + m^T_{it} + \tilde{\tau}_t}}\right) Y_t & \text{if} \quad m^T_{it} + \tilde{\tau}_t \geq 0, \quad m^T_{-it} + \tilde{\tau}_t < 0, \quad m^D_{it} + m^T_{it} + \tilde{\tau}_t > 0 \\ 0 & \text{if} \quad m^T_{it} + \tilde{\tau}_t \geq 0, \quad m^T_{-it} + \tilde{\tau}_t < 0, \quad m^D_{it} + m^T_{it} + \tilde{\tau}_t \leq 0 \\ \left(1 - \frac{1}{\lambda^{m^C_{it} - m^T_{it} - \tilde{\tau}_t}}\right) Y_t & \text{if} \quad m^T_{it} + \tilde{\tau}_t < 0, \quad m^T_{-it} + \tilde{\tau}_t \geq 0, \quad m^C_{it} - m^T_{it} - \tilde{\tau}_t > 0 \\ 0 & \text{if} \quad m^T_{it} + \tilde{\tau}_t < 0, \quad m^T_{-it} + \tilde{\tau}_t \geq 0, \quad m^C_{it} - m^T_{it} - \tilde{\tau}_t \leq 0 \\ \left(1 - \frac{1}{\lambda^{m^D_{it}}}\right) Y_t & \text{if} \quad m^T_{it} + \tilde{\tau}_t < 0, \quad m^T_{-it} + \tilde{\tau}_t < 0, \quad m^D_{it} > 0 \\ 0 & \text{if} \quad m^T_{it} + \tilde{\tau}_t < 0, \quad m^T_{-it} + \tilde{\tau}_t < 0, \quad m^D_{it} \leq 0 \end{cases}$$

$$
y_{it} = \begin{cases}
\frac{Y_t}{w_t} q^C_{-it} & \text{if} \quad m^T_{it} + \tilde{\tau}_t \geq 0, \quad m^T_{-it} + \tilde{\tau}_t \geq 0, \quad m^C_{it} > 0 \\
\frac{Y_t}{2w_t} q^C_{-it} & \text{if} \quad m^T_{it} + \tilde{\tau}_t \geq 0, \quad m^T_{-it} + \tilde{\tau}_t \geq 0, \quad m^C_{it} = 0 \\
0 & \text{if} \quad m^T_{it} + \tilde{\tau}_t \geq 0, \quad m^T_{-it} + \tilde{\tau}_t \geq 0, \quad m^C_{it} < 0 \\
\frac{Y_t}{w_t(1+\kappa\tau_t)} q^D_{-it} & \text{if} \quad m^T_{it} + \tilde{\tau}_t \geq 0, \quad m^T_{-it} + \tilde{\tau}_t < 0, \quad m^D_{it} + m^T_{it} + \tilde{\tau}_t > 0 \\
\frac{Y_t}{2w_t(1+\kappa\tau_t)} q^D_{-it} & \text{if} \quad m^T_{it} + \tilde{\tau}_t \geq 0, \quad m^T_{-it} + \tilde{\tau}_t < 0, \quad m^D_{it} + m^T_{it} + \tilde{\tau}_t = 0 \\
0 & \text{if} \quad m^T_{it} + \tilde{\tau}_t \geq 0, \quad m^T_{-it} + \tilde{\tau}_t < 0, \quad m^D_{it} + m^T_{it} + \tilde{\tau}_t < 0 \\
\frac{Y_t}{w_t} q^C_{-it} & \text{if} \quad m^T_{it} + \tilde{\tau}_t < 0, \quad m^T_{-it} + \tilde{\tau}_t \geq 0, \quad m^C_{it} - m^T_{it} - \tilde{\tau}_t > 0 \\
\frac{Y_t}{2w_t} q^C_{-it} & \text{if} \quad m^T_{it} + \tilde{\tau}_t < 0, \quad m^T_{-it} + \tilde{\tau}_t \geq 0, \quad m^C_{it} - m^T_{it} - \tilde{\tau}_t = 0 \\
0 & \text{if} \quad m^T_{it} + \tilde{\tau}_t < 0, \quad m^T_{-it} + \tilde{\tau}_t \geq 0, \quad m^C_{it} - m^T_{it} - \tilde{\tau}_t < 0 \\
\frac{Y_t}{w_t(1+\kappa\tau_t)} q^D_{-it} & \text{if} \quad m^T_{it} + \tilde{\tau}_t < 0, \quad m^T_{-it} + \tilde{\tau}_t < 0, \quad m^D_{it} > 0 \\
\frac{Y_t}{2w_t(1+\kappa\tau_t)} q^D_{-it} & \text{if} \quad m^T_{it} + \tilde{\tau}_t < 0, \quad m^T_{-it} + \tilde{\tau}_t < 0, \quad m^D_{it} = 0 \\
0 & \text{if} \quad m^T_{it} + \tilde{\tau}_t < 0, \quad m^T_{-it} + \tilde{\tau}_t < 0, \quad m^D_{it} < 0
\end{cases}
$$

$$
l_{it} = \begin{cases}
\frac{Y_t}{w_t}\frac{1}{\lambda^{m_{it}^C}} & \text{if} \quad m_{it}^T + \tilde{\tau}_t \geq 0, \quad m_{-it}^T + \tilde{\tau}_t \geq 0, \quad m_{it}^C > 0 \\
\frac{Y_t}{2w_t}\frac{1}{\lambda^{m_{it}^C}} & \text{if} \quad m_{it}^T + \tilde{\tau}_t \geq 0, \quad m_{-it}^T + \tilde{\tau}_t \geq 0, \quad m_{it}^C = 0 \\
0 & \text{if} \quad m_{it}^T + \tilde{\tau}_t \geq 0, \quad m_{-it}^T + \tilde{\tau}_t \geq 0, \quad m_{it}^C < 0 \\
\frac{Y_t}{w_t(1+\kappa\tau_t)}\frac{1}{\lambda^{m_{it}^D + m_{it}^T}} & \text{if} \quad m_{it}^T + \tilde{\tau}_t \geq 0, \quad m_{-it}^T + \tilde{\tau}_t < 0, \quad m_{it}^D + m_{it}^T + \tilde{\tau}_t > 0 \\
\frac{Y_t}{2w_t(1+\kappa\tau_t)}\frac{1}{\lambda^{m_{it}^D + m_{it}^T}} & \text{if} \quad m_{it}^T + \tilde{\tau}_t \geq 0, \quad m_{-it}^T + \tilde{\tau}_t < 0, \quad m_{it}^D + m_{it}^T + \tilde{\tau}_t = 0 \\
0 & \text{if} \quad m_{it}^T + \tilde{\tau}_t \geq 0, \quad m_{-it}^T + \tilde{\tau}_t < 0, \quad m_{it}^D + m_{it}^T + \tilde{\tau}_t < 0 \\
\frac{Y_t}{w_t}\frac{1}{\lambda^{m_{it}^C - m_{it}^T}} & \text{if} \quad m_{it}^T + \tilde{\tau}_t < 0, \quad m_{-it}^T + \tilde{\tau}_t \geq 0, \quad m_{it}^C - m_{it}^T - \tilde{\tau}_t > 0 \\
\frac{Y_t}{2w_t}\frac{1}{\lambda^{m_{it}^C - m_{it}^T}} & \text{if} \quad m_{it}^T + \tilde{\tau}_t < 0, \quad m_{-it}^T + \tilde{\tau}_t \geq 0, \quad m_{it}^C - m_{it}^T - \tilde{\tau}_t = 0 \\
0 & \text{if} \quad m_{it}^T + \tilde{\tau}_t < 0, \quad m_{-it}^T + \tilde{\tau}_t \geq 0, \quad m_{it}^C - m_{it}^T - \tilde{\tau}_t < 0 \\
\frac{Y_t}{w_t(1+\kappa\tau_t)}\frac{1}{\lambda^{m_{it}^D}} & \text{if} \quad m_{it}^T + \tilde{\tau}_t < 0, \quad m_{-it}^T + \tilde{\tau}_t < 0, \quad m_{it}^D > 0 \\
\frac{Y_t}{2w_t(1+\kappa\tau_t)}\frac{1}{\lambda^{m_{it}^D}} & \text{if} \quad m_{it}^T + \tilde{\tau}_t < 0, \quad m_{-it}^T + \tilde{\tau}_t < 0, \quad m_{it}^D = 0 \\
0 & \text{if} \quad m_{it}^T + \tilde{\tau}_t < 0, \quad m_{-it}^T + \tilde{\tau}_t < 0, \quad m_{it}^D < 0
\end{cases}
$$

$$
e_{it} = \begin{cases}
0 & \text{if} \quad m_{it}^T + \tilde{\tau}_t \geq 0 \\
\frac{\kappa Y_t}{w_t}\frac{1}{\lambda^{m_{it}^C - m_{it}^T}} & \text{if} \quad m_{it}^T + \tilde{\tau}_t < 0, \quad m_{-it}^T + \tilde{\tau}_t \geq 0, \quad m_{it}^C - m_{it}^T - \tilde{\tau}_t > 0 \\
\frac{\kappa Y_t}{2w_t}\frac{1}{\lambda^{m_{it}^C - m_{it}^T}} & \text{if} \quad m_{it}^T + \tilde{\tau}_t < 0, \quad m_{-it}^T + \tilde{\tau}_t \geq 0, \quad m_{it}^C - m_{it}^T - \tilde{\tau}_t = 0 \\
0 & \text{if} \quad m_{it}^T + \tilde{\tau}_t < 0, \quad m_{-it}^T + \tilde{\tau}_t \geq 0, \quad m_{it}^C - m_{it}^T - \tilde{\tau}_t < 0 \\
\frac{\kappa Y_t}{w_t(1+\kappa\tau_t)}\frac{1}{\lambda^{m_{it}^D}} & \text{if} \quad m_{it}^T + \tilde{\tau}_t < 0, \quad m_{-it}^T + \tilde{\tau}_t < 0, \quad m_{it}^D > 0 \\
\frac{\kappa Y_t}{2w_t(1+\kappa\tau_t)}\frac{1}{\lambda^{m_{it}^D}} & \text{if} \quad m_{it}^T + \tilde{\tau}_t < 0, \quad m_{-it}^T + \tilde{\tau}_t < 0, \quad m_{it}^D = 0 \\
0 & \text{if} \quad m_{it}^T + \tilde{\tau}_t < 0, \quad m_{-it}^T + \tilde{\tau}_t < 0, \quad m_{it}^D < 0
\end{cases}
$$

C.2 Law of motion Ψ and μ

C.2.1 Ψ

Add law of motion ψ here.

C.2.2 μ

The laws of motion for μ when the effective technology gap is either very small or very large are as follows for CC sectors (note that technology gaps is CC and DD sectors are always integers),

$$\begin{aligned}
\frac{\mu^{CC}_{0,t+\Delta t}-\mu^{CC}_{0,t}}{\Delta t} &= \mu^{CC}_{1,t}x_{-1,t}+\delta\sum_{n\in\mathcal{M}_{>0,t}}(\mu^{CC}_{n,t}+\mu^{CD}_{n,t})-2\mu^{CC}_{0,t}x_{0,t}+\frac{o(\Delta t)}{\Delta t},\\
\frac{\mu^{CC}_{1,t+\Delta t}-\mu^{CC}_{1,t}}{\Delta t} &= 2\mu^{CC}_{0,t}x_{0,t}+\mu^{CC}_{2,t}x_{-2,t}-\mu^{CC}_{1,t}(x_{1,t}+x_{-1,t}+\delta)+\frac{o(\Delta t)}{\Delta t},\\
\frac{\mu^{CC}_{\overline{m},t+\Delta t}-\mu^{CC}_{\overline{m},t}}{\Delta t} &= \mu^{CC}_{\overline{m}-1,t}x_{\overline{m}-1,t}-\mu^{CC}_{\overline{m},t}(x_{\overline{m},t}+x_{-\overline{m},t}+\delta)+\frac{o(\Delta t)}{\Delta t},
\end{aligned}$$

where $\mathcal{M}_{>0,t}$ is the set of all possible effective technology gaps greater than 0.

Similarly, the laws of motion for DD sector shares are

$$\begin{aligned}
\frac{\mu^{DD}_{0,t+\Delta t}-\mu^{DD}_{0,t}}{\Delta t} &= \mu^{DD}_{1,t}x_{-1,t}+\delta\sum_{n\in\mathcal{M}_{>0,t}}(\mu^{DD}_{n,t}+\mu^{DC}_{n,t})-2\mu^{DD}_{0,t}x_{0,t}+\frac{o(\Delta t)}{\Delta t},\\
\frac{\mu^{DD}_{1,t+\Delta t}-\mu^{DD}_{1,t}}{\Delta t} &= 2\mu^{DD}_{0,t}x_{0,t}+\mu^{DD}_{2,t}x_{-2,t}-\mu^{DD}_{1,t}(x_{1,t}+x_{-1,t}+\delta)+\frac{o(\Delta t)}{\Delta t},\\
\frac{\mu^{DD}_{\overline{m},t+\Delta t}-\mu^{DD}_{\overline{m},t}}{\Delta t} &= \mu^{DD}_{\overline{m}-1,t}x_{\overline{m}-1,t}-\mu^{DD}_{\overline{m},t}(x_{\overline{m},t}+x_{-\overline{m},t}+\delta)+\frac{o(\Delta t)}{\Delta t}.
\end{aligned}$$

The laws of motion of the technology gap distribution look as follows for CD and DC sectors with either the smallest or the largest possible effective technology gap,

$$\begin{aligned}
\frac{\mu^{CD}_{p,t+\Delta t}-\mu^{CD}_{p,t}}{\Delta t} &= \mu^{CD}_{p+1,t}x_{-p-1,t}+\mu^{DC}_{1-p,t}x_{p-1,t}-\mu^{CD}_{p,t}(x_{p,t}+x_{-p,t}+\delta)+\frac{o(\Delta t)}{\Delta t},\\
\frac{\mu^{CD}_{r,t+\Delta t}-\mu^{CD}_{r,t}}{\Delta t} &= \mu^{CD}_{r-1,t}x_{r-1,t}-\mu^{CD}_{r,t}(x_{r,t}+x_{-r,t}+\delta)+\frac{o(\Delta t)}{\Delta t},\\
\frac{\mu^{DC}_{q,t+\Delta t}-\mu^{DC}_{q,t}}{\Delta t} &= \mu^{DC}_{q+1,t}x_{-q-1,t}+\mu^{CD}_{1-q,t}x_{q-1,t}-\mu^{DC}_{q,t}(x_{q,t}+x_{-q,t}+\delta)+\frac{o(\Delta t)}{\Delta t},\\
\frac{\mu^{DC}_{s,t+\Delta t}-\mu^{DC}_{s,t}}{\Delta t} &= \mu^{DC}_{s-1,t}x_{s-1,t}-\mu^{DC}_{s,t}(x_{s,t}+x_{-s,t}+\delta)+\frac{o(\Delta t)}{\Delta t},
\end{aligned}$$

with $p=m^D_{Lt}+m^T_{Lt}+\tilde{\tau}_t\in(0,1)$, and $r=m^D_{Lt}+m^T_{Lt}+\tilde{\tau}_t\in(\overline{m}-1,\overline{m})$, $q=m^C_{Lt}-m^T_{Lt}-\tilde{\tau}_t\in(0,1)$, and $s=m^C_{Lt}-m^T_{Lt}-\tilde{\tau}_t\in(\overline{m}-1,\overline{m})$. These are the gaps in sectors with one clean and one dirty firm that are closest to 0 or the maximum possible gap. Note that $p=1-q$. Note further that a CD (DC) sector with the smallest possible gap becomes DC (CD) when the laggard innovates (and becomes leader).

The technology combination thresholds θ change as follows,

$$\frac{\theta_{1,t+\Delta t} - \theta_{1,t}}{\Delta t} = \delta \sum_{n \in \mathcal{M}_{>0,t}} \mu_{n,t}^{DC},$$

$$\frac{\theta_{2,t+\Delta t} - \theta_{1,t}}{\Delta t} = \delta \sum_{n \in \mathcal{M}_{>0,t}} (\mu_{n,t}^{DC} - \mu_{n,t}^{CD}) + \mu_{1-p,t}^{DC} x_{p-1,t} - \mu_{1-q,t}^{CD} x_{q-1,t},$$

$$\frac{\theta_{3,t+\Delta t} - \theta_{3,t}}{\Delta t} = -\mu_{r,t}^{CD} x_{r,t},$$

where p, q, r, s are as above.

D Calibration and solution algorithm

D.1 Initial conditions

I use the same micro data that was used in Section 2 to set the initial Ψ_{mt} matrices. To do so, I proceed in the following steps. First, I select all firms that have applied for at least one clean or dirty patent and that are active in a country-sector for which I have at least 20 firms in my data set for the year 2018 (the most recent year with good patent coverage). This is the exact same set of firms that was used for Figure 2. Again, I split the firms into leaders (top 10) and laggards (rank 11-20). I also compute a measure of concentration for each sector, namely total revenue of the top 10 firms divided by total revenue of the top 20 firms. Next, I compute the share of firms that have a (weakly) larger clean than dirty knowledge stock, where the knowledge stock is defined in (1). This is the case for 40.1% of sectors, which is an input for the calibration. Next, I convert real world patent stocks into model technology steps. Given that I set $\overline{m}^T$ to 16, I have 33 possible values for m^T (all integers in $[-16, 16]$). I use the following rule to set m_T:

$$m_i^T = \begin{cases} -16 & \text{if} \quad \sinh^{-1}(K_{it}^C) - \sinh^{-1}(K_{it}^D) < -5 \\ -15 & \text{if} \quad \sinh^{-1}(K_{it}^C) - \sinh^{-1}(K_{it}^D) \in [-5, \frac{-14}{3}) \\ ... & \\ -1 & \text{if} \quad \sinh^{-1}(K_{it}^C) - \sinh^{-1}(K_{it}^D) \in [\frac{-1}{3}, 0) \\ 0 & \text{if} \quad \sinh^{-1}(K_{it}^C) - \sinh^{-1}(K_{it}^D) = 0 \\ 1 & \text{if} \quad \sinh^{-1}(K_{it}^C) - \sinh^{-1}(K_{it}^D) \in (0, \frac{1}{3}] \\ ... & \\ 15 & \text{if} \quad \sinh^{-1}(K_{it}^C) - \sinh^{-1}(K_{it}^D) \in (\frac{14}{3}, 5] \\ 16 & \text{if} \quad \sinh^{-1}(K_{it}^C) - \sinh^{-1}(K_{it}^D) > 5 \end{cases} . \tag{56}$$

While the precise specification is to some degree arbitrary, several considerations have gone into this conversion. First, it is not obvious how many patents constitute an innovation step. Most patent applications are shared by multiple firms, and many firms thus only have a fraction of a patent. Hence, it is not necessarily so that a step is equivalent to (at least) one full patent. On the other hand, large firms often file many patents per year (sometimes hundreds of them), and it is not clear that each of them constitutes a step forward. Using the inverse hyperbolic sine transformation on the knowledge stocks mitigates this issue to some extent. Second, the distribution of the technology gap in Figure 2 is quite smooth and well behaved between the values of -5 and 5.

The next step in setting the initial conditions is to find the distribution of own technology gaps (m^T) for each possible level of the effective technology gap (m^E). I proceed by splitting the sample into clean firms ($m^T \geq 0$) and dirty firms ($m^T < 0$). I then sort sectors based on their degree of concentration (top 10 revenue/top 20 revenue), and divide them into 16 groups (the maximum gap $\overline{m} = 15$, so the possible gaps are all integers from 0 to 15). The size of each group depends on the balanced growth path effective gap distribution, which I obtain from the internal calibration (discussed below, it requires only the share of clean and dirty sectors and not the fully specified set of Ψ_{mt} matrices). For instance, 12.3% of sectors have a technology gap of 0 along the initial BGP, so I use the m^T distribution of the firms that are active in the 12.3% of sectors with the lowest concentration index to fill in $\Psi_{m=0,t=0}$. In the model each sector only has two firms. In the data, some sectors have multiple leaders and laggards that have applied for clean or dirty patents, and some sectors only have a single firm (leader or laggard) that has patented. In the former case I take all possible combinations of the leaders' m^T gaps and the laggards' m^T gaps and weight them such that each sector has an equal weight. In the latter case I set the missing firm's m^T such that both firms are equally productive in the technology that they are not using (unless this is restricted by the $\overline{m}^T$ bounds, in which case I set m^T equal to the bound).

The procedure above yields a fully specified joint distribution of m^E, m_L^T and m_F^T based on micro data. It shows exactly what share of sectors see a switch to clean production by the leader, the laggard or both as the result of a change in the carbon tax. Moreover, it reflects the main finding from Section 2, namely that leaders tend to be more invested in dirty technologies than laggards.

D.2 Solution algorithm for the transition

To solve for the transition after an increase in the carbon tax, I largely follow the solution algorithm in Akcigit and Ates (2023). An important difference between my exercise and theirs is that the shock to the carbon tax immediately affects the effective technology gap distribution, whereas the shocks considered by Akcigit and Ates (2023) affect the new BGP and the transition but not the gap distribution at the time of the shock. Firms' own technology gaps m_{it}^T determine whether firms switch from dirty to clean when the tax is increased. If (i) either the leader or the laggard switches or (ii) both firms switch, the effective technology gap changes according to (17). Note that a tax increase does not affect sectors in which both firms are already clean (other than through general equilibrium changes to output and wages), as neither firm is paying the tax.

An tax increase from τ to $\tau' > \tau$ can affect the effective technology gap between firms. The tax change affects the effective gap if one firm uses clean and the other uses dirty, or when one or both firms switch from dirty to clean. Distinguish the following 9 cases, where L and F denote the <u>initial</u> leader and laggard, respectively, and time subscripts are dropped for convenience:

1. $m_{Lj}^T < -\tilde{\tau}' \wedge m_{Fj}^T < -\tilde{\tau}'$: Both firms remain dirty, $m_{Lj}^E{}' = m_{Lj}^E$, $q_j' = q_j$.

2. $m_{Lj}^T \in (-\tilde{\tau}', -\tilde{\tau}) \wedge m_{Fj}^T < -\tilde{\tau}'$: Leader switches to clean, laggard does not. Leader remains leader. New effective gap: $m_{Lj}^E{}' = m_{Lj}^E + m_{Lj}^T + \tilde{\tau}' > m_{Lj}^E$. New productivity $q_j' = q_j \lambda^{m_{Lj}^T} < q_j$.

3. $m_{Lj}^T > -\tilde{\tau} \wedge m_{Fj}^T < -\tilde{\tau}'$: Leader remains clean and leader, laggard remains dirty. New effective gap: $m_{Lj}^E{}' = m_{Lj}^E - \tilde{\tau} + \tilde{\tau}' > m_{Lj}^E$. Productivity stays the same: $q_j' = q_j$.

4. $m_{Lj}^T < -\tilde{\tau}' \wedge m_{Fj}^T \in (-\tilde{\tau}', -\tilde{\tau})$: Leader remains dirty, laggard switches to clean. New effective gap: $m_{Lj}^E{}' = m_{Lj}^E - m_{Fj}^T - \tilde{\tau}'$.
   - If $m_{Lj}^E{}' > 0$: L remains leader, $q_j' = q_j$.
   - If $m_{Lj}^E{}' \leq 0$: F becomes leader, $q_j' = q_j \lambda^{m_{Fj}^T - m_{Lj}^E} < q_j$.

5. $m_{Lj}^T \in (-\tilde{\tau}', -\tilde{\tau}) \wedge m_{Fj}^T \in (-\tilde{\tau}', -\tilde{\tau})$: Both firms switch from dirty to clean. New effective gap: $m_{Lj}^E{}' = m_{Lj}^E + m_{Lj}^T - m_{Fj}^T$.
   - If $m_{Lj}^E{}' \geq 0$: L remains leader, $q_j' = q_j \lambda^{m_{Lj}^T} < q_j$.
   - If $m_{Lj}^E{}' < 0$: F becomes leader, $q_j' = q_j \lambda^{m_{Fj}^T - m_{Lj}^E} < q_j$.

6. $m_{Lj}^T > -\tilde{\tau} \wedge m_{Fj}^T \in (-\tilde{\tau}', -\tilde{\tau})$: Leader stays clean, laggard switches to clean. New effective gap: $m_{Lj}^E{}' = m_{Lj}^E - m_{Fj}^T - \tilde{\tau} > m_{Lj}^E$. Productivity stays the same: $q_j' = q_j$.

7. $m_{Lj}^T < -\tilde{\tau}' \wedge m_{Fj}^T > -\tilde{\tau}$: Leader remains dirty, laggard remains clean. New effective gap: $m_{Lj}^E{}' = m_{Lj}^E + \tilde{\tau} - \tilde{\tau}' < m_{Lj}^E$.
   - If $m_{Lj}^E{}' > 0$: L remains leader, $q_j' = q_j$.
   - If $m_{Lj}^E{}' \leq 0$: F becomes leader, $q_j' = q_j \lambda^{-m_{Lj}^E - \tilde{\tau}} < q_j$.

8. $m_{Lj}^T \in (-\tilde{\tau}', -\tilde{\tau}) \wedge m_{Fj}^T > -\tilde{\tau}$: Leader switches to clean, laggard stays clean. New effective gap: $m_{Lj}^E{}' = m_{Lj}^E + m_{Lj}^T + \tilde{\tau} < m_{Lj}^E$.
   - If $m_{Lj}^E{}' > 0$: L remains leader, $q_j' = q_j \lambda^{m_{Lj}^T} < q_j$.
   - If $m_{Lj}^E{}' \leq 0$: F becomes leader, $q_j' = q_j \lambda^{-m_{Lj}^E - \tilde{\tau}} < q_j$.

9. $m_{Lj}^T > -\tilde{\tau} \wedge m_{Fj}^T > -\tilde{\tau}$: Both firms remain clean, $m_{Lj}^E{}' = m_{Lj}^E$, $q_j' = q_j$.

To solve for the transition after a shock I first discretize the model and let one period be one year. I then assume that the economy is on a BGP when the shock occurs, and that it converges to a new BGP at time T_{BGP}.[8] The solution method for the transition is as follows. First, knowing the initial distribution Ψ_{mt} and the change in the tax τ_t, I compute the technology gap distribution Ψ_{mt}, which also gives μ_{mt}, right after the shock. Second, I solve for the terminal BGP distribution of μ_{mt}, as well as v_{mt} and x_{mt}, in the same way I solved for the initial BGP in the calibration.[9] Third, I guess a path for the interest rate r_t that ends at its terminal BGP value, as well as for the wage. Fourth, I use the Euler equation (7) to find the path of Y_t and solve for the firm values V_{mt} on the terminal BGP given the guesses. Fifth, I solve backwards in time for V_{mt}, v_{mt} and x_{mt} using the discretized version of value functions (27). Sixth, I use the path of x_{mt} to solve forwards in time for output and wages and I compute the implied interest rate, again using the Euler equation. I then update my guesses for the paths of r_t and w_t and repeat the procedure from step 4 onward until the guess and the implied path of r_t converge.

[8] I set T_{BGP} such that the economy has reached the new BGP in 100 periods.

[9] Note that τ_t features in ω_t (see equation 44), meaning that the effective gap distribution is not the same on the initial and the terminal BGP.

E Figures and tables

Add table with the exact IPC and CPC codes here. For now see Jee and Srivastav (2023).

Table E1: Path dependence in innovation

	(1)	(2)	(3)	(4)
	Clean patents	Dirty patents	Innovation gap (clean-dirty)	
Log K^C	0.664***	-0.114***	0.041***	
	(0.028)	(0.020)	(0.005)	
Log K^D	-0.063**	0.882***	-0.065***	
	(0.030)	(0.020)	(0.003)	
Technology gap (clean-dirty)				0.112***
				(0.004)
Estimator	Poisson	Poisson	OLS	OLS
(Pseudo) R^2	0.53	0.54	0.11	0.11
Observations	6,624,288	6,624,288	4,341,408	4,235,520

Notes: All independent variables are first lags. OLS regressions include country-sector-year fixed effects. The innovation and technology gaps are defined in (2) and (3), respectively. Further controls in columns 1 through 3 are the stock of patents in any category and dummies that are 1 if the stock variables equal zero (one dummy for each stock). Further controls in column 4 are the stock of patents in any category, a dummy that is 1 if the stock of patents is zero, and a dummy that is 1 if the technology gap is zero. Standard errors are clustered at the firm level. The sample covers the years 1978-2018.

Table E2: Heterogeneity in technology gaps and innovation

	(1)	(2)	(3)	(4)
	Technology gap (clean-dirty)		Clean patents	Dirty patents
Log revenue	-0.015***		0.783***	0.579***
	(0.003)		(0.086)	(0.082)
Log employment	0.004		0.010	0.195**
	(0.003)		(0.087)	(0.090)
Profit margin	-0.000		0.014**	0.011*
	(0.000)		(0.006)	(0.006)
Log age	-0.005		-0.012	-0.167
	(0.004)		(0.072)	(0.109)
Leader		-0.059***		
		(0.014)		
Laggard		-0.006		
		(0.011)		
Estimator	OLS	OLS	Poisson	Poisson
(Pseudo) R^2	0.08	0.08	0.47	0.44
Observations	223,093	401,587	26,227	31,630

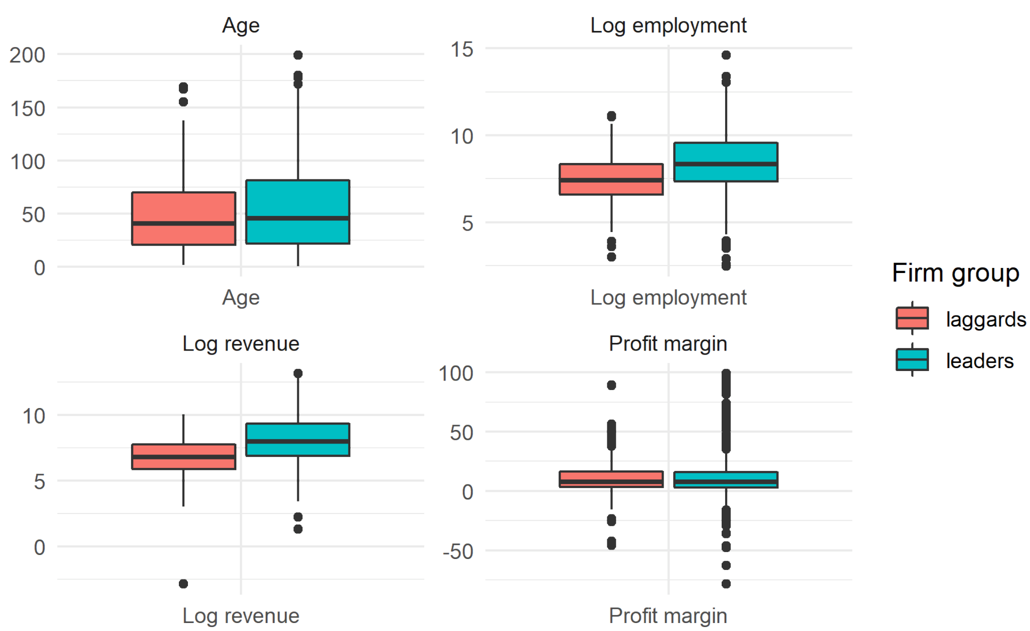
Notes: All regressions include country-sector-year fixed effects. All independent variables are contemporaneous values. The technology gap is defined in (3). Further controls in column 2 are the stock of patents in any category and a dummy that is 1 if the stock of patents is zero (both lagged). Standard errors are clustered at the firm level. The sample covers the years 2010-2018. The sample in column 3 (4) consists of all firms that have applied for at least one clean (dirty) patent between 1978 and 2018 and for which the financial variables were available.

Table E3: Heterogeneity in technology gaps

	(1)	(2)	(3)	(4)
	Technology gap (clean-dirty)			
Log revenue	-0.008***			
	(0.001)			
Log employment		-0.011***		
		(0.001)		
Profit margin			-0.000***	
			(0.000)	
Log age				-0.016***
				(0.001)
R^2	0.06	0.06	0.07	0.05
Observations	393,702	360,840	281,366	862,867

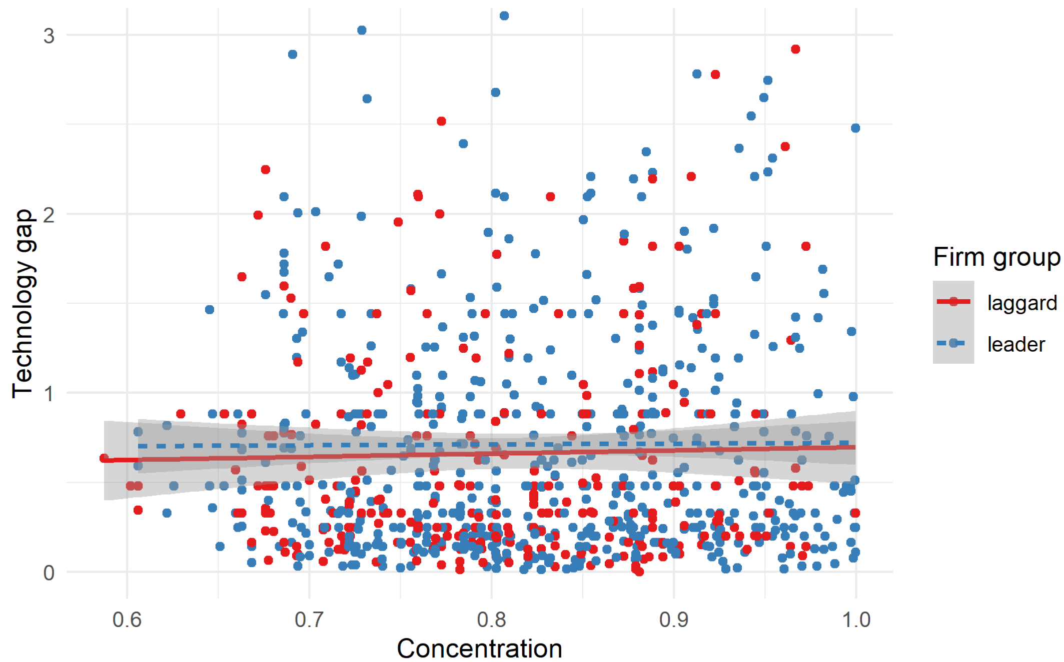
Notes: All independent variables are contemporaneous values. All regressions include country-sector-year fixed effects. Standard errors are clustered at the firm level. The sample covers the years 2010-2018.

Figure E1: The firm size distribution for leaders and laggards



Notes: Data sources: Orbis IP and Historical. Leaders are the top 10 firms in terms of revenue in their 2 digit NACE Rev. 2 industry and country, laggards are the firms ranked 11 until 20 in those same industries. Graph is for the year 2018 and includes only firms that applied for at least one patent in that year.

Figure E2: Competition and technology gaps for clean firms



Notes: Data sources: Orbis IP and Historical. This figure plots technology gaps as defined in (3) against industry concentration for leaders and laggards with a positive technology gap. The sample includes all firms that are classified as leader or laggard in the year 2018 and that applied for at least one patent in the past.

Table E4: Heterogeneity in technology gaps by competition and direction

	(1) High concentration	(2) Low concentration	(3) Clean sectors	(4) Dirty sectors
Leader	-0.070***	-0.043**	0.078	-0.103***
	(0.017)	(0.019)	(0.059)	(0.036)
Laggard	-0.013	-0.000	0.015	0.035
	(0.016)	(0.015)	(0.067)	(0.037)
Constant	-0.044***	-0.038***	0.058***	-0.192***
	(0.003)	(0.003)	(0.016)	(0.008)
R^2	0.10	0.06	0.11	0.15
Observations	173000	227573	19112	66327

Notes: All regressions include country-sector-year fixed effects. All independent variables are contemporaneous values. The technology gap is defined in (3). High (low) concentration sectors are those with above (below) median industry concentration (defined as top 10 revenue over top 20 revenue). Clean (dirty) sectors are those in which the median firm's technology gap is strictly positive (negative). Further controls in all columns are the stock of patents in any category and a dummy that is 1 if the stock of patents is zero (both lagged). Standard errors are clustered at the firm level. The sample covers the years 2010-2018.

Figure E3: Welfare levels for different values of τ

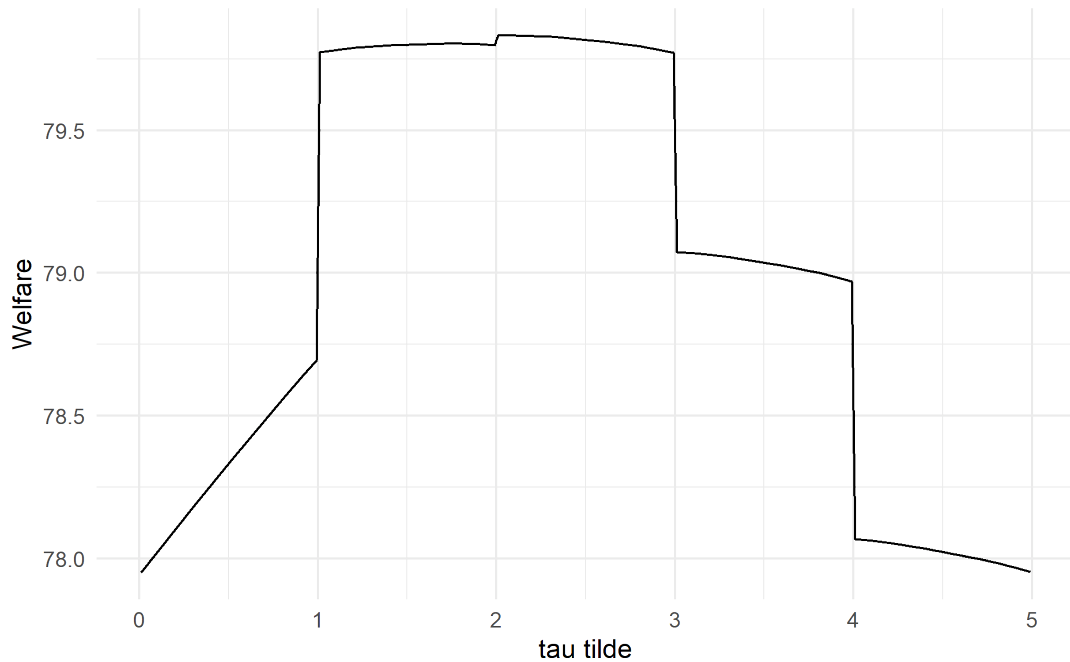


Figure E4: Transition with no immediate effect on market power ($\tilde{\tau} = 2.99$)