Waiting for the transition: The role of expectations in the decarbonisation of the electricity sector[1]

Unfinished work in progress version – please do not quote or circulate!

Louison Cahen-Fourot[a], Emanuele Campiglio[a,b], Louis Daumas[a,c], Michael Gregor Miess[a,d,e], Andrew Yardley[a]

[a] *Institute for Ecological Economics, Vienna University of Economics and Business, Vienna, Austria*
[b] *Grantham Research Institute, London School of Economics*
[c] *Centre International de Recherche en Environnement et Développement*
[d] *Environment Agency Austria, Vienna*
[e] *Institute for Advanced Studies, Vienna, Austria*

Abstract

We develop a macroeconomic model of the low-carbon transition of the electricity sector in a closed economy in discrete time. The structure of our modelling framework is based on the representation of the physical and financial stocks and flows of heterogeneous macroeconomic sectors. Our approach extends the existing literature in macroeconomic modelling of transition dynamics by developing a novel bounded rational but forward-looking expectation formation mechanism based on the notion of *fictional expectations*. This allows us to combine long-term views on system change with more short-term and profit-oriented investment choices in a coherent framework. For investment decisions, we incorporate discrete choice (DC) theory based on probabilistic distributions of future utilization of capital stocks and profits to derive choices of technology. We apply this DC framework to derive investment choices for heterogeneous capital goods today that have different environmental implications for the future and are subject to path dependency. Using our approach, we can provide insights on a wide range of issues that concern transition dynamics. This includes a novel analysis on how stranding of physical assets can occur as a phenomenon resulting from coordination problems on a macroeconomic level due to dissent and different beliefs about the future. Further potential applications of our framework are numerous: we can create taxonomies of transition dynamics following different levels of dissent, determine the role of opinion conflict for the low-carbon transition, relate energy demand growth to dissent, and finally simulate different forms of policy interventions including a carbon price in relation to different levels of dissent.

1. Introduction

The transition to a low-carbon electricity system progresses slower than necessary to meet the targets of the Paris agreement (IAEA, 2019; DNVGL, 2020; ETC, 2018). The same applies for the required electrification of all economic sectors worldwide, which is necessary to reach net carbon emissions by 2050. This large-scale electrification will require a 4-5 fold increase of global electricity production by mid-century (ETC, 2018). Certainly, recent encouraging market signals, like the increasing competitivness of renewables, have led to high upward trends of investments in renewable electricity production (IRENA, 2020; IEA, 2020). Yet, 2019 data from the BP Statistical Review indicate that renewable electricity (including hydroelectricity, solar, wind, geothermal and biofuels) still amounts to less than a third (26%) of worldwide electricity generation, while carbon-free electricity[2] makes up for 37% of total electricity. Conversely, the share of

[1]Thankfully supported by the Jubiläumsfonds of the Oesterreichische Nationalbank (Austrian Central Bank, Anniversary Fund, project number: 17641).

[2]I.e. electricity produced from renewable, nuclear, and other non-carbon energy sources.

renewables in worldwide primary energy consumption remains very low, at 5% (BP Review, 2019). As such, waiting for the transition to a renewable energy system seems akin to waiting for Godot (Beckett, 1956)—i.e. waiting for an event that is expected soon but never actually occurs.

The way to effectively carry out the low-carbon transition is still not completely carved out or agreed upon by both scientists and policy makers, but certainly it will involve both governmental and market actors. In particular, achieving net-zero will require aligning agents' expectations of possible future developments. The transition is indeed fraught with an important number of uncertainties, for instance on the timing of low-carbon policy. Further, the low-carbon transition will also entail sweeping technological, but also societal and lifestyle changes that can hardly be predicted, because it includes novelty, change, creativity (Shackle, 1970). Hence the low-carbon transition seems radically uncertain regarding its precise shape and modalities. Radical, fundamental or "deep" uncertainty describes this state of fact in which it is not possible to attribute reliable probability distributions to future outcomes. This inability is not due to a lack of coordination or cognitive ability. Rather, it either flows from the phenomenon under scrutiny, or the horizon of analysis simply put us in a situation of near-total ignorance. In other words, radical uncertainty is an ontological assumption about the nature of the world (Davidson, 1991), in which knowledge is conventional and subjective, and economic structures stabilised by institutions, norms and conventions. Expectations are part and parcel of this conventional and subjective knowledge, as the only possible way conceivable to the human mind to fill the void created by fundamental uncertainty through imagining—rather than predicting—possibles futures. Agents do not form their expectations in a historical and institutional vacuum. Imagined futures are shaped by the historical experience and the narrower set of relevant behaviours delimited by institutions.

Economics has long emphasised the role of agents' expectations in driving real-world outcomes (Lachmann, 1943). However, the embedding of expectations within a broader social context and its narrative aspects have so far entered economic thinking only to a limited extent. Common approaches to expectations in economics either assume away fundamental uncertainty or do not allow for forward-looking narratives, and they rarely account for broader institutional contexts. As a matter of fact, looking into the long term future under fundamental uncertainty in a transition process poses deep methodological problems to conceive agents' expectations formation and the coordination of these long-term anticipations. Yet, given its far-reaching policy and societal implications, this methodological challenge is worth taking up. In today's world, where decision centres are dispatched, and where private actors have large autonomy, expectations can be contradictory, and be motivated by various economic and political agendas. In a word, discourses and projections into the futures based on beliefs and imagined futures are part of a broader political economy, in which credibility, conviction and sentiment of certainty are key in furthering societal projects in a situation of radical uncertainty. Such narratives can engage agents on peculiar transition paths, more or less virtuous, and with different technological and institutional contents. Hence, there is full room to study how economic agents with conflicting views about the future and with possibly different overarching references engage in a transition towards a low-carbon economy, and how opinion conflicts arising from these views influence transition paths.

The research question of this article is, therefore, to explore under what conditions such opinion conflicts can harm transition dynamics, possibly down to the point of hampering it completely. To this effect, we develop a **system-dynamics model** of the low-carbon transition (Hafner et al., 2020) with forward-looking fictional expectations in discrete time. The model is meant to simulate transition trajectories for the electricity market, over which the energy mix will change through time. Our formulation of the expectation-formation process is the main contribution to the literature. It consists in a proposition for a *middle-ground between adaptive, backward-looking expectations, and forward-looking, perfect-foresight and model-consistent anticipations*, which are the two cases usually found in the literature. While we keep a forward-looking structure, we reject model consistency and perfect foresight by assuming away an infinite expectation horizon. We formalise a finite and rolling planning horizon of a given time length. We also suppose that agents formulate "fictional expectations" or "imagined futures" about some future outcomes, that do not necessarily match the law of motion the modeller herself chose for her framework for a given economic variable. Our agents are therefore boundedly (Simon, 1955), or ecologically rational (Gigerenzer, 2015), in the sense that they use simple indicators, based on their expectations, to guide their actions rather than optimisation procedures.

The remainder of this article is set out as follows. In section 2, we review the existing literature on expectations to further motivate and present our methodological choice. In Section 3, we describe the novel modelling methodology we have developed, including its technical implementation. Section 4 illustrates our choice of scenarios and shows model simulations as applications of our modelling framework. Section 5 concludes and provides an outlook of further research. Moreover, Appendix A provides further details on the electricity sector, Appendix B elaborates on the role of stranding expectations, Appendix C illustrates the probit module in detail, and Appendix D provides information on the calibration of this model to the European Union (EU).

2. Literature Review

In this section we discuss the various approach to expectations in economics and locate our article in this literature.

2.1. Expectations in economics

The literature on expectations formation can be broadly organized in three main branches following their methodological grounding.

Backward-looking expectations rely on simple expectation rules: *(i)* the future is assumed largely identical to the past at all point in time (purely myopic expectations); *(ii)* beliefs are updated according to past expectation mistakes (adaptive expectations); or *(iii)* by projecting past trends (extrapolative behaviour). They leave little room for patterns going beyond the projection of past trends, such as imagining innovative course of events, structural change, or non-linear dynamics. Therefore, backward-looking expectations appear ill-suited to investigate our research question.

Most short and medium-run **rational expectations** (RE) models assume a representative agent with model-consistent expectations about the future. These expectations are true when averaged out over the multiple possible futures and any expectation mistake can be attributed to random shocks that cannot be anticipated around the model's equilibrium (steady state growth) path. Long-run models are usually not interested in short-run fluctuations arising from random shocks: A non-stochastic world is assumed in which the representative agent derives a model-consistent dynamic growth path by optimising present and future anticipated utility. The assumption has been strongly criticised after the Great Financial Crisis (GFC) — mostly on the grounds that RE-based macroeconomic models had conveyed a false sense of economic stability by assuming away "tail-events" (Taleb, 2010). Perfect foresight in models of the low-carbon transition has also raised important criticisms (Stern, 2016). The heuristic interest of rational expectations for our research question looks quite limited as they assume away any kind of radical uncertainty.

Information ambiguity models refer to a situation in which information is not clear enough to derive a stable and firm probability distribution for shocks, due to cognitive limitations or a limited set of observables (Bachmann et al., 2013; Ilut and Saijo, 2020). Other authors interpret Knightian uncertainty as the generation of expectation *intervals* instead of single values, with applications mostly in the field of financial economics (Frydman et al., 2015). The *diagnosis expectation* strand, by contrast, emphasises cognitive biases in explaining deviations from RE (Bordalo et al., 2018, 2019). Yet, most of these endeavours remain attached to a stochastic-shock representation of uncertainty, and only tweak the reaction to such shocks. Moreover, they have remained confined to the study of the business-cycle, and seem therefore ill-suited to a long-run study such as ours.

"Fast and frugal" (Dosi et al., 2017) heuristics used by "ecologically rational" agents (Gigerenzer, 2015), that change their approach to the future according to the performance of these rules-of-thumb. Agents use heuristics (usually forecasting rules of different flavours (Anufriev and Hommes, 2012)) *on top* of the model's behaviour which are then evaluated *ex post* (Hommes, 2018). A difference between *actual* and *perceived* laws of motion results. Within this framework, expectations can then be highly model-inconsistent but can become closer and closer to the rational expectation benchmark. A second approach pits different population of agents using different heuristics, and who can switch heuristics, such that the prevalence of a given rule-of-thumb changes along the model run. Such models allow for waves of optimism and pessimism during which agents under- and overshoot in their expectations, triggering real under- and overreactions (De Long et al., 1990; De Grauwe and Macchiarelli, 2015; De Grauwe, 2012; Lavoie and Daigle, 2011; Franke and Westerhoff, 2017; Fagiolo and Roventini, 2017; Dosi et al., 2019). However, heuristic-based approaches are heavily centered around forecasting techniques, and rarely feature non-linear extrapolation.

All methodologies mentioned up to now do not fit with our purpose for three reasons: forward looking or fundamental uncertainty are assumed away; they are circumscribed to the study of short-term fluctuations; expectations cannot be reduced to a data-acquisition process but include elements of imagination and creativity (Shackle, 1972; Dequech, 1999). A more comprehensive understanding of expectations seems therefore necessary.

2.2. The concept of fictional expectations and its relevance to the low-carbon transition

2.2.1. Defining fictional expectations

We draw inspiration from economic sociology, more precisely from the concepts of *fictional expectations* and *expectation regimes* (Beckert, 2013a, 2016; Boyer, 2018; Beckert and Bronk, 2018). They are sometimes touched upon in interdisciplinary macroeconomic literature (Haldane, 2016; Haldane and Turrell, 2018, 2019; Bronk, 2009). In a context of radical uncertainty, each agent has in mind their own model of the economic system but also an underlying value system and hopes for the future. Such prospective models are dubbed "fictional", in that they can be understood as agent-specific narratives which cannot be considered *ex ante* as true or false ("model consistent" or not). They rather rely on “as if” assumptions and tools to apprehend the future that may go beyond reality as currently observed and used as motifs for engaging in potentially profitable but ultimately incalculable outcomes. On top of this, those “fictions” are embedded within a broader social and institutional context, and ultimately depend on cultural frames, dominant theories, the stratification structures of a society, social networks (Beckert, 2013b). These imaginaries are indeed are then shared, and discussed amongst agents. Consequently, different fictions are more or less credible, shared by different people from distinct groups, and more or less “enforceable”, that is, subject to relations of power between different agents or groups of agents. Agents, in turn, formulate their own personal narrative, more or less close to the one they autonomously believe in, know of, or are convinced by through others’ discourses, opening the way for opinion conflicts, but also power relations in the realm of ideology, representations, and so on. It should not be understood, however, that this concept is only one of “guesstimates” and “fantasies”. Fictions are shaped and told through collective devices, ranging from a convincing speech to potential fundraisers to a complex and sophisticated econometric model yielding some forecasts, which are given more or less ex ante credibility. Nowadays’ trust in figures and “hard” proof gives naturally more weight to statistics and quantitative tools than in other periods, when they were less accurate, and therefore “objectively” less reliable. The important point, for Beckert, is that these devices are tools for creating some certainty in a world of perfect uncertainty and at different horizons of analysis. That such devices fall short of reality in some instances is not important as such, but too often deceived expectations can lead to economic disturbances, such as financial crises (Beckert, 2016). We consider the fictional expectation framework as being fruitful for macroeconomic. It provides a sound theoretical basis to endow agents with fictions of different kinds, that can be made as heterogenous as desired. It can also justify the design design exogenous “overarching narratives” in the spirit of Boyer (2018), that can serve as “focal points” for agents, but whose credibility evolves through time, opening the possibility for credibility conflicts over several such narratives. Yet, to the best of our knowledge, despite their apparent ability to trace how agents in an economy deal with fundamental uncertainty, fictional expectations have not yet been applied for expectations formation in formalized aggregate macroeconomic models.

2.2.2. Relevance for the low-carbon transition

We nonetheless hold that fictional expectations are a very relevant framework to study the low-carbon transition. As a matter of fact, the very epistemology of the Intergovernmental Panel on Climate Change (IPCC) has come to rely on societal scenarios going beyond purely technical considerations: the Shared Socioeconomic Pathways (SSPs). These narratives help dealing with the high long-term uncertainty it has to face when producing the Sixth Assessment Report (AR6) on global warming.[3] These SSPs are narratives describing plausible alternative socio-economic developments—including sustainable development, regional rivalry, inequality, fossil-fueled development, and middle-of-the-road development—that are intended to cover wide ranges of uncertainty (Riahi et al., 2017). Many other agencies, such as the International Energy Agency (IEA), the OECD, or even independent think tanks, like the Shift Project (Salomon et al., 2005) in France, propose transition scenarios with different hypotheses, models and scopes than that of the IPCC. Even private actors such as Shell (2017) or ExxonMobil (2019) provide such projections. All differ or even enter in conflict in their technological assumptions and climate ambitions, in particular on the extent of changes necessary to achieve a given climate target. The question is therefore to what extent and under what such diversity of narrative can hamper the coordination over a well-defined transition path, and maybe harm its good course. This paper aims at providing a stylised framework to study this question.

2.2.3. Our transition narrative

As a central narrative we have chosen the S-curve shaped transition dynamics of technological diffusion, which is a well agreed-on empirical regularity for transition processes that is used for modelling purposes (Meade and Towhidul, 2006; Geroski, 2000). In particular for the case of the energy system, different kinds of energy have often experienced S-curve and inverted S-curve trajectories—or both combined in an inverted U-curve trajectory. Such trajectories are well-known in the literature on technological transitions. Figure 1 illustrates this stylized fact: coal, oil and primary electricity clearly follow S-curve and inverted U-curve trajectories since the early years of the Industrial Revolution in England and Wales, France and Germany. Secondly, we cater to expectations of physical assets stranding, which are also a widely agreed-upon prospective view on the future (Campiglio et al., 2018; Caldecott, 2017), where we hinge expectations of stranding on the future capacity utilization rate of high-carbon capital stock.

3. The Model

We build on our prior model of heterogeneous expectations of asset stranding (Cahen-Fourot et al., forthcoming [INSERT REF to ANALYTICAL PAPER]. The model describes an investment process in which investors form expectations of stranding and, on aggregate, allocate a share of investment to low and high-carbon technologies.

Periods are explicitly one-year long, and our simulations are 50-ticks long to represent a transition process. This model figures the competition between two types of capital on the electricity-production market: low- and high-carbon capital. The model core is therefore a structural shift model: The path of future energy demand is known but the precise energy mix will change through time due to changing investment behaviours of economic agents. At any point in time, some agents will invest in high-carbon capital and others will invest in the low-carbon capital. As the model runs, these agents will switch between these two options according to forward-looking expectations they shape period after period.

These expectations have two components: a commonly shared view about what the future will be, and an agent-specific idiosyncratic part. The population of agents will therefore exhibit a distribution of opinions at each point in time. Those believing that low-carbon capital will be more profitable will switch from high- to low-carbon investment and vice versa. The dynamics determines the progressive penetration of the low-carbon option depending on the agents' expectations.

Although our agents maximise their expected payoff, they do not derive an optimal investment path through an optimal control procedure. This extra degree of bounded rationality is intellectually close to Mercure (2015) and Mercure et al. (2019, 2014) and other proposals for technological sub-modules in integrated assessment modelling (Bond-Lamberty et al., 2020). The aim of the model is, therefore, to *complement* usual technological-progress-oriented analyses of the low-carbon transition by a formalisation of cognitive obstacles to the transition with a particular emphasis on expectations. To emphasise our point,

[3] The planned publication dates for the different sub-reports of the AR6 are the years 2021/2022, see `https://www.ipcc.ch/assessment-report/ar6/` for further information.

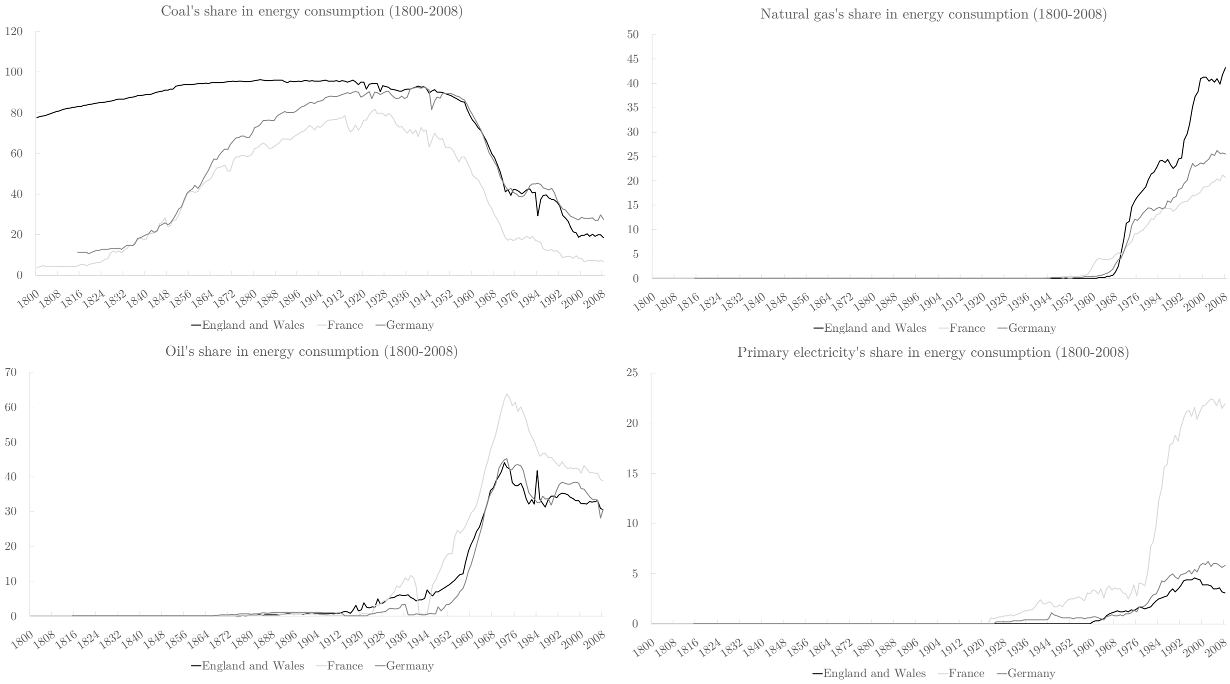


Figure 1: Historical shares of coal, oil, natural gas and primary electricity in energy consumption for England & Wales, France and Germany from 1800 to 2008. Primary electricity refers to electricity from geothermal, hydropower, nuclear power, wind, photovoltaics, tidals, wave and solar thermal. Shares do not sum up to 100% as other energy sources are not displayed here. Source: Kander (2013).

we assume away technological progress or any other pecuniary incentive like a carbon tax, at least as a first step, and only focus on the different parameters ruling our expectation-formation process. Our main outcomes for analysis are therefore the dynamics of low-carbon penetration, through the share of low-carbon capital in the total and that of low-carbon energy production. We also propose a calibration for the model, based on the European Union in the year 2017, a detailed description of which is provided in Appendix D.

In this section we outline the integration of behavioural conflicts and of rolling and fictional expectations to the initial model but leave its longer derivations to the appendices.

3.1. Expectations in The Electricity Sector

To describe a change of expectations made for a horizon of S years, we introduce a rolling planning horizon in which investors re-examine their forecast of the next S years at every time t. As a result, at each time t, taking the actual value as starting point, agents will formulate an expectation vector of length S. This is displayed visually in Figure 2. In formal terms, expectations take the form of a matrix with T rows and $T + S$ columns, and expected values are therefore indexed with two numbers. If X^e is an expected value for a given variable X at time t and at a distance s from the current period, we write:

$$\mathbb{E}_t(X) = \mathbb{E}_t(X_s) \tag{1}$$

We thus adopt a middle-ground between infinite-horizon, perfect-foresight expectations and myopic or adaptive ones: we assume that agents only formulate expectations over a planning horizon of a given length S that represents the *psychological time* of individuals (Zellner, 1979). This framework also relates to the work of Spiro (2014) and in consistent with the literature on limited foresight in investment decision (Hedenus et al., 2006).

In each period electricity is produced and sold according to both an exogenous energy demand e_d, which is assumed to be increasing over time, and a merit order, which means renewable energy is sold first due to

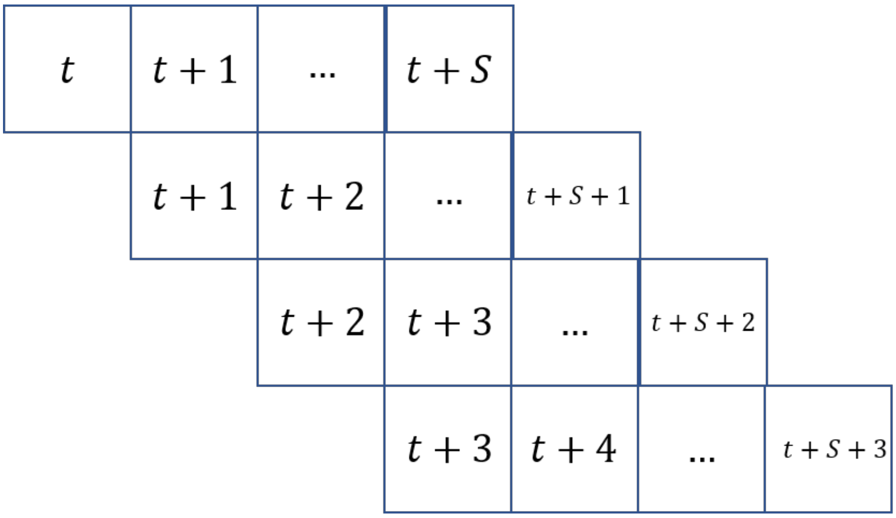


Figure 2: A stylised representation of our expectation structure (adapted from Spiro (2014))

its low marginal costs. If some share of non-renewable capital K_H[4] are not used then they may be considered under-utilised (i.e. stranded), which is signified by a low utilisation rate u_H. In order to make decisions, investors will also consult costs, productivity and profits of each type of capital. We outline the derivations of these variables in Appendix A.

We consider that agents formulate expectations on stranding of non-renewable assets on two levels.

3.2. The Probit module

3.2.1. Expectation dispersion

The first innovation of this framework lies in our modelling of the dispersion of expectation around a given anchor. Our approach relies on a reinterpretation of random utility theory (McFadden, 1980) and its application to discrete choice modelling in econometrics (Ben-Akiva and Lerman, 1985) and formalised economics (Azari et al., 2012). We therefore develop below a "Probit Module" relating discrete-choice theory with the formation of long-run expectations and, ultimately, investment choices. This proposal is formally similar to formulations suggested by Train (2009) and Acemoglu and Jensen (2018) for behavioural neoclassical growth models, building on Luce (2005) and Simon (1991). It is however conceptually distinct.

By making a few simple assumptions - namely, that future demand and productivity are known and companies to not buy more capital than they need - one can estimate the expected under-utilisation that non-renewable units will suffer. We show this process in Appendix B. We suppose that agents formulate *independent* projections about future outcomes around an anchor that represents the most common projection amongst agents. This anchor is given by a law of motion in psychological time. It can be more or less complex depending on the research question. And be more or less related to the model's actual behaviour. In our particular case, the variable in question is the utilisation rate of non-renewables u_H, the central projection of which μ_H is calculated based on the expected speed of the transition as mentioned above and derived in Appendix B. We include an error term around the expected utilisation rate which is a random variable with a given variance. This variance represents the dispersion of particular opinions around the central projection, and is our way to model heterogeneity in utilisation expectations. We thus assume the expectations of non-renewable utilisation at a time t to be the sum of the deterministic part $\mu_{t,t+s}$ (calculated as the expected utilisation rate in Appendix B) and an error term $\varepsilon_t \sim \mathcal{N}(0, \sigma_t^2)$.

$$\mathbb{E}_t(u_s) = \mathbb{E}_t(u_s^*) + \varepsilon_s \tag{2}$$

[4] A subscript of H indicates that a variable is referring to non-renewable capital which a subscript of L infers renewable capital.

Remains to determine a law of motion for the standard errors of the idiosyncratic opinions in psychological time. How this dissent evolves in psychological time requires some assumptions. Building on Keynes's (1936) observation that long-run expectations are more unstable and dispersed than short-run ones but that there is little difference between long- and "very long"-run expectations (see also Moureau and Rivaud-Danset, 2004), we assume that in psychological time σ follows a sharply increasing logistic law of motion:

$$\sigma_t = \sigma_{u,min} \tag{3}$$

$$\forall\ t < s \leq S,\quad \sigma_s = \sigma_{s-1}\left(1 + b_\sigma(1 - \frac{\sigma_{s-1}}{\sigma_{u,max}})\right) \tag{4}$$

That is, we consider a minimal level dissent for expected profit rates $\sigma_{u,min}$ that is present *ex ante*, and which increases sharply at a rate $b_\sigma$and up to $\sigma_{u,max}$. Dissent is first rather weak, with agents mostly agreeing on the mean projection, and increases quickly up to $\sigma_{u,max}$ before it reaches a plateau, signalling that agents disagree to the same extent in the late periods of the planning horizons. Note that nothing prevents $\sigma_{u,min}$ and $\sigma_{u,max}$ from being equal, or to be very close. In such a configuration, the level of heterogeneity amongst opinions changes very little over time.

3.2.2. Investment Choice

In light of heterogeneous expectations, investment is divided between the two technologies based on the probability that a random investor believes one technology will be more profitable than the other. Given that we have a distribution of opinions about future stranding, this probability can be evaluated by creating a proxy variable to indicate that the profit rate of renewables is higher ($r_L > r_H$) and using a probit model to evaluate its probability ζ_L (see Appendix C). The probability ζ_L (and the probability for non-renewables ζ_H) can be interpreted as a share of investors that would prefer to invest in that technology.

Based on the probabilities we just defined, investors decide on the share of investment they dedicate to one or the other type of capital. Denoting ℓ_{I_t} to be the share of investment going to low-carbon capital, we define it recursively as in Equation 8, where ℓ_{I_1} is a calibrated first-period value. This function takes inspiration from Mercure's (2015; 2019; 2014) demographics-based treatment of technological change in energy economics, and is commonplace in population dynamics, including in the economics of technological change (Kucharavy and De Guio, 2015) and competition (Marasco et al., 2016).

$$\ell_{I_1} = \ell_{I_1} \tag{5}$$

$$\forall\ t \geq 1, \ell_{I_{t+1}} = \ell_{I_t} + (1 - \ell_{I_t})\zeta_{L_t} - \ell_{I_t}\zeta_{L_t} \tag{6}$$

Let us interpret ℓ_{I_t} as the share of investors or investment projects dedicated to low-carbon capital in each period, and $1 - \ell_{I_t}$ the share dedicated to high-carbon capital. We consider that at each point in time, a fraction of these shares, respectively ζ_{L_t} and ζ_{H_t} will change course. In other words, a portion ζ_{L_t} of high-carbon investment projects will be abandoned, or a fraction ζ_{L_t} of agents investing in high-carbon capital will invest in low-carbon capital in the next period, hence that this number is added to the previous share ℓ_{I_t}. Conversely, a fraction ζ_{H_t} of investment dedicated to low-carbon projects will switch back to high-carbon capital. As a result, this fraction is subtracted from the current share of low-carbon investment.

ζ_{L_t} and ζ_{H_t} can be considered as *transition probabilities* between two behaviours: investing in high-carbon capital and investing in low-carbon capital. In most related works, more or less complex versions of a logistic distribution is assumed as functional forms for this probabilities. As we will see below, we rely on normal distributions for computational ease instead since the two distributions are relatively close and interchangeable in practice.

It is worth noting that they have no *a priori* reason to be symmetrical (*i.e*, $\zeta_{H_t} = 1 - \zeta_{L_t}$). If this is the case, then Equation 8 reduces to:

$$\ell_{I_{t+1}} = \zeta_{L_{t+1}} \tag{7}$$

This would suggest that all agents choose all the time between the two options, not allowing for the possibility that some will stick to their previous behaviour without considering change. We will therefore introduce some asymmetry between ζ_{H_t} and ζ_{L_t} to give rise to some stickiness in behaviour. Considering that this proportion represents the share of agents switching from high- to low-carbon investment, we have that:

$$\zeta_{L_t} = \chi \left(P \left(r_{L_t} - \ r_{H_t} > 0 \right) \right) \tag{8}$$

We introduce an asymmetry between ζ_{L_t} and ζ_{H_t} in by assuming that reverting from low-carbon investment to high-carbon investment while the low-carbon is ongoing requires a minimal payoff to be undertaken. Formally, it writes:

$$\zeta_{H_t} = \chi \left(\ P \left(r_{L_t} - \ r_{H_t} \leq \left(\sum_{1}^{S} \beta^s \right) \rho_L \right) \right) \tag{9}$$

with $\beta = \frac{1}{1+\rho} =$ with ρ the discount rate, and $\rho_L > 0$, and is discounted accordingly. That is, r_{H_t} must be higher than r_{L_t} by an extra margin. The parameter ρ_L can be interpreted as a "transition premium", either enforced by the regulator, or a measure of the risk to revert to high-carbon capital in the context of the transition. It is set to 1 throughout.

Total investment in low-carbon capital I_{L_t} is defined as:

$$I_{L_t} = I_t^d \ell_{I_t} \tag{10}$$

And high-carbon investment I_{H_t}

$$I_{H_t} = \left(1 - \ell_{I_t}\right) I_t^d \tag{11}$$

Where I_t^d is total investment demand, that is determined through the following step.

To determine absolute investment, agents will first compute the total amount of capital that will be required in the next period to meet demand in $t+1$, $K^d(t+1)$, which is the targeted total capital stock. As a result, I_t^d is defined through an extended perpetual inventory rule, in equation 14. Here we suppose that total investment makes up for natural depreciation δ_{L_t} and δ_{H_t} of the low and high-carbon stocks, respectively. We also assume that no capital is destroyed before the end of its productive lifetime.

$$I_t^d = \max\left(K_{t+1}^d - \left(K_{L_t} + K_{H_t}\right), 0\right) \ + \delta_{L_t} + \delta_{H_t} \tag{12}$$

In the model, agents formulate imaginary futures about how stranded high-carbon capital may be. If they expect stranding, they will naturally want to hedge against it. For the sake of simplicity, we assume that agents always near-perfectly hedge against stranding, by ensuring that the conventional utilisation rate u_H^τ is reached. As a result, we suppose that agents adopt the following target function, where Γ_{1_t} and Γ_{2_t} are intermediate terms defined defined in Equations 16 and 17. When substituted in, one can see that the target capital buys capital to fulfill demand while maintaining full utilisation.

$$K_{t+1}^d = \frac{e_{t+1}^d - \ \xi_L \Gamma_{1_t} - u_H^\tau \xi_H \Gamma_{2_t}}{u_H^\tau \xi_H \left(1 - \ell_{I_t}\right) + \xi_L \ell_{I_t}} \tag{13}$$

$$\Gamma_{1_t} = \left(1 - \ell_{I_t}\right) K_{L_t} - \ell_{I_t} \left(\delta_{L_t} + \delta_{H_t} - K_{H_t}\right) - \delta_{L_t} \tag{14}$$

$$\Gamma_{2_t} = \ell_{I_t} K_{H_t} - \left(1 - \ell_{I_t}\right) \left(\delta_{L_t} + \delta_{H_t} - K_{L_t}\right) - \delta_{H_t} \tag{15}$$

This rule allows agents to always reach the conventional utilisation rate u_H^τ except when it would require capital destruction beyond the natural depreciation of capital. This highly simplifying assumption is meant to avoid that agents remain passive if stranding occurs[5].

[5] More complex investment rules allowing for imperfect hedging can be implemented in the current version of the model, but do not change the results significantly.

3.3. The narrative level

Our second innovation is that we consider that agents formulate expectations on how stranded high-carbon capital assets will be in the future based on narratives derived from expected shares of low-carbon energy.. This is how fictional expectations are currently modelled within our framework.

A narrative, like a policy schedule determined by a government, gives agents an expectation of the speed of an ensuing logistic transition. Based on this expected speed, and from current levels of capital, utilisation and demand, agents can estimate the amount of stranding they expect to occur. We thus first endow our agents with such a narrative in terms of renewable energy consumption share in the total, that we call $\ell_E^N(t, t+s)$, N standing for "narrative", and define as follows:

$$\forall\ 0 \le s \le t,\ \ \mathbb{E}_t(\ell^N_{E_s}) = \ell_{E_s} \tag{16}$$

$$\forall t \le s \le S,\ \ \mathbb{E}_t(\ell^N_{E_s}) = \mathbb{E}_t(\ell^N_{E_{s-1}})\left(1 + b_N\left(1 - \frac{\mathbb{E}_t(\ell^N_{E_{s-1}})}{\bar{\ell}}\right)\right) \tag{17}$$

where b_N is exogenous, and $\bar{\ell}$, which represents the long-run low-carbon energy share goal, is 1. As a result, only the starting value of the schedule, which is the observed low-carbon share, changes. These expectations allow agents to compute an expected high-carbon energy production share, by definition equal to $\mathbb{E}_t(h^N_{E_s}) = 1 - \mathbb{E}_t(\ell^N_{E_s})$.

The sigmoid shape is a quite natural functional for technology and product diffusion and structural change, and has actually strong empirical underpinnings for many technologies (Mercure et al., 2014; Perez, 2002), products (**?**), and is considered as a sound pattern for the future displacement of fossil fuels (Vandevyvere and Nevens, 2015).

Yet, it cannot be taken for granted that all agents will abide by or believe in the narrative of a schedule set by policy or public opinion. It is more likely that there will be some conflict of opinions between a more optimistic narrative and more pessimistic beliefs, that will change over time based on what actually occurs. We introduce a framework for modelling this conflict of opinions, building from (Franke and Westerhoff, 2017), Weidlich and Haag (1983) and Lux (2009). Our goal is to study how a norm changes through time, with possible social lock-in effects. Our agents do not switch between optimism and pessimism as such, but between different norms depending on their overall credibility.

These social beliefs will take the form of two transition narratives, formalised by schedules of low-carbon capital penetration ℓ_E in psychological time. The first narrative ℓ_E^N is the one given above. The other belief ℓ_E^a is an endogenous, trend-following projection of the share of low-carbon capital in psychological time. At each point in time, agents will gauge the credibility of both beliefs through an evaluation process close to the approach proposed by de Grauwe and co-authors (De Grauwe and Ji, 2018). From this, agents will derive an aggregate narrative ℓ_E^o that will constitute the high-carbon utilization rate schedule in time S, henceforth diminishing stranding expectations if the adaptive narrative's schedule is below that of the regulator's overarching narrative.

The alternative narrative is defined as follows:

$$\forall -t \le s \le 0,\ \ \mathbb{E}_t(\ell^A_{E_s}) = \ell_{E_s} \tag{18}$$

$$\forall t < s \le S,\ \ \mathbb{E}_t(\ell^A_{E_s}) = \mathbb{E}_t(\ell^A_{E_{s-1}})\left(1 + \mathbb{E}_t(b_{a_t})\left(1 - \frac{\mathbb{E}_t(\ell^A_{E_{s-1}})}{\bar{\ell}}\right)\right) \tag{19}$$

The limit $\bar{\ell} = 1$ as before for the sake of simplicity. b_{a_t} moves in time t and is defined as the observed growth rate between the current and the previous period, corrected to turn into a logistic intrinsic growth rate:

$$\mathbb{E}_t(b_{a_1}) = b_N \tag{20}$$

$$\forall t > 1 \;\; \mathbb{E}_t(b_{a_t}) = \frac{\ell_{E_t} - \ell_{E_{t-1}}}{\ell_{E_t}\,(1 - \ell_{E_t})} \tag{21}$$

We finally define the "aggregate-opinion" narrative $\ell^o_{E_s}$ as the weighted average of both narratives:

$$\forall 0 \leq s \leq t, \;\; \ell^o_{_{E_s}} = x_t \ell_{_{E_s}} + (1 - x_t)\,\ell_{_{E_s}} = \ell_{E_t} \tag{22}$$

$$\forall t < s \leq S, \;\;\; \ell^o_{_{E_s}} = \; x_t\,\mathbb{E}_t(\ell^N_{_{E_s}}) + (1 - x_t)\,\mathbb{E}_t(\ell^A_{_{E_s}}) \tag{23}$$

The underlying story is as follows. The regulator adverts a transition narrative entailing a certain goal to be achieved in the years to come. However, this narrative is not univocal: only a share of the population, x_t takes it for granted and credible. The other share, $1 - x_t$ considers this narrative as not credible and relies on a simple forecast technique to project low-carbon capital shares over the planning horizon. The creation of this balanced new forecast from two narratives is displayed in figure 3 below.

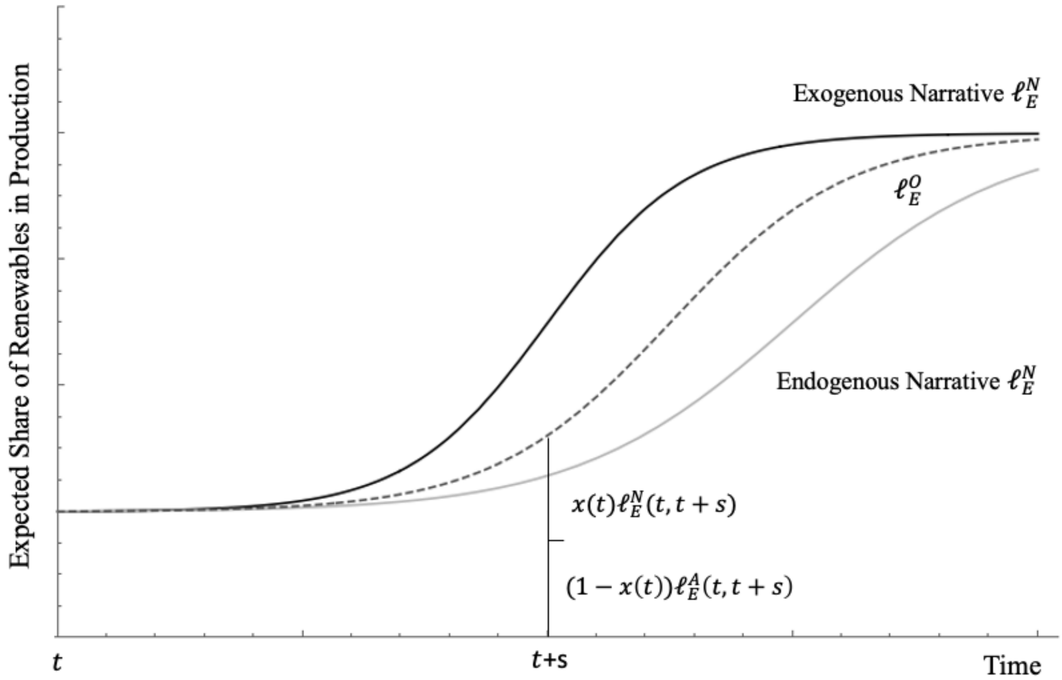


Figure 3: Creation of a balanced forecast ℓ^O_E from an exogenous narrative ℓ^N_E and an endogenous prediction ℓ^A_E

These shares will evolve in chronological time according to an evaluation process. We first define a variable w_{N_t} as follows:

$$w_{N_t} = \begin{cases} 0 \; if \; t = 1 \\ \sum_{k=1}^{t-1} \omega^{1-k} \frac{\left(E_k(\ell^N_{K_t}) - \ell_{K_t}\right)}{\ell_{K_t}} \; if \; t < 1 \leq M + 1 \\ \sum_{k=t-M}^{t-1} \omega^{1-k} \frac{\left(E_k(\ell^N_{K_t}) - \ell_{K_t}\right)}{\ell_{K_t}} \; if \; t > M + 1 \end{cases} \tag{24}$$

And another, w_{A_t} in the same way:

$$w_{A_t} = \begin{cases} 0 \; if \; t = 1 \\ \sum_{k=1}^{t-1} \omega^{1-k} \frac{\left(E_k(\ell_{K_t}^A) - \ell_{K_t}\right)}{\ell_{K_t}} \; if \; t < 1 \leq M+1 \\ \sum_{t-M}^{t-1} \omega^{1-k} \frac{\left(E_k(\ell_{K_{K_t}}^A) - \ell_{K_t}\right)}{\ell_{K_t}} \; if \; t > M+1 \end{cases} \tag{25}$$

And finally, w_t, the weighted average of both:

$$w_t = x_t w_{N_t} + (1 - x_t) \, w_{A_t} \tag{26}$$

That is, at each point in time, agents will take the relative spread (normalised by ℓ_{K_t}) between the actual share of ℓ_{K_t} and the share given by previous norms between the previous period and time M, fixed for simplicity equal to S. Past relative spreads are given an exponentially decreasing weight whose starting value is ω, fixed to 0.85, a commonly found value for such memory parameters in behavioural macroeconomics (De Grauwe and Ji, 2018; Hommes, 2018). This evaluation process is aimed at modelling a degree of reflexivity that agents have with respects to their beliefs, the w's being interpretable as credibility indicators, the farther from zero a given w, the less credible a given norm.

To derive this opinion dynamics, let us first define an opinion indicator j_t taking its values in $[0, 1]$ implicitly defining x_t as follows:

$$x_t = \frac{(1 + j_t)}{2} \tag{27}$$

$$1 - x_t = \frac{(1 - j_t)}{2} \tag{28}$$

That is, the closer to 1 j_t is, the higher the share of the population believing in the regulator's narrative, and the closer to -1, the higher the share believing in the alternative narrative. The opinion indicator j_t is itself defined recursively:

$$j\,(1) = j_0 \tag{29}$$

$$j_{t+1} = j_t + (1 - j_t)\, p_t^{AN} - (1 + j_t)\, p_t^{NA} \tag{30}$$

This equation is the opinion equivalent of the investment function defined in Equation 8. The variable $(1 - j_t)$ is a measure of the population believing in the alternative narrative, while $(1 + j_t)$ is a proxy for that of those believing in the regulator's narrative. The variable p_t^{AN} is therefore the transition probability from the alternative to the regulator's narrative, and p_t^{NA} that from the regulator's to the alternative narrative. They are given by:

$$p_t^{AN} = \nu \exp\left(-w_t\right) \tag{31}$$

$$p_t^{NA} = \nu \exp\left(w_t\right) \tag{32}$$

That is, the higher w_t (the more agents will have *over*estimated their potential performances in aggregate), the more they will shift towards the alternative narrative, and reciprocally. The value ν is a "natural switching" parameter meant to figure a probability for agents to switch between norms even though w_t is equal to zero. Given that we are dealing with a long-run model, we set it to 0.001 (that is one agent out of 1000 will "naturally" switch behaviour) to figure a certain natural stickiness in the evolution of norms[6].

[6]One could wonder why the aggregate credibility w_t should be the argument of both transition probabilities instead of the norm-specific ones w_{N_t} and w_{A_t}. This is indeed a legitimate specification. However, using the aggregate credibility allows agents to have a look at both norms at the same time, and therefore avoid that they switch towards a less credible alternative by only looking at the credibility of their own norm. We implemented both specifications, and results are qualitatively the same: we therefore keep the specification using the aggregate credibility measure.

Our aim through is to explore more consistently the role of social beliefs in giving rise to cognitive and social lock-ins within a modelling framework (*cf supra*). We indeed assume away, at least for now, any kind of technical inertia, and only take interest in how expectations and their formation influence the dynamics of our stylised low-carbon transition dynamics.

We finally use the difference between the two narratives and the fact the total energy demand schedule is perfectly known to derive two expected low-carbon energy demands in a recursive way as follows:

$$\mathbb{E}_t(e^N_{L_{s+1}}) = \mathbb{E}_t(e^N_{L_s})(1+g_E)\left(1+b_N\left(1-\frac{\mathbb{E}_t(e^N_{L_s})}{\bar{\ell}^d_s\,\mathbb{E}_t(e^d_s)}\right)\right) \tag{33}$$

$$\mathbb{E}_t(e^a_{L_{s+1}}) = \mathbb{E}_t(e^a_{L_s})(1+g_E)\left(1+\mathbb{E}_t(b_{a_t})\left(1-\frac{\mathbb{E}_t(e^a_{L_s})}{\bar{\ell}^d_s\,\mathbb{E}_t(e^d_s)}\right)\right) \tag{34}$$

And an aggregate low-carbon energy demand:

$$\mathbb{E}_t(e^o_{L_s}) = x_t\,\mathbb{E}_t(e^N_{L_s}) + (1-x_t)\,\mathbb{E}_t(e^a_{L_s}) \tag{35}$$

Given that the future energy schedule is perfectly known, we can derive the expected high-carbon energy demand:

$$\mathbb{E}_t(e^o_{H_s}) = \mathbb{E}_t(e^d_s) - \mathbb{E}_t(e^o_{L_s}) \tag{36}$$

We then determine what would be the "optimal" capital stock K^d_H that would be demanded if the high-carbon technology utilization rate were to be kept at the target rate u^τ_H:

$$\mathbb{E}_t(K^d_{H_s}) = \frac{\mathbb{E}_t(e^o_{H_s})}{u^\tau_H \xi_H} \tag{37}$$

If expected high-carbon energy demand is expected to decrease, $K^d_{H_s}$ will be lower than the current capital stock. Because we assume that the capital stock cannot be scrapped beyond the natural depreciation rate δ_H, the expected capital stock is given by:

$$\mathbb{E}_t(K_{H_s}) = \max\left(\mathbb{E}_t(K^d_{H_s}), (1-\delta_H)^s K_{H_t}\right) \tag{38}$$

Finally, the expected utilization rate is given by:

$$\mathbb{E}_t(u_{H_s}) = \frac{\mathbb{E}_t(e^o_{H_s})}{\mathbb{E}_t(K_{H_s})} \tag{39}$$

Intuitively, if the optimal capital stock is below the capital stock that would prevail would the current capital stock be left to depreciate naturally, then the expected utilization rate will be lower than the target rate, signalling expected stranding.

In what follows, we apply this framework to various experiments in a order to illustrate the functioning of the model and draw some key insights.

4. Scenarios and simulations

We calibrate our model to the electricity market of the European Union (EU) for the year 2017 (see Appendix D). However, the simulations below are intended at illustrating the assumptions behind our model, not at sketching concrete energy transition paths for the EU. We show that different expectations shaped by narratives about future transition dynamics and different levels of uncertainty about future utilization rates, stranding and profitability of technologies take concrete effects by shaping these transition dynamics.

4.1. Stranding expectations drive model results

Expectations of stranding are key to dynamic change in the model, but this does not necessarily mean that stranding will occur. If agents are more confident about the possibility of stranding, they will of course do what they can to avoid it. Below we outline the actual utilisation rate (i.e. the stranding that occurs) over time, along with the minimum expected utilisation at that time that will impact the investment choice in that period. We also show the trend of expected future utilisation in periods 1, 20 and 40.

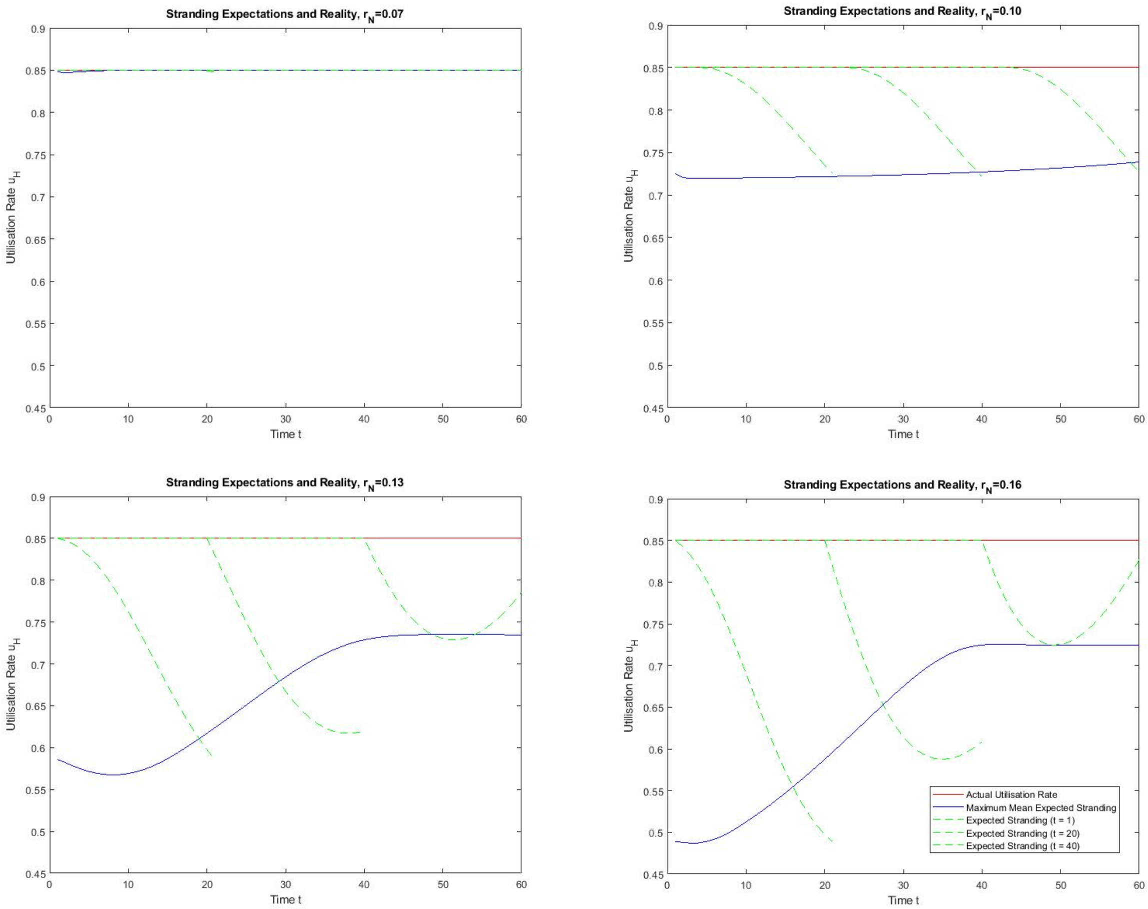
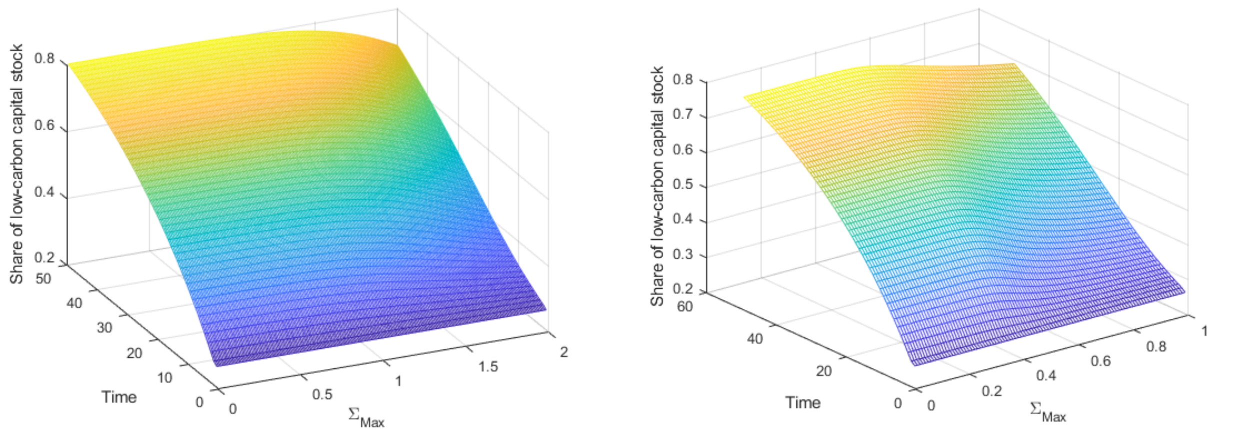


Figure 4: Utilisation rate (red) and maximum expected stranding (blue) over time, and expected utilisation rate for t=1, 20 and 40 (green) for expected transitions with speeds b_N 0.07, 0.10, 0.13 and 0.16.

One process within the model which is made clear here is that agents hedge against stranding in order to maintain higher profit rates, thus forwarding the transition and leading to more optimistic expectations in the next step. For the weaker expectations of stranding, there is little cause to fear, meaning little is done to change investment patterns and later expectations. For the stronger expectations in the bottom two figures, drastic expected dips in utilisation (indicated by the dashed green lines) causes companies to hedge against expectations and divert some investment to renewables in order to avoid units becoming under-utilised. In the more drastic expected transitions, companies are forced to hedge completely to avoid stranding, bringing about a full transition. Such a situation would require a strong narrative projected towards producers which convinces those investing in non-renewables that they will be left behind in the next 20 years if they do not change their investment strategies soon.

4.2. Narrative stances for transition dynamics: The single-narrative benchmark

We assume that there is only one narrative prevailing between agents, i.e. that the growth rate b_N of renewable electricity according to official communications by the European commission (2017) of about 12 % per year is upheld as the only relevant benchmark for transition dynamics (the *overarching narrative* by the regulator). Accordingly, we see a smooth and gradual shift to renewable electricity also for varying levels of maximum dispersion of opinion (**dissent**), $\sigma_{u,max}$, about stranding of high-carbon capital, see Figure 5a. In fact, within a model time horizon of about 50 years, 80 % of total capital stock in the electricity sector is renewable, i.e. ℓ_K closes in on 0.8.



(a) The *one* narrative benchmark case with varying dissent about profitability and utilization of the two technologies

(b) Results stay robust also with increasing growth rates of opinion dispersion r_{log_σ}

Figure 5: A benchmark with one narrative and varying levels of dissent

Results in Figure 5a are steered by agents expecting stranding of high-carbon capital stock and at the same time anticipating a premium in the profitability for low-carbon investment due to the increased penetration of this technology triggering network effects and economies of scale.

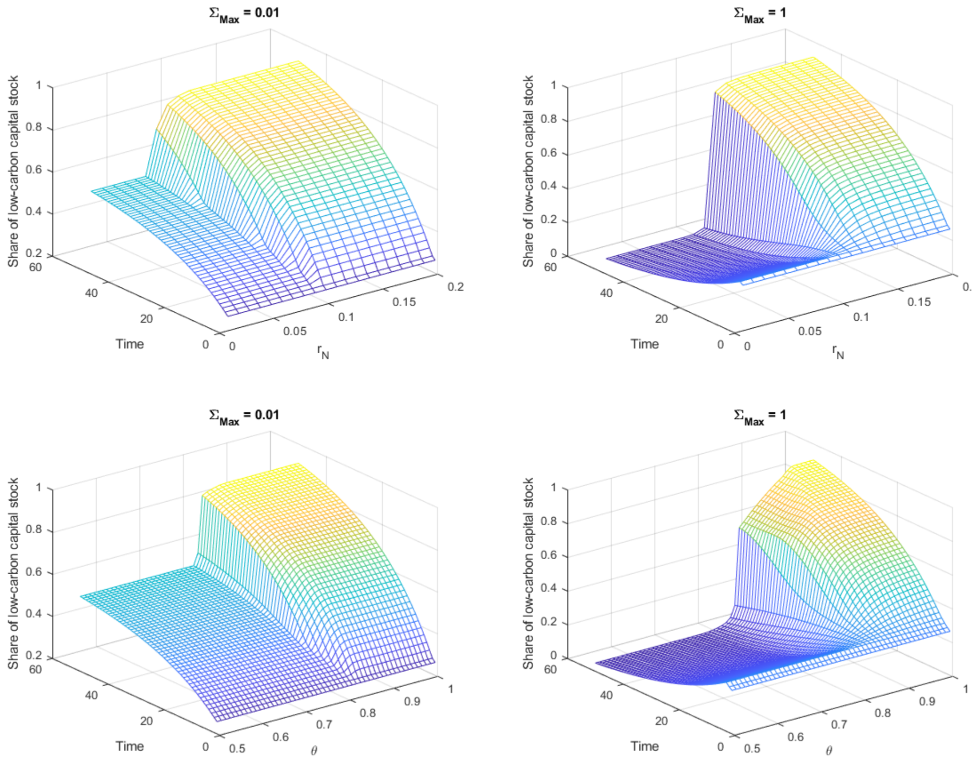


Figure 6: Varying ambitions and levels of dissent: at high levels of dissent and low levels of ambition, the transition may not take place

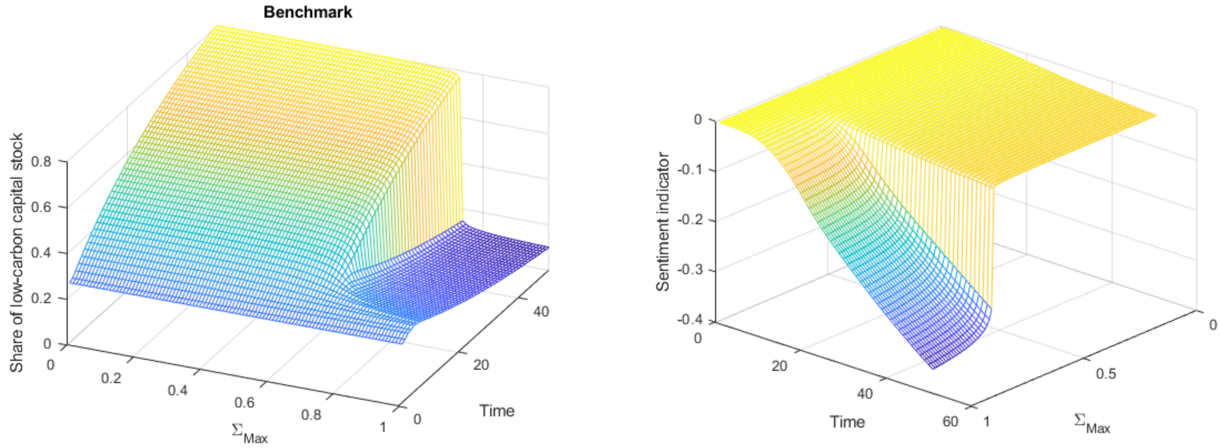
For high amount of dissent, renewable investments are limited especially in earlier periods. However, both the ambition as well as the belief of agents in the overarching narrative lead to virtuous cycles for investments in renewables. As agents observe whether the narrative holds true without forming dissenting opinions, the transition path self-validates: agents believing in the narrative invest in renewables, which leads others to follow and thus locks in the transition path. This result stays robust also if the growth rate of dissent in agents' expectations, b_σ varies, see Figure 5b.

Without dissenting opinions, however, too timid ambitions by the regulator can lead to the lock-in of an equilibrium where the transition essentially does not take place. This can be observed in figure 6: Both the growth rate b_N as well as the long-run target for the share of renewables in total capital stock $\bar{\ell}$ as decreed by the regulator vary. Below an ambition threshold and high levels of dissent about utilization rates and future profitability of the two technologies, agents revert to a portfolio of high-carbon capital stock.

4.3. Narrative stances for transition dynamics: opinion conflict and competing narratives

Given high ambitions of the regulator even widely varying levels of maximum dissent $\sigma_{u,max}$ could not phase out the transition. However, Figure 7a illustrate that in case of diverging opinions about the future, the transition path given by the regulator can lose its credibility, leading to a reversion of the transition for high levels of dissent. Here, a high dispersion of opinion around mean stranding expectations implies that a significant fraction of agents does not believe that stranding of high-carbon assets will take place due to the transition. Therefore, a large fraction of agents will revert to high-carbon investment.

The results of Figure 7a are further evident from the dynamics of opinion conflict shown in Figure 7b. That is, after a level of dissent for $\sigma_{u,max}$ of about 0.7, a large part of the agent population does not adhere to the overarching narrative ℓ_E^N, but rather to the alternative narrative ℓ_E^A, shifting the average opinion ℓ_E^O rather to the trend-dependent alternative narrative. This effect is self-reinforcing over time: Reductions in



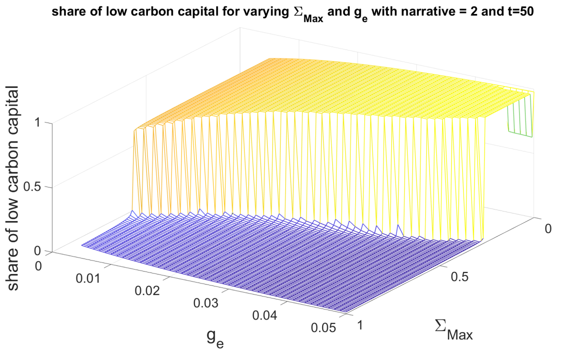
(a) The *two* narrative benchmark case with varying $\sigma_{u,max}$

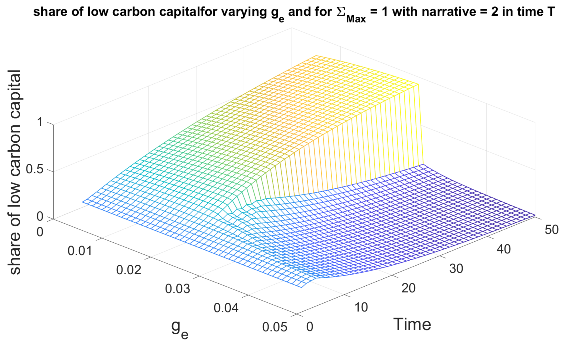
(b) How opinion conflicts affect the transition

Figure 7: A *two* narrative benchmark case shows a phase transition at higher levels of dissent

renewable investments tend to further decrease the validity of the overarching narrative as well as stranding expectations for the high-carbon technology. Conversely, if agents commit strongly to low-carbon investment in the early period of the run, the alternative narrative will in fact come close to that of the regulator's.

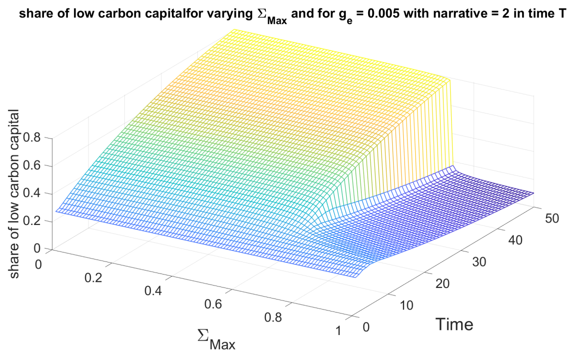
These insights from our model shed light on how fictional expectations interact with dynamics realized partly based on these narratives. Indeed, agents in our model are still forward-looking, in the sense that they formulate prospective projections into the future under conditions of fundamental uncertainty. The way they do it is nonetheless partly backward-looking, as they consider the past trend to be relevant for their investment decision. Assuming a mix between backward- and forward-looking expectations is enough to obtain behavioural stickiness and path dependency in opinion formation.

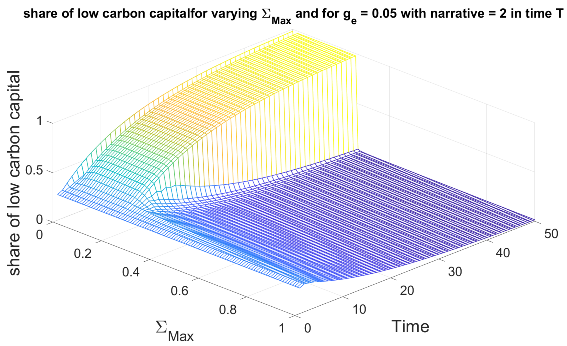




(a) Interaction between growth rate and SigmaMax for final model period

(b) Regime change for different growth rates at high dispersion of opinions

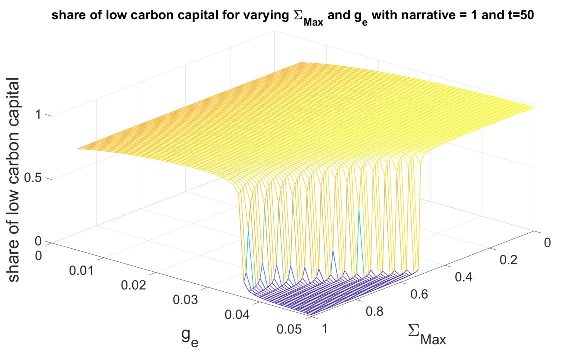


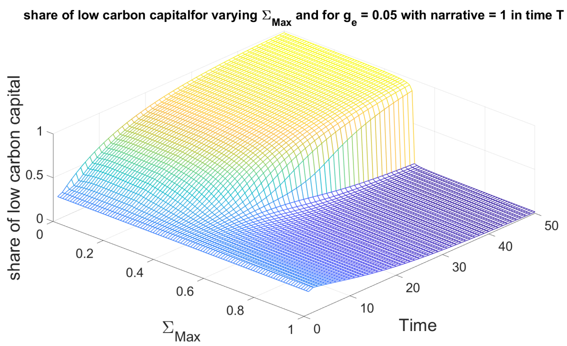


(c) Lower growth rates allow higher dispersion of opinions

(d) Higher growth rates lead to earlier regime shifts for dispersion of opinions

Figure 8: Interaction between electricity demand growth and conflicting opinions with alternative narrative regimes





(a) Interaction between growth rate and $\sigma_{u,max}$ for final model period - transition more robust with only one regime

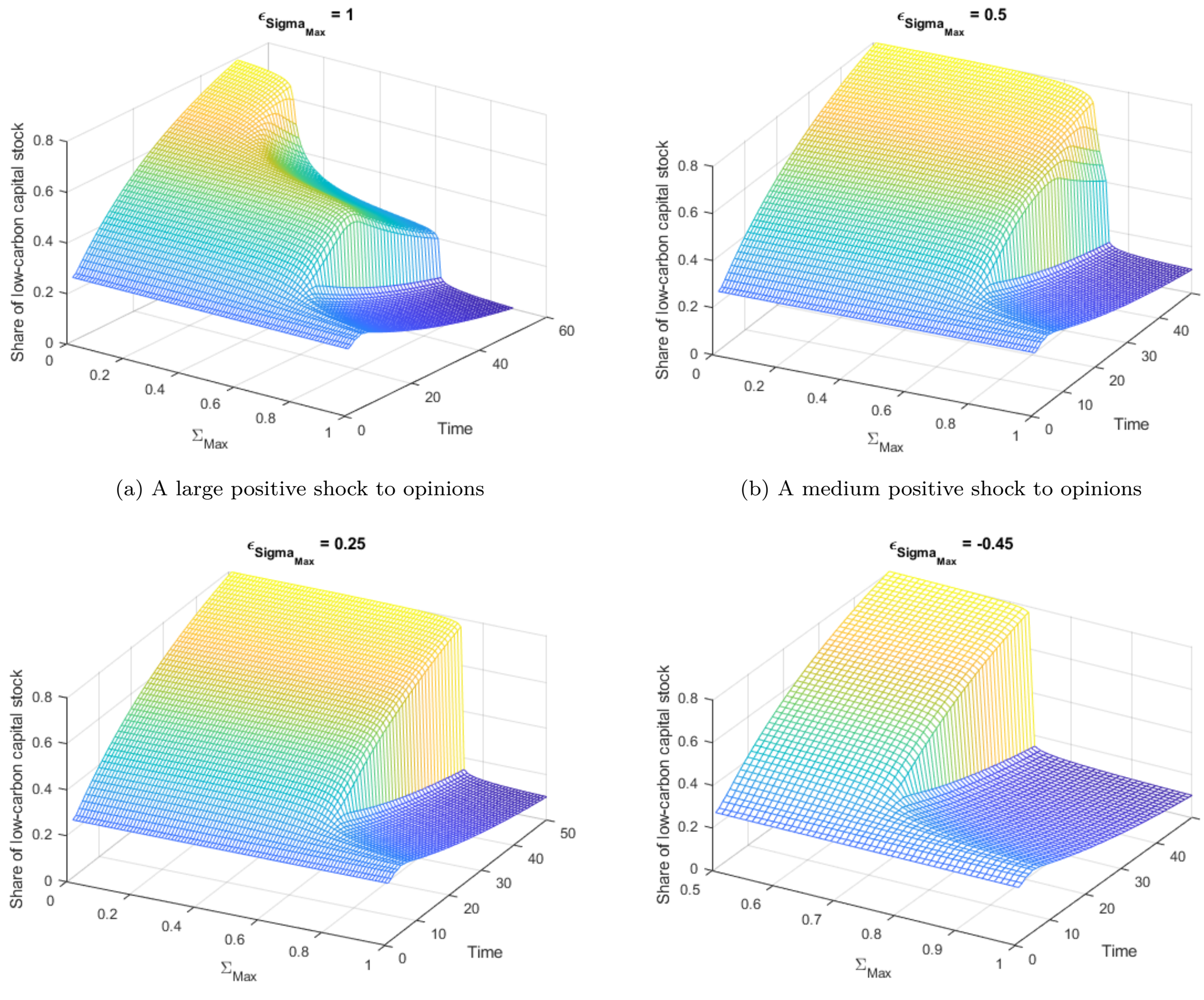
(b) At higher growth rates, also with only one prevailing narrative, a regime shift occurs

Figure 9: Interaction between electricity demand growth and conflicting opinions with only one narrative regime

This result calls for great policy caution. In the presence of behavioral stickiness a gap between the regulator's ambitions and realized investments or a less steady trend over even a relatively short time period may bear significant long-run consequences on the dynamics of the transition.

4.4. Changes in electricity demand growth and narratives about the future

We can observe a clear correlation between growth rates of energy demand and opinion dispersion in the model. For the transition to take place in a situation with conflicting opinions about the future, opinion dispersion needs to be lower for higher growth rates, as can be observed in Figure 8. In the upper left panel (8a), we can see how the phase transition between a no-transition lower equilibrium and an almost complete transition higher equilibrium varies for different values of opinion dispersion and growth rates. Similarly, in the two lower panels (8c and 8d), we can see how the regime shift between a transition and a no-transition scenario varies for lower (left lower panel) and higher (right lower panel) growth rates.



(a) A large positive shock to opinions

(b) A medium positive shock to opinions

(c) A small positive shock

(d) A negative shock to opinions

Figure 10: fig:Different shocks to dissent in the model

Again, here one can clearly see the effects of alternative narratives in the model. If only one narrative prevails as depicted in figure 9, the interaction between growth and the amount of dissent is much less pronounced. Only for very high rates of growth and dissent as shown in Figures 9a and 9b, we can observe a regime shift to an equilibrium where the transition does not take place.

4.5. Shocks to dissent and changes in beliefs

Figure 10 shows different shocks $\epsilon_{Sigma_{Max}}$ to dissent, from a large positive shock in Figure 10a to a negative shock in Figure 10d. These "dissent shocks" represent any external event influencing the divergence of opinions about the future, such as uncertainty about policy decisions regarding subsidies for renewables or the apparition of a breakthrough-technology. In our simulations only a large positive shock as depicted in Figure 10a suffices to negatively influence transition dynamics beyond a (rather low) threshold for dissent Σ_{Max}. Smaller shocks leave transition dynamics virtually untouched. Conversely, a negative shock to dissent as in Figure 10d simply moves the level of dissent where the regime change between the point where a transition still takes places and where it does not.

4.6. Policy simulations: regulator's norms updates and carbon tax

Figure 11 below illustrates different policy options that we can cover with our modelling framework. First, figure 11a shows the benchmark model run with two conflicting narratives without additional policy options. Second, in Figure 11b we assume that the regulator *updates* its norm according to the investors' performance through the following rule

$$\begin{aligned} &b_{N_1} = b_N \\ \forall t > 1 \ &b_{N_t} = b_{N_{t-1}} + \eta_{b_N} \left(b_{N_{t-1}} - r_{L_t}^{eff} \right) \end{aligned} \tag{40}$$

With $\eta_{b_N} = 0.5$.

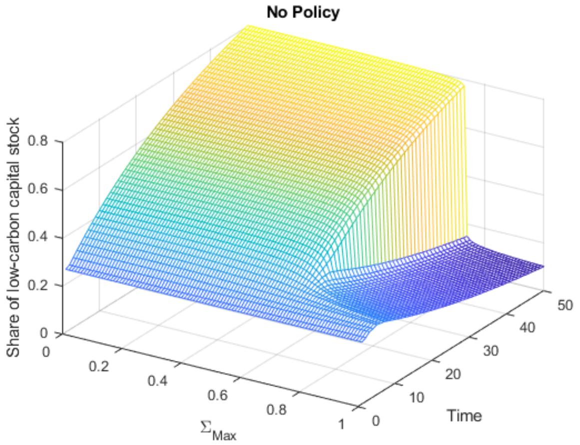
Third, we assume that the regulator sets more ambitious policy targets as in Figure 11c (with $b_N = 0.2$). Being ambitious right from the outset allows for sufficiently high stranding expectations to trigger significant commitment to low-carbon investment regardless of dissent. Updating the narrative is also slightly more efficient than the no policy benchmark. There is therefore an improving potential for some *active* forward-guidance in terms of transition objective so as to anchor agents' expectations and bypass dissent on economic outcomes, but the effects seem to be limited.

Implementing a univocal[7] carbon price of 300 \$ per ton CO_2[8] as in Figure 11d leads to a transition for all levels of dissent. The introduction of a univocal carbon tax does not prevent adverse dynamics down to a certain point but allows for a rebound of low-carbon investment after some time, regardless of the effect of dissent. This is because the price signal, rightly anticipated by investors, becomes sufficiently high to make low-carbon investment profitable. Agents come to invest more and more in renewable energy sources, increasing at the same time the share of low-carbon electricity in total demand. Therefore, the alternative narrative will adapt upwards, propelling virtuous opinion dynamics, themselves favouring low-carbon investment through higher stranding expectations. We can see here, all in all, a synergy between carbon pricing and norms, which suggests that economic instruments can give rise to long-lived social habits (or at least corporate governance norms) that are in line with the transition.

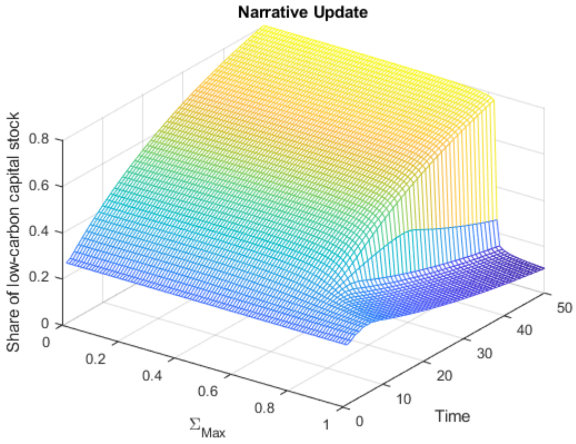
However, remains the question of what would happen if this price were not univocal. To explore this possibility, let us assume that only the fraction $x(t)$ of agents – those adhering to the regulator's narrative – actually incorporate the carbon price path in their expectations. The other share considers that the carbon price will not increase in subsequent periods. Results are displayed in Figure 11e.

[7]i.e. agents perfectly incorporate future carbon prices in their expectations.

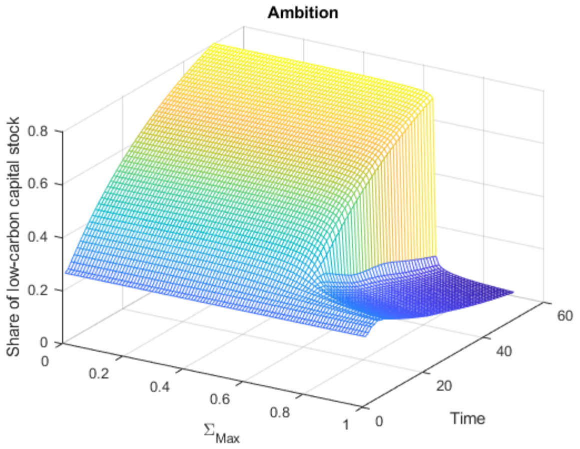
[8]A pricing consistent with French (Centre d'Analyse Stratégique, 2012) and OECD estimates (OECD, 2011) for developed countries.



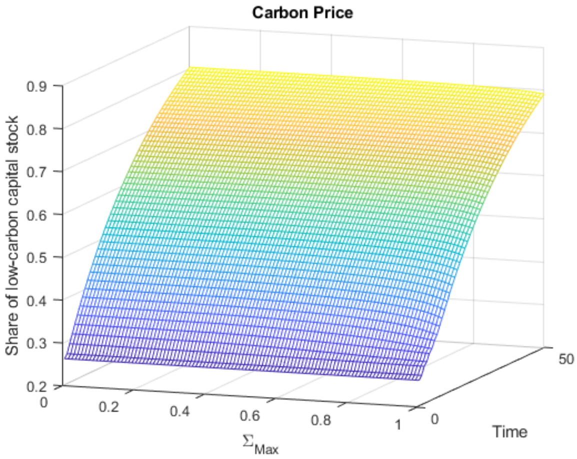
(a) No-policy benchmark



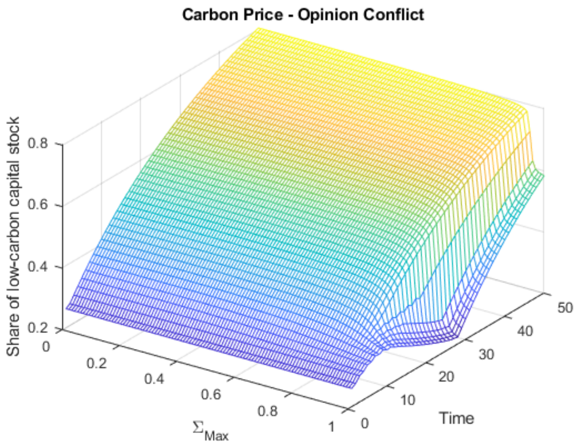
(b) Narrative update



(c) More ambitious targets for renewables



(d) Univocal carbon price



(e) Carbon Price - Opinion Conflict

Figure 11: Different policy options in the model

As can be seen, even for relatively high levels of dissent, applying a carbon price is sufficient to ensure a successful transition towards a low-carbon electricity system. However, for high dissent levels, carbon pricing is first inefficient to propel a transition, until a point from which the tide turns: When the direct —- not expected — carbon price becomes high enough to shift investment behaviours. The mechanism is as follows: Since, for high dissent levels, investment goes more timidly towards low-carbon energy sources in early periods of the run, the regulator's credibility is harmed, hence reducing the share of agents believing in a sustained carbon price. As a result, agents only incorporate the pecuniary incentive with a lag —- from a technical standpoint, their carbon price expectations become myopic —- reducing its effect through the expectation channel. At some point, the carbon price becomes strong enough for investment behaviours to shift towards low-carbon energy sources.

These effects seem to arise for a reduced constellation of parameters but they are nonetheless of interest. As can be seen, the share of low-carbon energy sources when carbon pricing is inefficient is lower than when it is by a rough 15-20 %. This is significant, as this means that in a situation of strong dissent, an economy can fall short of ambitious targets in the long run by a sizeable margin. More precisely, the carbon pricing policy, if driven consistently in spite of agents' expectations, is efficient with a significant lag, putting in jeopardy the sustainability of the economy if targets must be fulfilled within a well-defined time window. What's more, such results were obtained with relatively fluid investment behaviours, and with no limitation to the technical incorporation of low-carbon (renewable) energy sources into the system. In fact, investment behaviours may suffer from incumbency biases (Unruh, 2000) and renewables may face technical constraints, e.g. grid management and batteries that may force the prolonging of some high-carbon infrastructures to ensure baseload servicing of demand.

5. Conclusion

Achieving the transition to a low-carbon electricity system will require coordinated efforts by investors, electricity producers, and policy makers alike. Despite encouraging market signals of increased profitability of renewables, efforts need to be greatly intensified worldwide respect the Paris Agreement. The importance of competing narratives about the future, of beliefs in targets set by the regulator, and of dissent about how the future sequence of economic events will unfold has been largely neglected in macroeconomic modelling efforts so far.

To fill this gap in the literature, we have constructed a novel macroeconomic model where we explicitly use fictional expectation formation mechanisms with agents imagining the future according to different narratives. These narratives are at the intersection of backward and forward-looking expectations. These narratives are then allowed to compete with each other, and interact with official targets set by the regulator. To sharpen our analysis, we have decided to construct an analytical framework where future profitability considerations are more driven by expected stranding of capital stock and the overall state of the transition than by explicit cost development paths. We show that competition of different narratives, dispersion of opinion and the speed of the transition co-determine whether the transition takes place, or whether agents revert to a portfolio focused on high-carbon capital stock.

The main lesson that can be drawn from this broad exercise is that dissent on key economic outcomes linked to the transition (such as asset stranding), and conflicts over broad transitional narratives can significantly hamper the penetration of adverse technologies. While setting ambitious targets is, within our framework, sufficient to solve any dissent issue, redemption from a cognitive or normative lock-in situation is complex. As a result, other kind of policies can serve as complements to carbon pricing aimed at defusing biases. It is also to be noted that adverse dynamics only emerged for relatively high levels of dissent. However, this conclusion depends heavily on the pace of the growth of electricity demand, and on the fact that the European low-carbon transition is assumed to be already well under way. Assuming a higher growth in energy demand and starting from lower shares of low-carbon capital and/or investment yield adverse dynamics much more easily in the model, i.e. for lower levels of dissent. Generally speaking, dissent can be highly disturbing for social dynamics and the consistency of the regulator's stance matters enormously. The current covid-19 crisis provides a striking example: Social adherence to mask wearing, respect of barriers gestures and social distancing is a key determinant of the efficiency of health policies.

The reception of these public health advice is in turn determined to a great extent by the consistency of experts and government communication to shape an adequate narrative in a context of uncertainty about the pandemic Rajkhowa (2020); Morgan (2020).

Of course, our present work is not without limitations. Our assumption of perfectly anticipated total demand schedule is at odds with our emphasis on uncertainty and less-than-rational expectations. Lifting this hypothesis, originally made for the sake of simplicity, is therefore an obvious future direction. The same goes for our hypotheses on price and technology coefficient stability. A contemplated solution is to explicitly write a mental model of the economy for our agents, which would be able to generate complex and possibly conflicting expectations about the future. Finally, other obvious avenues of research would be the integration of this electricity-sector module within a broader macroeconomic model, or porting our forward-looking expectation methodology to other economic questions, related or not to the low-carbon transition. At any rate, this calls further exciting and, we hope, relevant developments.

References

Anufriev, M. and C. Hommes (2012, November). Evolutionary Selection of Individual Expectations and Aggregate Outcomes in Asset Pricing Experiments. *American Economic Journal: Microeconomics 4*(4), 35–64.

Bachmann, R., S. Elstner, and E. R. Sims (2013). Uncertainty and Economic Activity: Evidence from Business Survey Data. *American Economic Journal: Macroeconomics 5*(2), 217–49.

Beckert, J. (2013a). Imagined futures: Fictional expectations in the economy. *Theory and Society 42*(3), 219–240.

Beckert, J. (2013b). Imagined Futures: Fictional Expectations in the Economy. *Theory and Society 42*(3), 219–240.

Beckert, J. (2016). *Imagined Futures* (First ed.). Cambridge, Massachusetts: Harvard University Press.

Beckert, J. and R. Bronk (2018). *Uncertain Futures: Imaginaries, Narratives, and Calculation in the Economy.* Oxford University Press.

Beckett, S. (1956). *Waiting for Godot* (1st English ed.). London: Faber and Faber.

Bordalo, P., N. Gennaioli, R. L. Porta, and A. Shleifer (2019). Diagnostic Expectations and Stock Returns. *The Journal of Finance 74*(6), 2839–2874.

Bordalo, P., N. Gennaioli, and A. Shleifer (2018). Diagnostic Expectations and Credit Cycles. *The Journal of Finance 73*(1), 199–227.

Boyer, R. (2018). Expectations, Narratives, and Socio-Economic Regimes. In *Uncertain Futures. Imaginaries, Narratives and Calculation in the Economy (Chapter 2).* Oxford University Press.

Bronk, R. (2009). *The Romantic Economist: Imagination in Economics.* Cambridge: Cambridge University Press.

Caldecott, B. (2017, January). Introduction to special issue: Stranded assets and the environment. *Journal of Sustainable Finance & Investment 7*(1), 1–13.

Campiglio, E., Y. Dafermos, P. Monnin, J. Ryan-Collins, G. Schotten, and M. Tanaka (2018, June). Climate change challenges for central banks and financial regulators. *Nature Climate Change 8*(6), 462–468.

Davidson, P. (1991, February). Is Probability Theory Relevant for Uncertainty? A Post Keynesian Perspective. *Journal of Economic Perspectives 5*(1), 129–143.

De Grauwe, P. (2012). Booms and busts in economic activity: A behavioral explanation. *Journal of Economic Behavior & Organization 83*(3), 484–501.

De Grauwe, P. and C. Macchiarelli (2015). Animal spirits and credit cycles. *Journal of Economic Dynamics and Control 59*, 95–117.

De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann (1990). Noise Trader Risk in Financial Markets. *Journal of Political Economy 98*(4), 703–738. Publisher: University of Chicago Press.

Dequech, D. (1999, March). Expectations and Confidence under Uncertainty. *Journal of Post Keynesian Economics 21*(3), 415–430.

DNVGL (2020). Energy Transition Outlook 2020. Technical report, DNV GL.

Dosi, G., M. Pereira, A. Roventini, and M. Virgillito (2017, August). When more flexibility yields more fragility: The microfoundations of Keynesian aggregate unemployment. *Journal of Economic Dynamics and Control 81*, 162–186.

Dosi, G., A. Roventini, and E. Russo (2019, April). Endogenous growth and global divergence in a multi-country agent-based model. *Journal of Economic Dynamics and Control 101*, 101–129.

ETC (2018). Reaching net-zero carbon emissions: Mission Possible. Technical report, Energy Transitions Commission (ETC).

Fagiolo, G. and A. Roventini (2017). Macroeconomic Policy in DSGE and Agent-Based Models Redux: New Developments and Challenges Ahead. *Journal of Artificial Societies and Social Simulation 20(1).*

Franke, R. and F. Westerhoff (2017, December). Taking Stock: A Rigorous Modelling of Animal Spirits in Macroeconomics. *Journal of Economic Surveys 31*(5), 1152–1182.

Frydman, R., M. D. Goldberg, and N. Mangee (2015). Knightian uncertainty and stock-price movements: Why the REH present-value model failed empirically. *Economics 9*(1). ISBN: 1864-6042 Publisher: Sciendo.

Geroski, P. (2000). Models of technology diffusion. *Research Policy 29*, 603–625.

Gigerenzer, G. (2015). *Risk Savvy: How to Make Good Decisions.* Allen Lane.

Hafner, S., A. Anger-Kraavi, I. Monasterolo, and A. Jones (2020). Emergence of New Economics Energy Transition Models: A Review. *Ecological Economics 177*, 106779.

Haldane, A. G. (2016). The dappled world. *GLS Shackle Biennial Memorial Lecture*.

Haldane, A. G. and A. E. Turrell (2018). An interdisciplinary model for macroeconomics. *Oxford Review of Economic Policy 34*(1-2), 219–251.

Haldane, A. G. and A. E. Turrell (2019, March). Drawing on different disciplines: Macroeconomic agent-based models. *Journal of Evolutionary Economics 29*(1), 39–66.

Hommes, C. H. (2018). Behavioral & experimental macroeconomics and policy analysis: A complex systems approach. *Journal of Economic Literature (forthcoming)*.

IAEA (2019). Transitions to low carbon electricity systems: Key economic and investment trends. Technology Report, International Atomic Energy Agency.

IEA (2020). World Energy Investment 2020. Flagship Report, International Energy Agency.

Ilut, C. and H. Saijo (2020, January). Learning, Confidence, and Business Cycles. *Journal of Monetary Economics*.

IRENA (2020). Global Renewables Outlook: Energy transformation 2050. Technical report.

Kander, A., P. Malanima, and P. Warde (2013). *Power to the people: energy in Europe over the last five centuries*. The Princeton economic history of the Western world. Princeton, N.J: Princeton University Press. OCLC: ocn841050211.

Lachmann, L. M. (1943). The role of expectations in economics as a social science. *Economica 10*(37), 12–23. ISBN: 0013-0427 Publisher: JSTOR.

Lavoie, M. and G. Daigle (2011, July). A Behavioural Finance Model of Exchange Rate Expectations within a Stock-Flow Consistent Framework: Exchange Rate Expectations in SFC Model. *Metroeconomica 62*(3), 434–458.

Lux, T. (2009, November). Rational forecasts or social opinion dynamics? Identification of interaction effects in a business climate survey. *Journal of Economic Behavior & Organization 72*(2), 638–655.

Meade, N. and I. Towhidul (2006). Modelling and forecasting the diffusion of innovation - A 25 year review. *International Journal of Forecasting 22*, 519–545.

Morgan, M. (2020). Why Meaning-Making Matters: The Case of the UK Government's COVID-19 Response. *American Journal of Cultural Sociology 8*(3), 270–323.

Rajkhowa, A. (2020). COVID-19 Dissensus in Australia: Negotiating Uncertainty in Public Health Communication and Media Commentary on a Novel Pandemic. *Pacific Journalism Review : Te Koakoa 26*(1), 253–263.

Riahi, K., D. P. van Vuuren, E. Kriegler, J. Edmonds, B. C. O'Neill, S. Fujimori, N. Bauer, K. Calvin, R. Dellink, O. Fricko, W. Lutz, A. Popp, J. C. Cuaresma, S. Kc, M. Leimbach, L. Jiang, T. Kram, S. Rao, J. Emmerling, K. Ebi, T. Hasegawa, P. Havlik, F. Humpenöder, L. A. Da Silva, S. Smith, E. Stehfest, V. Bosetti, J. Eom, D. Gernaat, T. Masui, J. Rogelj, J. Strefler, L. Drouet, V. Krey, G. Luderer, M. Harmsen, K. Takahashi, L. Baumstark, J. C. Doelman, M. Kainuma, Z. Klimont, G. Marangoni, H. Lotze-Campen, M. Obersteiner, A. Tabeau, and M. Tavoni (2017, January). The Shared Socioeconomic Pathways and their energy, land use, and greenhouse gas emissions implications: An overview. *Global Environmental Change 42*, 153–168.

Shackle, G. L. (1972). Economics and Epistemics. *New Brunswick, NJ: Transaction Publishers*.

Shackle, G. L. S. (1970). *Expectation, Enterprise and Profit.* London: George Allen and Unwin.

Simon, H. A. (1955, February). A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics 69*(1), 99.

Stern, N. (2016). Economics: Current climate models are grossly misleading. *Nature News 530*(7591), 407.

Taleb, N. N. (2010, May). *The Black Swan: The Impact of the Highly Improbable: The Impact of the Highly Improbable: With a New Section: "on Robustness and Fragility"* (2nd edition ed.). New York: Random House Trade Paperbacks.

Appendix A. The Electricity Sector

Consistently with many technological change frameworks (Keramidas et al., 2017) the evolution of energy demand e_d is purely exogenous:

$$\forall t > 1, \; e_t^d = (1 + g_e)\, e^d\,(t-1) = \mathbb{E}_t(e^d), \tag{A.1}$$

where g_e is the constant exogenous growth rate of energy demand and t denotes the time period. The model has a conventional time step of one year and electricity demand is labelled in TWh per year.. We assume that agents know with certainty the path of future energy demand.

We assume two types of electricity-production capital: a high-carbon K_H and a low-carbon K_L one. They are characterised by Leontief fixed-coefficient technologies. In the model, the capital stock is labelled in GW of energy capacity.

The two capital stocks are not always fully utilised: There will be generally some slack in generating capacities. This allows us to define a utilization rate for the two types of capital:

$$u_{L_t} = \frac{e_{L_t}}{e^s_{L_t}} u_{H_t} = \frac{e_{H_t}}{e^s_{H_t}}, \tag{A.2}$$

where e_{L_t} and e_{H_t} are, respectively, the energy effectively produced from low- and high-carbon energy sources, and $e^s_{L_t}$ and $e^s_{H_t}$ are full-capacity output. The model is calibrated to ensure that $u_{L_t} = \overline{u_L} = 1 \; \forall t$. Indeed, in our calibration the low-carbon capital stock is identified to renewable energy sources (mainly wind and solar), whose utilisation cannot be controlled at will. We therefore assume that ξ_{L_t}, the amount of TWh a GW of low-carbon capacity can produce per year is corrected by a capacity factor averaging out the variability of modern renewable availability over a year.

This allows us to define full-capacity output for low-carbon capital as:

$$e^s_{L_t} = \; \xi_{L_t} K_{L_t} \tag{A.3}$$

Similarly, full-capacity output for high-carbon capital is defined as:

$$e^s_{H_t} = \xi_{H_t} K_{H_t}, \tag{A.4}$$

where ξ_{H_t} is defined as the amount of amount of TWh a GW of high-carbon capacity can produce per year. As the production of these types of capital can be modulated at near-will, we assume that agents can react to changes in demand, and that the utilization rate of high-carbon capital can vary. We suppose for the sake of simplicity that $\xi_{L_t} = \xi_L$ and $\xi_{H_t} = \xi_H$: The Leontief coefficients are constant.

The quantity of energy demanded is then dispatched according to a merit-order mechanism: Low-carbon technologies are mobilised in priority, and high-carbon capital absorbs the remnant. Labelling e_L and e_H the energy effectively produced by low- and high-carbon technology respectively, we define:

$$e_{L_t} = \begin{cases} 0 \; if \; K_{L_t} = 0 \\ e_{d_t} \; if K_{L_t} > 0 \; and \; e_{d_t} \leq e^s_{L_t} \\ e^s_{L_t} \; if K_{L_t} > 0 \; and \; e_{d_t} > e^s_{L_t} \end{cases} \tag{A.5}$$

and:

$$e_{H_t} = \begin{cases} 0 \; if \; K_{H_t} = 0 \\ 0 \; if \; K_{H_t} > 0 \; and \; e^d_t \leq e^s_{H_t} \\ e^d_t - e^s_{L_t} \quad if \; K_{H_t} > 0 \; and \; e^s_{L_t} \leq e^d_t < \left(e^s_{L_t} + e^s_{H_t}\right) \\ e^s_{L_t} \; if \; K_{H_t} > 0 \; and \; e^d_t \geq \left(e^s_{L_t} + e^s_{H_t}\right) \end{cases} \tag{A.6}$$

Low-carbon electricity serves the market first and as mentioned above, the model is calibrated so that $e_{d_t} \geq e^s_{L_t}$ at any point in time. This ensures full capacity utilization for low-carbon capital.

We assume away any labour input for the production of electricity and suppose that low-carbon technologies do not need marketed inputs to generate electricity. Conversely, high-carbon capital requires a fossil-fuel input to work. We define the "thermal efficiency" of the high-carbon stock ξ_{f_t} as the amount of non-renewable input (labelled in trillion British Thermal Units (BTUs)) needed to generate a TWh of electricity with high-carbon energy. This non-renewable input includes fossil inputs as well as uranium and biomass. We hold ξ_{f_t} as constant ($\xi_{f_t} = \xi_f$).

Concerning pricing, we drift from the usual merit-order process by assuming a unique, administered electricity price $p_{E_t} = p_E$. It is not derived from marginal pricing but on a public-private partnership calibrated from EU data (see below).

This allows firms to generate profits, computed as income minus expenses per unit of production. Production costs are first intermediary inputs for high-carbon electricity production: p_{f_t} per unit of non-renewable input. We hold this price constant throughout ($p_{f_t} = p_f$).

Second, We also consider two kinds of capital cost: financing and stranding costs. Firms must take up on loans to fund their investment expenditures that we measure per unit of capital. We therefore define the costs for both types of capital:

$$d^r_{L_t} = \alpha_{L_t}\psi_{L_t}c^k_{L_t} \tag{A.7}$$

$$d^r_{H_t} = \alpha_{H_t}\psi_{H_t}c^k_{H_t} \tag{A.8}$$

The variables $\alpha_{L_t} = \alpha_L$ and $\alpha_{H_t} = \alpha_H$ are a measure of yearly debt repayment including interest and principal. $\psi_{L_t} = \psi_L$ and $\psi_{H_t} = \psi_H$ are calibrated debt-to-capital ratios. Finally, $c^k_{L_t} = c^k_L$ and $c^k_{H_t} = c^k_H$ are the fixed costs of a unit of capital, that is the fixed cost of installing a GW of capacity.

Second, we include a metrics of capital losses incurred due to capital stranding. By assuming that a GW of capacity is valued at its market cost, capital losses, that only exist for high-carbon assets, are defined as:

$$c^l_{H_t} = \max\left(0, c^k_{H_t}\left(u^\tau_H - u_{H_t}\right)\right), \tag{A.9}$$

where u^τ_H is a conventional utilization rate, set to .85, a level roughly consistent with the structural utilization rates in advanced economies (Setterfield, 2019). No capital gains are incurred if the utilization rate happens to go higher than the conventional rate.

As a result, profit rates are defined as:

$$r_{L_t} = p_E\xi_L u_{L_t}\xi_L - d^r_{L_t} = p_E\xi_L - d^r_{L_t} \tag{A.10}$$

$$r_{H_t} = \left(p_E - \frac{p_f}{\xi_f}\right)\xi_H u_{H_t} - d^r_{H_t} - c^l_{H_t} \tag{A.11}$$

Appendix B. Expectations of Stranding

At any point in time, the actual share h_{E_t} is also equal to:

$$h_{E_t} = \frac{e^d_{H_t}}{e_{d_t}} = \frac{\xi_H u_{H_t} K_{H_t}}{e_{d_t}} \tag{B.1}$$

In their expectations, agents can use this equation and its equivalences:

$$h^N_{E_s} = \frac{\xi_H\,\mathbb{E}_t(u_{H_s})K^e_{H_s}}{e^e_{d_s}} \tag{B.2}$$

$$\Leftrightarrow u^e_{H_s} = \frac{h^N_{E_s}e^e_{d_s}}{\xi_H K^e_{H_s}} \tag{B.3}$$

$$\Leftrightarrow \; K^e_{H_s} = \frac{h^N_{E_s}e^e_{d_s}}{\xi_H\,\mathbb{E}_t(u_{H_s})} \tag{B.4}$$

These equations can be used to derive a high-carbon capital stock and utilization schedule period by period according to certain behavioural rules. That is, taking the narrative $h^N_{E_s}$ as a reference, agents will check at every point in time whether leaving the capital stock to depreciate naturally is consistent with the narrative's high-carbon energy share if the utilization rate is left at its current level. Formally, this writes:

1. If $\frac{\xi_H\,\mathbb{E}_t(u_{H_s})(1-\delta_H)\,\mathbb{E}_t(K_{H_s})}{e^e_{d_s}} < h^N_{E_s}$, there is possibly a need for more capital. More precisely, we assume that agents consider u^τ_H as a conventional anchor, and effectuate another case disjunction:
   (a) If $\mathbb{E}_t(u_{H_s}) < u^\tau_H$, $\mathbb{E}_t(K_{H_s}) = (1-\delta_H)\,\mathbb{E}_t(K_{H_s})$, that is agents expect that capital will be left to depreciate naturally
   (b) If $\mathbb{E}_t(u_{H_s}) \leq u^\tau_H$, $\mathbb{E}_t(K_{H_s}) = \frac{h^N_{E_s} e^e_{d_s}}{\xi_H u^\tau_H}$, that is, agents modulate their capital stock in order to stay at the target utilization rate or return to it after a period above it.
2. If $\frac{\xi_H\,\mathbb{E}_t(u_{H_s})\,\mathbb{E}_t(K_{H_s}}{e^e_{d_s}}) = \ h^N_E\,(t, t+s+1)$, then the capital stock is just right, and is left to depreciate naturally
3. $\frac{\xi_H\,\mathbb{E}_t(u_{H_s})\,\mathbb{E}_t(K_{H_s}}{e^e_{d_s}})\ h^N_E\,(t, t+s+1)$, agents cannot but leave the capital stock to depreciate naturally.

Note that we make here three assumptions:

1. The future demand schedule is known
2. Productivity is known and known to be constant
3. Companies do not buy more capital than what they need, that is, they always invest consistently with the expected utilization rate.

This allows us to determine a high-carbon utilization rate schedule in time S at any chronological time t that we plug back into Equation $(3.24b)$.

Appendix C. The Probit Model

We consider an underlying population of agents that we assume to be numerous. Each of them formulates expectations on future profit rates (that is, profits per unit of *capital*) for the two types of capital over their planning horizon for each period in psychological time s. These expectations are made of a common component that all agents share, and of an idiosyncratic part that represents the agent's specific opinion about future states of the world. Formally, the most general case for agent i is as follows:

$$\mathbb{E}_t(\pi_{L_{i_s}}) = \mathbb{E}_t(p_{E_s})\,\mathbb{E}(\xi_{L_s})\,\mathbb{E}_t(u_{L_s}) - \mathbb{E}_t\,\psi_L\,\mathbb{E}_t(c_{k_{L_s}}) + \epsilon_{L_{i_s}} \tag{C.1}$$

$$= \mathbb{E}(\gamma_{L_t})\,\mathbb{E}(u_{L_s}) - \mathbb{E}(dr_{L_s}) + \ \epsilon_{L_{i_s}} \tag{C.2}$$

$$\mathbb{E}_t(\pi_{H_i})_s = \left(\mathbb{E}_t(p_{E_s}) - \frac{\mathbb{E}_t(p_{f_s})}{\xi^e_{f_s}}\right)\ (\xi_{H_s})\,\mathbb{E}_t(u_{H_s}) - \mathbb{E}_t(\alpha_{H_s})\psi_H\,\mathbb{E}_t(c_{k_{H_s}})$$
$$-\ \max\left(0, \mathbb{E}_t(c_{k_{H_s}})\,(u^\tau_H -\ \mathbb{E}(u_{H_s}))\right)\ \ + \epsilon_{H_{i_s}} \tag{C.3}$$
$$= \gamma_{H_{i_s}}\,\mathbb{E}_t(u_{H_s}) - dr_{H_s} -\ \max\left(0, \mathbb{E}_t(c_{k_{H_s}})\,(u^\tau_H -\ \mathbb{E}_t(u_{H_s}))\right)$$
$$+\ \epsilon^e_{H_{i_s}} \tag{C.4}$$

For both technologies, we assume that energy prices p_E, fossil fuel prices p_{FF} andcapital costs c_{k_i} and capital recovery factors are the same for all investors, and constant through time s and t. They are therefore perfectly anticipated by agents. As a result, $\mathbb{E}(\gamma_{H_{i_s}}) = \gamma_{H_{i_s}}$ and $\mathbb{E}(\gamma_{L_{i_s}}) = \gamma_{L_{i_s}}$.

We further assume that the idiosyncratic components $\epsilon^e_{L_{i_s}}$ and $\epsilon^e_{H_{i_s}}$ are normally distributed across the population, with mean zero and given variances. To come up with a synthetic metrics, agent consider the discounted sum of expected profit rates, which gives them a measure of expected unit profit flows per unit of capital:

$$r_{L_t} = \sum_{s=1}^{S} \beta^s \left(\gamma_{L_t} u^e_{L_s} - dr_{L_s} + \epsilon^e_{L_{i_s}}\right) \tag{C.5}$$

$$r_{H_t} = \sum_{s=1}^{S} \beta^s \begin{pmatrix} \gamma_{H_s} u^e_{H_s} - dr_{H_s} \\ -\max\left(0, c^e_{k_{H_s}}\left(u^\tau_H - u^e_{H_s}\right)\right) + \epsilon^e_{H_{i_s}} \end{pmatrix} \tag{C.6}$$

We insist right from the outset that ϵ should *not* be understood as normally distributed disturbances around an equilibrium as in RBC or DSGE strands of modelling. Our framework is not stochastic: These terms cannot be interpreted as uncertain exogenous shocks. Rather, they figure dissent or consensus about how a given economic variable, here the profit rate, will behave in the future. If they are all equal to zero, that is they are not random and their σ are all nil, agents are perfectly coordinated around the deterministic component. Conversely, the higher the σ, the less coordinated they are around the deterministic component, and the more dissent there is. When uncertainty is radical, as with long-run expectations (Dequech, 2004), it is best figured by opinion dissent around a given projection rather than with stochastic exogenous disturbances that are more related to short-term, business-cycle considerations (Christiano et al., 2018; Dullien, 2017; Smets and Wouters, 2007).

Since we know the underlying distribution of opinions, we can derive the proportion of the population for which the aggregate profit rate for the low-carbon capital is higher than for the high-carbon capital:

$$P\left(r_{L_t} - r_{H_t} > 0\right) = \tag{C.7}$$

$$P\begin{pmatrix} \sum_{s=1}^{S}\left(\frac{1}{(1+\rho)^s}\right)\left(\epsilon^e_{L_{i_s}} - \epsilon^e_{H_{i_s}}\right) \\ > \sum_1^S\left(\frac{1}{(1+\rho)^s}\right)\begin{pmatrix} \gamma^e_{H_s} u^e_{H_s} - dr^e_{H_s} \\ -\max\left(0, c^e_{k_{H_s}}\left(u^\tau_H - u^e_{H_s}\right)\right)\end{pmatrix} - \begin{pmatrix}\gamma^e_{L_s} u^e_{L_s} \\ -dr^e_{L_s}\end{pmatrix}\end{pmatrix}$$

By making the assumption that the ϵ are not serially correlated in psychological time[9], we can state that $\sum_{s=1}^{S}\left(\frac{1}{(1+\rho)^s}\right)\left(\epsilon^e_{L_{i_s}} - \epsilon^e_{H_{i_s}}\right)$ follows a normal distribution of mean 0 and variance $\sum_{s=1}^{S}\left(\frac{1}{(1+\rho)^s}\right)\left(\left(\sigma^e_{L_s}\right)^2 - \left(\sigma^e_{H_s}\right)^2\right)$, this will give us the proportion of agents considering that the low-carbon type of capital will be more profitable than the high-carbon energy source.

Although the normality of expectation distributions is documented for some key economic variables, both in the short- and long-run (Gillingham et al., 2018), the choice of a normal distribution for the idiosyncratic term is an important simplification of what may happen in reality. This was mainly motivated by the additive stability of this distribution: This allows us to easily derive a discounted unit profit metrics over the agents' planning horizon. Further work will consist in adopting more realistic opinion distributions that are additively stable. In particular, the Lévy distribution, already used in the econophysics of wealth distribution (Levy and Solomon, 1997; Mandelbrot, 1960) and that has a workable functional form is a promising candidate.

This methodological choice is not without posing conceptual issues. Indeed, a normal distribution of mean zero and variance σ is defined on $\mathbb{R}$, and takes 99% of its values within a $[-3\sigma,\ 3\sigma]$ interval. As such, too high values for the $\sigma^e_{L_s}$ and $\sigma^e_{H_s}$ may figure agents having conceptually unsound expectations. For instance, if $\sigma^e_{L_s} = 1$, which would be the commonplace case of a standard normal distribution, we would model a population in which 99% of agents expect low-carbon profit rates ranging between $\gamma_{L_t} u^e_{L_s} - dr^e_{L_s} - 3$ and $\gamma_{L_t} u^e_{L_s} - dr^e_{L_s} + 3$. Extremes would therefore expect profit rates to be above 300% profit rates and slightly above -300% profit rates, hardly a realistic outcome.

[9] In the general case, agents therefore will not have monotonous expectations in psychological time.

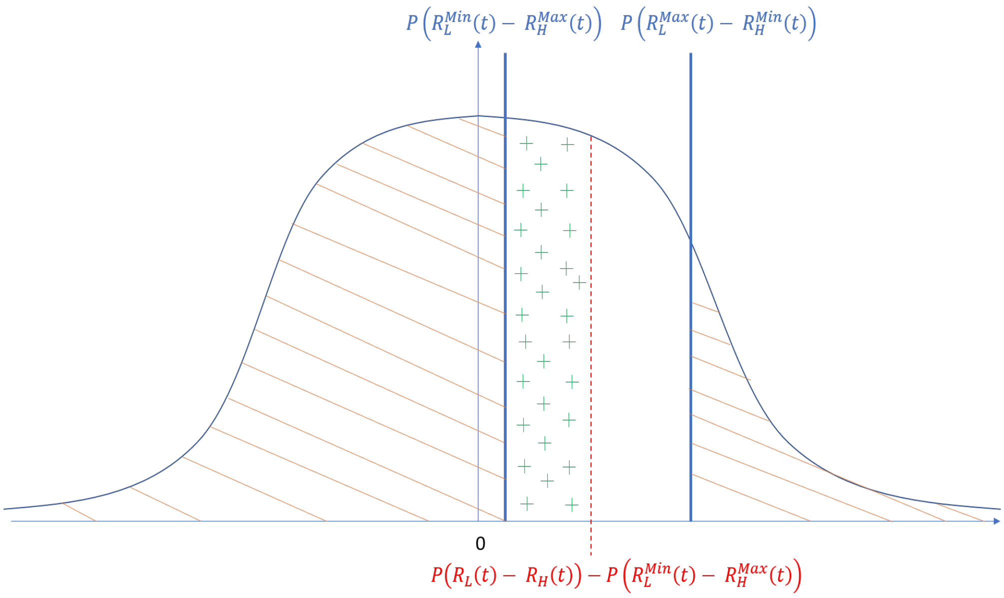


Figure C.12: A stylised censoring process

To make sure that expectations remain within conceptually sound boundaries, we therefore need to slightly complexify our treatment of the normal distribution. A first-best solution would be to consider a truncated normal distribution, which would take all its values between two conceptually sound values. Yet, a truncated normal distribution loses the additivity property of a full normal distribution, hence hampering the analytical tractability of our model. We therefore apply a second-best solution by *censoring* the normal distribution.

We derive minimum ($R_{L_t}^{Min}$ and $R_{H_t}^{Min}$) and maximum ($R_{L_t}^{Max}$ and $R_{H_t}^{Max}$) expected profit rates, and therefore minimum and maximum spreads between the two synthetic profit rates, respectively $R_{L_t}^{Min} - R_{H_t}^{Max}$ and $R_{L_t}^{Max} - R_{H_t}^{Min}$. Censoring the distribution consists in not considering expectations lying outside of the interval $A = \left[\, R_{L_t}^{Min} - R_{H_t}^{Max}, R_{L_t}^{Max} - R_{H_t}^{Min}\right]$. Let us define first:

$$\forall x, \quad 1_A(x) = \begin{cases} 0 \; if \; x \notin A \\ 1 \; if \; x \in A \end{cases} \quad \text{and} \quad 1_{A|}(x) = \begin{cases} 0 \; if \; x \leq R_{L_t}^{Max} - R_{H_t}^{Min} \\ 1 \; if \; x > R_{L_t}^{Max} - R_{H_t}^{Min} \end{cases} \tag{C.8}$$

We can now define the final proportion $\chi\left(P\left(r_{L_t} - r_{H_t} > 0\right)\right)$:

$$\chi_t = \frac{\begin{pmatrix} P\left(r_{L_t} - r_{H_t} > 0\right) - 1_A\left(r_{L_t} - r_{H_t}\right) P\left(R_{L_t}^{Min} - R_{H_t}^{Max} > 0\right) \\ - 1_{A|}\left(r_{L_t} - r_{H_t}\right) P\left(R_{L_t}^{Min} - R_{H_t}^{Max} > 0\right) \end{pmatrix}}{P\left(R_{L_t}^{Max} - R_{H_t}^{Min}\right) - P\left(\, R_{L_t}^{Min} - R_{H_t}^{Max}\right)} \tag{C.9}$$

The $P\left(R_{L_t}^{Max} - R_{H_t}^{Min}\right) - P\left(\, R_{L_t}^{Min} - R_{H_t}^{Max}\right)$ term is meant to normalise the number in the numerator to obtain a proportion between 0 and 1. Intuitively, the χ function can therefore be understood as the transformation of the basic proportion in (18*a*) into a conditional probability over a given interval, that can be decided based on past data of the relevant variable. This transformation allows us to be totally free in the choice of σ. Figure C.12 shows a graphical representation of the process.

These elements represent the most general version possible of the model. It would entail a strong degree of complexity, render the model less analytically tractable and make simulations less intuitive. We thus chose to simplify our treatment of uncertainty by focusing on high-carbon capital stranding expectations. Therefore, in the simulations in section 4, we only consider the maximum level of uncertainty of high-carbon capital stranding and term it $\sigma_{u,max}$.

We further simplify this treatment of expectations by assuming no heterogeneity surrounding the expectations of low carbon utilisation rate. We indeed assume that agents only formulate expectations on future utilisation rates for the high carbon sector and that idiosyncratic expectations are only related to high-carbon utilisation. Formally, this allows us to rewrite expected profit rates as follows:

$$r_{L_{i_s}} = \gamma_{L_t} - dr_{L_s} \tag{C.10}$$

$$r_{H_{i_s}} = \gamma_{H_t} \left(u^*_{H_s} + \epsilon_{H_{i_s}} \right) - dr_{H_s} \tag{C.11}$$

We assume that the $\epsilon^e_{H_{i_s}}$ are normally distributed through a random variable $\epsilon^e_{H_s}$ with mean zero and variance $\left(\sigma_{H_{i_s}}\right)^2$. These simplifications made, the principle of the module is exactly the same. After summing up expected profit rates and discounting, we consider the share of agents expecting high-carbon capital to be less profitable than low-carbon capital using the fact that the $\epsilon_{H_{i_s}}$ are normally distributed.

Furthermore, since consensus/dissent regards only stranding degrees within this framework, we assume that agents do not consider utilisation rates going above the long-run conventional rate u^τ_H. We therefore consider the following boundaries for r_{H_s}:

$$r_{H_{Min\,s}} = -dr_{H_s} \quad (3.25a) \qquad r_{H_{Max\,s}} = u^\tau_H \gamma_{H_t} - dr_{H_s} \quad (3.25b)$$

and apply the χ operator defined above.

Appendix D. Calibration

Appendix D.1. parameter values

Parameter	Meaning	Value	Source and justification
g_e	Growth of energy demand	0.0048	Energy Brainpool reference scenario (2019), amounts to a 17% increase in energy consumption by 2050
ξ_H	Leontief coefficient for high-carbon energy	3.689	EIA (2017)
xi_L	Leontief coefficient for low-carbon energy	2.628	EIA (2017)
ξ_F	Leontief coefficient for fossil fuel	0.1143	EIA (2017)
c_H^k	Unit capital cost for high-carbon energy	0.7	AIE (2017)
c_H^k	Unit capital cost for low-carbon energy	1.5	AIE (2017)
u_H^τ	Conventional utilization rate	0.85	FRED (2018)
δ_{KH}	Depreciation rate for high-carbon capital	0.02	Convention
δ_{KL}	Depreciation rate for low-carbon capital	0.02	Convention
ψ_H	Debt-to-investment ratio for high-carbon investment	0.5	Convention
ψ_L	Debt-to-investment ratio for high-carbon investment	0.8	Convention
LT	Loan tenure rate	20	
c_o^{price}	Rescaled carbon price (starting value)	0.003	Conseil d'analyse Stratégique (2018), OCDE (2018)
ϵ_f	Growth rate of carbon price	0.0689	Conseil d'analyse Stratégique (2018), OCDE (2018)
ρ	Discount rate	0	Convention
lb_{uh}	Lower bound for utilization rate in censorship process	0	Convention
hb_{uh}	Higher bound for utilization rate in censorship process	0.85 (+	Convention
S	Planning horizon length	20	Convention
M	Evaluation process range	20	Convention
ℓ	Long-run transition goal	1	Convention
ν	Natural switching rate	0.001	Convention
β_σ	Link between opinion variance and stranding expectation variance	1	Convention
b_N	Regulator's narrative intrinsic growth rate	0.1213	European Commission (2017)
L_H^{Markup}	Markup on loans to high-carbon investment	0.03	Convention
L_L^{Markup}	Markup on loans to low-carbon investment	0.03	Convention
r_F^int	Interest rate	0.02	Convention
ω	Memory parameter	0.85	Usual value in related literatures
r_{log_μ}	Intrisinc growth rate for dissent in time S	1	Keynes (1936)
ρ_L	Transition premium value	1	Convention
p_f	Price of the fossil fuel input	0.005	EIA (2017)
p_E	Price of the electricity output	0.2	European Commission (2017)

Appendix D.2. Initial values

Starting values			
Variable	Meaning	Starting Value	Source and justification
e_d	Energy demand	3000	European Commission (2017)
K_H	High-carbon capital stock	738.8325	Own calculations from low-carbon stock of capital
K_L	Low-carbon capital stock	260	European Commission (2017)
ℓ_I	Share of low-carbon investment	0.6	European Commission (2017)
u_L	Low-carbon utilization rate	1	Convention
u_H	High-carbon utilization rate	0.85	FRED (2018)
j	Opinion dynamic indicator	0	Convention

Table D.1: Initial values for EU calibration