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

Unfinished work in progress version for Momentum2020:Republik – please do not quote or circulate!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] *Institute for Advanced Studies, Vienna, Austria*
[e] *International Institute for Applied Systems Analysis, Laxenburg, Austria*

Abstract

To study the decarbonisation of physical capital, 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 along several dimensions. Firstly, we innovate the view on the role of expectations in transition processes by developing a novel bounded rational but forward-looking expectation formation mechanism based on the notion of *fictional expectations*. Secondly, this allows us to combine long-term views on system change with more short-term and profit-oriented investment logics in a coherent framework. Thirdly, we incorporate discrete choice theory based on probabilistic distributions of future utilization of capital stocks and profits 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 analytical view on how stranding of physical assets can occur as a phenomenon resulting from coordination problems on a macroeconomic level due to dissensus 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 dissensus, determine the role of opinion conflict for the low-carbon transition, relate energy demand growth to dissensus, and finally simulate different forms of policy interventions including a carbon price in relation to different levels of dissensus.

1. Introduction

The currently ongoing transition to a low-carbon electricity system is slower than necessary to meet the targets of the Paris agreement (IAEA, 2019; DNVGL, 2020; ETC, 2018). This is especially true for the need for an almost-complete electrification of our economies to reach net carbon emissions by 2050: It will require a 4-5 fold increase of global electricity production from slightly above 20,000 TWh in 2020 to 85,000-115,000 by mid-century (ETC, 2018). Certainly, recently encouraging market signals—e.g. renewables' profitability improving above conventional fossil-powered electricity— 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 energy (including hydroelectricity, solar, wind, geothermal and biofuels) amount for less than a third (26%) of worldwide electricity generation. Carbon-free electricity (renewable + nuclear + other non-carbon fuels) amounts for 37% of total. In 1990 and 2015, the share of renewable energy in electricity output amounted for 19.4% and 22.9% (World Development Indicators). This shows that while renewable electricity growth might have increased in the past five years, the share of renewables in worldwide primary energy consumption remains very low at a share of 5% (BP Review, 2019). In spite (or maybe because) of these signs of hope, waiting for the transition to a renewable energy system is then akin to waiting for Godot (Beckett, 1956). We spend our lives waiting for something to happen that never actually comes, or that is on its way so slowly that we cannot anticipate when it will arrive. While we are waiting, we seesaw between hope and despair, constantly bemoaning that we are waiting for something that never comes. Taking this fictional narrative from literature to economics and putting it in terms of the energy transition might translate as follows: We should not *wait* for the transition to happen, we must *make* it happen. Making the transition happen will involve both the government and market actors. Indeed, the transition cannot be reduced to a set of relevant expectations by individual actors. Rather, it is about the emergence of collective expectations, which policies will shape to a great extent. We have to align agents' expectations of possible future developments to reach certain outcomes. However, even when aligned within a collective narrative, expectations have often failed to deliver transformations (Boyer, 2018). Nevertheless, such alignment of expectations are necessary to manage the fundamental uncertainty of transition processes, maintaining profitability of investments for businesses and serving societal goals (e.g. the mitigation of climate change) at the same time. The energy transition thus certainly increases the perceived importance of the fundamental uncertainty under which actors have to make their decisions. Expectations are the only possible way conceivable to the human mind to fill the void created by fundamental uncertainty through imagining—rather than predicting—possibles futures. However, agents do not form their expectations in a historical and institutional vacuum. The imagined futures are shaped by the historical experience and the narrower set of relevant behaviours delimited by institutions.

Fundamental uncertainty and narrative structures. The term **fundamental uncertainty**[1] refers to situations "in which at least some essential information about future events *cannot be known at the moment of decision* because the information does not exist and cannot be inferred from any existing data set" (Dequech, 1999, p. 415, emphasis by the authors). The notion has received significant support throughout the history of economic thought (Keynes, 1936; Knight, 1921; Gigerenzer, 2015; Arrow, 2013), and has also been given firm philosophical underpinnings. In particular, the complexity of the world makes it impossible to assign "correct" probabilities to future events. We can estimate those probabilities, and we might be at least roughly right quite often due to path dependencies and the deeper structures of our system governing its dynamics. However, David Hume (Hume, 1748) demonstrated that inference from the past is only a poor guide to the future, especially when the "rules of the game", i.e. the underlying structural determinants of complex systems' change. This, of course, is especially true for processes of socio-economic transition as are the focus of this manuscript. For these reasons, the assumption of fundamental uncertainty of the future is appropriate in the context of as sweeping a social change as the low-carbon transition. After all, the precise course and endpoint of the transition are by no means unequivocal, and future developments can hardly be reduced to a probabilistic inventory of possible outcomes known to agents (Hafner et al., 2020). However, traditional approaches in economics have mostly shunned this notion, and resorted to environments that can be fully described by probability distributions. Although the assumption of the existence of probability distributions is heuristically justified in some instances, the main reason for the rejection of the fundamental uncertainty assumption is that it could represent a path towards nihilism. Indeed, if nothing—or very little—can be said on the future, how relevant can the economic discourse be? There is nonetheless a positive interpretation of fundamental uncertainty. As put by McKenna and Zannoni (2000), assuming that agents face a a fundamentally uncertain future begs the question of they cope with such a state of face. Collectively, it is the role of institutions and money contracts (Keynes, 1936; Minsky, 1986) to provide agents with enough stability to guide their actions, which poses in turn the question of how such institutions, norms, etc. can emerge. At the individual level, agents design heuristics as rules of thumb enabling them to deal with their immediate environment. However, these rules of thumb largely ignore the

[1]Fundamental uncertainty has many names in the literature, which we all see and treat as synonymous in this manuscript. Such alternative terms include "radical uncertainty" (Kay and King, 2020), "deep uncertainty" (Knight, 1921) or even "unknown unknowns" and black swans (Taleb, 2010, 2005).

role of expectations, that is, of how agents formulate hopes and projections about future outcomes. More precisely, institutions and heuristics bring stability, but hardly explain how agents can hope for change, be it social or technical. This question is even more pressing in the context of the low-carbon transition, which will most likely entail sweeping changes across the board. Societies and individuals have been dealing with such a fundamental uncertainty of the future as a basic predicament of the human condition by conceiving, operationalizing and implementing *narrative structures* on an ever grander scale throughout the ages. These narratives coordinate behavior of astonishing numbers of individuals up to the billions and give rise to socially and economically effective outcomes (Becker, 1973; Graeber, 2012; Harari, 2015, 2017, 2019). Collectively conceived and individually re-created, prospective narrative imagination is thus the response by economic agents to fundamental uncertainty. In this respect, Boyer (2018) suggests that the dynamics of capitalism has progressively come to be organised around broad narrative structures: *expectation regimes*. An important point for Boyer is that expectations cannot be reduced to a data-acquisition process, but include elements of imagination and creativity (Shackle, 1972; Dequech, 1999) that frame and delineate the horizon of expectation of agents in terms of technical and social change. From this standpoint, agents collectively organise around broad guidelines that either emerge autonomously, or are provided by key institutions.

A telling demonstration of this principle is the epistemology of the Intergovernmental Panel on Climate Change (IPCC), which 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.[2] 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). The SSPs are then coupled to their energy, land use, and greenhouse gas emissions implications to relate them to their effects on climate change. They represent an excellent example of how the fundamental uncertainty of the future is dealt with by the development of (potentially large-scale) narrative structures that imagine rather than predict the future. Such broad narratives are *focal points* (Schelling, 1960) that agents use as reference points to create their own expectations. Beckert (2016) suggests in this respect that agents *individually* resort to such focal points as *fictions* to guide their own actions, these fictions being discussed, disputed, confronted, and possibly giving rise to Boyer's expectation regimes.

Employing fictional expectations in macroeconomic modelling. Looking into the long term future under conditions of fundamental uncertainty in a transition process thus poses deep methodological problems of how to conceive expectation formation for agents in macroeconomic modeling. In particular, there are very few models that tackle the issue in a similar fashion as proposed by our approach. In such a situation of a lack of a suitable methodology in economics, it might help to look to other disciplines for inspiration. *Fictional expectations* based on different *expectation regimes* have been widely discussed in economic sociology and related disciplines (Beckert, 2013, 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). The existence and importance of narratives is the fundamental reason why we chose to employ ***fictional expectations* explicitly as the expectation formation mechanism for our macroeconomic model** of the transition to renewable electricity. To the best of our knowledge, despite their apparent ability to trace how agents in an economy deal with the fundamental uncertainty, fictional expectations have not yet been applied to formalized economic models as a methodology for expectation formation in aggregate macroeconomic models.

For the reasons sketched above, we explicitly chose to construct a macroeconomic model where agents form imagined future states the economy, which they then apply to their decision making procedures. As central narratives we have chosen firstly 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

[2] 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.

(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 asset 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.

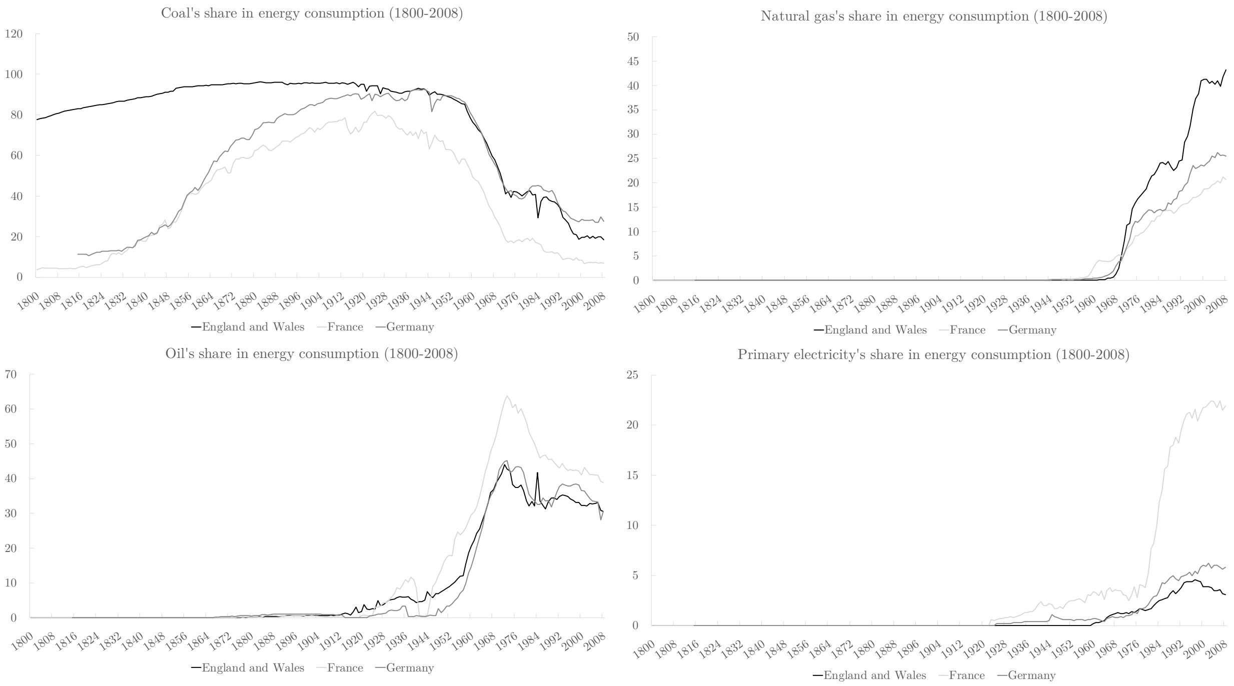


Figure 1: Historical shares of coal, oil, natural gas and primary electricity in energy consumption for England & Wales, France and Germany from 1800 to 2008. Primary electricity refers to electricity from geothermal, hydropower, nuclear power, wind, photovoltaics, tidals, wave and solar thermal. Shares do not sum up to 100% as other enerngy sources are not displayed here. Source: [TO ADD IN BIBFILE] Kander, A., Malanima, P., Warde, P., 2013. Power to the people: energy in Europe over the last five centuries, The Princeton economic history of the Western world. Princeton University Press, Princeton, N.J.

The model. Along these lines, we develop below a **system-dynamics model** (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. Periods are explicitly one-year long, and our simulations are 50-ticks long. This model figures the competition between two types of capital, one polluting and the other low-carbon in penetrating the electricity-production market. The model core is therefore a structural shift model, in which the path of future energy demand is known, but where the precise energy mix will change through time due to changing investment behaviours on the part of economic agents. At any point in time, some agents will invest in polluting, high-carbon capital, and others will invest in the green, low-carbon option. 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 idiosyncratic part that is agent specific. The population of agents will therefore draw, at each point in time, a distribution of opinions. 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. Our formulation of the expectation-formation process is our 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 extremes found in the literature (*cf supra*). 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. This process is meant to emphasise the fundamentally uncertain nature of long-run outcomes (Davidson, 1991; Keynes, 1936) agents can only cope with through beliefs, fictions, and rules of thumb (Beckert, 2016). Our agents are therefore boundedly (Simon, 1991), or ecologically rational (Gigerenzer and Todd, 1999; Todd and Gigerenzer, 2012), in the sense that they use simple indicators, based on their expectations, to guide their actions rather than optimisation procedures. Also, although our agents maximise their expected payoff, they do not derive an optimal investment path through an optimal control procedure. This endeavour to model an extra degree of bounded rationality is intellectually closer to the recent developments by Mercure and coauthors (Mercure, 2015; Mercure et al., 2019, 2014) and other proposals for technological submodules in integrated assessment modelling (Bond-Lamberty et al., 2020) than to traditional linear-programming models of optimal energy mix and technology penetration (Keramidas et al., 2017). The aim of the model is to **complement traditional, technological-progress-oriented analyses of the low-carbon transition by a formalisation of cognitive obstacles to the transition**. To emphasise our point, 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 proposed in Annex Appendix A.

2. Expectations, transition processes, and the history of economic thought (unfinished draft)

Ever since the "rational expectations revolution" (Lucas, 1976), economics has become the social science most explicitly oriented towards the future in its modelling efforts. However, this expectations revolution failed to account for the fact that the future is fundamentally uncertain (Knight, 1921; Keynes, 1936; Shackle, 1972). Rather, this literature was pushing a particular narrative, i.e. the rational expectations hypothesis (REH) (Muth, 1961), as the only way of perceiving the future (Beckert, 2016). This very rigid assumption was only slowly broken up by various approaches to arrive at a more multi-dimensional view of how economic agents might imagine future economic developments. Moreover, the founders and key originators was very clear in that the REH cannot (Lucas, 1981, p. 224) "be applicable in situations in which one cannot guess which, if any, observable frequencies are relevant: situations which Knight called 'uncertainty."' Transition processes as in our model certainly belong to this class of uncertain future developments.

Fictional expectations. **Fictional expectations** (Beckert, 2016, 2013; Beckert and Bronk, 2018) make the underlying *narratives* that govern agents' expectations explicit, accounting both for a multitude of heterogeneous potential expectations of the future, which are *forward-looking* and subject to *change* and *competition* of different discourses. Here, in economic action, "fictions" refer to the images of some future state of the world or a predicted sequence of events that are accessible to agents in the present through mental representation. Agents are motivated for their actions by these imagined future states of the world and organise their activities accordingly. These mental representations of future states are what is termed "fictional expectations", which are operated in narrative form as stories, theories, and discourses (Beckert, 2013). It is important to point out here that fictional expectations are conceived as a very general phenomenon, which is not limited to the economy, but all spheres of human action. It is due to the fundamental uncertainty of the future that fictional expectations are the only possible way of agents to form their expectations about the future, including economic theory itself (Beckert, 2016, p. 274): "Because the economic future is uncertain, economic theory is actually a narrative—a commitment to a specific interpretation of the economy." Similarly, Boyer (2018) points to the importance of **expectation regimes**, which reflect widely believed discursive structures that shape the fictional expectations ("design fantasies" (Beckert, 2013)) of agents in an economic system, and which have gained in importance in recent years: "In the modern era when the efficacy of macroeconomic institutions is declining, the heterogeneity of interests is increasing, and intense structural change is obscuring the future, the time of narratives has come." (Boyer, 2018, p. 41)

The importance of narrative structures in dealing with the future as a paradigm shift. The development of fictional expectations is part of a much broader paradigm shift in the human and social sciences, taking much more into consideration that living and social systems are inherently oriented toward the future (Poli, 2014). Harari (2019, 2017, 2015), which has gained high popularity in recent years, convincingly argues that the ability for collective conception and belief of narratives is the central organizing principle and major evolutionary advantage of humankind *sui generis*. It is clear that these insights also should shape the way we conceive expectation formation in economics as part of this general societal coordination via narratives. Moreover, thinkers in the neo-institutionalist tradition (DiMaggio, 1998) have pointed to the importance of social norms, formal and informal social rules, procedures and strategic interaction, and the importance of corporate actors such as companies and crowds to determine the expectation regimes whithin which economic agents will tend to operate. Several decades ago, Becker (1973) pointed to the importance of narratives in structuring human societies on a deep level, including insights from psychology, sociology, and cultural anthropology. Even before this, the French philosopher Gaston Berger (Berger, 1964) created the notion of *prospective*. Post-War history tended to go faster, be more disruptive and was more characterised by radical uncertainty than before. Therefore, over long horizons, the future could not be considered as a prolongation of the past, be it by extrapolating past trends or assuming structural constancy. He therefore suggested that decision-makers, basing themselves on a deep analysis of underlying factors and causalities, try and invent possible futures in order to guide present decisions:

> We thus must prepare our children to be *inventors*, and show both the courage required to take initiatives and the *imagination* indispensable for the discovery of new solutions [...] We need today much more than yesterday in order to shape bold hypotheses, pose again in new terms former problems, try new paths yet to be trodden. *The future is not to be awaited but to be built and, consequently, to be first imagined* [3]. (Berger, 1964)

Berger's philosophy invites decision-makers to rely on or create imaginaries, which are not pure fantasies, but exercises using all available information in order to shape plausible fictions for the future. Certainly, Friedrich Nietzsche (Nietzsche, 1873) was very early in realizing how imaginative re-creation of metaphors in a process of artistic transference shapes our thinking in language as an imaginative structure. This without clear relation to an objective reality (to which we have no access) besides our collective agreement that we give meanings to these words as imaginations that refer to certain entities in this world we (collectively and individually) call objects, concepts, feelings, and the like. Altogether, several strands of literature have been laying the ground for the concepts of fictional expectations and expectation regimes. It might be a good time at the present moment to take these important insights to the core discipline of economics, especially macroeconomic modelling, and re-frame some of the expectation formation mechanisms and theories that have been developed in economics as a core discipline in the more or less recent past.

Adaptive expectations in economic thought. **Adaptive expectations** is a form of fictional expectations that takes humans as governed by "animal spirits", who use the past (in lack of better predictors) to infer the future by assuming that "existing state of affairs will continue indefinitely, except in so far as we have specific reasons to expect a change" (Keynes, 1936, p.152). This falls short of acknowledging that expectations of the future that are not necessarily (linear or close to linear) extrapolations of the past with little room for imagination, optimization, or even change. However, it does allow for multiple conflicting heterogeneous

[3] Translation from the authors, emphasis added

belief systems of this world that are subject to overall macroeconomic coordination problems (Keynes' beauty contest, fallacy of composition, paradox of thrift, etc.). Compared to "rational expectations", **backward-looking expectation** structures consider that past and present are the best guides to the future, absent the possibility to unveil or even assume "deep" parameters or probability distributions allowing for a forward-looking stance. In such situations, agents tend to rely on simple expectation rules, for instance considering that the future will be identical to the present at all point in time (purely myopic expectations), by updating beliefs according to past expectation mistakes (adaptive expectations), or by prolonging past trends (extrapolative behaviour). Such structures have been conserved in many Keynesian (De Grauwe and Macchiarelli, 2015) and most post-Keynesian (Godley and Lavoie, 2007) models so as to figure fundamental uncertainty (Kappes and Milan, 2019).

Adaptive expectations were famously criticised by Lucas (1976) in the wake of the "rational-expectation revolution" on the grounds that they were subject to systematic and statistically traceable mistakes, and were henceforth a poor representation of individual behaviours. Such formulations were nonetheless given some empirical support for key economic variables, such as inflation, and were given back some theoretical credit in the behavioural macroeconomic literature (Hommes, 2018). Kappes and Milan (2019) nonetheless argue that such formulations, albeit intuitive, are fraught with indeterminacy, making them ultimately difficult to wield in many modelling framework . From the standpoint of fictional expectations, such expectation heuristics or extrapolation tools can be considered as fictions, and indeed relate for instance to the forecasts made by key economic institutions which serve as focal points but have rarely proven accurate (**?**). The importance of the present in formulating expectations corresponds also, for Keynes, to the importance of current conventional judgments in guiding agents' behaviours, in that they all rely, to different extents, on existing market *norms*. Although Keynes was mostly concerned with short-run dynamics, and by financial markets as stages for this "beauty contest", Boyer (2018) provides a sound generalisation to this intuition to the long-run through his notion of expectation regimes within a broader economic context, made of fictions widely believed in by economic agents.

Animal spirits and backward-looking rules. **Animal spirits** were proposed by Keynes as a complement to backward-looking rules in his famous Chapter 12 discussion:

> Most, probably, of our decisions to do something positive, the full consequences of which will be drawn out over many days to come, can only be taken as a result of animal spirits — of a spontaneous urge to action rather than inaction, and not as the outcome of a weighted average of quantitative benefits multiplied by quantitative probabilities [...] Thus if the animal spirits are dimmed and the spontaneous optimism falters, leaving us to depend on nothing but a mathematical expectation, enterprise will fade and die; — though fears of loss may have a basis no more reasonable than hopes of profit had before.It is safe to say that enterprise which depends on hopes stretching into the future benefits the community as a whole. But individual initiative will only be adequate when reasonable calculation is supplemented and supported by animal spirits, so that the thought of ultimate loss which often overtakes pioneers, as experience undoubtedly tells us and them, is put aside as a healthy man puts aside the expectation of death.

In other words, simple forecasting rules are not sufficient to back in investment decision, but must be completed by an urge to act according to those calculations. Absent such a positive impetus, mathematical tools, however "rational" they may be, do not have a sufficient weight as arguments to convince investors to move forward. More precisely,as put by Dequech (1999), animal spirits act conjointly with backward-looking forecasting tools to create a general "state of expectations" by influencing the confidence agents consider their own forecasts with. Shunned for a long time by economists, animal spirits have found audience anew, especially from scholars in behavioural macroeconomics (De Grauwe and Macchiarelli, 2015; **?**).

In terms of fictional expectations, animal spirits can be viewed as enforced and reinforcing positive or negative states of confidence (Beckert, 2013), up to exhilaration as was the case in the run-up to the 2008 (Boyer, 2018) crisis, or down to sheer pessimism in the case of the Eurocrisis .

UNFINISHED DRAFT - to be continued

3. The Model

We develop here, step after step, our whole model architecture. We begin by describing its overall structure and detailing its real-economy core. We then move on to the presentation of our treatments of long-run expectations, our main contribution.

3.1. The Model Core

Production and electricity demand. In the model, as in many technological change, "bottom-up", frameworks (Keramidas et al., 2017), the evolution of energy demand e_d is purely exogenous:

$$\forall t > 1,\ e^d(t) = (1 + g_e)\, e^d(t-1)\ \ \#(3.1)$$

Where g_e is the constant, exogenous growth rate of energy demand, and t denotes, intuitively, the time period. The model has a conventional time step of one year. For the sake of simplicity, we suppose that agents know with certainty the path of future energy demand. Electricity demand is labelled in TWh per time period, that is TWh per year.

We assume two types of electricity-production capital, a high-carbon K_H and a low-carbon K_L one, characterised by Leontief, fixed-coefficient technologies, again consistently with many "bottom-up" frameworks. In the model, the capital stock is labelled in GW of energy capacity.

The two capital stocks are not mandatorily utilised in full. In other words, there will be in general some slack in generating capacities, allowing us to define a utilization rate for the two types of capital:

$$u_L(t) = \frac{e_L(t)}{e_L^s(t)} \#(3.2a)$$

$$u_H(t) = \frac{e_H(t)}{e_H^s(t)} \#(3.2b)$$

Where $e_L(t)$ and $e_H(t)$ are respectively the energy effectively produced thanks to the low- and the high-carbon energy source, and $e_L^s(t)$ and $e_H^s(t)$ full-capacity output. The present version of the model is calibrated to ensure that $u_L(t) = \overline{u_L} = 1\ \forall t$. We indeed identify in our calibration the low-carbon capital stock to modern renewable energy sources (mainly wind and solar), whose utilisation cannot be controlled at will. We therefore assume that $\xi_L(t)$, 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:

$$e_L^s(t) = \ \xi_L(t)\, K_L(t)\ \ \#(3.3a)$$

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

$$e_H^s(t) = \xi_H(t)\, K_H(t) \#(3.3b)$$

Where $\xi_H(t)$ is defined as the amount of amount of TWh a GW of low-carbon capacity can produce per year. As the production of these types of capital can be modulated a near-will, we assume that agents can react to changes and 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 $\xi_H(t) = \xi_H$, that is, that the Leontief coefficients are constant.

The quantity of energy demanded is then dispatched according to a merit-order mechanism, according to which 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:

$$e_L(t) = \begin{cases} 0\ if\ K_L(t) = 0 \\ e_d(t)\ \ ifK_L(t) > 0\ and\ e_d(t) \le e_L^s(t) \\ e_L^s(t)\ \ ifK_L(t) > 0\ and\ e_d(t) > e_L^s(t) \end{cases} \ \#(3.4a)$$

and:

$$e_H(t) = \begin{cases} 0 \ if \ K_H(t) = 0 \\ 0 \ if \ K_H(t) > 0 \ and \ e^d(t) \leq e^s_H(t) \\ e^d(t) - e^s_L(t) \ \ if \ K_H(t) > 0 \ and \ e^s_L(t) \leq e^d(t) < (e^s_L(t) + e^s_H(t)) \\ e^s_L(t) \ \ if \ K_H(t) > 0 \ and \ e^d(t) \geq (e^s_L(t) + e^s_H(t)) \end{cases} \#(3.4b)$$

That is, low-carbon electricity serves the market first. As mentioned above, the model is calibrated so that $e_d(t) \geq e^s_L(t)$ at any point in time, ensuring full capacity utilization for low-carbon capital.

We assume away any labour input for the production electricity and suppose that low-carbon technologies do not need marketed inputs to generate electricity, consistently with their reliance on wind and solar instead of fossil fuels. Conversely, high-carbon capital requires a fossil-fuel input to work. We define the "thermal efficiency" of the high-carbon stock $\xi_{FF}(t)$ 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 $\xi_{FF}(t)$ as constant $(\xi_{FF}(t) = \ \xi_{FF})$.

Prices and profits. Concerning pricing, we drift from the usual merit-order process by assuming a unique, administered electricity price $p_E(t) = p_E$, that 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 proceeds minus expenses per unit of production. We model several production costs, beginning by intermediary input expenses which, as mentioned above, only exist for high-carbon electricity production, as a certain price $p_{FF}(t)$ per unit of non-renewable input. We hold this price as constant throughout $(p_{FF}(t) = \ p_{FF})$.

We also assume several measures of capital costs. First, firms must take up on loans to fund their investment expenditures, constituting a user cost of capital, that we measure per unit of capital. We therefore define for both types of capital:

$$d^r_L(t) = \alpha_L(t)\,\psi_L(t)\,c^k_L(t) \ \#(3.5a)$$

$$d^r_H(t) = \alpha_H(t)\,\psi_H(t)\,c^k_H(t) \ \#(3.5b)$$

The cost of capital for each type of capital stock. $\alpha_L(t) = \alpha_L$ 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:

$$c^l_H(t) = \max\left(0, c^k_H(t)\left(u^\tau_H - \ u_H(t)\right)\right) \#(3.6)$$

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.

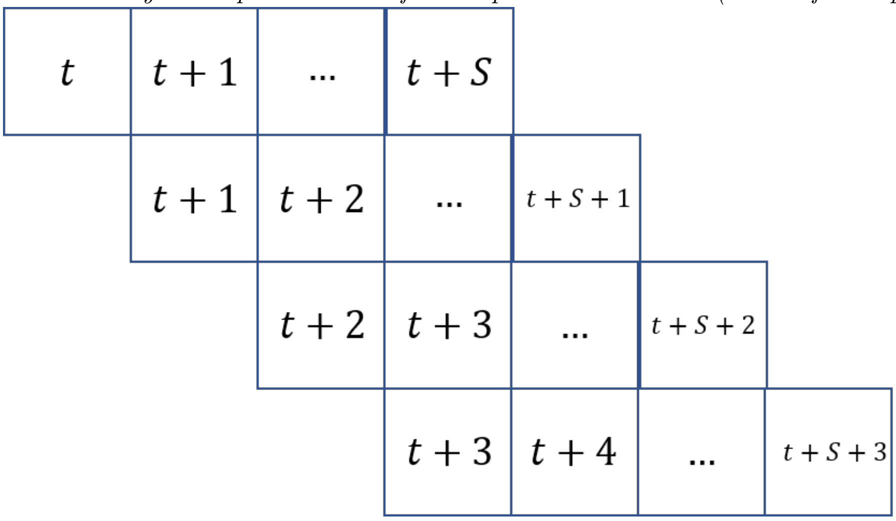
As a result, profit rates are defined as:

$$r_L(t) = p_E \xi_L u_L(t)\,\xi_L - d^r_L(t) = \ p_E \xi_L - d^r_L(t) \ \#(3.7a)$$

$$r_H(t) = \left(p_E - \frac{p_{FF}}{\xi_{FF}}\right)\xi_H u_H(t) - d^r_H(t) - c^l_H(t) \#(3.7b)$$

Expectations: Chronological and Psychological times. Consistently with our main object of study, agents will formulate expectations on economic quantities, but in a different way to most formalisms adopted in most E3 models. We indeed adopt a middle-ground between infinite-horizon, perfect-foresight expectations and myopic or adaptive ones by assuming 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 was inspired by the work of Spiro (2014) and in consistent with the literature on limited foresight in investment decision (Hedenus et al., 2006).

Figure 5-1 A stylised representation of our expectation structure (Drawn from Spiro (2014))



As a result, at each time t, taking the actual value as starting point, agents will formulate an expectation vector of length S. 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:

$$X^e = X^e(t, t+s) \# (3.8)$$

These expectation vectors will serve for several decisions that will involve different expectation distances.

Investment decisions. Our model being aimed at studying structural change, we model agents' investment choices that will progressively lead to changes in the economy's energy mix. Their decision is two-stepped.

Investment shares First, they decide on the share of the investment flow they dedicate to one or the other type of capital. Denoting $\ell_I(t)$ the share of the investment flow going to low-carbon capital, we define it recursively as follows:

$$\ell_I(1) = \ell_{I_0} \# (3.9)$$

Where ℓ_{I_0} is a calibrated first-period value.

$$\ell_I(t+1) = \ell_I(t) + \ (1 - \ell_I(t))\,\zeta_L(t) - \ell_I(t)\,\zeta_H(t) \ \# (3.10a)$$

This function takes inspiration from Mercure's (Mercure, 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).

Let us interpret $\ell_I(t)$ 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 $\zeta_L(t)$ and $\zeta_H(t)$ will change course. In other words, a portion $\zeta_L(t)$ of high-carbon investment projects will be abandoned, or a fraction $\zeta_L(t)$ 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 $\ell_I(t)$. Conversely, a fraction $\zeta_H(t)$ of investment dedicated to low-carbon projects will switch back to high-carbon capital. Hence that this fraction is subtracted from the current share of low-carbon investment.

Conceptually, $\zeta_L(t)$ and $\zeta_H(t)$ 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. The two distributions being relatively close and interchangeable in practice, this simplification is relatively legitimate, as least in this precise respect.

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, the Equation (**??**) reduces to:

$$\ell_I(t+1) = \zeta_L(t) \,\#\, (3.10b)$$

Which suggests that all agents choose all the time between the two options, not allowing for the possibility that some will stick to their previous behaviour. We will therefore introduce some asymmetry between $\zeta_H(t)$ and $\zeta_L(t)$ to give rise to some stickiness in behaviour. How $\zeta_L(t)$ and $\zeta_H(t)$ are determined precisely will be discussed in Section 3.3.

Total investment in low-carbon capital $I_L(t)$ will therefore be defined as:

$$I_L(t) = I^d(t)\,\ell_I(t) \;\#\, (3.11a)$$

And high-carbon investment $I_H(t)$

$$I_H(t) = (1 - \ell_I(t))\, I^d(t) \,\#\, (3.11b)$$

Where $I^d(t)$ is total investment demand, that is determined through the following step.

Investment shares 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, $I^d(t)$ is defined through an extended perpetual inventory rule:

$$I^d(t) = \max\left(K^d(t+1) - (K_L(t) + K_H(t)), 0\right) + dep_L(t) + dep_H(t) \;\#\, (3.12)$$

We therefore suppose that total investment makes for natural depreciations $dep_L(t)$ and $dep_H(t)$, and we assume away capital destruction.

In the model, agents formulate imaginary futures about how stranded high-carbon capital may be. It 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:

$$K^d(t+1) = \frac{e^d(t+1) - \xi_L \Gamma_1(t) - u_H^\tau \xi_H \Gamma_2(t)}{u_H^\tau \xi_H (1 - \ell_I(t)) + \xi_L \ell_I(t)} \,\#\, (3.13)$$

Where $\ell_I(t)$ is the share of the total investment flow dedicated to low-carbon capital, and $\Gamma_1(t)$ and $\Gamma_2(t)$ are intermediate terms defined as follows:

$$\Gamma_1(t) = (1 - \ell_I(t))\, K_L(t) - \ell_I(t)\,(dep_L(t) + dep_H(t) - K_H(t)) - dep_L(t) \;\#\, (3.14a) \quad \Gamma_2(t) = \ell_I(t)\, K_H(t) - (1 - \ell_I(t)$$

Where $dep_L(t)$ and $dep_H(t)$ are the respective depreciations of the low and high-carbon capital stocks.

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[4].

We know need to explain how $\zeta_L(t)$ and $\zeta_H(t)$, the transition probabilities governing the pace of structural change in our economy. This requires explaining more into detail how we model expectations of future probabilities.

[4] 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.

3.2. Imagined Futures: The Probit Module

The core innovation of this framework lies in our modelling of conflicting expectations about (a) given economic metrics. 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.

The general case.

Base 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, her "imagined future".

Formally, the most general case writes for agent i:

$$r^e_{L_i}(t,t+s) = p^e_E(t,t+s)\,\xi^e_L(t,t+s)\,u^e_L(t,t+s) - \alpha^e_L(t,t+s)\,\psi_L c^e_{k_L}(t,t+s) + \epsilon^e_{L_i}(t,t+s) \quad \#(3.15a)$$

$$= \gamma_L(t)\,u^e_L(t,t+s) - dr^e_L(t,t+s) + \epsilon^e_{L_i}(t,t+s) \quad \#(3.15b)$$

$$r^e_{H_i}(t,t+s) = \left(p^e_E(t,t+s) - \frac{p^e_{FF}(t,t+s)}{\xi^e_F(t,t+s)}\right)\xi^e_H(t,t+s)\,u^e_H(t,t+s) - \alpha^e_H(t,t+s)\,\psi_H c^e_{k_H}(t,t+s) - \max\left(0, c^e_{k_H}(t,t+s)\left(u^\tau_H - u^e_H(t,t+s)\right)\right) + \epsilon^e_{H_i}(t,t+s) \quad \#(16a)$$

$$= \gamma_{H_i}(t,t+s)\,u^e_H(t,t+s) - dr_H(t,t+s) - \max\left(0, c^e_{k_H}(t,t+s)\left(u^\tau_H - u^e_H(t,t+s)\right)\right) + \epsilon^e_{H_i}(t,t+s) \quad \#(3.16\text{b})$$

We assume that the idiosyncratic components $\epsilon^e_{L_i}(t,t+s)$ and $\epsilon^e_{H_i}(t,t+s)$ 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:

$$R_L(t) = \sum_{s=1}^{S} \frac{1}{(1+\rho)^s}\left(\gamma_L(t)\,u^e_L(t,t+s) - dr_L(t,t+s) + \epsilon^e_{L_i}(t,t+s)\right) \quad \#(3.17a)$$

$$R_H(t) = \sum_{s=1}^{S} \frac{1}{(1+\rho)^s}\left(\begin{array}{c}\gamma_H(t,t+s)\,u^e_H(t,t+s) - dr_H(t,t+s) \\ -\max\left(0, c^e_{k_H}(t,t+s)\left(u^\tau_H - u^e_H(t,t+s)\right)\right) + \epsilon^e_{H_i}(t,t+s)\end{array}\right) \quad \#(3.17b)$$

We insist right from the outset that our ϵ's should *not* be understood as normally distributed disturbances around an equilibrium as in RBC or DSGE strands of modelling. Our framework not being 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's are all nil, agents are perfectly coordinated around the deterministic component. Conversely, the higher the σ's, the less coordinated they are around the deterministic component, and the more dissent there is. This formalism is meant to figure that when uncertainty is radical, as it is the case in the context of 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).

Now, 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 type of capital is higher than that of the of high-carbon type of capital. It is simply given by:

$$P\left(R_L(t) - R_H(t) > 0\right) \ \#(3.18\text{a})$$

$$P\left(\begin{array}{c}\sum_{s=1}^{S}\left(\frac{1}{(1+\rho)^s}\right)\left(\epsilon^e_{L_i}(t,t+s) - \epsilon^e_{H_i}(t,t+s)\right) \\ > \sum_1^S\left(\frac{1}{(1+\rho)^s}\right)\left(\begin{array}{c}\gamma^e_H(t,t+s)\,u^e_H(t,t+s) - dr^e_H(t,t+s) \\ -\max\left(0, c^e_{k_H}(t,t+s)\left(u^\tau_H - u^e_H(t,t+s)\right)\right)\end{array}\right) - \left(\begin{array}{c}\gamma^e_L(t,t+s)\,u^e_L(t,t+s) \\ -dr^e_L(t,t+s)\end{array}\right)\end{array}\right) \#(3.18b)$$

By making the assumption that the ϵ's are not serially correlated in psychological time[5], we can state that $\sum_{s=1}^{S}\left(\frac{1}{(1+\rho)^s}\right)\left(\epsilon^e_{L_i}(t,t+s) - \epsilon^e_{H_i}(t,t+s)\right)$ follows a normal distribution of mean 0 and variance $\sum_{s=1}^{S}\left(\frac{1}{(1+\rho)^s}\right)\left(\left(\sigma^e_L(t,t+s)\right)^2 - \left(\sigma^e_H(t,t+s)\right)^2\right)$, 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, allowing 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), which luckily has a workable functional form, is a promising candidate. We nonetheless stick here to a normal distribution.

But this 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 $\sigma^e_L(t,t+s)$ and $\sigma^e_H(t,t+s)$ may figure agents having conceptually unsound expectations. For instance, if $\sigma^e_L(t,t+s) = 1$, 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 $\gamma_L(t)\,u^e_L(t,t+s) - dr^e_L(t,t+s) - 3$ and $\gamma_L(t)\,u^e_L(t,t+s) - dr^e_L(t,t+s) + 3$. Extremes would therefore expect profit rates to be above 300% profit rates and slightly above -300% profit rates, hardly a realistic outcome.

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

$$\forall x,\quad 1_A(x) = \begin{cases} 0\ if\ x \notin A \\ 1\ if\ x \in A\end{cases} \#(3.19\text{a}) \quad \forall x,\quad 1_{A|}(x) = \begin{cases} 0\ if\ x \le R^{Max}_L(t) - R^{Min}_H(t) \\ 1\ if\ x > R^{Max}_L(t) - R^{Min}_H(t)\end{cases} \#(3.19b)$$

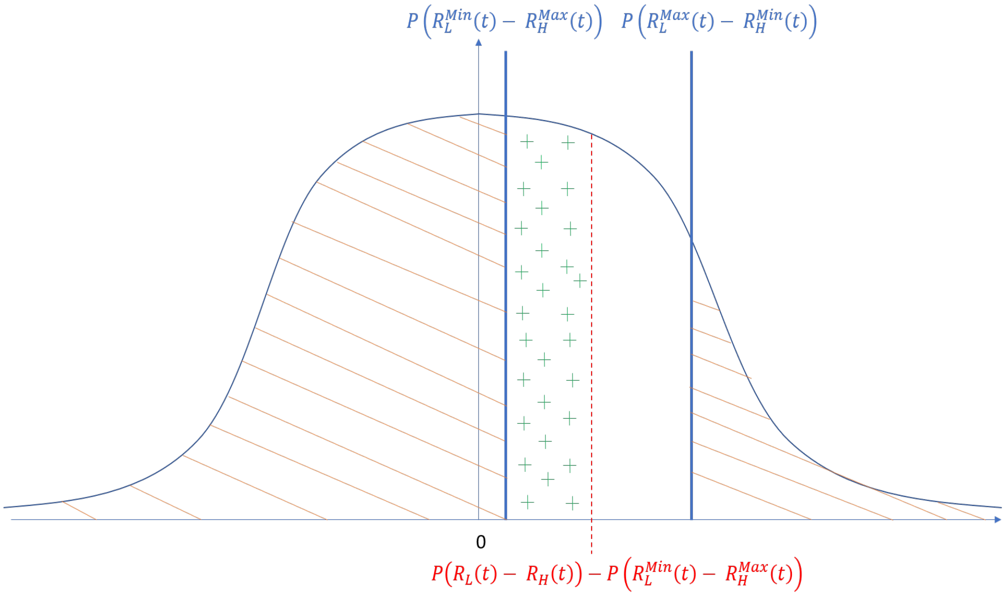
We can now define the final proportion $\chi\left(P\left(R_L(t) - R_H(t) > 0\right)\right)$:

$$\chi(t) = \frac{\left(\begin{array}{c}P\left(R_L(t) - R_H(t) > 0\right) - 1_A\left(R_L(t) - R_H(t)\right)P\left(R^{Min}_L(t) - R^{Max}_H(t) > 0\right) \\ -1_{A|}\left(R_L(t) - R_H(t)\right)P\left(R^{Min}_L(t) - R^{Max}_H(t) > 0\right)\end{array}\right)}{P\left(R^{Max}_L(t) - R^{Min}_H(t)\right) - P\left(R^{Min}_L(t) - R^{Max}_H(t)\right)} \#(19c)$$

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

The $P\left(R_L^{Max}(t) - R_H^{Min}(t)\right) - P\left(R_L^{Min}(t) - R_H^{Max}(t)\right)$ 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 the basic proportion in $(18a)$ into a conditional probability over a given interval, that can be decided based on past data of the relevant variable. Figure 3-2 shows a graphical representation of the process:

Figure 5-2 A stylised censoring process



This transformation allows us to be totally free in the choice of our σ's[6].

Now considering that this proportion represents the share of agents switching from high- to low-carbon investment, we can write:

$$\zeta_L(t) = \chi\left(P\left(R_L(t) - R_H(t) > 0\right)\right) \# (3.20a)$$

As mentioned above, we introduce an asymmetry between $\zeta_L(t)$ and $\zeta_H(t)$ 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:

$$\zeta_H(t) = \chi\left(P\left(R_L(t) - R_H(t) \leq \left(\sum_1^S \frac{1}{1+\rho}\right)\rho_L\right)\right) \# (3.20b)$$

With $\rho_L > 0$, and is discounted accordingly. That is, $R_H(t)$ must be higher than $R_L(t)$ by an extra margin. For results to be meaningful, the censoring procedure must be adapted accordingly, by expanding the support $\left[R_L^{Min}(t) - R_H^{Max}(t), R_L^{Max}(t) - R_H^{Min}(t)\right]$ to $\left[R_L^{Min}(t) - R_H^{Max}(t) - \rho_L, R_L^{Max}(t) - R_H^{Min}(t) - \rho_L\right]$. 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.

The deterministic components Contrarily to most models of the energy system dealing with transitional topics, we assume away model-consistent expectations (perfect foresight) and simplistic adaptive formulations. Rather, we suppose that agents formulate *independent* projections about future outcomes around an anchor that represents the mean projection. This anchor is given a law of motion in psychological time that can be more or less complex depending on the research question, correspond or not to the model's actual law of motion or being fully autonomous or closely related to the actual law of motion.

[6] A full demonstration is provided in Annex.

Let us consider an example. In the current version of the model, c_{k_L} is held constant for the sake of simplicity, but nothing prevents us from giving it a law of motion in actual time. We could, for instance, consider network externalities linking the cost of a GW of capacity to the share of low-carbon capital in the economy:

$$c_{k_L}(t) = c_{k_L}(1)\,\frac{(1+\beta_{NE}\ell_K(t))}{1+\beta_{NE}\ell_K(1)}\,\#\,(3.21a)$$

Where β_{NE} is a sensitivity parameter, and $\ell_K(t) = \frac{K_L(t)}{K_H(t)+K_L(t)}$. However, agents cannot observe network externalities, and will therefore rely on simple projection methods in their expectations, for instance:

$$\forall 0 \leq s \leq S,\ \ c_{k_L}(t,t+s) = c_{k_L}(t)\left(1 - r_{c_{k_L}}(t)\right)^s \#\,(3.21b)$$

With $r_{c_{k_L}}(t)$ an assumed constant decay either invariant or changing in every period. Agents can then more or less learn from their past mistakes by adapting their constant decay to what they observe in reality, or stick to this heuristic, etc. For instance, one can write the following adaptive learning process:

$$r_{c_{k_L}}(t) = \ r_{c_{k_L}}(t-1) + \ \eta_r\left(r_{c_{k_L}}(t-1) - \frac{c_{k_L}(t) - c_{k_L}(t-1)}{c_{k_L}(t-1)}\right)\#\,(3.21c)$$

This method is directly inspired by the recent literature in behavioural macroeconomics which emphasises the need to model simpler heuristics than rational expectations that get updated and improved on an ongoing basis by learning agents until near-complete convergence towards model-consistent expectations.

In our framework, however, as mentioned at length above, talking of learning is somewhat uneasy, as we model a "unique" transitional experiment, in which the structures of the economy change and in which learning is not possible through the repetition of close experiments. That is why we propose to go one step further than the formulation drawn from the behavioural macroeconomics literature and consider, building on Beckert's concept of *fictional expectations* (Beckert, 2016, 2013), *totally autonomous imagined futures*. Such imagined futures differ from the kind of heuristics developed just above in that they do not try to predict more and more accurately a given variable. In the words of Shackle (1979), such expectations are not data-acquisition processes, but pure acts of imagination aimed at figuring deep structural changes. As such, they reflect the agents' *prospective* attitude towards the future, which, taken "as if" they were true, guide the behaviour agents. These narratives are of course dynamic and updated through, but more based on the overall evolution of the model rather than on forecasting performances – although forecasting performances could well be included. Formally, we write for a variable X:

$$X^e(t,t+s) = f(t, Y(t),\ s)\,\#\,(3.22)$$

Where $Y(t)$ is a vector of variables of interest. We will develop an explicit functional form for this law of motion when we describe the particular implementation of this general treatment of expectations in the model.

Standard errors and dissent Remains to determine a law of motion for the standard errors of the idiosyncratic opinions in psychological time. As sketched earlier, these σ's are meant to figure consensus or dissent about the outcome of the economic variable at stake. 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 one, but that there is little difference between long- and "very long"-run expectations (see also Moureau and Rivaud-Danset, 2004), we assume that the σ's follow in psychological time a sharply increasing logistic law of motion writing as follows:

$$\sigma_H^e(t,t) = \sigma_{Min_H}\ \#\,(3.23a)$$

$$\forall s \leq S\ \sigma_H^e(t,t+s) = \ \sigma_H^e(t,t+s-1)\left(1 + r_{log_{\sigma_H}}\ \left(1 - \frac{\sigma_H^e(t,t+s-1)}{\Sigma_{Max_H}}\right)\right)\ \#\,(3.23b)$$

$$\sigma_L^e(t,t) = \Sigma_{Min_L}\ \#\,(3.23c)$$

$$\forall s \leq S \ \sigma_L^e\left(t,t+s\right) = \ \sigma_L^e\left(t,t+s-1\right)\left(1+r_{log_{\sigma_L}}\ \left(1-\frac{\sigma_L^e(t,t+s-1)}{\Sigma_{Max_L}}\right)\right)\ \#\left(3.23d\right)$$

That is, we consider a minimal level of uncertainty or dissent for each profit rate σ_{Min_H} and σ_{Min_L} that is present *ex ante*, and which increases sharply at a rate $r_{log_{\sigma_H}}$ and $r_{log_{\sigma_L}}$ up to σ_{Max_H} and σ_{Max_L}. dissent is first rather weak, agents mostly agreeing on the mean projection, and increases quickly up to σ_{Max_H} and then reaches a plateau, signalling that agents disagree to the same extent in the late periods of the planning horizons. Note that nothing prevents σ_{Min_H} and σ_{Max_H} from being equal, or σ_{Max_H} to be very close to σ_{Min_H}. In such configuration, the dissent is either the same whatever the period in psychological time or increases very little.

These elements represent the most general version possible of the model, which would entail a strong degree of complexity and render the model less tractable analytically and make simulations less intuitive. As such, we chose to simplify our treatment of uncertainty by focussing, as sketched above, on high-carbon capital stranding expectations. Therefore, in the simulation in chapter 4, we only the maximum level of uncertainty of high-carbon capital and term it Σ_{Max}.

The model's case. In the model, we simplify greatly this treatment of expectations. 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:

$$r_{L_i}^e\left(t,t+s\right) = \gamma_L\left(t\right) - dr_L\left(t,t+s\right)\#\left(3.24a\right)$$

$$r_{H_i}^e\left(t,t+s\right) = \gamma_H\left(t\right)\left(u_H^e\left(t,t+s\right)+\epsilon_{H_i}^e\left(t,t+s\right)\right) - dr_H\left(t,t+s\right)\ \#\left(3.24b\right)$$

We assume that the $\epsilon_{H_i}^e\left(t,t+s\right)$ are normally distributed through a random variable $\epsilon_H^e\left(t,t+s\right)$ with mean zero and variance $\left(\sigma_{H_i}^e\left(t,t+s\right)\right)^2$. 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 $\epsilon_{H_i}^e\left(t,t+s\right)$ are normally distributed.

Finding this proportion requires a law of motion for $u_H^e\left(t,t+s\right)$ in time s, to which we turn to now.

Standard errors and dissent We conserve the definition in Equations ($21a$-$21b$) and leave aside that of Equations ($21c$-$21d$).

Censoring Since consensus/dissent regards only stranding degrees within this framework, we start assume that agents do not consider utilisation rates going above the long-run conventional rate u_H^τ. We therefore consider the following boundaries for $r_H^e\left(t,t+s\right)$:

$$r_{H_{Min}}^e\left(t,t+s\right) = -dr_H\left(t,t+s\right)\#\left(3.25a\right) \quad r_{H_{Max}}^e\left(t,t+s\right) = u_H^\tau\gamma_H\left(t\right) - dr_H\left(t,t+s\right)\#\left(3.25b\right)$$

And apply the χ operator defined above.

Stranding expectations and narratives In the current version of the model, 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.

Stranded assets are usually defined as assets (natural resources, means of production or financial instruments) suffering from premature devaluation due to structural changes within the economic system as a whole, for instance as a result of "creative destruction" (Schumpeter, 1983).

The low-carbon transition most likely entailing such sweeping changes, and what's more mostly policy-induced, the concept has received much attention over the past few years. An important literature has indeed sprouted on the issue, ranging from the study of the foregoing of oil resources within Hotelling-like frameworks (van der Ploeg and Rezai, 2020) to climate stress-tests in the financial sector (Battiston et al., 2017) and going through the study of physical asset stranding.

The model focusses precisely on the latter aspect: that of physical asset stranding. Most studies on the matter either try to estimate the impact of climate policies onto the utilisation of productive capital (Cahen-Fourot et al., 2019; Jackson, 2018) or derive optimal transition paths with possible stranding using perfect foresight (Baldwin et al., 2020; Coulomb et al., 2019). Our model comes closer to the second type of study, by trying to relax the perfect-foresight assumption, and apply the concept of fictional expectations to future utilisation rates.

Benchmark with a single narrative

We 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:

$$\begin{aligned} \forall -t \le s \le 0,\ \ \ell_E^N(t,t+s) &= \ell_E(t+s) \\ \forall s>0,\ \ \ell_E^N(t,t+s) &= \ell_E^N(t,t+s-1)\left(1+r_N\left(1-\frac{\ell_E^N(t,t+s-1)}{\theta}\right)\right) \end{aligned} \# (3.26)$$

Where r_N is exogenous, and θ, which represents the long-run low-carbon energy share goal, is 1. As a result, only the starting value of the schedule, which the observed low-carbon share, changes. 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 (REF), and is considered as a sound pattern for the future displacement of fossil fuels (Vandevyvere and Nevens, 2015).

These expectations allow agents to compute a high-carbon energy share, by definition equal to $h_E^N(t,t+s) = 1-\ell_E^N(t,t+s)$. Now, at any point in time, the actual share $h_E(t)$ is also equal to:

$$h_E(t) = \frac{e_H^d(t)}{e_d(t)} = \frac{\xi_H u_H(t) K_H(t)}{e_d(t)} \# (3.27)$$

Which means that, in expectations, agents can use this equation and its equivalences:

$$\begin{aligned} h_E^N(t,t+s) &= \frac{\xi_H u_H^e(t,t+s) K_H^e(t,t+s)}{e_d^e(t,t+s)} \\ \Leftrightarrow u_H^e(t,t+s) &= \frac{h_E^N(t,t+s) e_d^e(t,t+s)}{\xi_H K_H^e(t,t+s)} \quad \# (3.28) \\ \Leftrightarrow\ K_H^e(t,t+s) &= \frac{h_E^N(t,t+s) e_d^e(t,t+s)}{\xi_H u_H^e(t,t+s)} \# \end{aligned}$$

These equations can be used to derive a high-carbon capital and a high-carbon utilization schedules period by period according to certain behavioural rules. That is, taking the narrative $h_E^N(t,t+s)$ 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 $\frac{\xi_H u_H^e(t,t+s)(1-\delta_H)K_H^e(t,t+s)}{e_d^e(t,t+s)} < h_E^N(t,t+s+1)$, there is possibly a need for more capital. More precisely, we assume that agents consider u_H^τ as a conventional anchor, and effectuate another case disjunction:
   (a) If $u_H^e(t,t+s) < u_H^\tau$, $K_H^e(t,t+s+1) = (1-\delta_H)K_H^e(t,t+s)$, that is agents expect that capital will be left to depreciate naturally
   (b) If $u_H^e(t,t+s) \le u_H^\tau$, $K_H^e(t,t+s+1) = \frac{h_E^N(t,t+s)e_d^e(t,t+s)}{\xi_H u_H^\tau}$, 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 $\frac{\xi_H u_H^e(t,t+s)K_H^e(t,t+s)}{e_d^e(t,t+s)} = h_E^N(t,t+s+1)$, then the capital stock is just right, and is left to depreciate naturally

3. $\frac{\xi_H u_H^e(t,t+s) K_H^e(t,t+s)}{e_d^e(t,t+s)} > h_E^N(t,t+s+1)$, agents cannot but leave the capital stock to depreciate naturally.

Note that we make here three assumptions:

1. The future demand schedule is known, as assumed in the preceding steps of the model
2. The 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).

A conflictive narrative add-on

I then introduce a belief conflict on how far and how intense the energy transition will be. Building on the work of Franke and Westerhoff (2017) on Keynes's animal spirits (1936), itself based the seminal works of Weindlish and Haag (Weidlich and Haag, 1983) and Lux. Our stance here is a little different, to the extent that we our goal is more to study how a norm changes through time, with possible social lock-in effects, consistently with Weindlich & Haag and Lux's original goal. As such, our agents do not switch between optimism and pessimism 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 will consist in 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 clause 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:

$$\begin{gathered}\forall -t \leq s \leq 0, \quad \ell_E^A(t,t+s) = \ell_E(t+s) \\ \forall s > 0, \quad \ell_K^A(t,t+s) = \ell_K^A(t,t+s-1)\left(1 + r_L^e(t)\left(1 - \frac{\ell_K^A(t,t+s-1)}{\theta}\right)\right)\end{gathered} \# (3.29)$$

With still $\theta = 1$ for the sake of simplicity. $r_L^e(t)$ 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:

$$\begin{gathered}r_L^e(1)\ ) = r_N \\ \forall t > 1 \ r_L^e(t) = \frac{\ell_K(t) - \ell_K(t-1)}{\ell_K(t)(1-\ell_K(t))} \#\end{gathered} \# (3.30)$$

We finally define the "aggregate-opinion" narrative $\ell_E^O(t,t+s)$ is the weighted average of both narrative:

$$\forall -t \leq s \leq 0, \quad \ell_E^O(t,t+s) = x(t)\,\ell_E(t+s) + (1-x(t))\,\ell_E(t+s) = \ell_E(t+s) \# (3.31a)$$

$$\forall s > 0, \quad \ell_E^O(t,t+s) = \ x(t)\,\ell_E^N(t,t+s) + (1-x(t))\,\ell_E^A(t,t+s) \# (3.31b)$$

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 as 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.

These shares will evolve in chronological time according to an evaluation process. We first define a variable $w_N(t)$ as follows:

$$w_N(t) = \begin{cases} 0\ if\ t=1 \\ \sum_{k=1}^{t-1} \omega^{1-k} \frac{\left(\ell_K^N(k,t)-\ell_K(t)\right)}{\ell_K(t)}\ if\ t<1\leq M+1 \\ \sum_{t-M}^{t-1} \omega^{1-k} \frac{\left(\ell_K^N(k,t)-\ell_K(t)\right)}{\ell_K(t)} if\ t>M+1 \end{cases} \#(3.32a)$$

And another, $w_A(t)$ in the same way:

$$w_A(t) = \begin{cases} 0\ if\ t=1 \\ \sum_{k=1}^{t-1} \omega^{1-k} \frac{\left(\ell_K^A(k,t)-\ell_K(t)\right)}{\ell_K(t)}\ if\ t<1\leq M+1 \\ \sum_{t-M}^{t-1} \omega^{1-k} \frac{\left(\ell_K^A(k,t)-\ell_K(t)\right)}{\ell_K(t)} if\ t>M+1 \end{cases} \#(3.32b)$$

And finally, $w(t)$, the weighted average of both:

$$w(t) = x(t)\,w_N(t) + (1-x(t))\,w_A(t)\,\#(3.33)$$

That is, at each point in time, agents will take the relative spread (normalised by $\ell_K(t)$) between the actual share of $\ell_K(t)$ 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:

$$x(t) = \frac{(1+j(t))}{2}\#(3.34a) \quad 1-x(t) = \frac{(1-j(t))}{2}\#(3.34b)$$

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. $j(t)$ is itself defined recursively:

$$j(1) = j_0\#(3.35a) \quad j(t+1) = j(t) + (1-j(t))\,p^{AN}(t) - (1+j(t))\,p^{NA}(t)\ \#(3.35b)$$

This equation is the opinion equivalent of the investment function defined in $(3.10a)$. $(1-j(t))$ 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. $p^{AN}(t)$ is therefore the transition probability from the alternative to the regulator's narrative, and $p^{NA}(t)$ that from the regulator's to the alternative narrative. They are given by:

$$p^{AN}(t) = \nu\exp(-w(t))\ \#(3.36a) \quad p^{NA}(t) = \nu\exp(w(t))\ \#(3.36b)$$

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. 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[7].

[7] 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.

My aim through this add-on 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. I also incidentally intend to illustrate the modularity of our modelling approach, showing how additional hypotheses can be added to explore specific research questions.

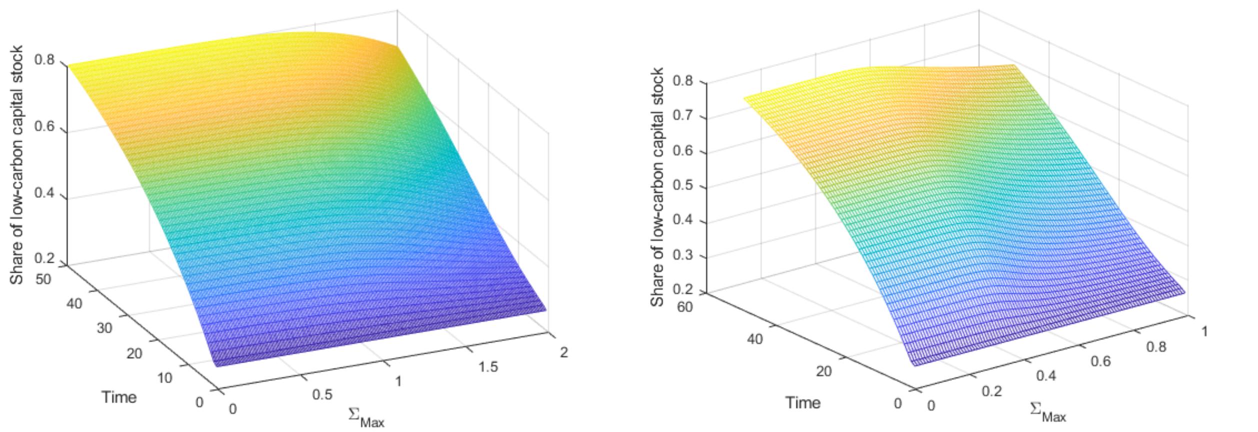
4. Scenarios and simulations

We roughly calibrate our model roughly to the electricity market of European Union (EU) for the year 2017, see section Appendix A for details. However, the simulations below are intended as a demonstration of the more general predicaments that underlie the assumptions behind our model, not to sketch concrete energy transition paths for the EU. That is, given the fundamental uncertainty of the future and the possibility of physical asset stranding via under-utilization of fixed capital, we show that different, narratively shaped expectations about future transition dynamics—no matter what exact form they might take—and different levels of uncertainty about future utilization rates, stranding and associated profitability of technologies take concrete effects by shaping these transition dynamics.

4.1. The importance of narrative stances for transition dynamics

4.1.1. The one-narrative benchmark

Certainly, our model is calibrated to the EU as a region in which the transition to renewable energy is already on its way. For this first model run, we assume that there is only one narrative prevailing between agents, i.e. that the growth rate r_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 (**dissensus**), Σ_{Max}, about stranding of high-carbon capital, see Figure 2a. 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 dissensus about profitability and utilization of the two technologies

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

Figure 2: A benchmark with one narrative and varying levels of dissensus

The underlying reason for the results depicted in Figure 2a is that agents expect stranding of high-carbon capital stock due to the transition, and at the same time anticipate a premium in the profitability for low-carbon investment due to the increased shift-in of this technology and associated network effects and economies of scale. For high amount of dissensus, 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 the investment in renewables. That is, as agents observe whether the narrative holds true without forming dissenting opinions, the transition path validates itself because agents that believe in the narrative invest in renewables, which leads others to follow their example and thus locks in the transition path. This result stays robust also if the growth rate of dissensus in agents' expectations, r_{log_σ} varies, see Figure 2b.

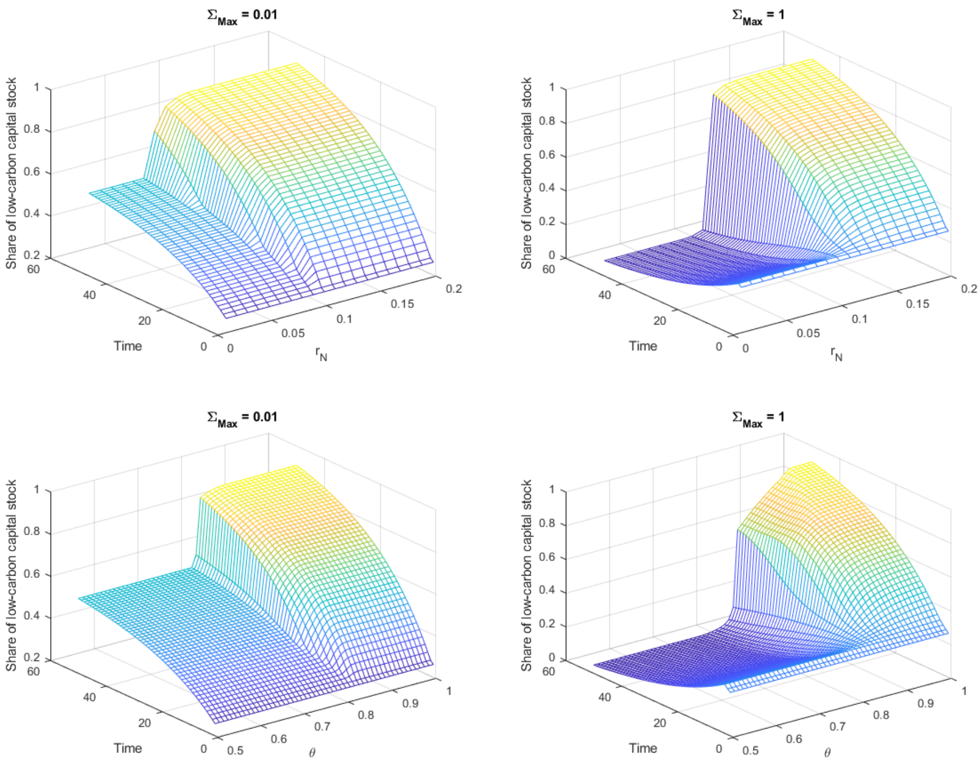
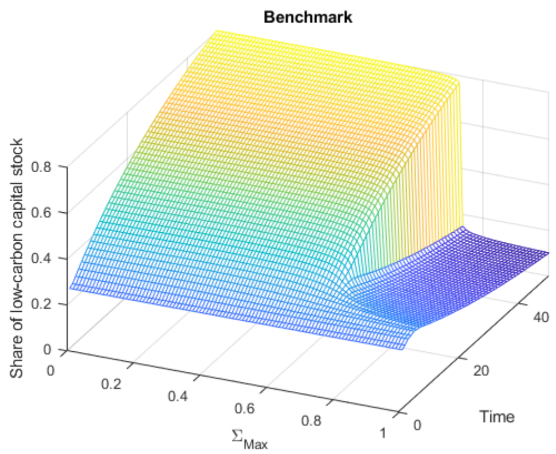


Figure 3: Varying ambitions and levels of dissensus: at high levels of dissensus and low levels of ambition, the transition may not take place

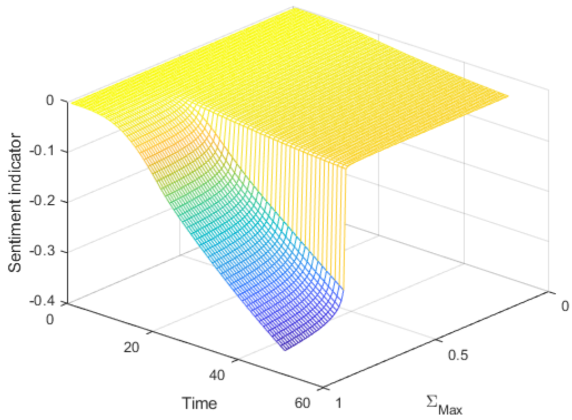
Also without dissenting opinions, however, too timid ambitions by the regulator can lead to the lock-in of an equilibrium where the transition essentially does not take place. This can be observed in figure 3, where both the growth rate r_N as well as the long-run target for the share of renewables in total capital stock θ as decreed by the regulator vary. Here, we can observe that below a certain threshold of ambition and high levels of dissensus about utilization rates and future profitability of the two technologies, agents revert to a portfolio of high-carbon capital stock.

4.1.2. The role of opinion conflict and competing narratives

Given high ambitions of the regulator, also widely varying levels of maximum dissensus Σ_{Max} could not phase out the transition. However, Figure 4a demonstrates 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 dissensus. Here, a high dispersion of opinion around mean stranding expectations implies that a significant fraction of agents does not believe that stranding of hiw-carbon assets will take place due to the transition. Therefore, a large fraction of agents will revert to high-carbon investment.



(a) The *two* narrative benchmark case with varying Σ_{Max}



(b) How opinion conflicts affect the transition

Figure 4: A *two* narrative benchmark case shows a phase transition at higher levels of dissensus

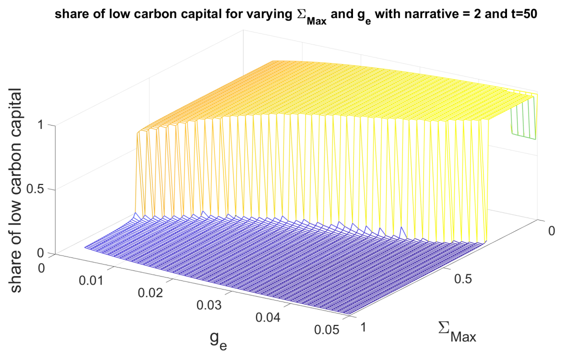
The results of Figure 4a are further evident from the dynamics of opinion conflict shown in Figure 4b. That is, after a level of dissensus for Σ_{Max} of about 0.7, a large part of the agent population does not adhere to the overarching narrative ℓ_E^N, but rather to the alternative narrative ℓ_E^A, shifting the average opinion ℓ_E^O rather to the trend-dependent alternative narrative. This effect is self-reinforcing over time, since reductions in renewable investments tend to further decrease the validity of the overarching narrative as well as stranding expectations for the high-carbon technology, moving more agents to the alternative narrative. 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, due to its adaptive structure, hence a positive lock-in.

These insights from our model runs are of primary importance in relation to the workings of fictional expectations, and how they 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. This reveals that it suffices to assume a mix between backward- and forward-looking expectations to obtain behavioural stickiness and path dependency in opinion formation. This result calls for great policy caution: if agents observe past performances—which is not a very strong assumption to make—and if these past performances at least partly influence how agents form their expectations, a gap between official goals and realized investment paths at some point in time, or a less steady trend over even a relatively short period, may bear significant long-run consequences on the dynamics of the transition.

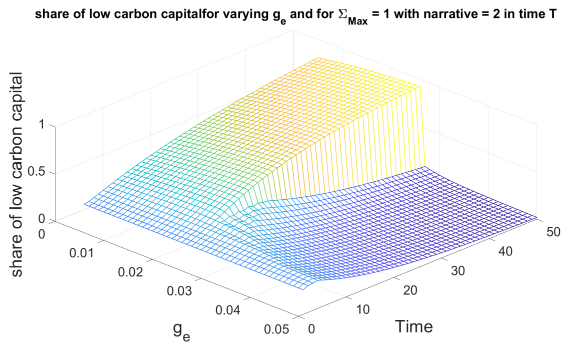
4.2. Changes in electricity demand growth and narratives about the future

We can observe a clear correlation between growth rates of energy demand and opinion dispersion in the model. This is especially relevant in a policy context, since satisfying growing electricity demand as potentially connected to an electrification of the European economy is widely held to be necessary to reach the goals of the Paris agreement (ETC, 2018). Such a widespreach electrification of our economies will require overall larger investment in renewables to reach the transition to a low-carbon economy. For this to take place, also in a situation with conflicting opinions about the future, opinion dispersion needs to be lower for higher growth rates to achieve the transition, as can be observed in Figure 5. Here, in the upper left panel (5a), 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.

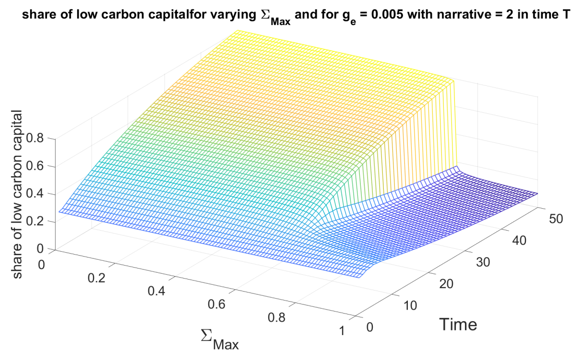
Similarly, in the two lower panels (5c and 5d), we can see how the regime shift between a transition and a no-transition scenario varies for lower (left lower panel) and higher (right lower panel) growth rates.



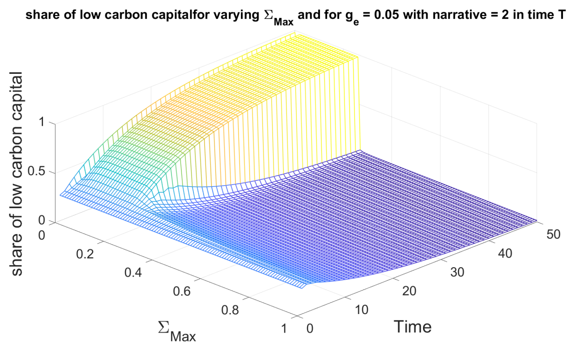
(a) Interaction between growth rate and SigmaMax for final model period



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



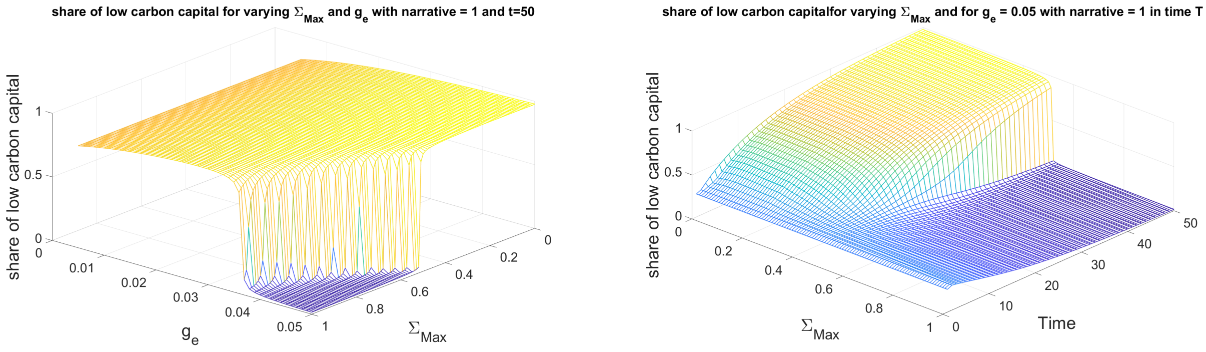
(c) Lower growth rates allow higher dispersion of opinions



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

Figure 5: Interaction between electricity demand growth and conflicting opinions with alternative narrative regimes

Again, here one can clearly see the effects of alternative narratives in the model. If only one narrative prevails as is the case for the simulations depicted in figure 6, also the interaction between growth and the amount of dissensus is much less pronounced. As it turns out, only for very high rates of growth and dissensus as shown in Figures 6a and 6b, we can observe a regime shift to an equilibrium where the transition does not take place.



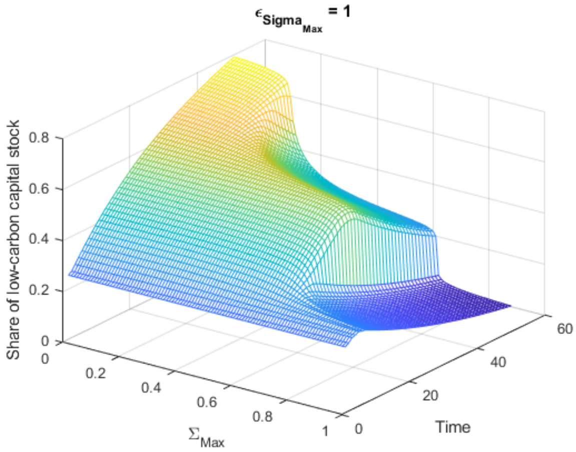
(a) Interaction between growth rate and SigmaMax for final model period - transition more robust with only one regime

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

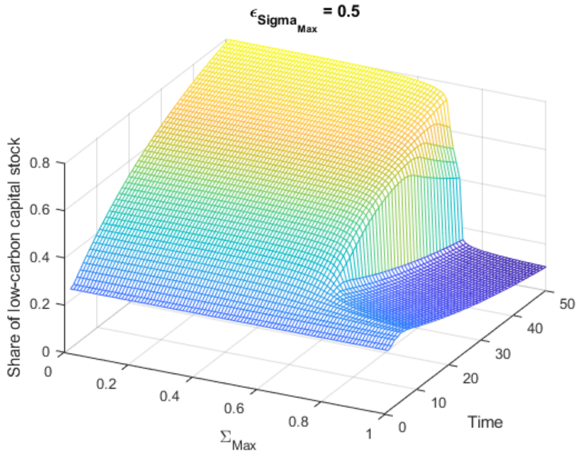
Figure 6: Interaction between electricity demand growth and conflicting opinions with only one narrative regime

4.3. Shocks to dissensus - changes in beliefs

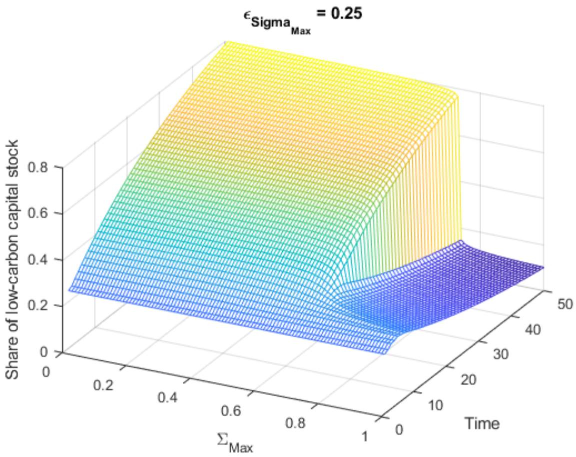
Figure 7 shows different shocks $\epsilon_{Sigma_{Max}}$ to dissensus, from a large positive shock in Figure 7a to a negative shock in Figure 7d. These "dissensus shocks" represent any external event influencing the divergence of opinions about the future, such as uncertainty about policy decisions regarding subsidies for renewables or the apparition of a breakthrough-technology. As it turns out in our simulations, only a large positive shock as depicted in Figure 7a suffices to negatively influence transition dynamics beyond a (rather low) threshold for dissensus Σ_{Max}. Smaller shocks leave transition dynamics virtually untouched. Conversely, a negative shock to dissensus as in Figure 7d simply moves the level of dissensus where the regime change between the point where a transition still takes places and where it does not.



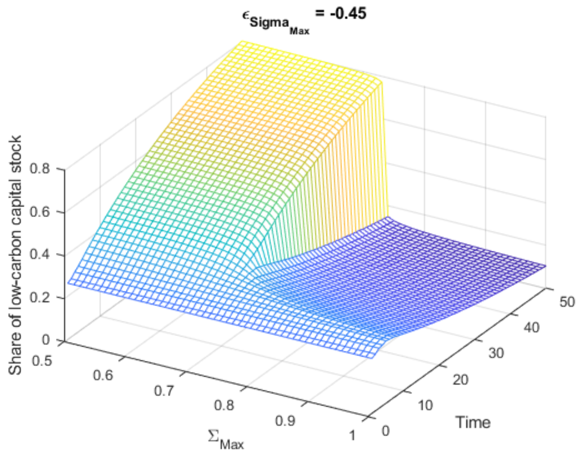
(a) A large positive shock to opinions



(b) A medium positive shock to opinions



(c) A small positive shock

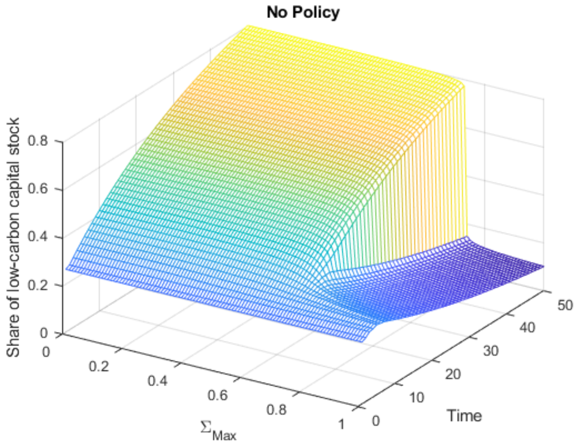


(d) A negative shock to opinions

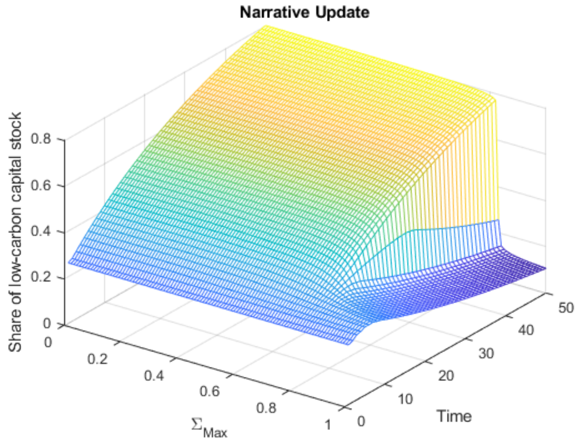
Figure 7: fig:Different shocks to dissensus in the model

4.4. Some policy simulations

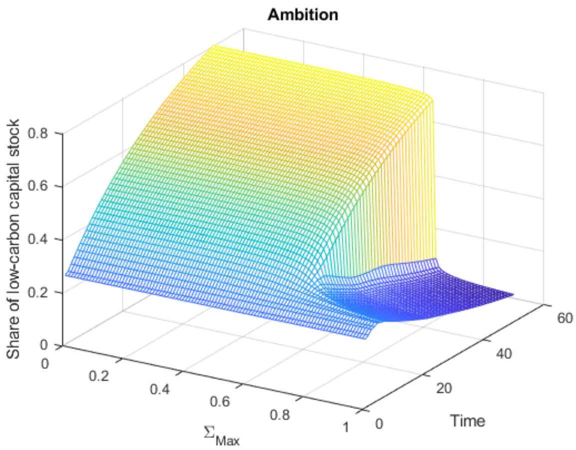
Figure 8 below demonstrates different policy options that we can cover with our modelling framework.



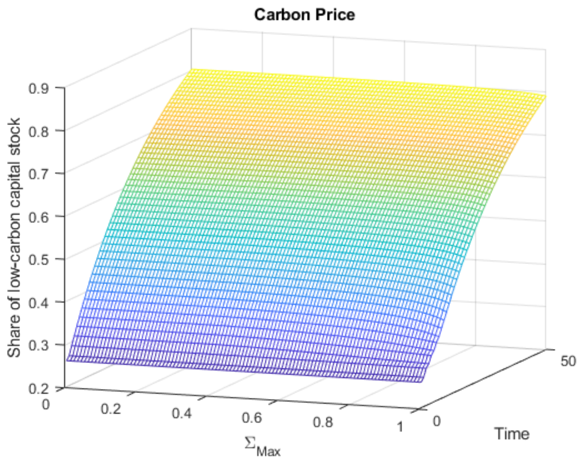
(a) No-policy benchmark



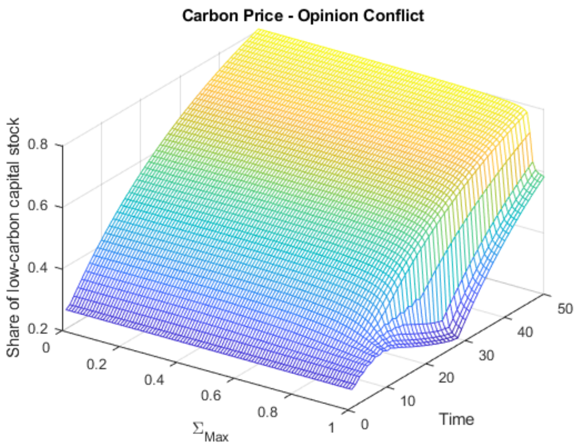
(b) Narrative update



(c) More ambitious targets for renewables



(d) Carbon price



(e) Carbon Price - Opinion Conflict

Figure 8: Different policy options in the model

Figure 8a shows the benchmark model run with two conflicting narratives without additional policy options. Second, in Figure 8b we assume that the regulator *updates* its norm according to the investors' performance through the following rule

$$r_N\left(1\right)=r_N$$
$$\forall t>1\ r_N\left(t\right)=r_N\left(1\right)+\eta_{r_N}\left(r_N\left(1\right)-r_L^{eff}\left(t\right)\right)\ \#\left(4.1\right)$$

With $\eta_{r_N}=0.5$.

Third, we assume that the regulator sets more ambitious policy targets as in Figure 8c (with $r_N=0.2$). As can be seen, being ambitious right from the outset allows for sufficiently high stranding expectations to trigger significant commitment to low-carbon investment regardless of dissensus. 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 dissensus on economic outcomes, but the effects seem to be limited. Implementing a univocal[8] carbon price of 300 \$ per ton CO_2[9] as in Figure 8d in leads to a transition for all levels of dissensus.

Univocal carbon tax. As can be seen, the introduction of a univocal carbon tax, although it does not prevent adverse dynamics down to a certain point, allows for a rebound of low-carbon investment after some time, regardless of the effect of dissensus. This is because the price signal, rightly anticipated by investors, becomes sufficiently high to make low-carbon investment profitable. As a result, 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.

Dissensus about carbon tax. 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, while the other share considers that the carbon price will not increase in subsequent periods. Results are displayed in Figure 8e.

As can be seen, even for relatively high levels of dissensus, applying a carbon price is sufficient to ensure a successful transition towards a low-carbon electricity system. However, for high dissensus levels, carbon pricing is first inefficient to propel a transition, until a point from which the tide turns, that is, when the direct – not expected – pecuniary incentive becomes high enough to shift investment behaviours. The mechanism is as follows: since, for high dissensus 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 direct pecuniary incentive becomes strong enough for investment behaviours to shift towards low-carbon energy sources. Although these effects seem to arise for a reduced constellation of parameters, they are nonetheless of chief 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 dissensus, 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 under scrutiny 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

[8] I.e. agents perfectly incorporate future carbon prices in their expectations.

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

incorporation of low-carbon (renewable) energy sources into the system, while such limits may exist. Indeed, investment behaviours may for instance suffer from incumbency biases (Unruh, 2000) and renewables may face technical constraints, for instance that of grid management and batteries, that may force the prolonging of some high-carbon infrastructures to ensure baseload servicing of demand.

5. Conclusion (unfinished draft)

The transition as a coordinated effort. Achieving the transition to a low-carbon electricity system will require coordinated efforts by investors, electricity producers, and policy makers alike. Most of all, despite encouraging market signals of increased profitability of renewables as compared to high-carbon electricity sources, efforts need to be greatly intensified worldwide to have the chance to meet the ambitious goals of the Paris Agreement. While profitability and cost considerations certainly are of primary importance for the low-carbon transition, the importance of competing narratives about the future, belief in targets set by the regulator, and dissensus 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 for agents when they imagine the future according to different narratives. These narratives are at the intersection of backward and forward-looking expectation formation mechanisms, and thus an ecologically rational decision making process under fundamental uncertainty.

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 that by explicit cost development paths. With this tool, 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. While our analysis is theoretical in nature, we contribute to the existing literature by showing some mechanisms at stake that were not highlighted to the fullest extent in the past.

Main conclusions. All in all, the main lesson that can be drawn from this broad exercise is that dissensus 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 dissensus 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 dissensus, which is a rather reassuring conclusion for our calibration area (Europe). 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 for lower shares of low-carbon capital and/or investment, would have yielded adverse dynamics much more easily. Applying this model to developing countries will therefore be an important area of application for further work.

Dissensus and social dynamics. Generally speaking, dissensus can be highly disturbing for social dynamics. The current covid-19 crisis provides a striking example: social adherence to mask wearing, respecting barriers gestures and social distancing is key in struggling against the virus and implementing the health policies. Essentially, social processes rely on social cooperation and consensus to a large extent...

Hope and despair. To end the cycle of our narrative view on the economics of transition processes in the energy sector, let us end with a quote from Beckett (1956). Here Vladimir might as well talk about the urgency of addressing the vital issues of humankind—before he plunges back into his usual mechanism of helplessness and dependency on something that is just not going to happen, i.e. that Godot might come one day. This might serve as a reminder that taking the right kind of action, if any at all, is not a simple thing to do, especially given the fundamental uncertainty of the future:

Let us not waste our time in idle discourse! Let us do something, while we have the chance! It is not every day that we are needed. Not indeed that we personally are needed. Others would meet the case equally well, if not better. To all mankind they were addressed, those cries for help still ringing in our ears! But at this place, at this moment of time, all mankind is us, whether we like it or not. Let us make the most of it, before it is too late! Let us represent worthily for one the foul brood to which a cruel fate consigned us! What do you say? It is true that when with folded arms we weigh the pros and cons we are no less a credit to our species. The tiger bounds to the help of his congeners without the least reflexion, or else he slinks away into the depths of the thickets. But that is not the question. What are we doing here, that is the question. And we are blessed in this, that we happen to know the answer. Yes, in the immense confusion one thing alone is clear. We are waiting for Godot to come—

Appendix A. Calibration

Appendix A.1. 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 A.1: Initial values for EU calibration

Appendix A.2. 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_{FF}	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
θ	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
r_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 dissensus in time S	1	Keynes (1936)
ρ_L	Transition premium value	1	Convention
p_{FF}	Price of the fossil fuel input	0.005	EIA (2017)
p_E	Price of the electricity output	0.2	European Commission (2017)

References

Arrow, K. J. (2013). Knowledge, Belief and the Economic System. *WIFO Monatsberichte (monthly reports) 86*(12), 943–951.

Becker, E. (1973). *The Denial of Death* (First Edition ed.). New York: Simon and Schuster.

Beckert, J. (2013). 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.

Berger, G. (1964, January). *Phénoménologie du temps et prospective.* Presses Universitaires de France.

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.

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

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

DiMaggio, P. (1998). The New Institutionalisms : Avenues of Collaboration. *Journal of Institutional and Theoretical Economics (JITE) / Zeitschrift für die gesamte Staatswissenschaft 154*(4), 696–705.

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

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

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

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

Godley, W. and M. Lavoie (2007). *Monetary Economics. An Integrated Approach to Credit, Money, Income, Production and Wealth.* Palgrave Macmillan, New York.

Graeber, D. (2012). *Debt: The First 5000 Years.* Penguin UK.

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.

Harari, Y. N. (2015). *Sapiens: A Brief History of Humankind* (First ed.). London: Vintage.

Harari, Y. N. (2017). *Homo Deus: A Brief History of Tomorrow* (First ed.). London: Vintage.

Harari, Y. N. (2019). *21 Lessons for the 21st Century* (First ed.). Vintage.

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

Hume, D. (1748). *An Enquiry Concerning Human Understanding* (First published ed.). Independently published.

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.

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

Kappes, S. A. and M. Milan (2019, July). Dealing with adaptive expectations in Stock-Flow consistent models. *Journal of Post Keynesian Economics 0*(0), 1–14.

Kay, J. and M. King (2020). *Radical Uncertainty: Decision-Making Beyond the Numbers.* New York, NY: W W NORTON & CO.

Keynes, J. M. (1936). *The General Theory of Employment, Interest and Money.* Palgrave Macmillan.

Knight, F. H. (1921). *Risk, Uncertainty and Profit.* Houghton Mifflin Company; Hart, Schaffner and Marx Prize Essays.

Lucas, R. (1976). Econometric Policy Evaluation: A Critique. In K. Brunner and A. Meltzer (Eds.), *The Phillips Curve and Labor Markets*, pp. 19–46. Carnegie-Rochester Conference Series on Public Policy 1. New York: American Elsevier.

Lucas, R. E. (1981). *Studies in Business-Cycle Theory.* Cambridge, Mass.: The MIT Press.

McKenna, E. J. and D. C. Zannoni (2000, December). Post Keynesian Economics and Nihilism. *Journal of Post Keynesian Economics 23*(2), 331–347.

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

Minsky, H. (1986). *Stabilizing an Unstable Economy.* Yale University Press, New Haven and London.

Muth, J. F. (1961). Rational Expectations and the Theory of Price Movements. *Econometrica 29*(3), 315–335.

Nietzsche, F. ([1976] 1873). Truth and lies in a nonmoral sense. In *The Portable Nietzsche* (translated by Walter Kaufmann, first ed.). New York: Viking Press.

Poli, R. (2014). Anticipation: What about turning the human and social sciences upside down? *Futures 64*, 15–18.

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.

Schelling, T. C. (1960). *The Strategy of Conflict*. Harvard University Press.

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

Taleb, N. N. (2005). *Fooled by Randomness: The Hidden Role of Chance in Life and in the Markets* (Second ed.). New York: Random House Trade Paperbacks.

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.