Journal of Economic Dynamics & Control 165 (2024) 104900

Contents lists available at ScienceDirect

Journal of Economic Dynamics and Control

journal homepage: www.elsevier.com/locate/jedc

Believe me when I say green! Heterogeneous expectations and climate policy uncertainty ☆

Emanuele Campiglio [a,b,c,d], Francesco Lamperti [b,c,e], Roberta Terranova [b,c,*]

[a] University of Bologna, Department of Economics, Bologna, Italy
[b] CMCC Foundation - Euro-Mediterranean Center on Climate Change, Lecce, Italy
[c] RFF-CMCC European Institute on Economics and the Environment, Milan, Italy
[d] LSE Grantham Research Institute on Climate Change and the Environment, London, United Kingdom
[e] Institute of Economics and L'EMbeDS, Scuola Superiore Sant'Anna, Pisa, Italy

ARTICLE INFO

JEL classification:
C63
D84
E71
Q54
Q58

Keywords:
Beliefs
Behavioural macroeconomics
Credibility
Investment decision-making
Heterogeneous expectations
Low-carbon transition
Policy uncertainty

ABSTRACT

We develop a dynamic model where heterogeneous firms take investment decisions depending on their beliefs on future carbon prices. A policy-maker announces a forward-looking carbon price schedule but can decide to default on its plans if perceived transition risks are high. We show that weak policy commitment, especially when combined with ambitious mitigation announcements, can trap the economy into a vicious circle of credibility loss, carbon-intensive investments and increasing risk perceptions, ultimately leading to a failure of the transition. The presence of behavioural frictions and heterogeneity - both in capital investment choices and in the assessment of the policy-maker's credibility - has strong non-linear effects on the transition dynamics and the emergence of 'high-carbon traps'. We identify analytical conditions leading to a successful transition and provide a numerical application for the EU economy.

1. Introduction

Transitioning to a low-carbon economy will require an expanding share of firms to allocate their physical capital investments to carbon-free technologies (IPCC, 2022). In market economies, investment choices are mainly driven by profitability expectations: firms adopt low-carbon investment strategies only if they expect them to be more convenient than carbon-intensive alternatives. In

☆ We would like to thank Christian Buelens, Louison Cahen-Fourot, Louis Daumas, Domenico Delli Gatti, Francesca Diluiso, Giorgos Galanis, George Krivorotov, Davide Pace, Piergiuseppe Pezzoli, Simon Quemin, Alessandro Spiganti, Marko Stojanovic, Frank Venmans, as well as participants to the EAERE 2022, EAEPE 2022, ESEE 2022, GRASFI 2022, EAEPE 2022, WEHIA 2021 and 2022, SURED 2022, EPSF 2022 and OCC 2022 conferences, and 2023 Behavioral Macroeconomics workshop in Bamberg. We would also like to thank seminar participants at the London School of Economics, University of Montpellier and Università Cattolica del Sacro Cuore. Further, we are grateful to Thomas Lubik as well as two anonymous reviewers for having provided useful comments. The research leading to these results has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Programme (Grant agreement No 853050 - SMOOTH) and under the Marie Skłodowska-Curie grant agreement No 870245.

* Corresponding author.
E-mail address: roberta.terranova@cmcc.it (R. Terranova).

https://doi.org/10.1016/j.jedc.2024.104900
Received 24 December 2023; Received in revised form 27 May 2024; Accepted 5 June 2024

Available online 11 June 2024

turn, their relative profitability expectations are affected by the expected strength and timing of future climate mitigation policies, such as carbon pricing initiatives (World Bank, 2022).

What do firms expect future carbon prices to be, and how do they formulate their expectations? It is reasonable to assume policy-makers' stated intentions to act as a key expectation anchor. For instance, a large number of countries in recent years have publicly pledged to reach 'net-zero' emissions by a certain date in the future.[1] However, policy announcements trigger the desired behavioural changes only if individuals believe these will be followed by actual policy actions. Recent history counts many examples of failed policy commitments or complete policy reversals.[2] Experience of policy volatility and uncertainty can lead individuals to doubt the credibility of their policy-makers' commitments.

Failure to meet policy targets is often motivated by the perception of excessively high transition risks, i.e. socio-economic costs generated by the process of structural change away from carbon-intensive technologies (Campiglio and van der Ploeg, 2022). In the context of the low-carbon transition, certain sectors and countries are likely to lose out because of mitigation policies. Systemic effects could ensue - unemployment, increase in energy prices, loss of competitiveness, capital stranding and financial instability - with potentially dire political implications for the implementing policy-maker. Despite the conflicting evidence on the socio-economic impacts of climate policies (see for instance Metcalf and Stock, 2023; Vona et al., 2018), governments in both high-income and emerging economies might ultimately succumb to the discontent of (parts of) the population and return on their steps, or just be replaced by new governments following different policy directions.

Confronted with multiple sources of uncertainty, individuals might develop heterogeneous beliefs regarding the credibility of the policy-maker and the future schedule of carbon prices, depending on the information they have access to, their ability to process it, their political preferences, the length of their planning horizon, and a number of other behavioural factors. Beliefs and expectations will also change in time with the arrival of new information, possibly amplified by herd behaviour dynamics or, to the contrary, restrained by deep-rooted convictions creating inertia in the updating process. The existence of heterogeneous, volatile and biased beliefs is well documented in economic analysis (see, among others, Afrouzi et al., 2023; Bordalo et al., 2022; Evans and Honkapohja, 2012; Hommes, 2021; Mankiw et al., 2003), and has been confirmed by the recent literature focusing on climate-related dimensions (Ardia et al., 2022; Bakkensen and Barrage, 2017; Engle et al., 2020; Gibson and Mullins, 2020; Giglio et al., 2023; Krueger et al., 2020; Noailly et al., 2021). The scarcer empirical evidence available on expectations concerning future climate policies also points in the same direction (Barradale, 2014; Nordeng et al., 2021).[3] However, despite the well-known relevance of behavioural factors in socio-technical transitions (Geels et al., 2016; Pollitt and Shaorshadze, 2013; Shogren and Taylor, 2008), the role of expectation dynamics for decarbonisation trajectories has yet to be studied in detail.

In this paper, we contribute to filling this gap by developing an analytically tractable model to study how heterogeneous and dynamic beliefs can affect the low-carbon transition, and how climate mitigation policies should be appropriately announced and implemented. Two key economic agents exist in our model: (i) a policy-maker in charge of setting the carbon price; and (ii) a continuum of firms producing output using a combination of low- and high-carbon capital stocks. The policy-maker publicly announces a forward-looking carbon price path. This is designed to mirror real-world long-term targets and mitigation pledges that should, in principle, guide the political process of short-term policy implementation. If pursuing a low-carbon transition is perceived to be at odds with socio-economic stability (that is, transition risks are perceived to be high), the policy-maker might decide to deviate from its plans and introduce a lower carbon price. As our analysis refers to a single jurisdiction with a single policy-maker, we assume the policy-maker to abstract from perspective climate damages.[4] We also assume the government to have an imperfect understanding of private firms' investment decision-making drivers.

At each point in time, firms choose the proportion of investments to allocate between the two available technological options depending on their expected relative costs, which in turn depend on their expectations of future carbon prices. Firms observe the public announcement, but they are not sure whether the policy-maker will actually enforce the planned tax.[5] For simplicity, we assume firms to be able to choose to belong to one of two populations: 'believers' trust the climate mitigation commitment of the policy-maker more than 'sceptics'.[6] Firms, faced with limited information about the government's decision-making process and its commitment to climate mitigation, choose their belief type by evaluating the policy-maker's past actions and its track-record in

[1] 'Net-zero' commitments have been announced by the European Union (climate neutrality by 2050), United States (2050), China (2060), Russia (2060) and India (2070), among others (Hale et al., 2022; Fankhauser et al., 2022).

[2] In the climate/energy policy sphere, examples include the rapid and in some cases retroactive withdrawal from clean energy subsidies in Europe (Sendstad et al., 2022); the introduction and subsequent repeal of a carbon price in Australia (Crowley, 2017); the US withdrawal from and return to the Paris Agreement (Urpelainen and Van de Graaf, 2018); the recalibration of French fiscal policy after the Gilet Jaunes movement (Douenne and Fabre, 2020); and the numerous cases of fossil fuel subsidy reform withdrawals following social unrest (McCulloch et al., 2022).

[3] A connected burgeoning literature looks at the public acceptability of climate policies (see Dechezleprêtre et al., 2022; Douenne and Fabre, 2020; Konc et al., 2022). However, these works usually do not include policy *expectations* in their analysis.

[4] Given our focus on the European Union, this seems a reasonable assumption. Indeed, the EU in 2021 was responsible for roughly only 7.5% of fossil CO2 emissions (GCP, 2022). Most emissions and associated future climate impacts are likely to be driven by forces originating elsewhere and following almost entirely exogenous trajectories. Our results should thus be interpreted as a study of transition dynamics, *given* a certain exogenous amount of climate damages. Fried et al. (2022) also abstract from climate damages in their main analysis of climate policy risks, as including them in their modelling framework does not significantly affect results.

[5] In other words, we assume government's transition risk perceptions to be private information and unavailable to firms. While our framework does not include electoral cycles, this assumption partially captures the inherent unpredictability of policy-making. The uncertainty firms face is thus a form of parametric, rather than stochastic, uncertainty, which they address, if necessary, by switching to an alternative system of beliefs.

[6] In the behavioural macroeconomics literature à la Brock and Hommes (1997, 1998), the assumption of two belief types has become the commonly accepted standard. This approach goes beyond the homogeneity assumption and enables firms to consider multiple belief options, providing an intuitive framing without overly complicating the modelling framework. However, a possible extension of this model involves the incorporation of multiple belief types, leading to what is

keeping its word. If one belief type is observed to forecast more precisely policy-maker's actions, agents learn from this and can switch their policy beliefs towards the better performing rule. In other words, if the policy-maker significantly deviates from its announced carbon price schedule due to transition cost concerns, a certain proportion of firms will lose faith in its plans and no longer use them to assess the relative technology convenience. The share of firms believing the policy-maker's announcements can thus be interpreted as a measure of the policy-maker's credibility, which is endogenous to past policy choices.

Building on the literature on behavioural macroeconomics (see Hommes, 2021, for a review), we allow firms' choices – (i) how to allocate investments across technologies, and (ii) whether to believe in the policy-maker announcements – to be subject to behavioural frictions. By behavioural frictions, we mean factors such as bounded rationality, cognitive limitations, incomplete information and any other behavioural dimension preventing firms from immediately choosing the optimal alternative among their available options and leading to heterogeneous choices across firms. To model such choices, we employ the discrete choice framework (Manski and McFadden, 1981) and introduce parameters for both the *investment* and *belief responsiveness* of firms, i.e. their ability to rapidly incorporate new available information and, if appropriate, switch to different investment/belief strategies.[7] We explore the full range of possible values for these behavioural parameters, going from zero (investment choices are made at random, as strong frictions fully debilitate firms' decision-making process) to infinity (the 'neoclassical limit': all firms immediately make the marginally most convenient choice, with no frictions).

We first derive analytical conclusions from a reduced version of the model. We show that, in the neoclassical limit without behavioural frictions and heterogeneity, two steady states (each fully dominated by one of the two technologies) can exist depending on i) the announced policy stringency; and ii) the policy-maker commitment level. We derive the conditions for existence of these steady states and find that the combination of ambitious mitigation plans and a weakly-committed policy-maker can lead to the emergence of multiple equilibria (i.e. a 'high-carbon trap'). When we introduce behavioural frictions and the associated heterogeneity of beliefs/expectations, we identify a set of 'behavioural premiums' that modify the long-run equilibria of the system. The conditions for existence of the low-carbon steady state become more challenging: heightened ambition and commitment from the policy-maker are required. Specifically, the minimum tax target necessary for the existence of the low-carbon steady state increases compared to the neoclassical limit case, and a new minimum commitment requirement emerges. Nevertheless, our findings indicate that behavioural frictions also limit the emergence of highly carbon-intensive steady states, helping the unambitious policy-maker to achieve 'mid-carbon' outcomes. Finally, we identify a level of the tax announcement that, contingent on the policy-maker's commitment and firms' responsiveness to beliefs, allows for a successful transition, as it ensures the existence and uniqueness of the low-carbon steady state.

We then calibrate the full version of the model to European data and run forward-looking numerical simulations. We distinguish two scenarios: (i) full commitment by the policy-maker, with climate policy targets always met regardless of transition costs; and (ii) less-than-full climate policy commitment. Under a fully-committed policy-maker and a sufficiently strong tax announcement, the economic system is almost always eventually reaching full decarbonisation, but the speed of the transition is significantly affected by behavioural dimensions (i.e. the investment and belief responsiveness of the firms' population). When allowing for the policy-maker to default on its commitment due to potential transition costs, we find that the decarbonisation can endogenously fail, getting trapped into a vicious circle of credibility loss, carbon-intensive investment choices and increasing transition risk perceptions. Our results suggest that, while the weakly-committed policy-maker could succeed in purposely overshooting its policy targets so to push the transition through before the credibility loss takes over, exceeding in deception can backfire and eventually compromise the transition process. Finally, we explore the role of belief polarisation, i.e. the distance between the belief systems of believers and sceptics, finding that, under certain conditions, a highly polarised belief system can lead to a more rapid transition than one with mildly polarised beliefs.

Our results suggest that policy-makers can, under certain conditions, benefit from a 'self-fulfilling prophecy effect', under which a sufficiently credible announcement manages to push through the low-carbon transition before firms fully realise that the policy-maker is not strongly committed to respecting their targets. However, if this deception strategy goes beyond certain thresholds with excessively ambitious strategies, it might backfire and lead to a complete failure of the transition. In this respect, less (but still sufficiently) ambitious credible policies appear to be better than strongly ambitious but not sufficiently credible ones.

Our work builds upon and connects two main streams of research, which have largely proceeded independently so far: (i) the literature on heuristic switching models, primarily employed in behavioural macroeconomic and finance modelling; and (ii) the analysis of transition dynamics, with a focus on the issue of climate policy uncertainty and credibility.

The first stream of work is particularly relevant from a methodological perspective, as our modelling framework with heterogeneous and dynamic expectations draws upon the one pioneered by Brock and Hommes (1997, 1998) and later developed by several contributions, mainly focusing on the interaction between heterogeneous inflation expectations and monetary policy decisions (Assenza et al., 2021; De Grauwe and Macchiarelli, 2015; Evans and Honkapohja, 2006; Hommes and Lustenhouwer, 2019a; Salle et al., 2013).[8] In this family of logit models, ultimately rooted in discrete choice theory (Manski and McFadden, 1981; McFadden, 1974),

known as the 'large type limit' proposed by, e.g., Anufriev et al. (2013), Brock et al. (2005) and (Hommes and Lustenhouwer, 2019b). This extension would allow for the representation of a continuous distribution of expectations, further enriching the dynamics and realism of the model.

[7] Given the well-documented significance of behavioural frictions in influencing the green transition, the discrete choice framework proves to be effective for capturing these aspects. It enables us to retain the standard assumption of utility maximisation while also accommodating potential deviations from the benchmark (Train, 2009). More specifically, it allows to characterise the importance of behavioural frictions through a succinct parametrisation.

[8] In monetary policy-making, the debate on the role of policy credibility – i.e., to what extent policy-makers should commit to a predetermined course of action or maintain discretion over time – has been central. Since Kydland and Prescott (1977), it has been emphasised that discretionary policies may be dynamically inconsistent and rules may be preferred to avoid non credible promises. When it comes to climate policy, one possible implication is that policy-makers should adopt

the choice between two alternative belief systems is governed by an *observable* component - typically their performance in predicting the actual evolution of the economic system - and an *unobservable* component, assumed to be distributed following a Gumbel distribution for analytical tractability.[9] This unobservable component captures factors such as bounded rationality and heterogeneity across agents, without overburdening the model with their explicit representations. By including this element, the ability of utility-maximising agents of consistently making optimal choices can be limited, resulting in heterogeneous decisions across agents based on heuristic rules (Gigerenzer and Brighton, 2009). We advance this literature by (i) applying its modelling framework to the analysis of the interactions between climate policy strategies, firms' behavioural responses and the transition to a low-carbon economy[10]; and (ii) developing a double logit framework, with a backward-looking choice on the policy-maker's credibility and a forward-looking choice on the technology to invest in.

The second stream of work we connect to is the one investigating the socio-economic dimensions of transitioning to a carbon-free economy. We are particularly interested in the issue of policy-makers' time inconsistency and the policy-related uncertainty firms face when making their investment decisions (Helm et al., 2003; Nemet et al., 2017). This research question has been addressed by the climate economics literature using multiple modelling approaches (see for instance Battiston et al., 2021; Fuss et al., 2008; Fried et al., 2022; Lemoine, 2017; van der Ploeg and Rezai, 2020) all pointing to the conclusion that uncertainties concerning future policy implementation - as well as their resolution - have a significant effect on investment decisions and the shape of the low-carbon transition.[11] However, none of these contributions allow policy-related expectations to be heterogeneous or subject to bounded rationality. Heterogeneity has been included and investigated in recent contributions in the field, but usually with a focus on the distributional implications of climate policies on households characterised by heterogeneous income, wealth or consumption levels (e.g. Douenne et al., 2023; Känzig, 2023). In parallel, a more empirical line of work has attempted to measure climate policy uncertainty and transition risks using bodies of text such as newspaper articles (Berestycki et al., 2022). Against such background, we introduce the innovation of forward-looking agents whose heterogeneous and dynamic climate policy expectations affect their physical capital investment decisions and, in turn, the overall dynamics of the low-carbon transition.[12]

The remainder of the paper is structured as follows. Section 2 presents the model. Section 3 derives some analytical conclusions using a reduced version of the model. Section 4 explains our calibration strategy for the full model. Section 5 presents and discusses our numerical results under the assumption of a fully committed policy-maker. Section 6 extends the numerical analysis to the case of a weakly committed policy-maker. Section 7 concludes.

2. The model

We consider an economy in discrete time, moving from t_0 to T. The system is populated by a continuum of firms producing a homogeneous final good Y. We assume demand of good Y to grow at an exogenous rate g_Y and supply to always be able to satisfy demand thanks to expansions of the capital stock, which is the only factor of production.[13] Two types of capital stocks exist: (i) high-carbon (fossil-fuelled) capital K_h, producing greenhouse gas emissions; and (ii) low-carbon capital K_l, with no production of emissions. The production cost θ_i of good Y, where $i \in \{l, h\}$ denotes the two technologies, varies between these two types of capital, with $\theta_l \geq \theta_h$. Aggregate capital stock is the sum of the two technology-specific stocks, i.e. $K = \sum K_i$. We define κ as the share of low-carbon capital over the aggregate capital stock, i.e. $\kappa_t \equiv \frac{K_{l,t}}{\sum_i K_{i,t}}$.

Fig. 1 illustrates the sequence of events in the model: at t_0, the policy-maker announces an exogenous future schedule of target carbon prices.[14] Before the policy-maker's announcement, the existing tax is denoted by τ_{-1}. Following the policy announcement, each time period $t \geq t_0$ unfolds through the following steps: (i) firms assess the credibility of the policy announcement and form

an aggressive stance from the outset to promptly establish credibility. Although this specific research question is not central for our analysis, we show in Appendix C.1 that, in our framework, the benefits of proposing a stringent initial policy can be offset by initially high transition risks, complicating efforts to address credibility concerns.

9 The Gumbel distribution assumption allows the cumulative distribution function of the difference between the two underlying distributions to be shaped as a logistic function. Probit models, which assume instead a Normal distribution of unobservables, can also be fruitfully employed in this context; see Cahen-Fourot et al. (2022) and Galanis et al. (2022a).

10 Its applications to environment- and climate-related questions have been limited so far. Cahen-Fourot et al. (2022) study how heterogeneous transition expectations affect investment decision choices in a forward-looking model with capital 'stranding'. Annicchiarico et al. (2022b) introduce the possibility for agents to switch among forecasting rules for output and inflation and study how this affects the impact of climate policies in a New Keynesian model. Zeppini (2015) and Mercure (2015) apply a logit framework to the choice on technology adoption. Cafferata et al. (2021) and Dávila-Fernández and Sordi (2020) focus on switching attitudes towards green policies. Galanis et al. (2022b) study country participation in international environmental agreements. Torren-Peraire et al. (2023) study the interaction between decarbonisation and cultural change in an agent-based model where the influence individuals have on each other recalls Brock and Hommes (1998) discrete choice models.

11 A related stream of work focuses more in general on climate-related uncertainty, with a stronger focus on the physical or economic dimensions of the problem, rather than the policy one (see, among others, Barnett et al., 2020; Cai and Lontzek, 2018; Campiglio et al., 2022; van den Bremer and van der Ploeg, 2021).

12 Expectations can play a key role also in shaping the carbon intensity of financial investment decisions (see Dunz et al., 2021).

13 Alternatively, without loss of generality, one might consider a linear production function where labour and capital are complements, and the former is abundant. We abstract from technical change.

14 While most of public climate policy announcements currently come in the form of net-zero dates, it is possible to derive the implicit associated optimal carbon price path, which has a quasi-exponential shape (Riahi et al., 2021). Also, given our longer-term perspective, we abstract here from the discussion of optimal carbon prices following business cycle fluctuations (see for instance Annicchiarico et al., 2022a; Doda, 2014). Finally, while a diversified portfolio of policies is necessary to decarbonise the economic system (see Campiglio, 2016; Fay et al., 2015), we restrict our analysis to carbon pricing, both for analytical simplicity and due to its predominant presence in the climate policy debate.

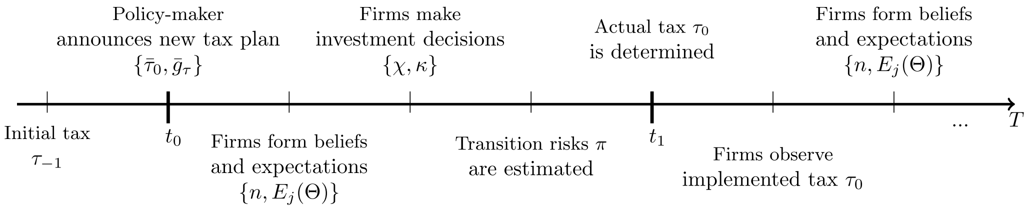


Fig. 1. Timeline of events.

expectations regarding the tax; (ii) based on these expectations, firms make investment decisions; (iii) the policymaker observes these investment decisions and assesses the transition risks associated with implementing the tax plan; (iv) subsequently, the actual tax rate is determined.

2.1. Policy announcements and expectation dynamics

Climate policy in our model takes the form of a tax on high-carbon capital production. At t_0, the policy-maker announces its intention to implement a new schedule of rising carbon tax rates $\bar{\tau}_t$, starting from an exogenous level $\bar{\tau}_0$ and increasing at a constant growth rate $\bar{g}_\tau$.[15] That is, at t_0 the policy-maker announces that, at each $t \in [t_0, T]$,

$$\bar{\tau}_t = \bar{\tau}_0(1 + \bar{g}_\tau)^t. \tag{1}$$

At each time t, firms formulate expectations on future carbon prices by 'discounting' the policy-maker announcement. We assume firms to be either 'believers' (b) or 'sceptics' (s), with $j \in \{b, s\}$ indicating the belief type. Belief-specific carbon price expectations are defined as

$$\begin{aligned} E_{b,t}(\tau_{t+r}) &= \bar{\tau}_0(1 + \epsilon_b \bar{g}_\tau)^{t+r} \\ E_{s,t}(\tau_{t+r}) &= \tau_{-1}(1 + \epsilon_s \bar{g}_\tau)^{t+r} \end{aligned} \tag{2}$$

where the operator $E_t(\cdot)$ denotes the expectations formulated at time t, τ_{-1} is the initial exogenous carbon price prevailing before the announcement of the new tax plan, and $\epsilon_j \in [0, 1]$ is a parameter indicating to what extent firms believe to the announced tax growth rate, with $\epsilon_b > \epsilon_s$ (i.e. believers trust policy announcements more than sceptics). The distance between ϵ_s and ϵ_b can be interpreted as a proxy for opinion polarisation: when $\epsilon_s = \epsilon_b$, sceptics' expectations are entirely aligned to those of believers (homogeneous population); when $\epsilon_s = 0$ and $\epsilon_b = 1$, the system of beliefs is instead heavily polarised.[16]

Similarly to Brock and Hommes (1997, 1998), firms switch belief type depending on their relative accuracy in predicting the policy-maker's behaviour in the past. The larger is the difference between forecast and actual policy, the higher is the likelihood of firms changing their belief. We define the fitness measure of expectation rule j in period t, $U_{j,t}$, as the weighted sum of the last observed absolute prediction error, and the previous value of the fitness measure:

$$U_{j,t} = (1 - \eta)\left|E_{j,t-1}(\tau_t) - \tau_t\right| + \eta U_{j,t-1}, \tag{3}$$

where $0 \leq \eta \leq 1$ is a memory parameter indicating to what extent firms update their evaluation of the expectation rule with new information. If η is set to zero, firms only consider the last prediction error, whereas if $\eta > 0$, firms also take into account past values of the fitness measure.

The switching mechanism between beliefs, based on the accuracy of predictions evaluated in the previous period, determines the share $n \in [0, 1]$ of firms adopting belief b (i.e. choosing to be 'believers') at time t:

$$n_t = \frac{\exp(-\beta U_{b,t-1})}{\sum_j \exp(-\beta U_{j,t-1})}, \tag{4}$$

where $\beta \geq 0$ represents 'belief responsiveness', i.e. the responsiveness of firms' beliefs to prediction errors and, consequently, to policy choices.[17] Low values of β indicate that firms' beliefs react mildly to prediction errors, with $\beta = 0$ indicating a population of firms evenly split between the two beliefs, regardless of U_j. If instead β tends to infinity, all firms immediately adopt the best performing expectation rule, even if the performance gap between the two is small (i.e. a 'bang-bang' solution). In our framework, since firms' expectations are driven by announced and implemented policies, β also indicates the speed of firms' response to policy choices. We interpret weak belief responsiveness as a consequence of high behavioural frictions, stemming from inherent biases in their trust or

[15] There is ample consensus in the climate economics literature that carbon prices should increase to stimulate emission mitigation and that, in many cases, the optimal price trajectory is exponential (e.g. Golosov et al., 2014; Nordhaus, 2017; Riahi et al., 2021; van den Bremer and van der Ploeg, 2021). For an alternative approach leading to declining optimal carbon prices, see Daniel et al. (2019).

[16] When using the expression 'opinion polarisation', we refer to the difference between the two system of beliefs, regardless of the size of the two populations.

[17] In the framework developed by Brock and Hommes (1997, 1998), β is referred to as the 'intensity of choice' parameter.

distrust of the policy-maker's announcements, or from their inability to accurately observe the policy-maker's behaviour.[18,19] Based on past policy choices and firms' response to them, the share of believers n_t is determined endogenously and measures the ex-post policy-maker's credibility, similarly to Hommes and Lustenhouwer (2019a).

2.2. Investment choices and capital dynamics

At each time t, the aggregate amount of investments can be expressed as a function of the existing capital stock, given exogenous and constant growth rate of output g_Y (also equivalent to growth rate of the capital stock) and capital depreciation rate δ:

$$I_t = (g_Y + \delta)K_t. \tag{5}$$

Building on their chosen beliefs, firms decide how to allocate their investments I across the two available technologies. They do so by evaluating the net present value of the expected production costs $\Theta_{i,t}$ associated to each technology $i \in \{l, h\}$ as

$$E_{j,t}(\Theta_{i,t}) = \begin{cases} \sum_{r=1}^{R} D^r \theta_{i,t+r} & \text{if } i = l, \\ \sum_{r=1}^{R} D^r \theta_{i,t+r}\left[1 + E_{j,t-1}(\tau_{t+r})\right] & \text{if } i = h, \end{cases} \tag{6}$$

where R is the length of firms' planning horizon, $D = \frac{1}{1+\rho}$ is a discount factor with discount rate ρ, and θ is the technology-specific pre-tax production cost, comprising both capital installation and operational costs. Once a firm, pertaining to a specific belief population j, has assessed the cost prospects of the two technologies, it allocates its investments to the technology it perceives to be the most convenient.

Not all firms pertaining to the same belief population will invest in the same technology. In line with the discrete choice theoretical framework, we assume that a number of other unobservable variables affect firms' investment decisions. Similarly to the belief adoption rule described in section 2.1, firms might be subject to preference biases that sway them towards favouring one technology over the other for reasons unrelated to their costs. These biases could stem from factors such as varying levels of environmental consciousness, or from an inability to accurately assess the costs associated with each technology.

We can thus define the j-specific share of low-carbon investment $\chi_{j,t} \equiv \frac{I_{l,j,t}}{\sum_i I_{i,j,t}}$, as

$$\chi_{j,t} = \frac{\exp(-\gamma E_{j,t}(\Theta_{l,t}))}{\sum_i \exp(-\gamma E_{j,t}(\Theta_{i,t}))} \tag{7}$$

where γ indicates 'investment responsiveness', that is, the responsiveness of firms' investment to the difference in expected costs. As cost expectations depend on policy announcements and implementation, investment responsiveness, similarly to belief responsiveness β, can be interpreted as the speed of firms' response to the policy-maker's behaviour. Also, it signals the degree of behavioural frictions affecting firms' decisions: $\gamma = 0$ implies that firms randomly choose the technology, leading to a clean investment share of 0.5, regardless of cost differentials; if instead γ tends to infinity, all firms choose the technology they expect to be marginally more convenient. This choice is belief-specific; that is, given their different carbon price expectations, believers ad sceptics might have a different perception of which technology is the best performing.

Building on firms' technology choice (equation (7)) and on the belief switching dynamics (equation (4)), we can derive the evolution of the aggregate low-carbon investment share over time, χ_t, as a weighted average of $\chi_{j,t}$

$$\chi_t = n_t \chi_{b,t} + (1 - n_t)\chi_{s,t}. \tag{8}$$

Finally, assuming the standard capital law of motion $K_{i,t+1} = (1-\delta)K_{i,t} + I_{i,t}$, and building on equation (5), we can define the evolution of the low-carbon share of capital κ as a function of χ[20]:

$$\kappa_t = \frac{\kappa_{t-1}(1-\delta) + \chi_t(g_y + \delta)}{1 + g_y}. \tag{9}$$

2.3. Transition risks and policy implementation

Once investment decisions are taken, the policy-maker decides if and to what extent the announced policy will actually be implemented. While the policy-maker's intentions might have been sincere at the time of the announcement, the potential costs related to the low-carbon transition, together with their electoral implications, might weaken its resolution and lead to a change of its mitigation strategy. Transition costs can be of various nature, ranging from higher energy bills for households and firms to unemployment in high-carbon industries and systemic financial disruptions.

[18] Weak belief responsiveness is therefore a different concept from asynchronous updating of choices whereby agents do not update their choices simultaneously (see for instance Hommes et al., 2005).

[19] We follow the common assumption of time-invariant parameters adopted in behavioural macroeconomic models by setting β and γ (introduced in next section) constant over time. To address its potential limitations, we conduct extensive sensitivity analysis on their values in sections 5 and 6.

[20] A related literature investigates the desirable dynamics of power generation capacity accumulation, see for instance (Kollenbach, 2017; Tsur and Zemel, 2011).

We assume our policy-maker to formulate a 'transition risk index' $0 \leq \pi \leq 1$, increasing in both the announced tax target $\bar{\tau}$ and in the carbon intensity of the economy's productive basis, $1 - \kappa$. The intuition is that governments will consider their economies more exposed to the risk of transition costs when climate objectives are more stringent (i.e. higher announced carbon prices) and when a larger part of the productive system relies on high-carbon technologies (e.g. oil and gas extraction; coal- and gas-fuelled electricity production; internal combustion engine vehicle production; etc.). This is in line with the approach used by Peszko et al. (2020) to calculate the index of countries' exposure to low-carbon transition risks.[21] When no policy is scheduled ($\bar{\tau} = 0$), or if the economy is already entirely decarbonised ($\kappa = 1$), we assume perceived transition risks to be equal to zero and to have no weight in the policy-maker's decisions. We thus write

$$\pi_t = 1 - \frac{1}{1 + a(1 - \kappa_t)\bar{\tau}_t}, \tag{10}$$

where a is a parameter capturing the vulnerability of the economy (as perceived by the policy-maker) to transition risks. This vulnerability might depend on several factors, such as the exposure of the banking and financial system to transition risks, the fragility of the welfare system and the vulnerability to social turmoil (see Appendix D for a graphical representation). One can interpret a as the inverse of the index of resilience to low-carbon transition impacts proposed by Peszko et al. (2020). While our concave functional form is the one we believe to be most representative of the current debate on transition costs, where even the announcement of relatively mild mitigation policies can lead to large protests by vocal minorities, we test for alternative formulations in Appendix D, showing that the qualitative results of the model remain untouched.

Once transition risks have been estimated, the policy-maker weighs them against its climate mitigation objectives in order to choose the carbon tax to actually implement. We formalise this policy choice as a weighted average between the announced tax target $\bar{\tau}_t$ and the tax target reduced by the transition risk index, $\bar{\tau}_t(1 - \pi_t)$:

$$\tau_t = c\bar{\tau}_t + (1 - c)\bar{\tau}_t(1 - \pi_t), \tag{11}$$

where $0 \leq c \leq 1$ is an exogenous parameter indicating the policy-maker commitment to climate mitigation objectives. A value of $c = 1$ indicates a policy-maker fully committed to mitigation, who will therefore always impose taxes as scheduled, regardless of the transition risks involved. On the other hand, $c = 0$ represents the case of a policy-maker fully committed to the reduction of transition risks and willing to reduce the tax to a floor level of $\bar{\tau}_t(1 - \pi_t)$.

3. Analytical results

In this section, we explore the analytical properties of a slightly simplified version of the model. We study the existence of multiple steady states and characterise their stability in order to assess the long-term behaviour of the low-carbon transition and its dependence on policy settings and firms' attitudes. We start by introducing three sets of assumptions that grant us tractability without affecting the qualitative behaviour of the system (we will remove all these assumptions and run numerical simulations for the full calibrated model in Sections 5 and 6).

Assumption 1. The announced carbon price is positive and constant, i.e. $\bar{\tau}_t = \bar{\tau} \in R_0^+$ for all t, and higher than the initial exogenous tax, i.e. $\bar{\tau} > \tau_{-1}$.

We assume the policy-maker to just announce a carbon tax $\bar{\tau}$, without including an increasing schedule of prices. In other words, a discrete positive jump from τ_{-1} to $\bar{\tau}$ is announced. The actual tax τ_t might however be different from what announced depending on transition risk perceptions and the policy-maker's commitment level, as in the full model. This setting simplifies our analysis by making our dynamic model autonomous, i.e. independent of time. At the same time, however, it allows us to obtain more general results compared to the case of a tax target growing over time according to a specific function.

Assumption 2. We set $\delta = 1$; $\eta = 0$; $E_{b,t}(\tau_{t+r}) = \bar{\tau}$ and $E_{s,t}(\tau_{t+r}) = \tau_{-1}$ for all t and r.

To simplify the derivation of analytical results, we assume full capital depreciation, i.e. $\delta = 1$. While this assumption affects the dynamics of the model, it does not modify steady states, as these are independent of the rate of capital depreciation. We set the memory parameter $\eta = 0$, thereby assuming agents to fully update the fitness measure of beliefs with the newly available information. Finally, we set $E_{b,t}(\tau_{t+r}) = \bar{\tau}$ and $E_{s,t}(\tau_{t+r}) = \tau_{-1}$, so to characterise the two beliefs types to their extreme version (i.e. believers entirely believe in policy announcements; sceptics don't believe in them at all). i.e. sceptics will expect the tax rate τ never to move from its initial level τ_{-1}, while believers will expect it to be equal to the announced rate $\bar{\tau}$.

Assumption 3. We impose $\tau_{-1} < \frac{\theta_l - \theta_h}{\theta_h}$; $\bar{\tau} > \frac{\tau_{-1}}{1 - a\tau_{-1}}$; and $\tau_{-1} < \frac{1}{a}$.

---

[21] More in detail, Peszko et al. (2020) uses four measures to compute the transition risk exposure index: carbon intensity of manufacturing exports; committed power emissions as a proportion of current annual power generation; fossil fuel export as a proportion of GDP; expected resource rents as a proportion of GDP.

The first condition on τ_{-1} implies that the tax expected by sceptics does not cover for the percentage cost difference between low- and high-carbon technologies. The second and third conditions assume sceptics' prediction errors to be positive, i.e. sceptics tend to underestimate actual climate policy, independently of c and k.[22]

Under these assumptions, the dynamics of the low-carbon capital share is given by

$$\kappa_{t+1} = (\chi_b - \chi_s)n_{t+1} + \chi_s \equiv f(\kappa_t), \tag{12}$$

where χ_b and χ_s are independent of time because the tax target is now a constant (see Assumption 1). The share of believers n_{t+1} depends on the carbon intensity of capital stock κ_t as follows:

$$n_{t+1} = \left[1 + \exp\left(-\beta\left\{2\bar{\tau}\left[c + \frac{1-c}{1+a(1-\kappa_t)\bar{\tau}}\right] - \tau_{-1} - \bar{\tau}\right\}\right)\right]^{-1}. \tag{13}$$

Since $n \in [0,1]$ and $\chi_b > \chi_s$, $f(\kappa)$ is bounded between the low-carbon investment shares of the two populations, χ_s and χ_b, both of which are bounded between 0 and 1. Combining equations (12) and (13), we can derive the following proposition:

Proposition 1. *The system has at least one stable steady state. If there is more than one steady state, those with an odd index are stable.*

Proof. Proof of Proposition 1 is provided in Appendix B.1. □

Remark 1. By numerically studying the system, we find that the maximum number of steady states is three. Fig. A1 in Appendix B.1 provides a graphical representation of the S-shaped function $f(\kappa)$ and of such steady states, for different values of $\bar{\tau}$ and β, under weak commitment ($c = 0.3$). Depending on the parameters, we observe the emergence of up to three steady states, identified by the points in which the curves intersect the 45-degrees line. The intermediate steady state defines the boundaries of the basins of attraction of the high- and the low-carbon equilibria. Hence, whenever the intermediate steady state is closer to the low-carbon one, for a policy-maker it is relatively more difficult to obtain a successful low-carbon transition. Fig. A1b shows that for low values of belief responsiveness β, the steady state is unique and decreasing in β. As the belief responsiveness crosses a certain threshold ($\beta > 0.75$), the steady states become three, where the high- and low-carbon ones are stable and the intermediate one is unstable.

To provide a better understanding of the system, we now evaluate $f(\kappa)$ in three illustrative cases wherein the model shows a unique stable steady state. First, we assume a fully committed policy-maker ($c = 1$), implying that the tax actually implemented equals the tax target in every t, independently of κ. In this case, $f(\kappa) = (\chi_b - \chi_s)\left[1 + \exp\left(-\beta\left(\bar{\tau} - \tau_{-1}\right)\right)\right]^{-1} + \chi_s$. Second, we set $\beta = 0$, that is, firms choose their belief type at random. This leads to $n = 0.5$ and thus $\kappa = \frac{1}{2}(\chi_b + \chi_s)$ at every point of time. Third, we assume $\gamma = 0$, i.e. believers and sceptics split their investment equally between high- and low-carbon technologies, i.e. $\chi_b = \chi_s = 0.5$ at each point of time. The resulting equilibrium level of the low-carbon capital share is $f(\kappa) = 0.5$.

In the more general case where the policy-maker is not fully committed ($c < 1$) and both belief and investment responsiveness are not equal to zero ($\beta > 0$ and $\gamma > 0$), the system may present multiple steady states. To explore them, we consider two settings that exhaustively describe agents' behavioural attitudes: (i) a benchmark case, where belief and investment responsiveness are infinite, i.e. $\beta = \gamma = +\infty$; and (ii) a scenario characterised by boundedly rational decisions, with finite β and γ. The first scenario proxies what we shall call a *neoclassical limit* case – mirroring the label used in Brock and Hommes (1997) – wherein all agents choose the superior belief type and the cheapest technology at each time step, even if by a small margin. In contrast, the second scenario includes *behavioural frictions* inversely proportional to β and γ (see more details in Appendix A).

3.1. Steady states in the neoclassical limit

In the neoclassical limit scenario ($\beta = \gamma = \infty$) the policy-maker can induce, via its announcements and level of commitment, three different types of systems, characterised by: i) a unique high-carbon steady-state; ii) a unique low-carbon steady state; or iii) multiple steady states with varying basins of attraction. The two lines depicted in Fig. 2 illustrate the thresholds on the tax target (black line) and on the commitment level (blue line) delimiting the parameter spaces in which low- and high-carbon steady states exist. Such conditions of existence are outlined in the following proposition.

Proposition 2. *Under the assumption that $\beta = \gamma = \infty$,*

(i) *The low-carbon steady state $\kappa_l^* = 1$ exists if*

$$\bar{\tau} > \frac{\theta_l - \theta_h}{\theta_h} \equiv \bar{\tau}^*; \tag{14}$$

(ii) *The high-carbon steady state $\kappa_h^* = 0$ exists if*

$$\bar{\tau} < \bar{\tau}^* \quad \text{or} \quad c < \frac{1}{2} - \mu_1 \equiv c^* \tag{15}$$

[22] Note that, by construction, believers' prediction errors are either zero or negative.

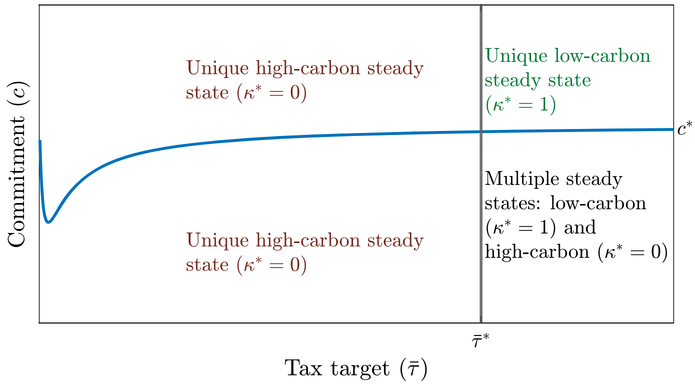


Fig. 2. Existence of low- and high-carbon steady states depending on commitment c and tax target $\bar{\tau}$ in the neoclassical limit case. (For interpretation of the colours in the figure(s), the reader is referred to the web version of this article.)

where $\mu_1 = \frac{\bar{\tau}-\tau_{-1}(1+a\bar{\tau})}{2a\bar{\tau}^2} > 0$.

Proof. Proof of Proposition 2 is provided in Appendix B.2. □

Condition (14) states that the low-carbon steady state exists if the announced tax target is higher than the percentage cost difference between low- and high-carbon production costs ($\bar{\tau}^*$, i.e. the black line in Fig. 2). The high-carbon steady state, instead, exists if either the announced tax target does not cover for the above mentioned cost difference $\bar{\tau}^*$, or if the policy-maker's commitment is sufficiently low (condition (15)). More specifically, the commitment threshold c^* (blue line in Fig. 2) is composed of two terms. The first one, $1/2$, is the upper bound of c^* and corresponds to a commitment equally split between meeting policy targets and reducing transition risks. The second term (μ_1) is decreasing in both the initial tax rate τ_{-1} and in the vulnerability to transition risks a. An increase in these two variables thus moves up the commitment threshold, expanding the parameter space where the high-carbon steady state exists. As evident in Fig. 2, the effect of the announced tax target $\bar{\tau}$ is instead non-linear, decreasing up to a value $\bar{\tau} = \frac{2\tau_{-1}}{1-a\tau_{-1}}$ and increasing afterwards.[23] Since the turning point takes place for very low values of $\bar{\tau}$, we can state that, for sufficiently high and reasonable values of announced carbon prices, an increase in ambition increases the likelihood of a high-carbon steady state emerging.

By combining the conditions outlined in Proposition 2, we obtain the different long-term states of the model shown in Fig. 2. A full decarbonisation of the economy (top-right quadrant) can be achieved by announcing a tax able to compensate for the relative backwardness of low-carbon technologies and by being sufficiently committed to such target. This happens because, with infinite investment responsiveness, even the smallest expected cost difference in favour of low-carbon capital leads believers to fully invest in the clean technology ($\chi_b = 1$). Moreover, under infinite belief responsiveness and a sufficiently high commitment, the policy-maker is able to convince every firm of its credibility ($n = 1$). Together, investment and belief dynamics determine the uniqueness of the low-carbon steady state.

To the contrary, an excessively low carbon tax target ($\tau < \tau^*$) generates a unique and stable high-carbon steady state (top- and bottom-left quadrants). Intuitively, in absence of behavioural frictions, if the carbon tax is insufficient to cover the cost advantage of the high-carbon technology, all firms will avoid investing in the inferior technology, leaving $\kappa^* = 0$ to be the only possible equilibrium level of low-carbon capital.

Finally, when the announced tax target is sufficiently high but the policy-maker is weakly committed to the decarbonisation process, the system exhibits multiple equilibria (bottom-right quadrant). While announcing ambitious policy targets stimulates believers' clean investment share χ_b, not meeting such targets hampers the policy-maker's credibility (decrease in n) and therefore the decarbonisation process. This results in the emergence of multiple equilibria, generating what we label as a *high-carbon credibility trap* for the economic system.

3.2. Steady states under behavioural frictions

We now move to the analysis of the scenario characterised by behavioural frictions and hence by heterogeneity in investment decisions. In such a context, the system never achieves a full decarbonisation ($\kappa = 1$) as, even under the most favourable conditions for low-carbon technologies, a small but positive amount of capital investments will flow to high-carbon capital. The low-carbon steady state under behavioural frictions can thus be defined as $\kappa_l^* = 1 - \lambda_l$, where λ_l represents the distance between the low-carbon steady state and full decarbonisation. Similarly, the high-carbon steady state will not be equal to zero, as a minor proportion of firms will always invest in low-carbon capital or believe in the policy-maker. We hence define $\kappa_h^* = \chi_s + \lambda_h$, where λ_h indicates the

[23] To have c^* increasing in $\bar{\tau}$, $\tau_{-1} < \frac{1}{a}$ also needs to be verified, which is always the case in our numerical examples.

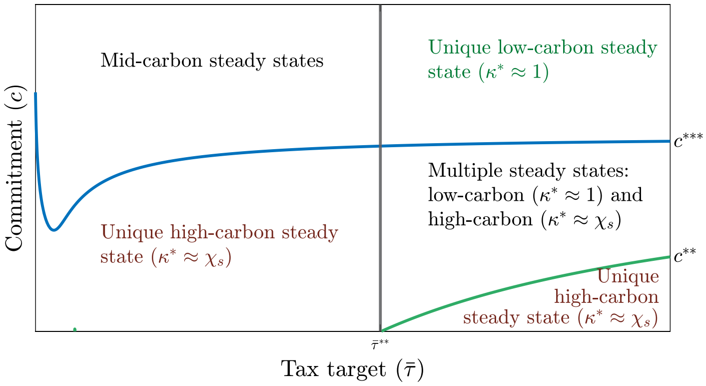


Fig. 3. Existence of low- and high-carbon steady states depending on commitment c and tax target $\bar{\tau}$ in the behavioural frictions case.

distance between the high-carbon steady state and sceptics' clean investment share. Similarly to the neoclassical limit case, we prove the following proposition illustrating the conditions of existence of the low- and high-carbon steady states.

Proposition 3. *Under the assumption that β and γ are finite,*

(i) *A low-carbon steady state $\kappa_l^* = 1 - \lambda_l$ exists if a positive real number $\tilde{\lambda}_l$ exists such that*

$$\bar{\tau} > \frac{\theta_l - \theta_h}{\theta_h} + \nu_{\tau l} \equiv \bar{\tau}^{**} \qquad \text{and} \qquad c > \frac{1}{2} - \mu_2 + \nu_{cl} \equiv c^{**} \tag{16}$$

where

$\tilde{\lambda}_l = \lambda_l + \varepsilon_l$, *with* ε_l *a small positive number and* $\tilde{\lambda}_l \in (0, \frac{1}{2})$,

$\nu_{\tau l} = -\ln\left(\frac{\tilde{\lambda}_l}{1-\tilde{\lambda}_l}\right)\rho\{\gamma\theta_l(1+\rho)\left[1-(1+\rho)^{-(R+1)}\right]\}^{-1} > 0$

$\nu_{cl} = -\ln\left(\frac{\chi_b - 1 + \tilde{\lambda}_l}{1-\tilde{\lambda}_l-\chi_s}\right)(2\bar{\tau}\beta)^{-1}\left(1+\frac{1}{a\tilde{\lambda}_l\bar{\tau}}\right) \gtrless 0$, *and*

$\mu_2 = \frac{\bar{\tau}-\tau_{-1}(1+a\tilde{\lambda}_l\bar{\tau})}{2a\tilde{\lambda}_l\bar{\tau}^2} > 0.$

(ii) *A high-carbon steady state $\kappa_h^* = \chi_s + \lambda_h$ exists if a positive real number $\tilde{\lambda}_h$ exists such that*

$$c < \frac{1}{2} - \mu_3 + \nu_{ch} \equiv c^{***} \tag{17}$$

where

$\tilde{\lambda}_h = \lambda_h + \varepsilon_h$, *with* ε_h *a small positive number and* $\tilde{\lambda}_h \in (0, \chi_b - \chi_s)$,

$\nu_{ch} = -\ln\left(\frac{\chi_b-\chi_s-\tilde{\lambda}_h}{\tilde{\lambda}_h}\right)(2\bar{\tau}\beta)^{-1}\left\{1+\frac{1}{a[1-(\chi_s+\tilde{\lambda}_h)]\bar{\tau}}\right\} \gtrless 0$, *and*

$\mu_3 = \frac{\bar{\tau}-\tau_{-1}\{1+a[1-(\chi_s+\tilde{\lambda}_h)]\bar{\tau})\}}{2a[1-(\chi_s+\tilde{\lambda}_h)]\bar{\tau}^2} > 0$

Proof. Proof of Proposition 3 is provided in Appendix B.3. □

Differently from the neoclassical limit scenario, the existence of the low-carbon steady state is now subject to two simultaneous conditions (16). The first condition is that the policy target needs to be sufficiently large (i.e. $\bar{\tau} > \bar{\tau}^{**}$, or to the right of the black line in Fig. 3). This condition, which guarantees that believers invest mostly in the low-carbon asset, corresponds to its analogue in the neoclassical limit case (14), but the tax threshold under behavioural frictions (τ^{**}) now has an additional term, $\nu_{\tau l}$. The presence of this 'behavioural premium' means that a tax target covering only the technology cost differential would now be insufficient to convince investors to decarbonise. The term $\nu_{\tau l}$ is positive and decreasing in γ, ρ, R, θ_h and λ_l. In other words, the threshold that the announced carbon price $\bar{\tau}$ needs to satisfy to allow for the existence of a low-carbon steady state becomes lower if the investment responsiveness γ increases, moving closer to the neoclassical limit without behavioural frictions. The threshold becomes instead harder to satisfy, the more myopic agents are in their investment choices (short planning horizon R and high discount rate ρ); the closer is the desired low-carbon steady state to full decarbonisation (low λ_l); and, similarly to the neoclassical limit case, the lower are high-carbon technology costs (low θ_h).[24]

[24] The constraint that $\tilde{\lambda}_l < 0.5$ implies that the desired long-term capital structure is one where low-carbon capital is more abundant than high-carbon capital (i.e. $\kappa* > 0.5$). For values of $\tilde{\lambda}_l > 0.5$, where the equilibrium clean capital share can be below 50%, there is an additional constraint to be satisfied ($\tau_{-1} < (\theta_l - \theta_h)/\theta_h - \nu_{\tau l}$). Moreover, under $\tilde{\lambda}_l > 0.5$, the effect of γ, ρ, R, and θ_h, on the tax target threshold changes sign. The effect of $\tilde{\lambda}_l$ on the tax target threshold, instead, remains negative.

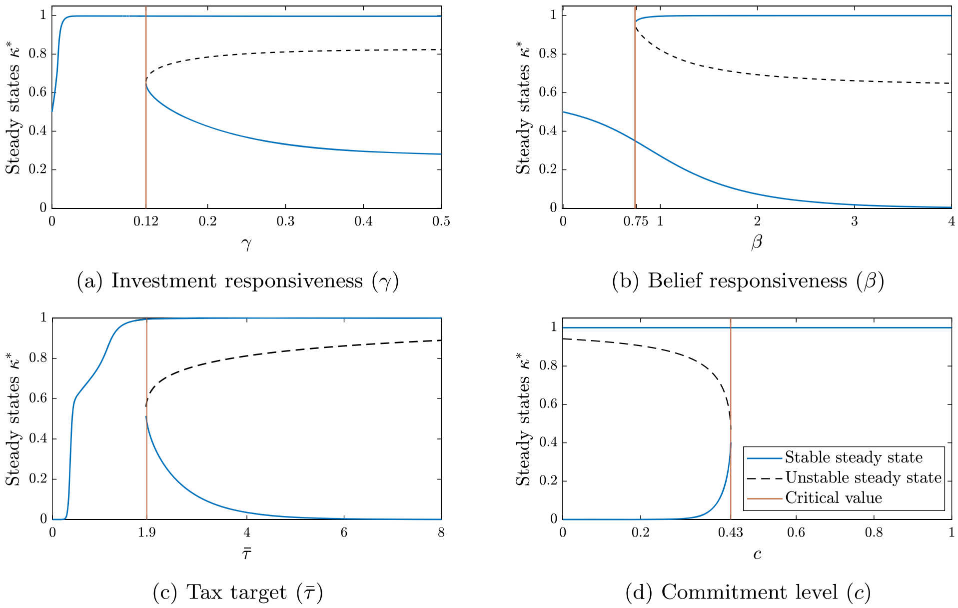


Fig. 4. Bifurcation diagrams. Default parameter values are $\bar{\tau} = 6$, $c = 0.3$, $\gamma = 1$ and $\beta = 3$.

The second condition for the existence of the low-carbon steady state in (16) is that the policy-maker's commitment needs to be sufficiently high (i.e. $c > c^{**}$, or above the green line in Fig. 3); this condition ensures that most firms believe in policy targets. Similarly to c^* in (15), the threshold c^{**} has an upper bound at $\frac{1}{2}$. The second term, μ_2, is equivalent to the term μ_1 in (15), although augmented by a term $\tilde{\lambda}_l$, while the third term ν_{cl} can be interpreted as an additional 'behavioural premium' on commitment levels. For sufficiently small values of $\tilde{\lambda}_l$ ($\tilde{\lambda}_l < 1 - \frac{\chi_b + \chi_s}{2}$), higher behavioural frictions (i.e. a lower β) increase the necessary commitment threshold, diminishing the area of existence of the low-carbon steady state. Likewise, lower clean investment shares for both sceptics (χ_s) and believers (χ_b) increase the commitment threshold c^{**}. The overall effect of the other parameters is instead harder to establish as they might have contrasting effects on the commitment thresholds.

The existence of the high-carbon steady state is determined only by condition (17) on commitment. In particular, the policy-maker's commitment must be below the threshold value c^{***} (blue line in Fig. 3), which, again, has an upper bound at $\frac{1}{2}$ and two other terms, μ_3 and a behavioural premium ν_{ch}. For small values of $\tilde{\lambda}_h$ ($\tilde{\lambda}_h < \frac{\chi_b - \chi_s}{2}$), the threshold c^{***} increases for higher values of belief responsiveness β (i.e. lower behavioural frictions), vulnerability to transition risks a and tax target $\bar{\tau}$, expanding the area of existence of the high-carbon steady state and thus increasing the likelihood of the high-carbon trap. The other parameters have instead ambiguous effects. By combining the conditions presented in Proposition 3, we can characterise how the policy-maker's behaviour influences the long-term behaviour of the model under behavioural frictions, as illustrated in Fig. 3. Some portions of the space are qualitatively similar to the ones identified in the neoclassical limit case (Fig. 2): an ambitious and highly committed policy-maker drives the economy towards decarbonisation (upper-right quadrant); lack of ambition and commitment causes instead the transition to fail (lower-left quadrant); ambitious policy targets announced by an uncommitted policy-maker generate multiple steady states (mid-right quadrant). However, the presence of behavioural frictions leads to the emergence of two additional long-run behaviours. First, a committed policy-maker consistently meeting a non-ambitious tax target entails the existence of one or multiple mid-carbon steady states (upper-left quadrant). In this area, the equilibrium low-carbon capital share is lower than $1 - \lambda_l$ but higher than $\chi_s + \lambda_h$, suggesting less carbon-intensive steady states than those in the same quadrant of Fig. 2. The reason is that, although under low tax targets high-carbon capital is cheaper than low-carbon capital, a portion of firms decide to invest in the more expensive low-carbon technology, due to particular environment-friendly preferences or to limited awareness of technology cost differences. Therefore, under a non ambitious but highly committed policy-maker, the presence of behavioural frictions is actually positive for the low-carbon transition. Second, non credible but ambitious tax announcements (bottom right corner) have worse effects on the low-carbon transition under behavioural frictions than in the neoclassical limit benchmark case. In fact, in the presence of behavioural frictions, the high-carbon sector never disappears entirely. As a result, transition risks remain positive and become considerable if policy targets are too ambitious, driving the actual policy away from targets, losing credibility and preventing the low-carbon steady state to exist.

3.3. High-carbon trap drivers

Fig. 4 offers more details on the role of key parameters (γ, β, $\bar{\tau}$ and c) in determining the number and nature of the system steady states. We develop a numerical example based on our wider model calibration (see section 4), where we assume weak mitigation commitment by the policy-maker (i.e. $c = 0.3$). The blue solid lines represent stable steady states, while the black dashed lines are unstable intermediate steady states. The vertical red lines correspond to the critical value where the bifurcation occurs.

Fig. 4a illustrates the impact of investment responsiveness γ on the position of steady states κ^*. For low values of γ, the model exhibits a unique stable equilibrium steeply increasing in the investment responsiveness. As γ passes the threshold indicated by the vertical red line ($\gamma \approx 0.12$ in this numerical example), a new stable high-carbon equilibrium emerges, together with an intermediate unstable one. That is, if investors are sufficiently responsive to expected cost differentials, the system could fall into either of the two steady states, depending on the initial conditions. If the economy is already sufficiently decarbonised (κ around 80%), then the policy announcement can push low-carbon investment strongly enough to ensure a full transition before the policy-maker can lose its credibility due to weak commitment. Otherwise, for initial values of κ lower than 0.8, despite a potential initial spur in low-carbon investments, the loss of credibility of the policy-maker eventually leads to a failure of the decarbonisation process. The relative sizes of the steady states' basins of attraction also move, in favour of the high-carbon one, as investment responsiveness γ increases. Because an initial condition of κ around 0.8 is quite high, it follows that high levels of γ can significantly hamper the success of the low-carbon transition.

Fig. 4b focuses on the impact of belief responsiveness β on long-term steady states of κ^*. It shows that for low values of belief responsiveness, the steady state is unique and decreasing in β, as a higher belief responsiveness leads weak commitment to be more punished by firms. As β passes a certain threshold ($\beta \approx 0.75$), two additional steady states emerge, one of which is low-carbon and stable. We thus confirm the importance of belief responsiveness in determining the nature of the system steady states, already pointed out in section 3.2 where higher β moves the green line in Fig. 3 down and the blue line up, expanding the area where the low- and high-carbon steady states coexist. Further, in the area to the right of the critical value β^* in Fig. 4b, an increase in belief responsiveness widens the low-carbon basin of attraction, but also leads to worse high-carbon steady states and better low-carbon steady states, i.e. low-carbon steady states closer to $\kappa^* = 1$. Concerning the initial conditions of κ, akin to the case of γ, the existence of multiple equilibria favours the high-carbon steady state, as its basin of attraction extends to high levels of the low-carbon capital share reaching up to κ around 0.6.

Fig. 4c shows the evolution of the long-term steady states as the tax target varies. As $\bar{\tau}$ grows, the unique steady state κ^* increases and the system approaches full decarbonisation. Under the current calibration, the nonlinear behaviour of κ^* with respect to the tax target is due to the fulfilment of conditions (16). When $\bar{\tau} > \bar{\tau}^{**}$ all believers fully invest in low-carbon capital but the overall share of believers grows slowly, as policy intensity is not sufficient to satisfy the second condition in (16); this corresponds to the first slowdown in Fig. 4c. The second steep increase in κ^* is due to the fulfilment of such second condition in (16), which boosts the switch of sceptic firms to believers, making κ^* more sensitive to the policy strength. Finally, as the announced tax target passes the critical value indicated by the vertical red line ($\bar{\tau} \approx 1.9$ in our numerical example), two additional steady states emerge, one stable and one unstable. The resulting long-term low-capital share thus fundamentally depends on the basins of attraction of such steady states. In particular, the higher the tax target, the larger the basin of attraction of the high-carbon steady state, up to the point where even values of κ around 0.8 are attracted to the carbon intensive equilibrium.

Finally, Fig. 4d illustrates the bifurcation diagram of c. When c is lower than a critical value ($c \approx 0.43$), two additional steady states emerge through a fold bifurcation. Similarly to the other cases, the majority of initial conditions leads to the high-carbon steady state, although less so for increasing levels of c. Once the threshold is passed by c, the two lower steady states disappear and κ_l^* remains the unique fixed point. The critical value of c fundamentally depends on the tax target level, as shown in Propositions 2 and 3. In this respect, an additional key feature of the equilibrium conditions, both in the neoclassical limit and behavioural frictions scenarios, is that for $\bar{\tau} \to \infty$, the threshold of c tends to $\frac{1}{2}$, meaning that for very ambitious policy objectives, the policy-maker must strictly prefer meeting them than reducing transition risks, in order to avoid the emergence of a bad equilibrium. Therefore, announcing a sufficiently high tax target is needed in order to achieve the low-carbon transition, but being too ambitious in policy objectives increases the required commitment to those objectives. In other words, ambitious climate policies must be credible, or a high-carbon trap might emerge, potentially leading the low-carbon capital share to lower levels than under a less ambitious tax target.

3.4. A safe threshold for the carbon price announcements

From the policy-maker's perspective, the conditions for ensuring the existence (and uniqueness) of the low-carbon steady state and avoiding the high-carbon trap might be hard to estimate and, therefore, to apply. In addition, if we consider commitment as an intrinsic characteristic of the policy-maker, the only policy choice concerns the appropriate target tax schedule, given a specific c. Hence, we present an additional, simpler rule identifying a sufficient condition for the uniqueness of the equilibrium.

Proposition 4. *Under behavioural frictions, a unique low-carbon steady state exists if conditions* (16) *are met and:*

$$\bar{\tau} < \frac{1}{\beta(1-c)}. \tag{18}$$

Proof. Proof of Proposition 4 is shown in Appendix B.4. □

Table 1
Parameters and initial values in the central scenario.

Parameter	Symbol	Quarterly value	Source
Production			
Output growth rate	g_Y	0.5%	Gomme and Rupert (2007); Lera and Sornette (2017)
Depreciation rate	δ	1.77%	Gomme et al. (2011)
Initial low-carbon capital share	κ_0	0.2	Eurostat (2021); IRENA (2021)
Low- to high-carbon production cost	$\frac{\theta_l}{\theta_h}$	1.36	IEA (2020c)
Beliefs and investment decisions			
Investment responsiveness	γ	1	Indirectly calibrated
Planning horizon	R	120	Farfan and Breyer (2017); Tran and Smith (2018)
Discount rate	ρ	1.7%	IEA (2020c)
Initial fitness measure differential	$U_{b,0} - U_{s,0}$	0.85	Refinitiv (2019)
Initial shares of belief types	n_0; $1 - n_0$	0.3; 0.7	Refinitiv (2019)
Belief responsiveness	β	3	Ellen et al. (2010)
Belief-specific discount on targets	ϵ_b; ϵ_s	1; 0	Sensitivity analysis
Memory parameter	η	0.5	Hommes and Lustenhouwer (2019a)
Policy			
Initial tax level	τ_{-1}	0.15	IEA (2019a,b); EEA (2021c)
Initial announced tax rate	$\bar{\tau}_0$	0.2	IEA (2020c); EEA (2021b); van der Ploeg and Rezai (2021)
Announced tax growth rate	$\bar{g}_\tau$	0.02	BEIS (2021); Stratégie (2019); Riahi et al. (2021)
Transition risk index parameter	a	2	Indirectly calibrated

Proposition 4 identifies a sufficient condition: it states that the lower the policy-maker's commitment and the higher firms' belief responsiveness β, the less ambitious the policy announcement can be to guarantee the uniqueness of the low-carbon steady state.[25] The tax target implied in Proposition 4 can be interpreted as a safe threshold, below which the equilibrium is unique and able to generate a smooth transition, provided that it satisfies condition (16). The policy-maker could therefore consider setting a tax slightly below the safe threshold to achieve full decarbonisation without risking the high-carbon credibility trap.

4. Calibration

We now remove the simplifying assumptions of the reduced version of the model used in section 3 and we calibrate and initialise the full model to European data. We use quarterly time steps, investigating the period 2020-2060 (with a total span of 160 time periods). Our baseline configuration is presented in Table 1. For all parameter values we deem as uncertain, we provide a sensitivity analysis in Appendix C.

4.1. Production

We study the transition to low-carbon capital in a growing economy where the growth rate of output, g_Y, is kept constant and equal to 0.5% (quarterly), corresponding to a yearly growth rate of approximately 2% (cfr. Lera and Sornette, 2017; van der Ploeg and Rezai, 2020; Gomme and Rupert, 2007). The quarterly depreciation rate, δ, is calibrated at 1.77%, consistent with Gomme et al. (2011) and corresponding to an annual depreciation of approximately 7.27% per period.

Firms' planning horizon, R, is set to 30 years, i.e. 120 quarters, corresponding to the average technical lifetimes of power plants (IEA, 2020c).[26] The yearly discount rate is set to 7%, as in IEA (2020c), corresponding to a quarterly discount rate ρ of approximately 1.7%. The implied quarterly discount factor D is 0.98.

We rely on data from the power sector to define the initial share of low-carbon capital κ_0, which is set to 0.2. In particular, we consider EU solar and wind installed electrical capacity in 2020 weighted by their capacity factors, i.e. 22289.32 and 64291.05 MW, respectively (Eurostat, 2021; IRENA, 2021).[27] The total installed capacity adjusted by capacity factors includes, besides wind and solar, also combustible fuels, hydro, nuclear and other sources (geothermal and bioenergy-fired power), and amounts to 422142.91 MW (Eurostat, 2021).[28] Hence, the share of solar and wind capacity over total capacity approximately equals 20%. Under slightly

[25] The same condition can be interpreted as an additional minimum commitment threshold, besides c^{**} in condition (16). This reads, given a certain announced tax rate, as follows: $c > 1 - \frac{1}{\bar{\tau}\beta}$.

[26] Farfan and Breyer (2017) show that an average power plant technical lifetime of about 40 years for coal, 34 years for gas and 34 years for oil-fired power plants. Concerning renewable energy technologies, Tran and Smith (2018) considers lifetime for solar and wind plant ranging between 15 and 35 years.

[27] Solar and wind installed capacity amount to 138443 and 176985 MW, respectively (Eurostat, 2021) These values are then adjusted by solar and wind global weighted-average capacity factors, i.e. 16.1% and 36.3% (IRENA, 2021). The wind capacity factor corresponds to the average of onshore (36%) and offshore (40%) capacity factors weighted by their shares of EU wind installed capacity in 2020.

[28] The source-specific installed capacities are the following. The combustible fuels installed capacity is 388223 MW, while their average capacity factor is estimated at 47,9%, i.e. the average between the 2021 US capacity factor of coal (49.3%) and natural gas (45.76%), weighted by their shares (EIA, 2021). Hydro and nuclear

different assumptions about the energy sources to include, the years to consider and the likes, we obtain values for the initial low-carbon capital share ranging between 0.15 and 0.21. We thus run a sensitivity analysis on κ_0, and present the results in Appendix C.

The ratio of low- to high-carbon capital costs, $\frac{\theta_l}{\theta_h}$, is initialised to 1.36. This quantity proxies the relative convenience of capital investment in high-carbon technologies (see also Acemoglu et al., 2012; Lamperti et al., 2020; van der Ploeg and Rezai, 2020). We calibrate it by computing the average power generation costs of high-carbon (coal and gas) and low-carbon technologies (solar PV, wind onshore and wind offshore), excluding nuclear plants, in the EU countries at the end of 2019. In particular, we consider the average levelised cost of electricity (LCOE) for high- and low-carbon sources, which is, respectively, 80.6 and 109.6\$/MWh,[29] indicating that high-carbon technologies exhibit a 36% advantage in cost-competitiveness.

4.2. Beliefs and investment decisions

Investment responsiveness, γ (cfr. equation (7)), is indirectly calibrated to a value that allows the low-carbon transition - as proxied by the low-carbon share of total capital stock approaching 100% - to occur by year 2050 in the benchmark scenario, i.e. with full commitment to climate objectives. After running a battery of experiments, we set it to 1 and provide a sensitivity analysis in Appendix C. Our choice is motivated by the willingness to use a benchmark scenario wherein the transition to low-carbon investments gradually happens. From a quantitative perspective, $\gamma = 1$ implies that, given the initial backwardness of low-carbon technologies and assuming no climate policy, the share of low-carbon investment is lower than 1%. Further, this choice allows capturing some of the observed inertia in investment rebalancing processes (see also Waisman et al., 2012; Bilias et al., 2010; Vogt-Schilb et al., 2012).

We initialise the difference between beliefs' fitness measures to $U_{b,0} - U_{s,0} = 0.85$, so that the resulting initial share of believers equals $n_0 = 0.3$, corresponding to the proportion of participants to the 2019 Refinitiv Carbon Market Survey (Refinitiv, 2019) expressing trust in the policy-maker's announced strategy regarding the Market Stability Reserve of the EU Emission Trading Scheme.[30]

Estimates of the intensity of choice governing switching mechanisms between various expectation rules (β, see equation (4)) have been at the core of intense debate. Though high values of such parameter predict strong responsiveness of economic agents towards more accurate expectations, several studies conclude in favour of relatively low values or even not significantly different from zero for β, especially when underlying data comes from financial markets (Boswijk et al., 2007; Kukacka and Barunik, 2017; Lamperti, 2018). For example, Chiarella et al. (2014) empirically assesses heterogeneous expectations in asset pricing, using a maximum likelihood approach on S&P500 data to estimate a structural model and finds the estimates to be positive but small. Further, Ellen et al. (2010) estimates an oil price dynamics model with fundamentalist and chartist expectation rules and reports values for the intensity of choice around 1. However, it is important to stress that the calibration of the intensity of choice depends on the unit of measurement of the fitness measure and therefore is model specific. Hence, we select a value of 3, which is reasonably close to the majority of estimates available in the relevant literature (see also Hommes, 2021), but experiment with an ample range of alternative values in sections 5 and 6.

Following Hommes and Lustenhouwer (2019a), we set the memory parameter, η, to 0.5, allowing agents to significantly update their evaluation of the heuristics when new information arises, but also to put considerable weight on the past.

Finally, our benchmark calibration assumes believers to fully trust the policy-maker's announcement, i.e. $\epsilon_b = 1$, and sceptics to discount the announced tax growth entirely, i.e. $\epsilon_s = 0$ and $E_s(g_\tau) = 0 \ \forall t$. This is the extreme case of sceptics expecting no carbon price increase at all. We perform extensive analyses on this parameter, whose exact value is - a priori - extremely difficult to identify without dedicated experiments.

4.3. Policy

Before the tax plan announcement, the prevailing tax level is denoted by τ_{-1}. In order to calibrate it, we convert the 2019 EU-ETS carbon price per ton of CO2 to a tax on production costs. In 2019, the average allowance price in the EU-ETS was approximately 28\$/tCO2 (EEA, 2021a). We convert it to a tax rate by considering data on EU emission intensity of electricity generation in 2018-2019, which ranges between 0.000475 tCO2/kW h (IEA, 2019b) and 0.00034 tCO2/kW h (IEA, 2019a). Hence, the carbon cost per kW h of electricity generated from fossil fuels varies between 0.01-0.013\$, corresponding to a tax of 12.4 - 16.1% on the high-carbon average LCOE (0.0806 \$/kW h). We thus set $\tau_{-1} = 0.15$ and let it vary between 0.1 and 0.2 in the sensitivity analysis illustrated in Appendix C.

The path of the announced climate policy is fully determined by the couple $\{\bar{\tau}_0, \bar{g}_\tau\}$, which is composed by the announced initial tax rate and its growth rate. To calibrate the initial announced tax rate at time 0 ($\bar{\tau}_0$) we rely on the baseline social cost of carbon as determined by van der Ploeg and Rezai (2021) for the year 2019, that is 64 \$/tCO2. This value also falls within the range of carbon

installed capacity equal 150771 MW and 106008 MW, with a capacity factor of 46% and 80.3%, respectively. Other sources such as geothermal and bioenergy-fired power amount to 2171 MW with a capacity factor of 76.5%.

[29] Data are taken from the Levelised Cost of Electricity Calculator of the International Energy Agency (IEA, 2020b). As we are interested in LCOEs under no carbon price, we first set the carbon price to zero and then compute the average LCOE for EU27. The geographic coverage of the calculator varies across technologies. Hence, not all European countries are included and the ones included vary across energy sources.

[30] Participants were asked whether they believed the permits' intake of the Market Stability Reserve (MSR) in case of excess allowances to be reduced at 12%, as planned at the time, or kept at the same level, i.e. 24%. Around 69% of respondents predicted the MSR intake rate to remain stable, with only 31% expecting the announced policy to actually be implemented. Although in this specific case the policy-maker announced a *less* stringent policy to be implemented in the future, we interpret this result as a proxy for the general trust in policy-makers' stated plans.

price suggested by Stiglitz et al. (2017) (40-80 \$/tCO2 in 2020). We convert it to a tax rate using the same procedure employed for the calibration of τ_{-1}, but relying on the EU emission intensity of electricity generation in 2020 (which varies between 0.0002 and 0.0003 tCO2/kW h (IEA, 2020a; EEA, 2021c)). Consequently, the social cost of carbon per kW h of electricity generated from fossil fuels fluctuates between \$0.013 and \$0.019, corresponding to a tax rate between 16-24% on the high-carbon average LCOE. Hence, we set the announced tax rate at t_0 to 0.2. To account for the variability in carbon price estimates, we perform a sensitivity analysis within Appendix C, considering values spanning from 0.1 to 0.35.

The tax growth rate $\bar{g}_\tau$ is calibrated to 0.02 (equivalent to 2% quarterly and 8.2% annually) such that the projected carbon prices are aligned with those resulting from IPCC scenarios and those employed by the British and French governments. BEIS (2021) estimates the 2050 carbon 'value' to be within the range of 189-568£/tCO2e, with a central value of 378£/tCO2e, while Stratégie (2019) proposes a shadow price of carbon between 600 and 900€/tCO2e for 2050. Under our model configuration, the carbon price reaches a value of approximately 590\$/tCO2 at the middle of the century. This calibration is also reasonably close to the average growth rate of the carbon price from mitigation pathway scenarios to 2100 taken from ENGAGE, an inter-model comparison project involving sixteen Integrated Assessment Models (IAMs) to design cost-effective pathways meeting the objectives of the Paris Agreement.[31] The average quarterly growth rate of the carbon price in the scenarios compatible with a 2 °C temperature constraint is approximately between 2.4% and 2.9% (9.95 and 12.11% annually), depending on whether temperature overshoot is allowed.[32] The carbon price growth rates suggested by cost-efficient IAMs can be considered unrealistically high and driven by the implicit acceptance of a second-best scenario where the policy-maker will not accept large immediate price jumps in the short-term Gollier (2022). In the context of our model, what the optimal price is however less relevant, as we are primarily interested in capturing the actual 'political' projections communicated to the public opinion, however they are chosen.

Vulnerability to transition risks a (see equation (10)) is set to 2, such that (i) the concavity of the transition risk index is smooth in the low-carbon share of capital and the tax target; (ii) transition risks are relatively low in 2020 ($\pi_0 \approx 0.25$); and (iii) would need a sharp increase of current policy stringency to rise above its mid-point value ($\pi_0 \approx 0.5$ for $\bar{\tau} \approx 0.7$, equivalent to a carbon price of 188.1 \$/tCO2), proxying a European economy relatively insensible to small variations of carbon prices (Metcalf and Stock, 2020) but potentially vulnerable to large changes (Känzig, 2021).[33,34] Appendix C illustrates how the low-carbon capital share and the believers' share vary with a.

Finally, being a particularly key but ineffable value to calibrate, we explore the implications of the policy-maker's commitment level c (see equation (11)) across its whole range of potential values $[0, 1]$. However, we set two reference values for the parameter in sections 5 and 6: $c = 1$ represents our fully-committed policy-maker; $c = 0.3$, a value below the threshold causing the emergence of multiple steady states (see Fig. 4), represents instead our weakly-committed policy-maker.

5. The transition dynamics under full commitment

By simulating the calibrated model from 2020 to 2060, we first study how the economic system behaves when the policy-maker does not deviate from its climate policy targets ($c = 1 \ \forall \ t$). The full commitment scenario, whereby announced objectives are always achieved, provides a benchmark against which we are able to assess the transition dynamics under different parameter configurations. In what follows, we first illustrate the evolution over time of a set of key variables and then move to studying how the transition dynamics responds to varying degrees of behavioural frictions and opinion polarisation. We will later explore the case of low climate mitigation commitment in section 6.

5.1. The transition path

Fig. 5 presents the evolution of a selection of key variables up to 2060. In this scenario, the carbon tax effectively imposed grows exponentially at the announced growth rate $\bar{g}_\tau$ (Fig. 5a). As the policy-maker consistently meets its policy targets, the share of firms believing in the announced carbon price trajectory progressively increases (Fig. 5b). The speed at which the policy-maker gains credibility depends both on the belief responsiveness, as measured by β, and on the relative accuracy of the two expectation rules (equation (4)). Since the tax target grows gradually, the difference between believers' and sceptics' prediction accuracy is initially small. Therefore, at first firms are unable to precisely assess the policy-maker's credibility and the increase in the share of believers is relatively slow. The shift later accelerates when it becomes clearer that believers are consistently correct in their expectations.

As illustrated in equation (8), the evolution of the aggregate clean investment share χ is determined by both the belief dynamics and the technology choices of each belief type χ_j (Fig. 5c). While sceptics expect no tax increase and thus have a constant low-carbon investment share close to zero, believers invest almost everything in the clean technology from the start of the simulation, with χ_b converging close to 1 from 2020. The resulting aggregate investment increases mainly thanks to the increase of the share of believers in the population. It eventually reaches a value close to one around 2050.

[31] All data about the scenarios we use can be retrieved at https://data.ene.iiasa.ac.at/engage/#/workspaces; see Riahi et al. (2021).

[32] More precisely, our set of scenarios assumes the pursue of Nationally Determined Contributions (NDC) until 2030, followed by the imposition a carbon budget between 1000 and 2000 GtCO2 IPCC (see also 2021).

[33] The impact of the parameter a on the implemented tax is given by $\frac{\partial \tau}{\partial a} = \frac{-(1-c)\bar{\tau}^2(1-\kappa)}{(1+a(1-\kappa)\bar{\tau})^2}$.

[34] As shown in Känzig (2021), the carbon policy surprise series employed is characterised by quite large variations in the carbon price, indicating large policy shocks.

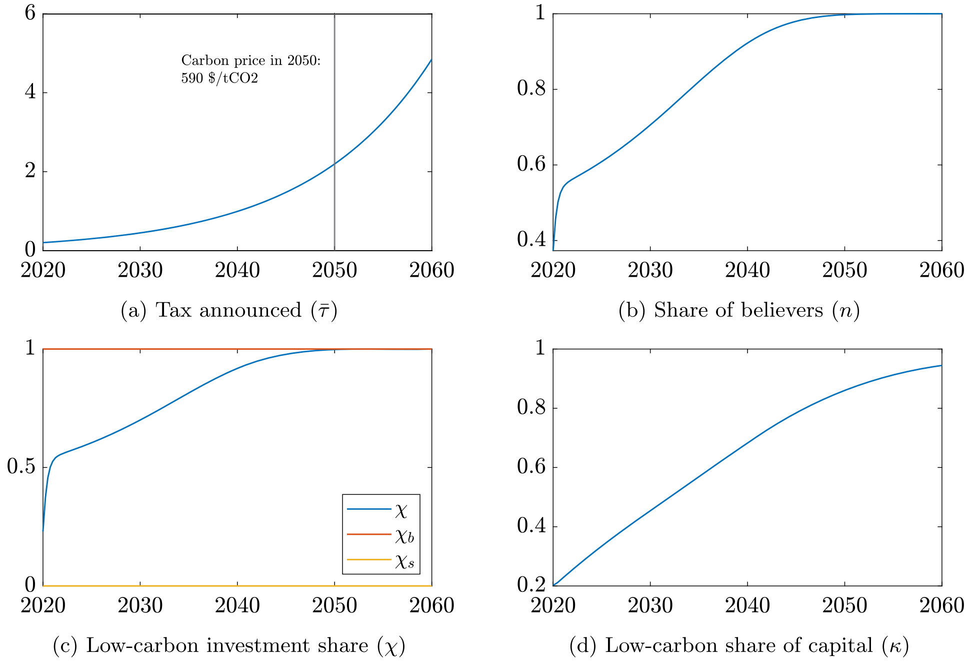


Fig. 5. Evolution over time of selected variables under full commitment ($c = 1$).

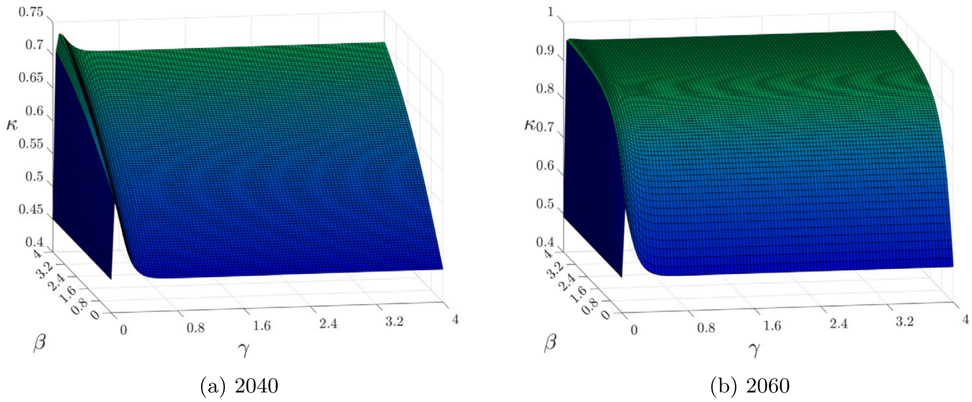


Fig. 6. Low-carbon capital share κ as a function of belief responsiveness β and investment responsiveness γ, under $c = 1$, in (a) 2040 and (b) 2060. All other parameters at their default value (Table 1).

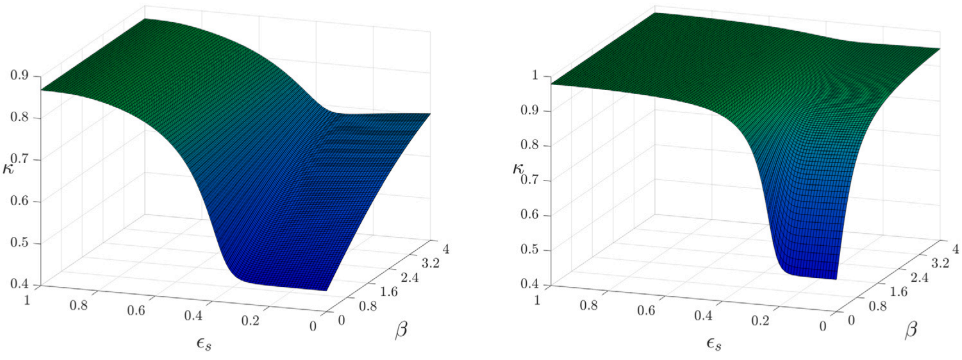
Low-carbon investment drives the dynamics of the clean capital share κ, as in equation (9). Fig. 5d illustrates its evolution through the years: with a policy-maker fully committed to climate objectives, more than 80% of capital is low-carbon by 2050 and the transition is achieved by the 2060.[35]

5.2. Behavioural frictions and opinion polarisation

The transition pathway depends not solely on policy targets and policy-maker's behaviour, but also on how firms respond to them. Even under full commitment, the transition might be hampered if firms fail to internalise long-term policy objectives and/or are reluctant to modify their investment choices. In what follows, we present snapshots of the clean capital share κ at selected years under different behavioural configurations.

Fig. 6 shows clean capital share κ in 2040 (panel (a)) and 2060 (panel (b)) for various degrees of belief and investment responsiveness (β and γ, respectively), which, as discussed in section 2, can be thought of as inversely related to behavioural frictions. A fully committed policy-maker is able to eventually achieve (almost) full decarbonisation under most behavioural configurations,

[35] Note that, as detailed in section 3, under finite belief and investment intensities of choice, an infinitesimal portion of high-carbon capital continues to exist even in the low-carbon steady state.



(a) 2040 (b) 2060

Fig. 7. Low-carbon capital share κ as a function of sceptics' discounting of the tax target growth rate ϵ_s and belief responsiveness β, under $c = 1$, in (a) 2040 and (b) 2060. All other parameters at their default value (Table 1).

except when firms are entirely unresponsive to its policy choices (i.e. if β and/or γ equal zero). Excluding these corner parameter values, the speed of transition greatly varies depending on behavioural frictions. A lower responsiveness of firms to prediction errors (lower β) is undesirable, as it hampers clean investment and delays decarbonisation. The impact of investment responsiveness, instead, is mixed. *Ceteris paribus*, when firms are more responsive to expected cost differences (higher γ), the biases of both believers and sceptics (towards clean and dirty choices, respectively) are amplified, leading them to allocate a higher investment share to their preferred technology. As a result, the impact of investment responsiveness on the transition pace varies with belief dynamics. In the first decades, an increase in γ first accelerates the transition, because of the positive impact on believers' clean investment, but later, as γ crosses a certain threshold, it disproportionately lowers sceptics' clean investment, hampering the transition. Over time, however, for sufficiently high values of belief responsiveness, sceptics rapidly disappear from the population and higher investment responsiveness does not slow down the transition.

Fig. 7 explores how the clean capital share varies with belief responsiveness and opinion polarisation. As mentioned in section 2.1, assuming a fixed ϵ_b, we can use the degree of trust of sceptics in the announced tax target growth rate, $\epsilon_s \in [0, 1]$, as our proxy for opinion polarisation. Through the entire transition, the absence of opinion polarisation ($\epsilon_s = \epsilon_b = 1$) leads to the highest low-carbon capital shares – above 87% in 2040 and close to 100% in 2060, regardless of β. When all firms believe in policy announcements, the transition dynamics depends solely on the tax schedule targeted by the policy-maker and belief responsiveness has no impact whatsoever. For low values of β, the low-carbon share of capital increases monotonously in ϵ_s: as firms are split almost equally between the two expectation rules, the more these converge towards believing the policy announcements, the higher the low-carbon share of capital reached in a certain period.

For higher values of β, opinion polarisation has a non-monotonous impact on the transition dynamics. In particular, as belief responsiveness crosses a certain threshold, a very high polarisation between beliefs (i.e. $\epsilon_s = 0$), leads to a faster low-carbon transition than intermediate values of ϵ_s. The reason lies in the belief switching mechanism. When sceptics expect the tax to be constant over time ($\epsilon_s = 0$) but the policy-maker is fully committed to its policy targets, sceptics' predictions end up being inaccurate soon. Thus, their prediction errors lead firms to switch to the believer expectation rule. Also, the higher β the sooner this belief switch occurs, as firms react rapidly to prediction errors. On the other hand, sceptics mildly discounting the tax target growth rate (values of ϵ_s closer to 1), lead sceptics' prediction errors to be not large enough to induce a rapid switch in beliefs, especially in the first part of the simulation, when the tax target is still low. Over time, this non-monotonicity is reduced because, eventually, sceptics disappear from the population of firms, leading to high values of clean capital share κ.[36]

6. The credible commitment problem

We now investigate transition patterns with varying policy-maker's mitigation commitment levels, as measured by parameter $c \in [0, 1]$. When the policy-maker is not fully committed to meeting its previously announced policy targets ($c < 1$), it might partly deviate from them in response to the perceived transition risks, as in equation (11). As shown in section 3, this departure from policy targets lies behind the emergence of multiple equilibria and might lead the economy into a high-carbon trap. In what follows, we illustrate the transition paths towards the various steady states of the model and the role played by behavioural frictions and opinion polarisation in the transition dynamics under weak commitment.

[36] In Appendix C.2, we extend the analysis presented in this subsection by exploring the combined impact of investment responsiveness and opinion polarisation ϵ_s, under full and weak commitment levels.

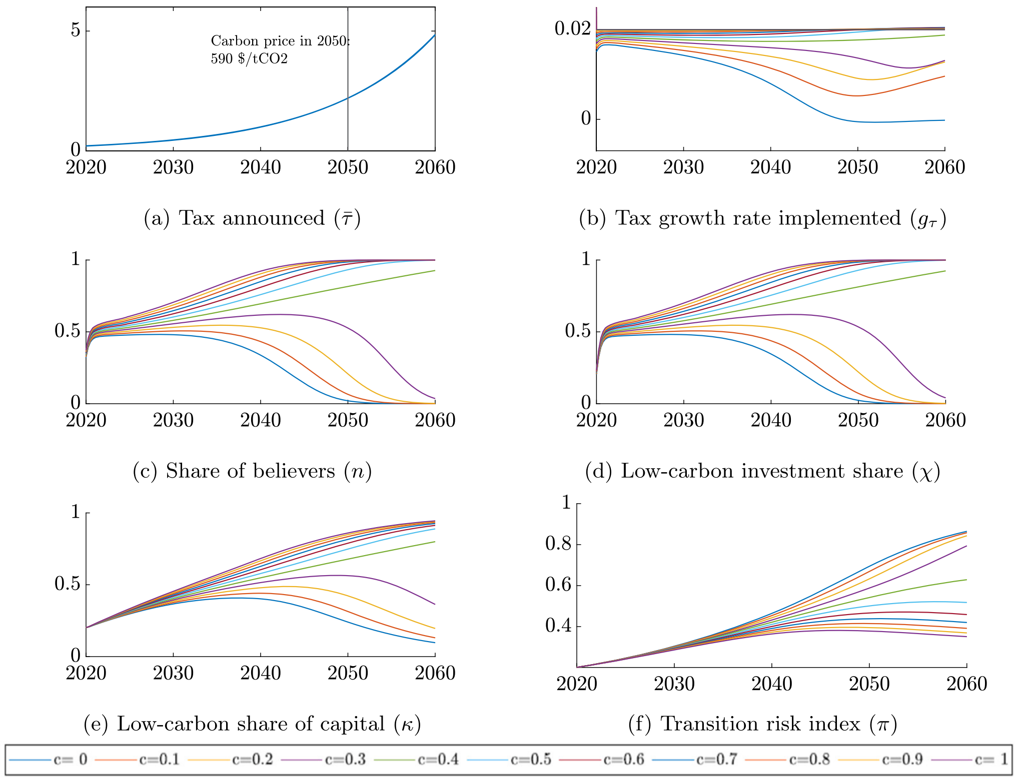


Fig. 8. Evolution over time of selected variables under various levels of commitment (c).

6.1. The transition path

Fig. 8 presents the evolution over time of a set of key variables. Panel (a) shows the tax target schedule, which is independent of commitment c. The tax growth rate actually implemented (panel (b)), instead, varies for lower levels of commitment, because of the policy-maker's attempt to reduce higher transition risks. The initial implemented tax rate (τ_0) ranges from 0.1516 (approximately 41 $/tCO2) for $c = 0$, to 0.1758 (around 47 $/tCO2) for mid-level commitment, and finally 0.2, i.e. the announced tax for $c = 1$.[37] As a result of mild policy stringency and default on its pledges, the non-committed policy-maker pays a cost in terms of credibility loss, reflected in the slower increase of believers' share n (panel (c)). This ultimately leads the low-carbon transition to slow down or, for low enough values of c, to fail entirely (panel (e)). In this case the system goes back to a dynamics characterised by a majority of sceptics ($n \approx 0$) and low clean investment and capital shares ($\chi \approx 0$; $\kappa \approx 0$).

The transition failure emerges only quite late in time: at first, even a weakly committed policy-maker cannot depart too much from policy targets, as these are initially mild and the associated transition risks low. In the first periods, therefore, the policy-maker's credibility is at worst more or less constant. As the tax target and transition risks grow over time, firms recognise the weakly committed policy-maker and eventually lose trust in its announcements, up to the point where low-carbon investment is significantly reduced. As this tipping point is reached, the transition dynamics reverses and κ converges to the sceptics' clean investment share, χ_s.

Furthermore, transition risks end up being significantly larger under weak rather than stronger commitment levels, because the high-carbon share of the economy affected by the tax is larger. Hence, the weakly committed policy-maker cannot really escape transition risks, but only postpone (and amplify) them to the future. The result is a higher cost of policy action, which further reduces policy-makers' options, up to the point where the low-carbon transition fails.

Fig. 9 shows how the low credibility trap illustrated in section 3 emerges over time under an ambitious but uncommitted policy-maker. For the initial decades the low-carbon capital share is increasing in the target tax growth rate even for low levels of policy-maker's commitment ($c > 0.2$), as transition risks are perceived to be mild and even the weakly committed policy-maker does not depart excessively from the announced targets (panel (a)). However, unmet ambitious objectives ultimately generate a credibility loss, leading to an increase in the population of sceptics and a decrease in the low-carbon capital share, ultimately causing a transition failure (panel (b)).

37 Although outside the scope of the paper, it is worth noting how the EU ETS price in 2020 and 2021 was closer to the ones predicted by a low commitment level.

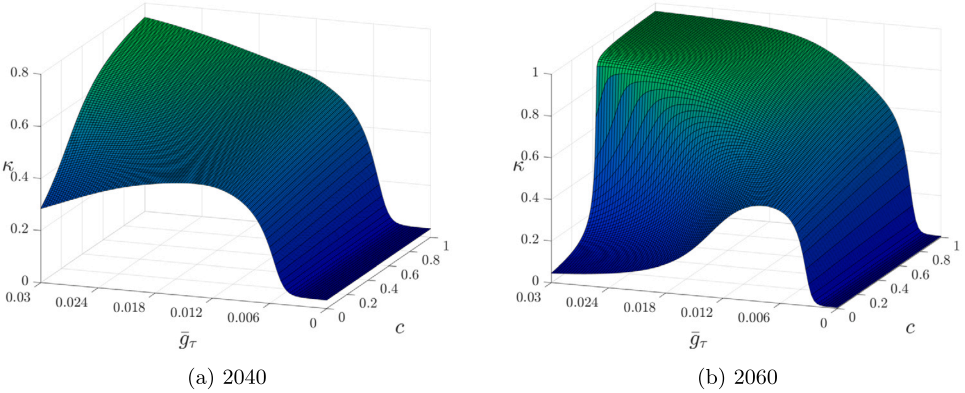


Fig. 9. Low-carbon capital share κ as a function of the tax target growth rate $\bar{g}_\tau$ and commitment c, in (a) 2040 and (b) 2060. All other parameters at their default value (Table 1).

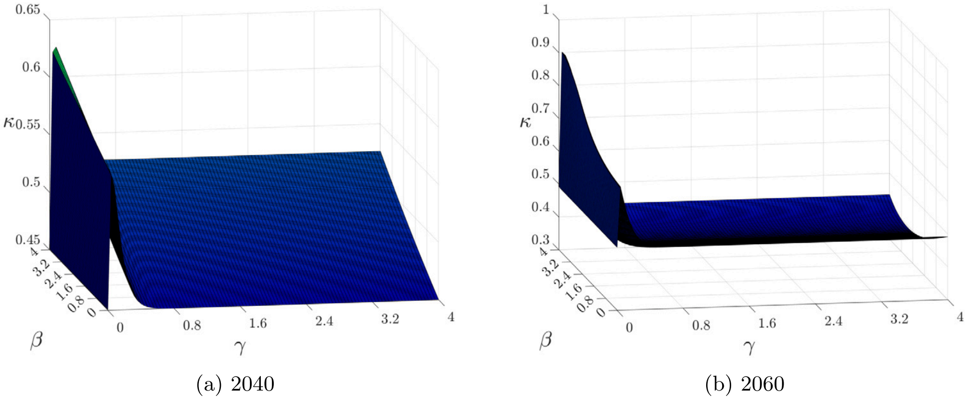


Fig. 10. Low-carbon capital share κ as a function of belief responsiveness β and investment responsiveness γ, under $c = 0.3$, in (a) 2040 and (b) 2060. All other parameters at their default value (Table 1).

6.2. Behavioural frictions and opinion polarisation

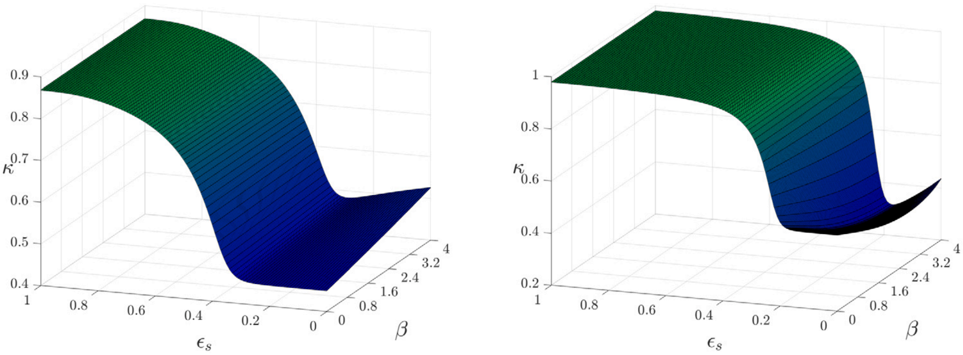
The policy-maker's behaviour is a key determinant of the transition success or failure. However, as already shown in section 5.2, firms' response to policy choices is equally important for rapidly achieving decarbonisation. In what follows, we explore the impact of behavioural factors on the transition dynamics, under weak commitment ($c = 0.3$).

Fig. 10 shows snapshots of the clean capital share κ at different points of time for various levels of belief and investment responsiveness. In 2040, the low-carbon capital share presents a similar relationship with β and γ to that under full commitment (Fig. 6), although at lower levels of κ. The picture is instead different two decades later (panel (b)). Indeed, belief responsiveness has a non monotonous effect on the clean capital share. Values of β slightly larger than 0 hamper the transition as firms realise the government is not keeping its word. As β crosses a certain threshold, the transition is slightly faster and involves lower transition risks faced by the policy-maker, who can thus implement a tax closer to the target.

Fig. 11 illustrates how κ evolves under various levels of belief responsiveness β and opinion polarisation, proxied by ϵ_s. For sufficiently high values of belief responsiveness, strong opinion polarisation (i.e. low ϵ_s) accelerates the transition, with respect to milder polarisation. While the same effect takes place under full commitment (Fig. 7), under weak policy-maker's commitment the non-monotonicity is much more pronounced. The reason is that, for low values of commitment c, a certain degree of scepticism produces even more accurate tax predictions than under full commitment, causing a delay in the disappearance of sceptics from the population of firms. Eventually, for high values of β, sceptics are proved wrong and the transition takes place, although at a low pace, as illustrated in panel (b). For low belief responsiveness and high opinion polarisation, instead, the negative feedback loop between policy-maker's weak credibility and firms' investment choices emerges, pushing the economy into a high-carbon trap.

7. Conclusions

In this paper, we model and analyse the dynamic interaction between heterogeneous expectations, investment decisions and climate mitigation policy-making. We develop a novel modelling approach rooted in discrete choice theory, able to account for dynamic beliefs and uncertain carbon price paths. We obtain four key results. First, a 'high-carbon credibility trap' - i.e. the con-



(a) 2040 (b) 2060

Fig. 11. Low-carbon capital share κ as a function of sceptics' discounting of the tax target growth rate ϵ_s and belief responsiveness β, under $c = 0.3$, in (a) 2040 and (b) 2060. All other parameters at their default value (Table 1).

vergence to a carbon-intensive steady state when an alternative low-carbon equilibrium is also present - might emerge when an ambitious plan is announced by a weakly committed policy-maker. This can trap the economic system in a vicious circle of carbon-intensive investment, increasing transition risks and policy-maker's credibility, eventually leading to a transition failure. Second, the presence of behavioural frictions - either in capital investment choices or in the assessment of the policy-maker's credibility - creates heterogeneous expectations and affects the conditions of existence of long-run system equilibria, making it harder to achieve full decarbonisation. However, higher responsiveness to the performance of belief/investment strategies makes firms' behaviour more volatile, increasing the likelihood to fall into a high-carbon trap. Third, even when the economic system is directed towards a low-carbon equilibrium (e.g. under a fully committed policy-maker), behavioural frictions affect the rapidity of the decarbonisation process in non-linear manners. Finally, belief polarisation can also have non-linear implications on decarbonisation, with higher belief polarisation being beneficial for the transition under certain circumstances.

Our work can benefit from a number of additional refinements. Our numerical model suffers from the scarce availability of systematic data concerning transition-related beliefs and expectations, making it hard to calibrate behavioural parameters (primarily γ and β). Consequently, the exact timing of our transition dynamics should mainly be interpreted in a qualitative manner, rather than a precise forecast. The complexity of the modelling framework could also be expanded, or directed towards additional research questions. For instance, we focus on carbon taxation as a prototypical climate policy instrument. While it shares similarities with other instruments such as emission trading systems or emission standards in terms of credibility, it would be important to delve deeper into how policy uncertainty might affect the effectiveness of various instruments, accounting for the specificities of each instrument. Moreover, we rely on an exogenous growth path to better focus on the investment allocation choice, but the transition dynamics is likely to have wider macroeconomic implications, suggesting additional insights could be obtained by making growth dynamics endogenous. Our commitment level c, now an intrinsic and immutable feature of a policy-maker, could also be made endogenous and variable in time, possibly jumping following electoral cycles. Another possible direction of research is to study financial, rather than capital, investment decisions, which would require incorporating a financial sector in the model. Finally, we abstract here from premature decommissioning, loss of capacity utilisation and costly capital reconversion ('stranding'), although including these dimensions, both in reality and in expectations, is likely to have implications on the overall transition dynamics (see for instance Cahen-Fourot et al., 2022; Campiglio et al., 2022).

Despite these limitations, our results offer several key insights for policy-making. We have shown how the direction and heterogeneity of expectations of future climate mitigation policies can significantly affect the dynamics of the climate policies themselves. It is thus absolutely key for the policy-maker to (i) be aware of what these expectations are and their distribution; and (ii) be able to orient them as desired. As mentioned above, the current availability of data on transition-related expectations, their drivers and their dynamic behaviour is very scarce. Public institutions should contribute to running surveys, experiments and empirical analysis that could, in combination, provide a more solid calibration basis in the future. The ability to orient expectations comes instead from credibility, itself a function of the past track-record in sticking to announced plans. In the context of climate policies, many jurisdictions have shown worrying signs of being unable to maintain their course for sufficiently extended periods of time. A wide societal debate on the most appropriate institutional configuration to achieve long-termist and credible policies is urgently needed.

Finally, our results warn policy-makers of the risks of having both too little and too much ambitions in their policy announcements. Too little ambition will clearly not provide a sufficient economic inventive to shift investment decisions towards the low-carbon option, which would remain less convenient than the incumbent. Too much ambition could instead be a winning strategy, under certain conditions. If the initial announcement is sufficiently credible, announcing a stronger carbon plan than what the government is actually planning to implement, could tip the economic system towards a low-carbon equilibrium before the deception is internalised by firms. However, if this strategy is brought too far, it will eventually push the system into a bad equilibrium characterised by low policy credibility and carbon-intensive investments, ultimately leading to a transition failure.

Appendix A. Microfoundation of logit model

We use a logit model to characterise firms' belief switching and investment choices. This model is based on a discrete choice framework McFadden (1974), which can be microfounded with the random utility framework Train (2009). According to this framework, agents attempt to maximise their utility, which depends on factors common to everyone and explicitly modelled and on idiosyncratic factors treated as random. These idiosyncratic factors can be interpreted as behavioural frictions that impede agents' ability to maximise their utility, leading to heterogeneous choices across agents. In what follows, we show how the random utility model leads to the aggregate belief and investment choices we employ in this paper.

Let us start from the belief switching process, where firm f faces two alternatives $j \in \{b, s\}$, each of which is characterised by a certain random utility Z^*_{fj}. The utility deriving from each alternative is decomposed into a part labelled V_j that is common to all firms, and an idiosyncratic part ϵ^*_{fj} that is treated as random:

$$Z^*_{fj} = V_j + \epsilon^*_{fj} \quad \forall f, j \tag{19}$$

where the common factor $V_j = -\beta^* U_j$ is a function of the fitness measure U_j (see equation (3)) and β^* is the effect of U_j on V_j.

The logit model is obtained by assuming that each ϵ^*_{fj} is independently, identically distributed Gumbel with variance $\sigma^2 \frac{\pi^2}{6}$, where σ is a scale parameter Train (2009). The Gumbel distribution is very similar to a normal, except that it is characterised by slightly fatter tails, thus allowing for slightly more aberrant behaviour than the normal.

By scaling the random utility by $\frac{1}{\sigma}$, the ordering of alternatives is unchanged and the probability of firm f choosing believers' expectation rule is given by the following cumulative density function:

$$P_{fb} = 1 - P_{fs} = \text{Prob}(-\beta U_b + \epsilon_{fb} > -\beta U_s + \epsilon_{fs}), \tag{20}$$

where $\beta = \frac{\beta^*}{\sigma}$ and $\epsilon_{fj} = \frac{\epsilon^*_{fj}}{\sigma}$. The parameter σ thus scales the coefficient β^* to reflect the variance of the idiosyncratic portion of utility. In fact, β indicates the effect of the common variable on the utility relative to the variance of the idiosyncratic factors. A larger variance of the latter factors leads to smaller values of β, even if the common factors have the same effect on utility.

Equation (20) is equivalent to:

$$P_{fb} = \text{Prob}(\epsilon_{fs} - \epsilon_{fb} < \beta U_s - \beta U_b), \tag{21}$$

Since the difference between two Gumbel variables is distributed logistic, then we can express P_{fb}, as well as the share of believers n, as follows:

$$n = P_{fb} = \frac{1}{1 + \exp(-\beta(U_s - U_b))}. \tag{22}$$

Based on expectation rule j, firm f chooses its investment allocation between low- and high-carbon technologies. Both technologies $i \in \{l, h\}$ provide a certain random utility to firm f:

$$Y^*_{fji} = W_{ji} + \epsilon^*_{fji} \tag{23}$$

where $W_{ji} = \gamma^* E_j(\Theta_i)$ is the factor common to all firms with belief type j and is a function of expected production costs of technology i. ϵ^*_{fji} is the factor idiosyncratic to firm f and is assumed iid Gumbel with variance $\sigma^2 \frac{\pi^2}{6}$. Similarly to the belief choice, we scale the random utility by $\frac{1}{\sigma}$ and obtain the probability of firm f with belief j choosing technology l:

$$P_{fjl} = \text{Prob}(\epsilon_{fjh} - \epsilon_{fjl} < \gamma E_j(\Theta_h) - \gamma E_j(\Theta_l)), \tag{24}$$

where $\epsilon_{fji} = \frac{\epsilon^*_{fji}}{\sigma}$ and $\gamma = \frac{\gamma^*}{\sigma}$. Under the assumption that the idiosyncratic factors are iid Gumbel, equation (24) corresponds to:

$$\chi_j = P_{fjl} = \frac{1}{1 + \exp(\gamma(E_j(\Theta_h) - E_j(\Theta_l)))}. \tag{25}$$

Appendix B. Proofs and derivation of analytical results

B.1. Proof of Proposition 1

Proof. First, let us rewrite $f(\kappa)$ for clarity:

$$f(\kappa) = (\chi_b - \chi_s) \frac{1}{1 + \exp\left(-\beta\left\{2\bar{\tau}\left[c + \frac{1-c}{1 + a(1-\kappa)\bar{\tau}}\right] - \tau_{-1} - \bar{\tau}\right\}\right)} + \chi_s.$$

The function $f(\kappa)$ is a composition of an exponential function and, therefore, is continuous in [0,1] provided that the denominator $1 + a(1-\kappa)\bar{\tau} \neq 0$, which is always verified given that $a(1-\kappa)\bar{\tau} \geq 0$.

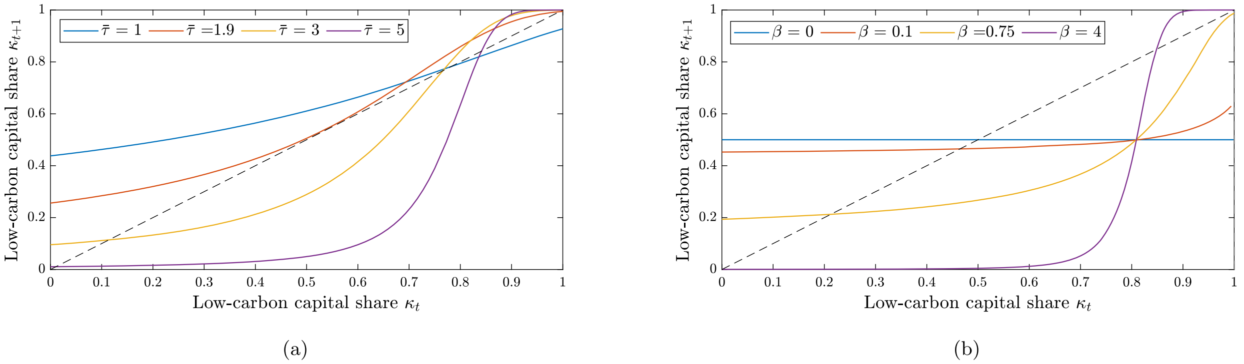


Fig. A1. Dynamics of the low-carbon share of capital κ for different levels of (a) the tax target $\bar{\tau}$ and (b) belief responsiveness β, under $c = 0.3$.

In the case of finite β and γ, $f(\kappa) \in (0,1)$. More specifically, we show that $f(0) > 0$. Indeed, $f(0) = (\chi_b - \chi_s)\frac{1}{1+\exp\left(-\beta\left(2\bar{\tau}\frac{ca\bar{\tau}+1}{1+a\bar{\tau}} - \tau_{-1} - \bar{\tau}\right)\right)} + \chi_s$, where

- $\chi_b, \chi_s \in (0,1)$ (see Equation (7)).
- $\chi_b - \chi_s > 0$ as we assume that believers' expected tax ($\bar{\tau}$) is higher than sceptics' (τ_{-1}) (see Assumption 1).
- $\frac{1}{1+\exp\left(-\beta\left(2\bar{\tau}\frac{ca\bar{\tau}+1}{1+a\bar{\tau}} - \tau_{-1} - \bar{\tau}\right)\right)} > 0$, because the denominator is positive.

Hence, $f(0) > 0$.

Further, it is possible to show that $f(1) < 1$. Indeed, $f(1) = (\chi_b - \chi_s)\frac{1}{1+\exp(-\beta(\bar{\tau} - \tau_{-1}))} + \chi_s$. We thus set the inequality:

$$(\chi_b - \chi_s)\frac{1}{1 + \exp\left(-\beta\left(\bar{\tau} - \tau_{-1}\right)\right)} < 1 - \chi_s, \tag{26}$$

which is always verified given that $\frac{1}{1+\exp(-\beta(\bar{\tau}-\tau_{-1}))} < 1$, and $\chi_b < 1$. Combining these elements, it follows that $(\chi_b - \chi_s)\frac{1}{1+\exp(-\beta(\bar{\tau}-\tau_{-1}))} < \chi_b - \chi_s < 1 - \chi_s$.

Finally, $f(\kappa)$ is an increasing function of κ. The first derivative $f'(\kappa)$ of the function is given by

$$f'(\kappa) = -\frac{2\,a\,\bar{\tau}^2\,\beta\,e^{\beta\left[\tau_{-1} - \bar{\tau}\left(c + \frac{c-1}{a\,\bar{\tau}\,(\kappa-1)-1}\right)\right]}\,e^{\beta\left[\bar{\tau} - \bar{\tau}\left(c + \frac{c-1}{a\,\bar{\tau}\,(\kappa-1)-1}\right)\right]}\,\tilde{X}_j\,(c-1)}{\left\{e^{\beta\left[\tau_{-1} - \bar{\tau}\left(c + \frac{c-1}{a\,\bar{\tau}\,(\kappa-1)-1}\right)\right]}\,e^{\beta\left[\bar{\tau} - \bar{\tau}\left(c + \frac{c-1}{a\,\bar{\tau}\,(\kappa-1)-1}\right)\right]} + 1\right\}^2 (a\,\bar{\tau} - a\,\bar{\tau}\,\kappa + 1)^2}, \tag{27}$$

where $\tilde{X}_j \equiv \chi_b - \chi_s$. For $c \neq 1$ and $\beta \neq 0$, all the elements of Equation (27) are positive, hence $f(\kappa)$ is increasing in [0,1].[38]

The continuity of $f(\kappa)$, combined with the fact that $f(0) > 0$, $f(1) < 1$ and $f(\kappa)$ is increasing, implies that the map starts above the 45 degree line and ends below the 45 degree line. Therefore, at least one stable steady state exists. If there exists a unique steady state, it is stable, as it intersects the 45-degree line from above, i.e. with a derivative smaller than one. If there are an odd number of steady states, excluding those cases where f is tangent to the 45 degree line, the odd ones intersect the 45-degree line with a derivative smaller than one, while the even ones intersect it from below, indicating a derivative greater than one and thus instability.

See Proof B.2 for the case in which β and γ are infinite. □

B.2. Proof of Proposition 2

Proof. Under $\beta = \gamma = \infty$, plugging $\kappa_t = 1$ into equation (13), leads to

$$n_{t+1} = \frac{1}{1 + \exp(-\beta(\bar{\tau} - \tau_{-1}))}. \tag{28}$$

Since $\bar{\tau} > \tau_{-1}$ by assumption, it follows that for $\kappa_t = 1$, $n_{t+1} = 1$ and therefore $\kappa_{t+1} = \chi_b$. The believers' low-carbon investment share χ_b can be expressed as[39]

[38] As mentioned in section 3, $c = 1$ and $\beta = 0$ represent two particular cases where the model features unique steady states that are determined solely by the exogenous parameters.

[39] After the sum of expected future production costs (equation (6)) in the absence of expected climate policy has been rearranged as $\theta_i \frac{1-D^{R+1}}{1-D}$ and simplified.

$$\chi_b = \frac{1}{1+\exp\left(-\gamma \frac{1-D^{R+1}}{1-D}[\theta_h(1+\bar{\tau}) - \theta_l)]\right)}. \tag{29}$$

In order for χ_b to equal 1, under $\gamma = \infty$, the announced tax target $\bar{\tau}$ must make the low-carbon technology more convenient than the high-carbon one ($\bar{\tau} > \frac{\theta_l - \theta_h}{\theta_h}$). If this condition is satisfied, the low-carbon steady state $\kappa_l^* = 1$ exists.

With respect to the high-carbon steady state, let us first note that, under the assumption that $\tau_{-1} < \frac{\theta_l-\theta_h}{\theta_h}$ and under infinite γ, sceptics do not invest at all in the low-carbon technology ($\chi_s = 0$). Also, from equation (12) it follows that $\kappa = 0$ is a steady state if, for $\kappa_t = 0$, $n_{t+1} = 0$ or $\chi_b = 0$. Concerning the former case ($n_{t+1} = 0$) let us plug $\kappa_t = 0$ into equation (13):

$$n_{t+1} = \frac{1}{1+\exp\left(-\beta\left[2\bar{\tau}\left(c+\frac{1-c}{1+a\bar{\tau}}\right) - \tau_{-1} - \bar{\tau}\right]\right)}, \tag{30}$$

which, under infinite β, equals zero if

$$c < \frac{1}{2} - \frac{\bar{\tau} - \tau_{-1}(1+a\bar{\tau})}{2a\bar{\tau}^2}. \tag{31}$$

Concerning the latter case, $\chi_b = 0$ if $\bar{\tau} < \frac{\theta_l-\theta_h}{\theta_h}$. □

B.3. Proof of Proposition 3

Proof. Concerning the low-carbon steady state, we assume that $\kappa_t = 1 - \tilde{\lambda}_l$, where $\tilde{\lambda}_l = \lambda_l + \varepsilon_l$, with ε_l a sufficiently small positive number, and we impose that $\kappa_{t+1} > 1 - \tilde{\lambda}_l$, meaning that κ is converging to a stable steady state $\kappa_l^* = 1 - \lambda_l$. Hence:

$$\kappa_{t+1} = \frac{\chi_b - \chi_s}{1+\exp\left(-\beta\left\{2\bar{\tau}\left[c+\left(\frac{1-c}{1+a\tilde{\lambda}_l\bar{\tau}}\right)\right]\right\}\right)} + \chi_s > 1 - \tilde{\lambda}_l, \tag{32}$$

which implies

$$\beta\left\{2\bar{\tau}\left[c+\left(\frac{1-c}{1+a\tilde{\lambda}_l\bar{\tau}}\right)\right]\right\} > -\ln\left(\frac{\chi_b - 1 + \tilde{\lambda}_l}{1-\tilde{\lambda}_l - \chi_s}\right) \tag{33}$$

and

$$c > \frac{1}{2} - \frac{\bar{\tau} - \tau_{-1}(1+a\tilde{\lambda}_l\bar{\tau})}{2a\tilde{\lambda}_l\bar{\tau}^2} - \ln\left(\frac{\chi_b - 1 + \tilde{\lambda}_l}{1-\tilde{\lambda}_l-\chi_s}\right)(2\bar{\tau}\beta)^{-1}\left(1+\frac{1}{a\tilde{\lambda}_l\bar{\tau}}\right). \tag{34}$$

In order for equation (34) to be well defined, we impose

$$\begin{aligned} &\chi_b > 1 - \tilde{\lambda}_l, \\ &\chi_s < 1 - \tilde{\lambda}_l. \end{aligned} \tag{35}$$

The former condition is satisfied if

$$\bar{\tau} > \frac{\theta_l - \theta_h}{\theta_h} - \frac{\ln\left(\frac{\tilde{\lambda}_l}{1-\tilde{\lambda}_l}\right)}{\frac{1-D^{R+1}}{1-D}\gamma\theta_h}, \tag{36}$$

where $D \equiv \frac{1}{1+\rho}$. The second condition is satisfied if

$$\tau_{-1} < \frac{\theta_l - \theta_h}{\theta_h} - \frac{\ln\left(\frac{\tilde{\lambda}_l}{1-\tilde{\lambda}_l}\right)}{\frac{1-D^{R+1}}{1-D}\gamma\theta_h}. \tag{37}$$

Since we assume that $\tau_{-1} < \frac{\theta_l-\theta_h}{\theta_h}$, if $\tilde{\lambda}_l < 0.5$, then condition (37) is always verified. If, instead, $\tilde{\lambda}_l > 0.5$, this is an additional constraint.

Concerning the high-carbon steady state, we assume that $\kappa_t = \chi_s + \tilde{\lambda}_h$, where $\tilde{\lambda}_h = \lambda_h + \varepsilon_h$, with ε_h a sufficiently small positive number, and we impose that $\kappa_{t+1} < \chi_s + \tilde{\lambda}_h$, meaning that κ is converging to a stable steady state $\kappa_h^* = \chi_s + \lambda_h$. Hence:

$$\kappa_{t+1} = \frac{\chi_b - \chi_s}{1+\exp\left(-\beta\left[2\bar{\tau}\left(c+\frac{1-c}{1+a(1-\chi_s-\tilde{\lambda}_h)\bar{\tau}}\right)\right]\right)} + \chi_s < \chi_s + \tilde{\lambda}_h \tag{38}$$

which implies

$$c < \frac{1}{2} + \frac{\bar{\tau} - \tau_{-1}\{1+a[1-(\chi_s+\tilde{\lambda}_h)]\bar{\tau})\}}{2a[1-(\chi_s+\tilde{\lambda}_h)]\bar{\tau}^2} - \frac{\ln\left(\frac{\chi_b-\chi_s-\tilde{\lambda}_h}{\tilde{\lambda}_h}\right)}{2\bar{\tau}\beta}\left\{1+\frac{1}{a[1-(\chi_s+\tilde{\lambda}_h)]\bar{\tau}}\right\}. \tag{39}$$

In order for equation (39) to be well defined, we impose $\tilde{\lambda}_h < \chi_b - \chi_s$. □

B.4. Proof of Proposition 4

Proof. The second derivative of $f(\kappa)$ is:

$$f''(\kappa) = -\frac{\tilde{x}\, e^{\tilde{\beta}} \left[\left(a\,\bar{\tau} - \bar{\tau}\,\beta + \bar{\tau}\,\beta\,c - a\,\bar{\tau}\,\kappa_t + 1 \right) + e^{\tilde{\beta}} \left(a\,\bar{\tau} + \bar{\tau}\,\beta - \bar{\tau}\,\beta\,c - a\,\bar{\tau}\,\kappa_t + 1 \right) \right]}{(e^{\tilde{\beta}} + 1)^3 \left(a\,\bar{\tau} - a\,\bar{\tau}\,\kappa_t + 1 \right)^4}, \tag{40}$$

where $\tilde{x} \equiv \left(\chi_b - \chi_s\right)(c-1)\,4\,a^2\,\bar{\tau}^3\,\beta < 0$ and $\tilde{\beta} \equiv \beta\left(\tau_{-1} - 2\tau_t + \bar{\tau}\right)$. Although we cannot find analytically the inflection points of κ where $f''(\kappa) = 0$, we observe that, for $\beta \neq 0$ and $c \neq 1$, if $c > 1 - \frac{1}{\bar{\tau}\beta}$, then $\left(a\,\bar{\tau} - \bar{\tau}\,\beta + \bar{\tau}\,\beta\,c - a\,\bar{\tau}\,\kappa_t + 1\right) > 0$ and $f''(\kappa) > 0$ for all $\kappa \in [0,1]$ and the fixed point is unique. If $c < 1 - \frac{1}{\bar{\tau}\beta}$, one or more fixed points may exist. □

Appendix C. Sensitivity analysis

C.1. Alternative tax plans

The sensitivity analysis presented in Table A1 provides insights into key model parameters, including the announced initial tax rate and tax growth rate. However, it is also interesting to vary these two parameters simultaneously and study the impact of different tax plans. Thus, we consider four distinct tax plans:

- Tax plan 1: starting from $\bar{\tau}_0 = 0.1$ and growing at $\bar{g}_\tau = 0.02$
- Tax plan 2: starting from $\bar{\tau}_0 = 0.3$ and growth rate announced $\bar{g}_\tau = 0.015$
- Tax plan 3: starting from $\bar{\tau}_0 = 0.5$ and growth rate announced $\bar{g}_\tau = 0.01$
- Tax plan 4: starting from $\bar{\tau}_0 = 0.7$ and growth rate announced $\bar{g}_\tau = 0.005$.

We select these tax plans because they embody contrasting approaches to policy setting. One strategy involves initiating with a low tax rate and progressively increasing it to incentivise decarbonisation swiftly. Conversely, another approach opts for an immediate imposition of a high tax rate, albeit with a slower subsequent rate of increase. This latter option seeks to mitigate the time-inconsistency issue (Kydland and Prescott, 1977) which can potentially lead to the erosion of policy credibility over time. The rationale here is that by imposing a high tax rate upfront, policy-makers establish credibility from the beginning while avoiding the need for steep increases in the subsequent periods.

Our findings are shown in Fig. A2, which illustrates the model dynamics under these tax plans, focusing on a scenario with a weakly committed policy-maker ($c = 0.3$). Our results suggest that even announcing a high initial tax rate fails to shield policymakers from the transition risks inherent in decarbonisation efforts. In fact, the risks appear even more pronounced initially due to the substantial presence of high-carbon sectors.

In our analysis, we observe that none of the considered tax plans succeed in facilitating a transition to a low-carbon economy. Despite variations in outcomes, all tax plans ultimately lead to a failed decarbonisation endeavour. This is due to high transition risks, loss of credibility and a subsequent decline in the proportion of believers across all scenarios. These results underscore the inherent risks associated with weak commitment in policy-making. Even in scenarios characterised by policy pathways presumed to mitigate time-inconsistency issues, our model yields insights into the risks of weak commitment.

C.2. Additional results on behavioural frictions and opinion polarisation

Fig. A3 shows how the low-carbon capital share evolves under different values of investment responsiveness (γ) and opinion polarisation (ϵ_s), assuming full policy-maker's commitment ($c = 1$).

We confirm the non-monotonous impact of investment responsiveness and opinion polarisation on the pace of decarbonisation. More specifically, we find that, when firms are highly responsive to expected cost differences (high γ), decreasing opinion polarisation (varying ϵ_s from zero to approximately 0.4) slightly decreases the transition speed. This non-monotonicity is hardly visible because it results from two balancing forces of roughly equal force. On the one hand, decreasing opinion polarisation under full commitment slows down the disappearance of sceptics from the population of firms. On the other hand, a higher ϵ_s diminishes the sceptics' bias towards high-carbon technologies. As ϵ_s passes a certain threshold, especially in 2040, the second effect strongly grows in importance, speeding up the transition. Further in time (panel (b)), if the policy-maker is committed to its policy targets, investment responsiveness and opinion polarisation, while still having a non monotonous impact, do not significantly hamper the transition.

Fig. A4 shows the impact of investment responsiveness γ and opinion polarisation ϵ_s on the speed of decarbonisation, assuming a weakly committed policy-maker. While in 2040 the figure is very similar to that under full commitment (see Fig. A3), in 2060, under high investment responsiveness γ and high opinion polarisation (i.e. small ϵ_s), the weakly committed policy-maker forces the economy into a high-carbon trap. The reason is that both these behavioural factors increase sceptics' bias towards high-carbon technologies, depressing sceptics' low-carbon investment share, up to the point where the policy-maker fails to decarbonise the economy.

Table A1
Sensitivity analysis.

	Full commitment ($c = 1$)				Poor commitment ($c = 0.3$)			
Parameter	κ in 2040	κ in 2060	n in 2040	n in 2060	κ in 2040	κ in 2060	n in 2040	n in 2060
Production								
Output growth rate g_Y								
0.003	0.6631	0.9323	0.9225	1.0000	0.5009	0.3643	0.6107	0.0357
0.005	0.6826	0.9447	0.9225	1.0000	0.5151	0.3633	0.6176	0.0357
0.007	0.6999	0.9546	0.9225	1.0000	0.5279	0.3617	0.6240	0.0355
0.009	0.7153	0.9625	0.9225	1.0000	0.5395	0.3597	0.6299	0.0354
0.011	0.7290	0.9689	0.9225	1.0000	0.5500	0.3576	0.6353	0.0352
Depreciation rate δ								
0.006	0.5168	0.7938	0.9225	1.0000	0.4002	0.3500	0.5656	0.0334
0.01	0.5862	0.8696	0.9225	1.0000	0.4469	0.3614	0.5856	0.0351
0.014	0.6402	0.9259	0.9225	1.0000	0.4845	0.3646	0.6029	0.0357
0.018	0.6826	0.9447	0.9225	1.0000	0.5151	0.3633	0.6176	0.0357
0.022	0.7161	0.9629	0.9225	1.0000	0.5401	0.3596	0.6302	0.0354
Initial low-carbon capital share κ_0								
0.1	0.6676	0.9423	0.9225	1.0000	0.4933	0.3482	0.6069	0.0337
0.2	0.6826	0.9447	0.9225	1.0000	0.5151	0.3633	0.6176	0.0357
0.3	0.6976	0.9470	0.9225	1.0000	0.5372	0.3801	0.6289	0.0381
0.4	0.7125	0.9494	0.9225	1.0000	0.5597	0.3992	0.6408	0.0412
0.5	0.7275	0.9517	0.9225	1.0000	0.5826	0.4211	0.6532	0.0453
Low- to high-carbon production costs $\frac{\theta_l}{\theta_h}$								
1	0.8801	0.9812	0.9225	1.0000	0.8800	0.9812	0.8436	1.0000
1.1	0.8665	0.9787	0.9225	1.0000	0.8610	0.9770	0.8304	1.0000
1.2	0.6962	0.9472	0.9225	1.0000	0.5459	0.4488	0.6331	0.0512
1.3	0.6827	0.9447	0.9225	1.0000	0.5152	0.3637	0.6177	0.0357
1.4	0.6826	0.9447	0.9225	1.0000	0.5151	0.3633	0.6176	0.0357
1.5	0.6825	0.9447	0.9225	1.0000	0.5149	0.3631	0.6175	0.0357
Investment responsiveness γ								
0	0.4551	0.4929	0.9225	1.0000	0.4551	0.4929	0.5897	0.0613
1	0.6826	0.9447	0.9225	1.0000	0.5151	0.3633	0.6176	0.0357
2	0.6826	0.9447	0.9225	1.0000	0.5151	0.3633	0.6176	0.0357
3	0.6826	0.9447	0.9225	1.0000	0.5151	0.3633	0.6176	0.0357
4	0.6826	0.9447	0.9225	1.0000	0.5151	0.3633	0.6176	0.0357
Planning horizon R								
80	0.6822	0.9446	0.9225	1.0000	0.5145	0.3629	0.6173	0.0356
100	0.6826	0.9447	0.9225	1.0000	0.5151	0.3633	0.6176	0.0357
120	0.6826	0.9447	0.9225	1.0000	0.5151	0.3633	0.6176	0.0357
140	0.6826	0.9447	0.9225	1.0000	0.5151	0.3633	0.6176	0.0357
160	0.6826	0.9447	0.9225	1.0000	0.5151	0.3633	0.6176	0.0357
Discount rate ρ								
0.01	0.6826	0.9447	0.9225	1.0000	0.5151	0.3633	0.6176	0.0357
0.02	0.6826	0.9447	0.9225	1.0000	0.5151	0.3633	0.6176	0.0357
0.03	0.6822	0.9446	0.9225	1.0000	0.5147	0.3638	0.6174	0.0357
0.04	0.6776	0.9439	0.9225	1.0000	0.5094	0.3630	0.6148	0.0356
Expectations								
Belief responsiveness β								
0	0.4544	0.4928	0.5000	0.5000	0.4544	0.4928	0.5000	0.5000
1	0.5453	0.8534	0.6954	0.9901	0.4729	0.4267	0.5330	0.2659
2	0.6231	0.9228	0.8390	0.9999	0.4932	0.3776	0.5721	0.1030
3	0.6826	0.9447	0.9225	1.0000	0.5151	0.3633	0.6176	0.0357
4	0.7255	0.9551	0.9645	1.0000	0.5385	0.3989	0.6692	0.0149
Memory parameter η								
0	0.6891	0.9461	0.9265	1.0000	0.5198	0.3562	0.6208	0.0306
0.2	0.6875	0.9458	0.9255	1.0000	0.5186	0.3580	0.6200	0.0318
0.4	0.6847	0.9452	0.9238	1.0000	0.5166	0.3610	0.6187	0.0339
0.6	0.6794	0.9440	0.9205	1.0000	0.5127	0.3666	0.6160	0.0385
0.8	0.6632	0.9404	0.9103	1.0000	0.5007	0.3809	0.6078	0.0545
1	0.0924	0.0756	0.0724	0.0724	0.0924	0.0756	0.0724	0.0724

(*continued on next page*)

Table A1 (*continued*)

	Full commitment ($c = 1$)				Poor commitment ($c = 0.3$)			
Parameter	κ in 2040	κ in 2060	n in 2040	n in 2060	κ in 2040	κ in 2060	n in 2040	n in 2060
Policy								
Initial tax τ_{-1}								
0.1	0.7024	0.9485	0.9325	1.0000	0.5521	0.4514	0.6701	0.0596
0.12	0.6947	0.9470	0.9287	1.0000	0.5374	0.4123	0.6495	0.0475
0.14	0.6867	0.9455	0.9246	1.0000	0.5226	0.3785	0.6283	0.0390
0.16	0.6784	0.9439	0.9203	1.0000	0.5076	0.3490	0.6068	0.0328
0.18	0.6700	0.9422	0.9158	1.0000	0.4927	0.3231	0.5852	0.0283
0.2	0.6613	0.9405	0.9110	1.0000	0.4790	0.300	0.5639	0.0247
Initial announced tax rate $\bar{\tau}_0$								
0.1	0.5203	0.8817	0.7336	0.9988	0.4509	0.5769	0.5735	0.5045
0.15	0.6205	0.9257	0.8513	1.0000	0.5004	0.4891	0.6090	0.1865
0.2	0.6826	0.9447	0.9225	1.0000	0.5151	0.3633	0.6176	0.0357
0.25	0.7292	0.9554	0.9611	1.0000	0.5213	0.2732	0.6091	0.0069
0.3	0.7638	0.9621	0.9809	1.0000	0.5200	0.2129	0.5852	0.0014
0.35	0.7892	0.9666	0.9907	1.0000	0.5121	0.1712	0.5466	0.0003
Tax target growth rate $\bar{g}_\tau$								
0.005	0.1517	0.5766	0.6093	0.7074	0.1317	0.4710	0.5040	0.5706
0.01	0.5380	0.8018	0.7069	0.9240	0.4537	0.6031	0.5672	0.6650
0.015	0.6241	0.9105	0.8212	0.9976	0.5028	0.6135	0.6054	0.5877
0.02	0.6826	0.9447	0.9225	1.0000	0.5151	0.3633	0.6176	0.0357
0.025	0.7322	0.9572	0.9797	1.0000	0.5154	0.1782	0.5841	0.0000
0.03	0.7673	0.9634	0.9974	1.0000	0.4939	0.1084	0.4573	0.0000
Transition risk index parameter a								
0.5	0.6826	0.9447	0.9225	1.0000	0.6401	0.9304	0.8650	1.0000
1	0.6826	0.9447	0.9225	1.0000	0.5959	0.9022	0.7874	0.9994
1.5	0.6826	0.9447	0.9225	1.0000	0.5533	0.7540	0.7005	0.7478
2	0.6826	0.9447	0.9225	1.0000	0.5151	0.3633	0.6176	0.0357
2.5	0.6826	0.9447	0.9225	1.0000	0.4823	0.2574	0.5466	0.0161
3	0.6826	0.9447	0.9225	1.0000	0.4548	0.2084	0.4890	0.0107

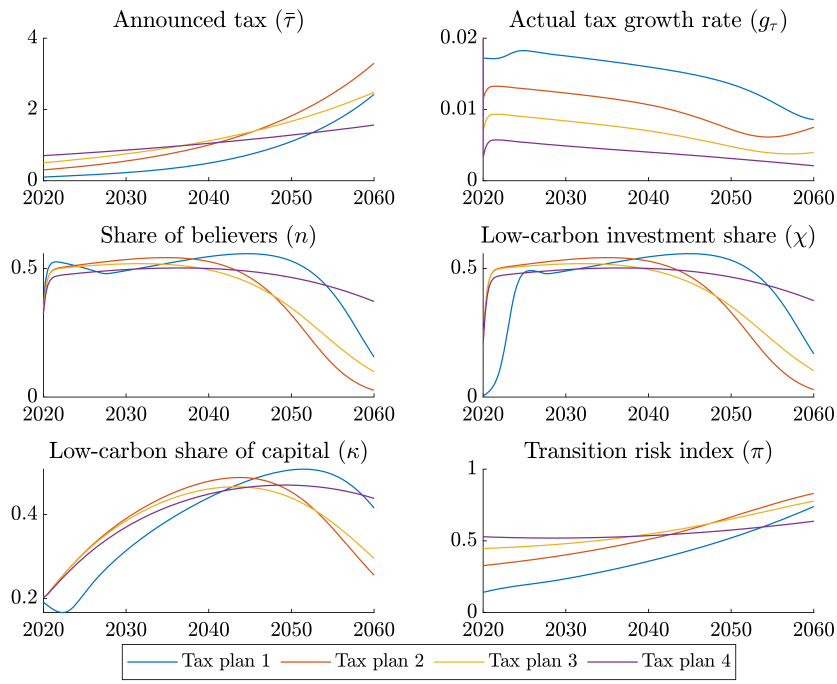
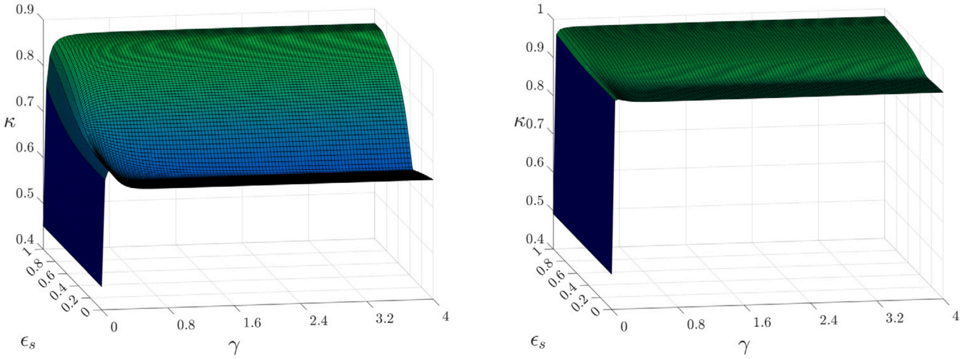
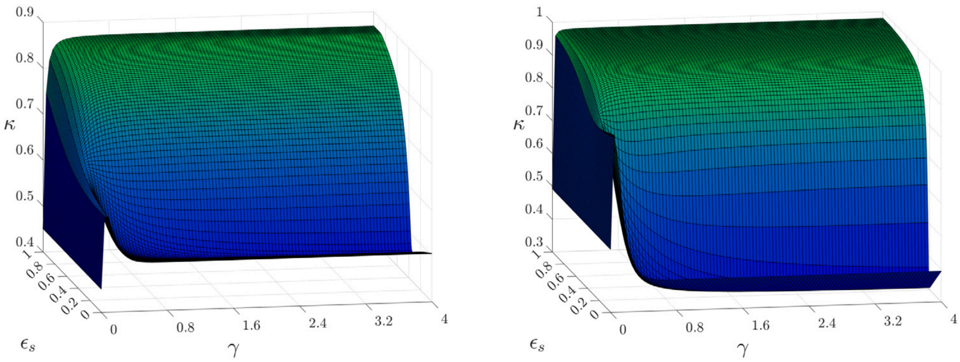


Fig. A2. Evolution over time of selected variables under various tax plans, assuming weak commitment ($c = 0.3$).



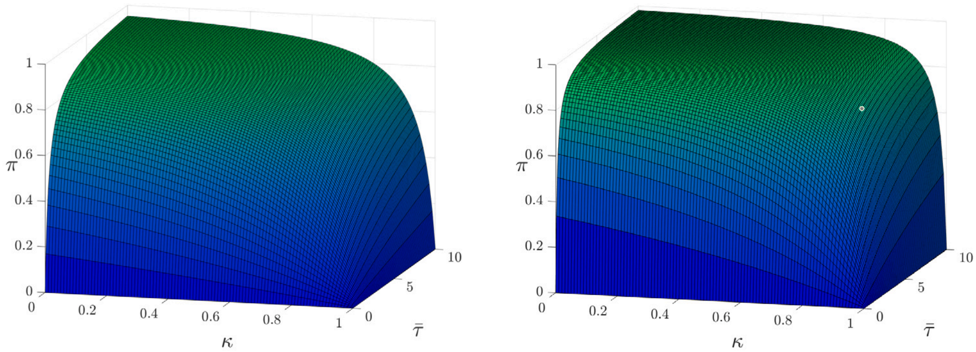
(a) 2040 (b) 2060

Fig. A3. Low-carbon capital share κ as a function of sceptics' discounting of the tax target growth rate ϵ_s and investment responsiveness γ, under $c = 1$, in (a) 2040 and (b) 2060. All other parameters at their default value (Table 1).



(a) 2040 (b) 2060

Fig. A4. Low-carbon capital share κ as a function of sceptics' discounting of the tax target growth rate ϵ_s and investment responsiveness γ, under $c = 0.3$, in (a) 2040 and (b) 2060. All other parameters at their default value (Table 1).



(a) $a = 2$ (b) $a = 5$

Fig. A5. Transition risk index π as a function of κ and $\bar{\tau}$, for two distinct levels of a.

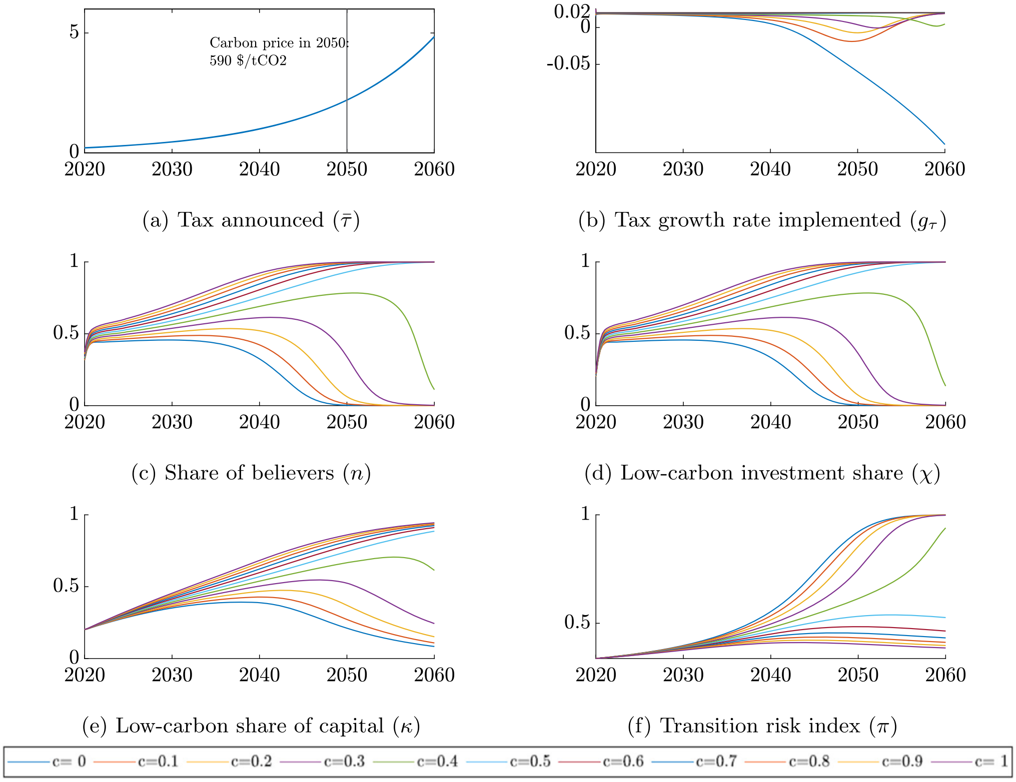
Appendix D. Transition risk index function: graphical representation and robustness

First, let us provide a graphical representation of the transition risk index function (see Fig. A5).

Second, we present the transition path of the model under the assumption of a logistic transition risk index function, defined as

$$\pi_t = \frac{1}{1 + \exp\left(-a[(1-\kappa_t)\bar{\tau}_t - \pi_0]\right)}, \tag{41}$$

where π_0 is the inflection point of the function and is calibrated here to 0.5. As shown in Fig. A6, the dynamics of the model are qualitatively similar to the baseline version of the model.



(a) Tax announced ($\bar{\tau}$)

(b) Tax growth rate implemented (g_τ)

(c) Share of believers (n)

(d) Low-carbon investment share (χ)

(e) Low-carbon share of capital (κ)

(f) Transition risk index (π)

Fig. A6. Evolution over time of selected variables under various levels of commitment (c), assuming a logistic transition risk index function.

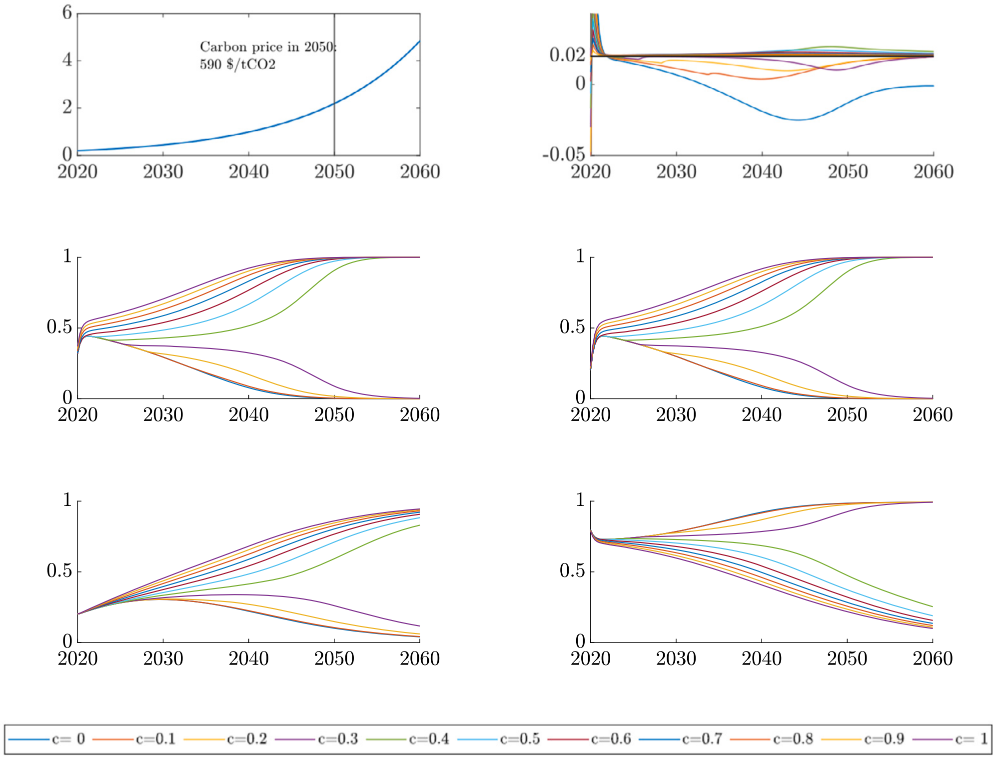


Fig. A7. Evolution over time of selected variables under various levels of commitment (c), under the transition risk index function specified in Equation (42).

Finally, we implement an alternative formulation of the transition risk index function, in which the policy-maker considers the alignment of firms' long-term tax expectations with the planned tax target, $\bar{\tau}$, rather than the tax target itself as in Equation (10). In practice, firms' expectations are elicited through methods such as surveys of professional forecasters. The rationale behind this revised transition risk index is that higher tax expectations among firms indicate greater preparedness for the planned policy, thereby reducing the transition risks associated with its implementation. Formally, the policy-maker considers the ratio between the discounted sum of the tax target schedule and the discounted sum of firms' expectations, as follows:

$$\pi_t = 1 - \frac{1}{1 + a(1-\kappa_t)\frac{\sum_r^R (D^r \bar{\tau}_{t+r})}{\sum_r^R (D^r (n_{t-1} E_{b,t+r} + (1-n_{t-1}) E_{s,t+r}))}}, \tag{42}$$

where $n_{t-1} E_{b,t+r} + (1 - n_{t-1}) E_{s,t+r}$ is the weighted average of firms' long-term tax expectations. The weight is the believers' and sceptics' shares in the population in the previous period, as the future shares are unknown.

The dynamics of this version of the model are shown in Fig. A7. As illustrated, this specification of the transition risk index does not alter the main qualitative results of the paper. Indeed, under low levels of commitment, the low-carbon transition is slowed down and even fails, because of the credibility loss in which the policy-maker incurs.

References

Acemoglu, D., Aghion, P., Bursztyn, L., Hemous, D., 2012. The environment and directed technical change. Am. Econ. Rev. 102, 131–166.

Afrouzi, H., Kwon, S.Y., Landier, A., Ma, Y., Thesmar, D., 2023. Overreaction in expectations: evidence and theory*. Q. J. Econ. 138, 1713–1764.

Annicchiarico, B., Carattini, S., Fischer, C., Heutel, G., 2022a. Business Cycles and Environmental Policy: A Primer. Environmental and Energy Policy and the Economy, vol. 3. The University of Chicago Press, pp. 221–253.

Annicchiarico, B., Di Dio, F., Diluiso, F., 2022b. Climate Actions, Market Beliefs and Monetary Policy. CEIS Working Paper 535. Centre for Economic and International Studies, Rome.

Anufriev, M., Assenza, T., Hommes, C., Massaro, D., 2013. Interest rate rules and macroeconomic stability under heterogeneous expectations. Macroecon. Dyn. 17, 1574–1604. Cambridge University Press.

Ardia, D., Bluteau, K., Boudt, K., Inghelbrecht, K., 2022. Climate change concerns and the performance of green vs. brown stocks. Manag. Sci. INFORMS.

Assenza, T., Heemeijer, P., Hommes, C.H., Massaro, D., 2021. Managing self-organization of expectations through monetary policy: a macro experiment. J. Monet. Econ. 117, 170–186.

Bakkensen, L.A., Barrage, L., 2017. Flood Risk Belief Heterogeneity and Coastal Home Price Dynamics: Going Under Water?.

Barnett, M., Brock, W., Hansen, L.P., 2020. Pricing uncertainty induced by climate change. Rev. Financ. Stud. 33, 1024–1066.

Barradale, M.J., 2014. Investment under uncertain climate policy: a practitioners' perspective on carbon risk. Energy Policy 69, 520–535.

Battiston, S., Monasterolo, I., Riahi, K., van Ruijven, B.J., 2021. Accounting for finance is key for climate mitigation pathways. Science 372, 918–920.

BEIS, 2021. Valuation of greenhouse gas emissions: for policy appraisal and evaluation. UK Department for Business, Energy, & Industrial Strategy, London. Policy paper.

Berestycki, C., Carattini, S., Dechezleprêtre, A., Kruse, T., 2022. Measuring and assessing the effects of climate policy uncertainty. Tech. rep. OECD, Paris.

Bilias, Y., Georgarakos, D., Haliassos, M., 2010. Portfolio inertia and stock market fluctuations. J. Money Credit Bank. 42, 715–742.

Bordalo, P., Gennaioli, N., Shleifer, A., 2022. Overreaction and diagnostic expectations in macroeconomics. J. Econ. Perspect. 36, 223–244.

Boswijk, H.P., Hommes, C.H., Manzan, S., 2007. Behavioral heterogeneity in stock prices. J. Econ. Dyn. Control 31, 1938–1970.

Brock, W.A., Hommes, C.H., 1997. A rational route to randomness. Econometrica 65, 1059–1096.

Brock, W.A., Hommes, C.H., 1998. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. J. Econ. Dyn. Control 22, 1235–1274.

Brock, W.A., Hommes, C.H., Wagener, F.O.O., 2005. Evolutionary dynamics in markets with many trader types. J. Math. Econ. 41, 7–42.

Cafferata, A., Dávila-Fernández, M.J., Sordi, S., 2021. Seeing what can(not) be seen: confirmation bias, employment dynamics and climate change. J. Econ. Behav. Organ. 189, 567–586.

Cahen-Fourot, L., Campiglio, E., Daumas, L., Miess, M.G., Yardley, A.A., 2022. Stranding Ahoy? Heterogeneous Transition Beliefs and Capital Investment Choices.

Cai, Y., Lontzek, T.S., 2018. The Social Cost of Carbon with Economic and Climate Risks. Journal of Political Economy 127, 2684–2734. The University of Chicago Press.

Campiglio, E., 2016. Beyond carbon pricing: the role of banking and monetary policy in financing the transition to a low-carbon economy. Ecol. Econ. 121, 220–230.

Campiglio, E., Dietz, S., Venmans, F., 2022. Optimal Climate Policy as If the Transition Matters. CESifo Working Paper 10139. CESifo, Munich.

Campiglio, E., van der Ploeg, F., 2022. Macrofinancial risks of the transition to a low-carbon economy. Rev. Environ. Econ. Policy, 000.

Chiarella, C., He, X.-Z., Zwinkels, R.C.J., 2014. Heterogeneous expectations in asset pricing: empirical evidence from the S&P500. J. Econ. Behav. Organ. 105, 1–16.

Crowley, K., 2017. Up and down with climate politics 2013–2016: the repeal of carbon pricing in Australia. WIREs Climate Change 8, e458.

Daniel, K.D., Litterman, R.B., Wagner, G., 2019. Declining CO2 price paths. Proc. Natl. Acad. Sci. 116, 20886–20891.

De Grauwe, P., Macchiarelli, C., 2015. Animal spirits and credit cycles. J. Econ. Dyn. Control 59, 95–117.

Dechezleprêtre, A., Fabre, A., Kruse, T., Planterose, B., Sanchez Chico, A., Stantcheva, S., 2022. Fighting Climate Change: International Attitudes Toward Climate Policies.

Doda, B., 2014. Evidence on business cycles and CO2 emissions. J. Macroecon. 40, 214–227.

Douenne, T., Fabre, A., 2020. French attitudes on climate change, carbon taxation and other climate policies. Ecol. Econ. 169, 106496.

Douenne, T., Hummel, A.J., Pedroni, M., 2023. Optimal Fiscal Policy in a Climate-Economy Model with Heterogeneous Households.

Dunz, N., Naqvi, A., Monasterolo, I., 2021. Climate sentiments, transition risk, and financial stability in a stock-flow consistent model. J. Financ. Stab. 54, 100872.

Dávila-Fernández, M.J., Sordi, S., 2020. Attitudes towards climate policies in a macrodynamic model of the economy. Ecol. Econ. 169, 106319.

EEA, 2021a. National emissions reported to the UNFCCC and to the EU Greenhouse Gas Monitoring Mechanism. Tech. rep. European Environment Agency.

EEA, 2021b. Share of energy consumption from renewable sources in Europe. Tech. rep. European Environment Agency.

EEA, 2021c. Trends and projections in the EU ETS in 2021 the EU Emissions Trading System in numbers. Tech. rep. European Environment Agency.

EIA, 2021. Electric Power Monthly - U.S. Energy Information Administration (EIA). Tech. rep. Energy Information Administration.

Ellen, S. t., Zwinkels, R.C.J., 2010. Oil price dynamics: a behavioral finance approach with heterogeneous agents. Energy Econ. 32, 1427–1434.

Engle, R.F., Giglio, S., Kelly, B., Lee, H., Stroebel, J., 2020. Hedging climate change news. Rev. Financ. Stud. 33, 1184–1216.

Eurostat, 2021. Electricity and heat statistics. Tech. rep. Eurostat, Luxembourg.

Evans, G.W., Honkapohja, S., 2006. Monetary policy, expectations and commitment. Scand. J. Econ. 108, 15–38.

Evans, G.W., Honkapohja, S., 2012. Learning and Expectations in Macroeconomics. Princeton University Press.

Fankhauser, S., Smith, S.M., Allen, M., Axelsson, K., Hale, T., Hepburn, C., Kendall, J.M., Khosla, R., Lezaun, J., Mitchell-Larson, E., Obersteiner, M., Rajamani, L., Rickaby, R., Seddon, N., Wetzer, T., 2022. The meaning of net zero and how to get it right. Nat. Clim. Change 12, 15–21.
Farfan, J., Breyer, C., 2017. Aging of European power plant infrastructure as an opportunity to evolve towards sustainability. Int. J. Hydrog. Energy 42, 18081–18091.
Fay, M., Hallegatte, Stephane, Vogt-Schilb, Adrien, Rozenberg, Julie, 2015. Decarbonizing Development: Three Steps to a Zero-Carbon Future. Tech. rep. World Bank, Washington, DC.
France Stratégie, 2019. The value for climate action. A shadow price of carbon for evaluation of investments and public policies. France Stratégie.
Fried, S., Novan, K., Peterman, W.B., 2022. Climate policy transition risk and the macroeconomy. Eur. Econ. Rev. 147, 104174.
Fuss, S., Szolgayova, J., Obersteiner, M., Gusti, M., 2008. Investment under market and climate policy uncertainty. Appl. Energy 85, 708–721.
Galanis, G., Kollias, I., Leventidis, I., Lustenhouwer, J., 2022a. Generalizing Heuristic Switching Models. Working Papers, number: 0715. University of Heidelberg, Department of Economics.
Galanis, G., Ricchiuti, G., Tippet, B., 2022b. The Global Political Economy of a Green Transition. Working Papers - Economics, wp2022_22.rdf. Universita' degli Studi di Firenze, Dipartimento di Scienze per l'Economia e l'Impresa.
GCP, 2022. Global Carbon Budget 2022. Tech. rep. Global Carbon Project.
Geels, F.W., Berkhout, F., van Vuuren, D.P., 2016. Bridging analytical approaches for low-carbon transitions. Nat. Clim. Change 6, 576–583.
Gibson, M., Mullins, J.T., 2020. Climate Risk and Beliefs in New York Floodplains. Journal of the Association of Environmental and Resource Economists, vol. 7. The University of Chicago Press, pp. 1069–1111.
Gigerenzer, G., Brighton, H., 2009. Homo heuristicus: why biased minds make better inferences. Top. Cogn. Sci. 1, 107–143.
Giglio, S., Maggiori, M., Stroebel, J., Tan, Z., Utkus, S., Xu, X., 2023. Four Facts About ESG Beliefs and Investor Portfolios.
Gollier, C., 2022. The cost-efficiency carbon pricing puzzle. TSE Working Papers 952. Toulouse School of Economics (TSE), pp. 18–952.
Golosov, M., Hassler, J., Krusell, P., Tsyvinski, A., 2014. Optimal taxes on fossil fuel in general equilibrium. Econometrica 82, 41–88.
Gomme, P., Ravikumar, B., Rupert, P., 2011. The return to capital and the business cycle. Rev. Econ. Dyn. 14, 262–278.
Gomme, P., Rupert, P., 2007. Theory, measurement and calibration of macroeconomic models. J. Monet. Econ. 54, 460–497.
Hale, T., Kuramochi, T., Lang, J., Mapes, B., Smith, S., Aiyer, R., Black, R., Boot, M., Chalkley, P., Hans, F., Hay, N., Hsu, A., Höhne, N., Mooldijk, S., Walsh, T., 2022. Net Zero Tracker. Tech. rep. Energy and Climate Intelligence Unit, Data-Driven EnviroLab, NewClimate Institute, Oxford Net Zero, Oxford.
Helm, D., Hepburn, C., Mash, R., 2003. Credible carbon policy. Oxf. Rev. Econ. Policy 19, 438–450.
Hommes, C., 2021. Behavioral and experimental macroeconomics and policy analysis: a complex systems approach. J. Econ. Lit. 59, 149–219.
Hommes, C., Huang, H., Wang, D., 2005. A robust rational route to randomness in a simple asset pricing model. J. Econ. Dyn. Control 29, 1043–1072.
Hommes, C., Lustenhouwer, J., 2019a. Inflation targeting and liquidity traps under endogenous credibility. J. Monet. Econ. 107, 48–62.
Hommes, C., Lustenhouwer, J., 2019b. Managing unanchored, heterogeneous expectations and liquidity traps. J. Econ. Dyn. Control 101, 1–16.
IEA, 2019a. CO2 emissions – Global Energy Review 2019 – Analysis. Tech. rep. International Energy Agency.
IEA, 2019b. Emissions – Global Energy & CO2 Status Report 2019 – Analysis. Tech. rep. International Energy Agency.
IEA, 2020a. Development of CO2 emission intensity of electricity generation in selected countries, 2000-2020 – Charts – Data & Statistics. Tech. rep. International Energy Agency.
IEA, 2020b. Levelised Cost of Electricity Calculator. Tech. rep. International Energy Agency.
IEA, 2020c. Projected Costs of Generating Electricity 2020. Tech. rep. International Energy Agency.
IPCC, 2021. Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. In: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S.L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M.I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J.B.R., Maycock, T.K., Waterfield, T., Yelekçi, O., Yu, R., Zhou, B. (Eds.). Cambridge University Press.
IPCC, 2022. Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Intergovernmental Panel on Climate Change.
IRENA, 2021. Renewable Power Generation Costs in 2020. Tech. rep. International Renewable Energy Agency.
Känzig, D.R., 2021. The Unequal Economic Consequences of Carbon Pricing. SSRN Scholarly Paper ID 3786030 Social Science Research Network, Rochester, NY.
Känzig, D.R., 2023. The Unequal Economic Consequences of Carbon Pricing.
Kollenbach, G., 2017. On the optimal accumulation of renewable energy generation capacity. J. Econ. Dyn. Control 77, 157–179.
Konc, T., Drews, S., Savin, I., van den Bergh, J.C.J.M., 2022. Co-dynamics of climate policy stringency and public support. Glob. Environ. Change 74, 102528.
Krueger, P., Sautner, Z., Starks, L.T., 2020. The Importance of Climate Risks for Institutional Investors. The Review of Financial Studies, vol. 33. Oxford Academic, pp. 1067–1111.
Kukacka, J., Barunik, J., 2017. Estimation of financial agent-based models with simulated maximum likelihood. J. Econ. Dyn. Control 85, 21–45.
Kydland, F.E., Prescott, E.C., 1977. Rules Rather than Discretion: The Inconsistency of Optimal Plans. J. Polit. Econ. 85, 473–491. The University of Chicago Press.
Lamperti, F., 2018. Empirical validation of simulated models through the GSL-div: an illustrative application. J. Econ. Interact. Coord. 13, 143–171.
Lamperti, F., Dosi, G., Napoletano, M., Roventini, A., Sapio, A., 2020. Climate change and green transitions in an agent-based integrated assessment model. Technol. Forecast. Soc. Change 153, 119806.
Lemoine, D., 2017. Green expectations: current effects of anticipated carbon pricing. Rev. Econ. Stat. 99, 499–513.
Lera, S., Sornette, D., 2017. Secular Bipolar Growth Rate of the Real US GDP per Capita: Implications for Understanding Past and Future Economic Growth. SSRN Scholarly Paper ID 2703882 Social Science Research Network, Rochester, NY.
Mankiw, N.G., Reis, R., Wolfers, J., 2003. Disagreement About Inflation Expectations. In: NBER Macroeconomics Annual, vol. 18. The University of Chicago Press, publisher, pp. 209–248.
Manski, C.F., McFadden, D., 1981. Structural Analysis of Discrete Data with Econometric Applications. MIT Press, Cambridge, MA.
McCulloch, N., Natalini, D., Hossain, N., Justino, P., 2022. An exploration of the association between fuel subsidies and fuel riots. World Dev. 157, 105935.
McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Ed.), Frontiers in Econometrics. Academic Press, New York, pp. 105–142.
Mercure, J.-F., 2015. An age structured demographic theory of technological change. J. Evol. Econ. 25, 787–820.
Metcalf, G.E., Stock, J.H., 2020. Measuring the macroeconomic impact of carbon taxes. AEA Pap. Proc. 110, 101–106.
Metcalf, G.E., Stock, J.H., 2023. The macroeconomic impact of Europe's carbon taxes. Am. Econ. J. Macroecon.
Nemet, G.F., Jakob, M., Steckel, J.C., Edenhofer, O., 2017. Addressing policy credibility problems for low-carbon investment. Glob. Environ. Change 42, 47–57.
Noailly, J., Nowzohour, L., Van Den Heuvel, M., 2021. Heard the News? Environmental Policy and Clean Investments. Graduate Institute of International and Development Studies. CIES Research Paper 70 Centre for International Environmental Studies, Geneva.
Nordeng, A., Kolos, M., Lin, Y., Liao, C., Rihel, A.S., Zelljadt, L., 2021. Refinitiv Carbon Market Survey 2021. Tech. rep., Refinitiv.
Nordhaus, W.D., 2017. Revisiting the social cost of carbon. Proc. Natl. Acad. Sci. 114, 1518–1523.
Peszko, G., van der Mensbrugghe, D., Golub, A., Ward, J., Zenghelis, D., Marijs, C., Schopp, A., Rogers, J.A., Midgley, A., 2020. Diversification and Cooperation in a Decarbonizing World: Climate Strategies for Fossil Fuel-Dependent Countries. World Bank, Washington, DC. accepted: 2020–06-30T19:11:44Z.
Pollitt, M.G., Shaorshadze, I., 2013. The role of behavioural economics in energy and climate policy. In: Handbook on Energy and Climate Change. Edward Elgar Publishing, pp. 523–546. section: Handbook on Energy and Climate Change.
Refinitiv, 2019. Refinitiv Carbon Market Survey 2019. Tech. rep. Refinitiv.

Riahi, K., Bertram, C., Huppmann, D., Rogelj, J., Bosetti, V., Cabardos, A.-M., Deppermann, A., Drouet, L., Frank, S., Fricko, O., Fujimori, S., Harmsen, M., Hasegawa, T., Krey, V., Luderer, G., Paroussos, L., Schaeffer, R., Weitzel, M., van der Zwaan, B., Vrontisi, Z., Longa, F.D., Després, J., Fosse, F., Fragkiadakis, K., Gusti, M., Humpenöder, F., Keramidas, K., Kishimoto, P., Kriegler, E., Meinshausen, M., Nogueira, L.P., Oshiro, K., Popp, A., Rochedo, P.R.R., Ünlü, G., van Ruijven, B., Takakura, J., Tavoni, M., van Vuuren, D., Zakeri, B., 2021. Cost and attainability of meeting stringent climate targets without overshoot. Nat. Clim. Change 11, 1063–1069.

Salle, I., Yıldızoğlu, M., Sénégas, M.-A., 2013. Inflation targeting in a learning economy: an ABM perspective. Econ. Model. 34, 114–128.

Sendstad, L.H., Hagspiel, V., Mikkelsen, W.J., Ravndal, R., Tveitstøl, M., 2022. The impact of subsidy retraction on European renewable energy investments. Energy Policy 160, 112675.

Shogren, J.F., Taylor, L.O., 2008. On Behavioral-Environmental Economics. Review of Environmental Economics and Policy, vol. 2. The University of Chicago Press, pp. 26–44.

Stiglitz, J.E., Stern, N., Duan, M., Edenhofer, O., Giraud, G., Heal, G.M., la Rovere, E.L., Morris, A., Moyer, E., Pangestu, M., Shukla, P.R., Sokona, Y., Winkler, H., 2017. Report of the High-Level Commission on Carbon Prices. Tech. rep.

Torren-Peraire, D., Savin, I., van den Bergh, J., 2023. An Agent-Based Model of Cultural Change for a Low-Carbon Transition. SSRN Scholarly Paper 4344035. Rochester, NY.

Train, K.E., 2009. Discrete Choice Methods with Simulation, 2 ed. Cambridge University Press, Cambridge.

Tran, T.T.D., Smith, A.D., 2018. Incorporating performance-based global sensitivity and uncertainty analysis into LCOE calculations for emerging renewable energy technologies. Appl. Energy 216, 157–171.

Tsur, Y., Zemel, A., 2011. On the dynamics of competing energy sources. Automatica 47, 1357–1365.

Urpelainen, J., Van de Graaf, T., 2018. United States non-cooperation and the Paris agreement. Climate Policy 18, 839–851.

van den Bremer, T.S., van der Ploeg, F., 2021. The risk-adjusted carbon price. Am. Econ. Rev. 111, 2782–2810.

van der Ploeg, F., Rezai, A., 2020. The risk of policy tipping and stranded carbon assets. J. Environ. Econ. Manag. 100, 102258.

van der Ploeg, F., Rezai, A., 2021. Optimal carbon pricing in general equilibrium: temperature caps and stranded assets in an extended annual DSGE model. J. Environ. Econ. Manag. 110, 102522.

Vogt-Schilb, A., Meunier, G., Hallegatte, S., 2012. How Inertia and Limited Potentials Affect the Timing of Sectoral Abatements in Optimal Climate Policy. Policy Research Working Papers. The World Bank.

Vona, F., Marin, G., Consoli, D., Popp, D., 2018. Environmental Regulation and Green Skills: An Empirical Exploration. J. Assoc. of Environ. Res. Econ. 5, 713–753. The University of Chicago Press.

Waisman, H., Guivarch, C., Grazi, F., Hourcade, J.C., 2012. The Imaclim-R model: infrastructures, technical inertia and the costs of low carbon futures under imperfect foresight. Clim. Change 114, 101–120.

World Bank, 2022. State and trends of carbon pricing 2022. Tech. rep. The World Bank, Washington, DC.

Zeppini, P., 2015. A discrete choice model of transitions to sustainable technologies. J. Econ. Behav. Organ. 112, 187–203.