THE ISSUE OF DE-IDEALIZATION IN ECONOMIC MODELING: A CRITICAL EXAMINATION

By
Ida Korth

Submitted to Central European University - Private University
Department of Philosophy

In partial fulfilment of the requirements for the degree of Bachelor in Philosophy, Politics and Economics

Supervisor: Professor Maria Kronfeldner

Vienna, Austria
2025

COPYRIGHT NOTICE

[1]

[1] Icon by Font Awesome.

AUTHOR'S DECLARATION

I, the undersigned, **Ida Korth**, candidate for the BA degree in Philosophy, Politics, and Economics declare herewith that the present thesis is exclusively my own work, based on my research and only such external information as properly credited in notes and bibliography. I declare that no unidentified and illegitimate use was made of the work of others, and no part of the thesis infringes on any person's or institution's copyright. I also declare that no part of the thesis has been submitted in this form to any other institution of higher education for an academic degree.

Vienna, 26 May 2025

Ida Korth

ABSTRACT

This thesis critically examines the process of de-idealization in modeling and its implications for the economic sciences. There is a wide debate about whether idealized models in economics can be used to draw conclusions about the real world, given that they represent a distorted and inexact picture of it. Philosophers and economists frequently claim that the method of de-idealization can constitute a defense for the use of idealization in economic modeling. However, this claim has been significantly challenged. Various philosophers of science claim that de-idealization can only enhance the explanatory and predictive power of idealized models if they essentially claim something true. This thesis will discuss this argument and its implications for the method of de-idealization as a way of proceeding in the economic sciences. It will argue that de-idealization is a relatively conservative way of proceeding in science and might be particularly inefficient in the economic context.

TABLE OF CONTENTS

1. INTRODUCTION

Economic theory is full of idealized and unrealistic assumptions – ranging from the purely rational agent to perfect equilibrium, zero transaction costs, or time-invariant preferences. It would be easy to find more examples of the same kind. Such simplifications are at the core of economic models that intend to explain and predict real-world economic phenomena. Economic models represent simplified versions of otherwise overly complex market structures by idealizing them through the abstraction and distortion of certain properties (Rol 2008). However, the use of idealized models in science, particularly in economics, has been subject to a wide debate. Critics question how far they can be used to draw conclusions about the real world, considering that idealized models involve deliberate distortions of real phenomena. Prominent sources of this critique are Earman and Roberts (1999). The authors argue that idealized models would solely bring about *ceteris paribus* conclusions that possess no explanatory and predictive power when applied to the real world, in which additional interferences are operating.

To defend the use of idealization in scientific methodology, philosophers of science such as McMullin (1985) and Nowak (1992) claim that the explanatory and predictive power of idealized models can be strengthened by the process of *de-idealization.* Thereby, de-idealization is understood as the reversal process of idealization as it adds back complexities that have been previously omitted for purposes of simplification. By gradually adding back more complex and realistic assumptions, scientists can return to an increasingly precise approximation of reality.

Despite the frequent use of de-idealization as a defense for idealized models among philosophers of science and economists, de-idealization as a topic on its own has not gained much of attention. Knuuttila and Morgan (2019) have dedicated some of their research to the issue in (economic) modeling. Using the example of the homo economicus, they found that de-idealization can be particularly problematic as a defense for idealized models if de-idealization, for the purpose of approximating reality, would require adding complexities that contradict the original model. Since de-idealization always operates within the "conceptual framework" (Knuuttila and Morgan 2019, 652) of the original model, there would be no room for adding contradictions without annulling the model in its content and structure.

This thesis seeks to build upon the argument raised by Knuuttila and Morgan (2019) on the issue in question. I argue that de-idealization, as traditionally conceptualized, constitutes a conservative way of proceeding in science that may be insufficient to capture the dynamic and evolving nature of economic conditions. Going back to Knuuttila's and Morgan's (2019) example of the homo economicus I demonstrate that economics often refuses to abandon or revise outdated models even in the face of mounting empirical contradictions. It is argued that particularly in the economic context, de-idealization necessarily has to be transformative or even discard the preceding model in order to approximate real conditions. By specifically looking at the method of de-idealization, as McMullin defined it, this thesis aims to address two questions: can de-idealization be a method for defending the explanatory and predictive power of idealized models, especially in economics? And also, to the question of what broader implications this method has for scientific progress in the economic sciences?

The thesis is structured as follows. In Section 2, I will explain the broader debate around the issue of idealization and de-idealization. What follows is an analysis of the argument that

has been raised by Knuuttila and Morgan (2019) against the process of de-idealization in Section 3. Next, I will outline that de-idealization as such can be considered a conservative way of proceeding in science. I suggest that, particularly in the economic context, de-idealization often has to be more transformative.

2. THE DEBATE THUS FAR

The status of idealized models has been extensively discussed in the literature of philosophy of science and economics. However, comparatively little attention has been drawn to the process of de-idealization despite its frequent use among economists and philosophers to defend idealization in modeling and its explanatory and predictive power. In the following section, I shall review and contextualize some of the literature on the method of de-idealization as a potential defense for the use of idealized models. In doing so, I will start with the broader debate around the status of idealized models in science, particularly in economic sciences, and continue with the debate around the role of de-idealization.

2.1 The status of idealized models in science

In general terms, a model is a surrogate version of something in the real world, whether it's an object, a system, or some other phenomenon. When seeking a more precise answer to the question of what models constitute, it becomes clear that they come in a multitude of forms, ranging from physical to abstract, non-physical objects. In scientific theory, we usually encounter non-physical models that can, generally speaking, be considered representations of systems in the real world. When working with a model, the scientist aims to learn something about the real system they stand for. This is only possible if they represent the target system to a considerable degree, as for example Frigg and Nguyen (2022) have argued.

What kind of cognitive functions those representations fulfill and how they help the observer to learn something about the target system is interpreted differently among views about the status of models in science. There are two central interpretations regarding the cognitive functions of models: the realist and the instrumentalist interpretation. In the following,

I will introduce the central characteristics of both views about the cognitive functions of models to demonstrate that idealization is a particular challenge for the realist view of models.

Reiss (2012) characterizes instrumentalists as viewing science striving to provide a 'toolbox'. This toolbox helps the scientists to organize the world they are observing. Models and theories are not attempts to describe the world. Instead, they are devices with pragmatic intentions to understand the target systems. Here, truth is of limited importance, if any, as even falsehoods can be helpful for understanding the system in question. Therefore, models have a heuristic value as they are solely a means for the scientists to understand and discover the system for themselves. In that way, models must not represent the world in a literal sense, meaning that they constitute more or less precise representations of real-world systems. Rather, they are a means to an end that can introduce deliberate falsehood, allowing for the construction of hypothetical systems that do not exist in this way in the real world, as long as they help to find explanations for real world phenomena.

Meanwhile, realists, such as Mäki (2005) and Rorty (1979), argue that the knowledge-bearing nature of models lies in the resemblance between the model and the real world (Reiss 2012).[2] This resemblance enables scientists to study parts of the real world within a simplified setting. The general belief is that the conclusions from the model must be held in a similar way in the real world, given that the model presents veridical features of the target system. Thus, proponents of this view hold that models should represent the world in a truth-like and precise manner. It quickly becomes clear why the use of idealization in economic modeling can pose a

[2] In the philosophy of science the term "realism" has far-reaching connotations, yet for this thesis, it is only relevant that the realists, compare to the instrumentalists, aim for a veridical representation of the real-world phenomena. To avoid confusions with other implications of realism, Knuuttila and Morgan (2012) and other philosophers frequently refer to the view as the representation view.

significant challenge to this view: by idealizing, standard economic models inevitably distort the representation of the target system.

Reiss (2012) posits that in the realist conception, science strives to discover what Rorty (1979) called a "mirror of nature". However, it is commonplace that all models are false. That is because models can never be perfect representations of the system they aim to reflect. Reiss compares models to maps. By their very nature, maps involve numerous simplifications and do not depict every detail. They only capture truthfully those parts that are essential for understanding the geographical room they aim to map. Which parts they capture truthfully depends on the purpose of the map. Similar to models, the model never precisely represents real aspects of the world as they are. Instead, it necessarily leaves out certain descriptions and specifications that can be considered as given. To avoid trivializing their matter in favor of instrumentalism, realists have to commit to specific aspects of the target system that the model is supposed to be a veridical representation of. In the philosophy of science, this approach is often referred to, following Reiss, as partial realism. Variants of partial realism include entity, structural, or causal realism. Lawson (1997) and Mäki (2005) argue that causal realism is most relevant for economics. According to them, economic models essentially aim to represent truthfully casual relations. However, which parts of the target system are aimed to be represented in a verisimilarly way is of no relevance for the following inquiry. Relevant is that the realists conceive those models as possessing explanatory and predictive power if they are partial truthful representations of real systems.

2.2 Idealization in economic modeling

As Reiss (2012) noted, models and idealizations tend to go hand in hand. From the literature reviewed, it appears that there are multiple conceptions about the characteristics and

implications of idealizations. However, it is commonplace that the use of idealizations creates models that are simplified and distorted representations of the systems they aim to explain. McMullin (1985) understands idealization generally as a product of abstraction and distortion. Abstraction is typically understood as the process of isolating one or multiple properties of a system from the others, allowing them to be studied independently, that is, without the interference of other properties. Through the abstraction of certain features, the scientist might neglect causal factors that would, in reality, operate and affect the target system. Not always, but often, abstraction intends to eliminate all those properties that are considered 'unimportant' while highlighting those that are thought to essentially explain the system at hand, e.g., essential causal relations (Portides 2021). Abstraction is thus a partial representation of reality (Knuuttila and Morgan 2019)[3]. Additionally, McMullin's understanding includes that, often, idealization in model-building results from the distortion of features of the target system. Distortion is thereby a deliberate misrepresentation of reality through the introduction of false facts.

In economics, one can find a multitude of examples that include both the abstraction and distortion of features of the target system. The model of the homo economicus (to which I will return frequently) is a prevalent example from the field of microeconomics that includes both distortional and abstract features in its construction. By isolating the ability of rational thinking in humans and subtracting all the other characteristics and conditions influencing human behavior, e.g., volatility, risk-aversion, or the socio-historical and economic context in which an individual is being socialized, the model examines economic decisions as if solely driven by individualistic rational optimization. By assuming that rational agents have perfect knowledge about the present and the future, Knight (1921) promulgated an axiom that

[3] There is disagreement about whether abstraction is a partial, although veridical representation of reality. Knuuttila and Morgan (2019) argue that it solely subtracts certain features of the target system, while leaving the rest unchanged and verisimilar. Portides (2021) argues it is a deliberate misrepresentation of reality because by omitting certain properties the target system is changed and its causal relations are distorted as a whole.

represents a distorted reality. In reality, humans typically do not have complete information about the present and, especially, not about the future. The use of idealization, as seen in the example of the homo economicus, has significantly contributed to the constant concern of philosophers and economists about the explanatory and predictive power of economic models.

The concern stems from the following question: If models ought to represent essential aspects of real economic systems truthfully, are idealized models, which distort potentially essential features of the target system, legitimate for drawing conclusions about the real world? The conclusions and generalizations that are drawn as a result are, as mentioned in the introduction, accused of being valid only ceteris paribus, meaning they remain true only under the same conditions assumed by the theory that produced them. Reutlinger et al. (2024) summarizes that ceteris paribus generalizations are generally vague and not testable. Their vagueness stems from the fact that every failure of the generalization can be explained by a different factor for which there is independent evidence. As a result, the generalization remains unfalsifiable, preserving its truth at the cost of explanatory and predictive power for other cases in which the conditions are different. For example, suppose one concludes from the model of the homo economicus that, all else being equal, humans will always choose the option that will give them the highest utility. If then there is a case where most humans act differently, say they act altruistically for some reason, one could argue that this was due to factors of this specific case for which the model does not account. If those factors had not interfered, the conclusion derived from the model would have been most likely true. As the conclusion was drawn under the ceteris paribus condition, it would still remain true, even when there is no empirical justification for it.

So far, I have outlined the broader debate on the cognitive functions of models. Instrumentalists regard models as heuristic tools, while realists believe the cognitive function of the model lies in the partial resemblance between the model- and the real world. I then emphasized that, particularly for the realists, idealization in modeling poses a problem as it results in a deliberate distortion of the target system's features. The conclusions drawn from the models are accused of being valid only ceteris paribus and thus potentially possess no explanatory and predictive power in the real world. I now turn to de-idealization as a method for defending the use of idealization. I thereby elaborate on how the realists employ the method and then continue with the implications of de-idealizations as a method studied on its own.

2.3 The role of de-idealization

Among realists, de-idealization is commonly used to defend the use of idealization in economic modeling, as seen in McMullin (1985), Mäki (2005), and Weisberg (2007). That is because they hold that models must, at least in part, truthfully represent the world and that it is the resemblance between model and reality that enables scientific inquiry. In the realist view, de-idealization is seen as particularly helpful for making progress with the model, as it can potentially create an increasingly truthful representation of reality, approximating step-by-step real conditions. It is more precisely understood as the gradual adding back of complexities that were eliminated for simplification purposes. According to the realist view, the predictive and explanatory power of idealized models enhances as the resemblance between model- and real-world is thought to become more accurate.

Nowak (1992) and McMullin (1985) regard idealized models solely as intermediary versions of theories that must undergo de-idealization in order to bring about scientific progress.

In this view, the idealized model is assumed to be only the starting point of continuing research. Continuing research would then employ the method of de-idealization. Using Bohr's model of the hydrogen atom, McMullin shows the extent to which the initial idealized features have become increasingly precise approximations of reality and how this brought about great progress in atomic physics. He explains that in the early 1900s, scientists were trying to understand how electrons behave and how atoms emit light. The hydrogen atom indicated certain patterns of light when heated, and the Bohr model was one attempt to explain why. The model proposed a simplified assumption in which the electron is in a circular orbit around the proton, and the proton is assumed to be at rest. This assumption did not affect the explanatory and predictive power of the equation that Bohr was able to derive from the model, except where the hydrogen was subjected to an intense electrical field. According to McMullin, in 1916, Schwartzchild de-idealized the circular orbit assumption, using the more accurate assumption that electrons are in elliptical orbits, resulting in the model to match exactly the experiments that were made to test it. McMullin argues that the Bohr model and the preceding models gradually de-idealizing it yielded a significant breakthrough in atomic physics, demonstrating the successful implementation of idealization followed by its reverse process.

McMullin perceives only those models containing explanatory and predictive power that can ultimately be de-idealized. For an effective de-idealization, the idealized model must, among other things, capture the true essence of the target system that it aims to describe (McMullin 1985, 259). Only then the model can acquire increasing realism through a process of de-idealization. This seems quite plausible: if the starting point does not resemble the real world in any particular way, one cannot expect that subsequent modifications will bring about increasing realism, so a more veridical representation of the model. If the model captures the essence of the target system, then any de-idealization processes will only enhance its

explanatory and predictive power. At the same time, the model's essential assumptions, which reflect the target system, should remain intact. This is why McMullin maintains a very specific conception of de-idealization. He understands it as a process that specifically targets the particular idealized assumptions and structures in the original model, reversing exactly those. The complexities that are added back correct and further develop the model's previous ideas that capture the "essence of the original problem situation" (McMullin 1985, 259). Given that they do so substantial scientific progress can be achieved.

Knuuttila and Morgan (2019) discuss four ways in which de-idealization can work in practice. Scientists could de-idealize by either recomposing, reformulating, situating, or concretizing the model. The authors criticize that especially concretizing an idealized model will, in many cases, not bring about the aspired aim, namely making a model a more truthful representation of the target system. It is falsely assumed that a more concrete version of the model is a more truthful representation of the real target system. Knuuttila and Morgan (2019) criticize this assumption, claiming that increasing concreteness does not necessarily equate to increasing realism. This criticism will become clearer when demonstrated in the example of the homo economicus in the following section.

Having outlined the broader debate on the issue of idealization and de-idealization in science, particularly in economics, I now move on to specific issues of de-idealization and its implications in the economic context.

3. THE ISSUE OF DE-IDEALIZATION PROCESSES

3.1 The contradiction argument

In what follows, I will outline the critique that has been raised against the process of de-idealization as a possible defense for the explanatory and predictive power of idealized models. I start by analyzing the argument that I am designating the contradiction argument, as my argument will partly build upon aspects of it. As previously mentioned, Knuuttila and Morgan (2019) argue that it is often falsely assumed that increasing concreteness will bring about increasing realism. De-idealization will correct the simplified assumptions and replace them with more concrete and thus verisimilar ones. However, this is not always the case, especially when the original model essentially conveys something false. The authors demonstrate this criticism using the example of the homo economicus. The notion of the economic man was first developed by Adam Smith in 'Wealth of Nations' (Morgan 2006). He intended to provide a detailed and philosophical description of man's propensities and instincts (Morgan 2006) characterized by their natural tendency to economic activities. This idea later became the inspiration for the model of the homo economicus, which was, among others, developed by John Stuart Mill in 1836, who depicted the economic agent as a lazy, egoistic, and wealth-seeking man. Nowadays, the homo economicus is still used in neoclassical theory, specifically in the field of microeconomics (Urbina and Ruiz-Villaverde 2019). Frank Knight, who further refined the homo economicus in "Risk, Uncertainty, and Profit" (1921), recognized that his model is similar to a slot machine, devoid of reasoning power and intelligence and whose actions are precisely calculable and predictive. The homo economicus has full information about the future and the present and, against this background, consistently acts to maximize personal benefit. However, reality shows us that this is not always the case. Indeed, humans are

intelligent beings whose interests go beyond the aspects the model accounts for. While intelligence is a rather abstract characteristic, it contrasts with the primitive rationality of the homo economicus by encompassing the ability to weigh motivations in a given situation that go beyond pure self-interest and utility maximization. Consequently, the behavior of an intelligent human being cannot be fully captured and predicted by a model that only accounts for a limited number of factors. The authors argue that concretizing the model by, e.g., refining it mathematically, would not bring about increasing realism. Instead, approximating real conditions would require representing humans as intelligent and, therefore, inherently unpredictable beings.

In theory, McMullin addresses the objection that increasing realism cannot come about when the model states something wrong. In his paper, he notes that: 'This technique [of de-idealization] will work only if the original model idealizes the real structure of the object. To the extent that it does, one would expect the technique to work' (McMullin 1985, 261). McMullin claims that in cases where the model does not capture the true essence of the target system, the model has to be either revised and transformed or replaced with a "more coherent one" (McMullin 1985, 261). In response to Knuuttila's and Morgan's example of the homo economicus, McMullin would counter that the model does not capture the true essence of the phenomena. Real-world observations reveal evidence that contradicts the model: humans are not perfectly predictable beings whose actions are guided solely by utility maximization. One cannot expect that any kind of de-idealization process, such as concretizing the model, can bring about increasing realism. Therefore, McMullin would argue that it must either be revised or replaced by a more coherent model, so that it can actually approximate real conditions through de-idealization.

Although McMullin addressed the contradiction argument in his theory, Knuuttila and Morgan (2019) made an important observation about the conceptual boundness of de-idealization to the previous models, as demonstrated by their example of the homo economicus. The model's conceptual framework constitutes humans as agents whose actions and decisions are perfectly calculable and predictive. Adding intelligence to the model would contradict its original assumptions, which represent consistent behavior. Thus, de-idealization leaves no room for making modifications that contradict the original model as they cannot be added to the model without invalidating it at the same time. By emphasizing that the model in its content and structures does not allow for certain modifications that the model would need to undergo to approximate real phenomena, they emphasized that de-idealization, the way McMullin suggested it, is restricted to the model's conceptual framework. This is also how McMullin intended de-idealization to operate; it should always target the preceding idealized assumptions, gradually developing them further.

Morrison (2005) sums up the idea of the conceptual boundness of de-idealization processes concisely and clearly: McMullin's conception indicates that a model that can be successfully de-idealized is a good, approximately true representation of reality. The more it is de-idealized, the better its explanatory power. Morrison argues that this is only possible if there is a "set of approximately true assumptions, that remains constant throughout the de-idealization process" (Jebeile and Kennedy 2015, 4). This is what McMullin refers to as the essence of the model. As the model of the homo economicus showed, there are frequently observations that appear in the world that the model in its structure and content does not account for. Integrating those observations would require to either revise what the model is essentially conveying and transform it or alternatively replace the model entirely with a more coherent model. Transforming the model, could mean that essential structures and features are being

replaced by more truthful ones. Since, in practice, it is almost never the case that real-world observation correspond to what the model is essentially stating, Morrison (2005) questions whether de-idealization can ever work so that the essence of the idealized model can be retained. The criticism seems to point towards a central question: If for purposes of approximating truth, de-idealization would require to transform or even discard the original model in its structure and content, what remains of the model? Or more specifically, does the original model retain its explanatory and predictive power, and if not, what epistemic value does it offer?

3.2 De-idealization as a conservative method

Knuuttila's and Morgan's (2019) observation that de-idealization is always bound to the conceptual framework of the model raises doubts about the practical feasibility of the method as defined by McMullin. Instead of discussing the critique in more depth, I want to explore the implications of the argument when de-idealization is applied as a method of proceeding in the economic sciences. I will claim that de-idealization is a rather conservative way of proceeding in science as it does not allow for the transformation of the model and defines scientific progression by gradually refining previous ideas that have hitherto stood the test of time. I claim that this approach is potentially inappropriate for the economic discipline.

If de-idealization is frequently employed as a method in economics, then scientific progress would involve progressively refining and developing idealized models by systematically correcting their distorted features. McMullin's method involves the idea that de-idealization should preserve the essence of the preceding model, ultimately implying that it cannot be transformed. From this idea follows a prevalent implication for the question of how the scientist can proceed when de-idealizing: the process and its including modifications are

restricted by what the model essentially conveys. As Knuuttila and Morgan (2019) and Morrison (2005) also showed, de-idealization is restricted by the conceptual framework of the preceding model that it attempts to modify. From this, I take it that the process, as understood by McMullin, can be considered a rather conservative approach to scientific progression. By "conservative", I refer to the idea that as science proceeds, the scientist should further develop the ideas of previous models rather than discard or transform them. Thereby, scientific progress is thought to come about gradually.

The notion of conservatism is inherent in the process of de-idealization, as it intends to improve and develop ideas of preceding models. At this stage, I would like to invoke Kuhn's theory of the development of science, as it partly resonates with the idea of de-idealization as a conservative approach to scientific progression. In *The structure of scientific revolutions* (1962), Thomas S. Kuhn develops a theory on the progression of scientific knowledge. Indeed, Kuhn (1962) also found that science is characterized by commitments to existing models and theories. He describes the state of the 'normal science' as a period in which scientists commit to a common intellectual framework that constitutes a set of puzzles that are aimed to be explained, instruments to do so, and even an underlying metaphysics (Bird 2022). This intellectual framework is what Kuhn calls a *paradigm*. Similar to McMullin's account of de-idealization, Kuhn's notion of normal science argues that scientists build upon established models and paradigms rather than radically transforming or discarding them. During the period of normal science, scientists work within the framework of the underlying paradigm and proceed all in a similar manner. Kuhn observes that, in the phase of normal science, there is a conservative resistance to the attempted refutation of guiding theories. Bird (2022) states scientists neither test nor seek to confirm the guiding models and theories of their paradigm. Nor do they regard anomalies, that is, observations that speak against their theories, as falsifying them. Rather, anomalies are ignored or explained away due to their scientific commitments and

a lack of critical reflection. I take it that de-idealization can be understood as an instrument of the period of normal science in which certain models, including their underlying paradigms, are being upheld despite potential anomalies. The preceding model's ideas are further developed and specified, upholding the essential features that the model established. One will expect that de-idealization can actually improve the explanatory and predictive power as long as the model is essentially true. But what if the target systems in question could be better explained by entirely different aspects that previous models have overlooked? While de-idealization may increase the model's concreteness by further developing its ideas, it might still miss the point if the model essentially fails to capture salient aspects that explain the system at hand. In such cases, even a more concrete model may fail to enhance explanatory and predictive power simply because it continues to overlook key (causal) factors.

As Knuuttila and Morgan (2019) argue real-world observations indicate something fundamentally different than what the model of the homo economics is essentially conveying. To increase the realism of the model, there are modifications that would have to be made but can not be made without discarding or transforming it in its essential content and structure. A model that essentially depicts humans as slot machines cannot be de-idealized into a model that represents humans as intelligent beings without significantly transforming it or potentially discarding the entire model. However, instead of transforming the model, the development of the model shows how it was frequently concretized while keeping the essential features intact. Morgan (2006) provides an overview of this development, which I will briefly summarize in the following.

From its origin in the eighteenth century until now, the essential idea of the economic man as a rational, utility-seeking man was left more or less intact. After Knight further

developed Smith's idea and endowed the homo economicus with perfect knowledge about the present and the future, Jevons used those assumptions to develop a mathematical formulation enabling economists to use the model for the calculation of economic behavior. Jevons's model is considered the basis for neoclassical economics in the twentieth century. From then onwards, the model itself did not undertake significant transformation, only the narratives and discussions surrounding it. Morgan (2006) observes that in the second half of the twentieth century, the homo economicus was considered no longer a truthful representation of economic behavior but rather a normative model for guiding economic action. After the mid-twentieth century, multiple opposing views arose that aimed to challenge the widely accepted and anchored notion of the rational economic man. Knight himself noted that there are clearly some inadequacies in his model, as humans are essentially intelligent rather than machines performing consistent behavior. Others, like institutional economists such as Faber et al. (2002) developed the *homo politicus*, whose behavior is constructed from a given set of institutional and social influences. The homo politicus posits that "human beings do not care solely about their private interests in respect of their own individual preferences, but they also want to receive the approval from their fellow citizens for what they say and for what they do" (Faber et al. 2002, 328). Despite the criticism that emphasized contradicting evidence from market societies, the notion of the homo economicus is still a fundamental pillar of neoclassical economics today as it is structured around and based on the model of the homo economicus and its implications. Urbina and Ruiz-Villaverde (2019) argue that most theories in micro- but also in macroeconomics, directly or indirectly, require the assumption that humans act in accordance with the homo economicus model. For instance, the theory of the firm, consumer choice theory, or the wage-setting models all rely on the assumption that economic behavior aims consistently at utility maximization.

According to Knuuttila and Morgan (2019) de-idealizing the model would require transforming or discarding the essential features of the model. Yet the model continues to be referenced in mainstream economics, particularly in the field of micro- and macroeconomics, without significant revisions and transformations. This persistence comes at the cost of the model's ability to represent actual economic behavior, as it is observed in the real world.

3.3 The implications for economics

Until this point, I have argued that the conceptual boundness of de-idealization, which Knuuttila and Morgan (2019) have identified, makes it a rather conservative method of proceeding in science. This is particularly problematic when models are still continued to be developed even though they have been challenged by contradicting observations. Particularly in economics, there is an issue when de-idealization conservatively preserves the essential features of preceding models and their underlying paradigm. The purpose of de-idealization as a means to enhance the explanatory and predictive power of idealized models, which, according to the realists, depends on truthfully representing parts of the target system, is thereby lost. In what follows, I will elaborate on this claim.

As Mill (1843) noted, in most cases in economics, scientists can merely derive tendency laws. That is because conclusions that are drawn in theoretical models can only be approximately true in the real world. The model constitutes a very simplified version of real-world phenomena. Are conclusions derived from the model traced back to the real world, it will become clear that those conclusions do not hold in the same way as they do in the model world. They are only true ceteris paribus and are thus rather tendency than universal laws. Philosophers such as Hardt (2020) imply that discovering universal laws is more attainable in the natural

sciences because their target systems tend to be relatively stable, unlike those in economics. While physics, for instance, deals with rigorous target systems such as motion or gravitation, economics deals with inquiries about human behavior or institutions. The nature of the economic sciences is such that their systems are rather dynamic compared to the natural sciences. Hirshleifer (1985) argues that even in economics, one can identify regularities that are more or less universal, such as the law of supply and demand, which states that when a good is scarce, and demand is high, its price increases. However, this law also has exceptions. For instance, governments may intervene to prevent price increases on goods that are considered vital. MacKenzie (2006) addresses this type of scenario through the concept of 'performativity' in economics. He argues that knowledge about economics can also change economic conditions because it often flows into policy advice or shapes economic behavior. Most economic concepts are common knowledge, and as economic agents ourselves, we have a good pre-understanding of basic economic concepts. This pre-understanding can influence our economic behavior, just as the knowledge of economists shapes policy decisions (Knuuttila and Morgan 2012, 50). This creates a reflexive dynamic in which the subject of study and the study itself interact, constantly changing the conditions that are aimed to be explained or predicted. According to MacKenzie (2006), economic models are not just representations of the systems that they aim to explain, but they actively shape the status quo. Economic systems are characterized by institutions, culture, and human behavior, all of which can change significantly. As these factors change across different times and places, so too can economic realities, sometimes even drastically.

The issue hinges around the possibility that de-idealization as a method of preceding the ideas of previous models fails to capture the dynamic nature of economic conditions. De-idealization as an instrument for upholding models and paradigms could result in continuing to

develop models that are no longer relevant, as economic conditions have changed. However, to capture the dynamic nature of economic conditions, often a transition or rejection of preceding models might be necessary to achieve a veridical representation. Particularly within the neoclassical theory, there are multiple examples, such as the homo economicus, that are proven to essentially convey something false. If de-idealization is aimed to be implemented for purposes of approximating real conditions, the process must transform or discard the original model, in the economic sciences, potentially quite frequently.

At this point, I would like to bring an example of a model from the field of macroeconomics. The example intends to show that, also in economics, there is an initiative to reconsider and transform preceding models for a more truthful representation. The Solow model was introduced by Robert Merton Solow in 1956 and was for a while considered as the guiding model that mathematically explains how economies grow. The Solow growth model explains long-term economic growth through primarily technological progress, population growth, and capital accumulation (Prescott 1988). Much like other models of economic growth, it treats the environment as a relatively stable, neutral backdrop that has no interference with economic growth. However, recent findings from climate science show that environmental degradation, such as irreversible damage to the climate system, can negatively affect the conditions for economic activity and long-term growth (Acheampong and Opoku, 2023). These climate-related observations have not been accounted for in the Solow model and would require a radical revision of the model that includes assumptions related to environmental degradation. Even though the environment interacts with the explanandum, that is, long-term economic growth, the Solow model essentially does not account for this variable. Rather, it treats the environment as stable and concludes that technical progress and population growth will positively affect economic growth in the long run. If one considers these essential claims and

attempts to de-idealize them progressively, the scientists might get a more concrete picture of how those factors positively affect growth; however, to incorporate the environmental aspect, they would have to significantly transform the model. Scientists would have to acknowledge that long-term economic growth is not just naturally given by demographic changes and technical advancements but that it is always embedded within a set of externalities that shape the conditions for economic growth. Yet, there are models that have questioned and revised the Solow model, aiming to include the negative impacts of increasing environmental degradation. For instance, the Green Solow model takes, in addition to traditional economic factors, also factors such as the flow of pollution and environmental quality into account (Brock and Taylor 2010), similarly, models on green growth. Much like alternative models to the homo economicus, those green growth models do not define the economic sphere, yet, they play an increasing importance in policy advice (European Commission 2020).

As the example of the Solow model has shown, depending on the socio-historical context, new phenomena can arise that potentially undermine the relevance of previous theories. Nowadays, environmental degradation is increasingly relevant for evaluating long-term economic growth; preceding models that do not account for it necessarily have to be transformed or discarded. There are various examples, such as the Green Solow model, that demonstrate that there is an initiative to revise previous models in order to represent the conditions for economic growth more truthfully. For models to retain their explanatory and predictive power, such transformations are often necessary in light of a dynamically evolving world.

3.4 Some objections

In what follows, I would like to address potential objections to the idea that de-idealization has to be transformative rather than conservatively preserving the ideas of preceding models. This conservatism is particularly problematic when the real world indicates contradicting evidence. What has been previously discussed implies that de-idealization could potentially be misused to further develop models that have actually been proven wrong due to a lack of critical reflection or scientific commitments to a paradigm. As explained in section 3.1 of this paper, one could oppose that McMullin's conception of de-idealization is meant to be applied only up to the point where contradicting evidence appears. In this case, what the model is essentially stating is proven wrong. Also, Kuhn's theory did not just end with the period of normal science in which scientists are dogmatically continuing previous ideas even though they have no empirical justification. When anomalies are accumulating, proving theories and models wrong, science enters the phase of a crisis during which scientists seek alternatives to explain the anomalies (Bird 2022). When a convincing alternative is found to explain the anomalies, a scientific revolution follows. According to Kuhn, then, scientists would employ the method of de-idealization only until contradicting evidence begins to accumulate. Since economic structures are dynamically changing, paradigms and their models will follow the same path, meaning that models are being frequently discarded and replaced by more representative ones.

Now, particularly in economics, there are multiple examples, like the homo economicus, of models that are continued to be referenced even though there are real-world cases that falsify them. There is a wide discussion about whether economics can be considered a dogmatic science (Slattery et al. 2013), especially due to its adherence to neoclassical views that continue to dominate the science even though there are views and evidence opposing it. Now, particularly in economics, as MacKenzie posits, theories do actively shape the economic conditions that are

being observed. This contributes to the frequent criticism that economics is not only tied to empirical accuracy but also influenced by ideological convictions (Slattery et al. 2013). Urbina and Ruiz-Villaverde (2019) posit that the model of the homo economicus, for instance, constitutes a theoretical base for the ideological and moral legitimation of our whole economic system. The underlying logic is that if humans act rationally, the system as a whole is rational and effective, markets are in equilibrium, and the resources of society are allocated in the most efficient way. Given the model's central role in providing ideological and moral legitimacy to neoclassical theory, it is unsurprising that it remains firmly upheld.

The concern is that in the attempt to de-idealize models, the scientist may either be ideologically committed to the conceptual framework of the preceding model or do not critically reflect on whether it represents something true, dogmatically continuing an outdated model. Should the scientists be ideologically committed to the model, de-idealization could be misused as a strawman to defend and further develop ideologically relevant theories rather than those of truly representative value. As a result, they risk overlooking, dismissing, or explaining away alternative or contradictory evidence at the cost of the model's ability to represent things veridical.

Another objection is a sort of semantic observation that nevertheless has important implications for the discussion. If the previous discussion suggests that de-idealization should be understood in a way so that it can also operate outside the model's conceptual framework and transform or even discard the model, can we still call it de-idealization? In such cases, are scientists not effectively starting from scratch constructing an entirely new model? The prefix de- implies that it is a negation or reversal of something that has been previously idealized. The term alone implies that de-idealization specifically reverses the idealized features of the model,

necessarily operating within its conceptual framework. If de-idealization can also involve discarding certain features or even the entire model, can one still apply the term de-idealization then? To that, I would respond partly in the affirmative. While idealization aims at simplifying conditions, thereby distorting them, de-idealization, in contrast, has the overarching aim of approximating reality again (McMullin 1985). Any sort of modification that is done to reach this aim could thus be called de-idealization. Unless the model is completely discarded and pronounced useless, transformative modifications still originate from the model, albeit involving essential revision.

Sometimes, approximating real conditions will require de-idealization to transform or partly discard instead of operating within the conceptual framework of preceding models. It should do so, especially in the economic sciences, where conditions are dynamically and sometimes drastically changing. The concern is that de-idealization in the economic context could potentially be misused to defend unjustified convictions.

4. CONCLUDING REMARKS

This thesis aimed to critically discuss the method of de-idealization and its implications for the economic context. By that, it intended to discuss two central questions. Firstly, can de-idealization be effectively utilized to defend idealized models and their explanatory and predictive power? To answer this question, I referred to Knuuttila's and Morgan's (2019) argument, which demonstrated that de-idealization can only be effectively used if the original model captures the true essence of the target system. In theory, this is also how de-idealization was intended to be used. However, Morrison (2005) has argued that, in practice, most models are too distorted to actually capture the essence of the target system, while the real world often indicates contradicting evidence. De-idealizing these models would potentially result in more concrete versions of the original, but it would not yield representations that reflect real conditions more truthfully.

Secondly, I also discussed the broader question of what are the implications of employing de-idealization as a method to proceed in the economic sciences. I found that de-idealization can be considered a conservative way of proceeding in science because of the conceptual boundness to the essential ideas of preceding models. This restricts the economists to solely continuing previous ideas, preventing contradicting or conceptually different modifications to be made to idealized models. Thus, the economist solely operates within the boundaries of preceding models, in economics also because of ideological commitments. However, the nature of economics is dynamic. In order to represent this dynamic truthfully, economists should frequently transform or discard preceding models. The example of the homo economicus and the Green Solow model showed that in economics a refusal to transform and an initiative to do so coexist.

What this thesis does not address is the question of when exactly it is appropriate to abandon or transform the original model. It was assumed that such a point arises upon the emergence of contradictory evidence. However, it remains an open question whether a single instance of contradicting evidence that is potentially an isolated anomaly is enough to discard or transform the model. In such a case, the model might still retain sufficient explanatory and predictive power for other cases. Alternatively, should the criterion be that anomalies must accumulate? If so, what precisely constitutes this accumulation? Those questions regarding the discussion of when a model is considered falsified remain beyond the scope of this thesis and should be addressed in another inquiry.

To conclude, de-idealization as a method of proceeding in the economic sciences might be too restrictive, preventing necessary modifications that contradict or differ from the model's previous ideas can be made. This stems from the fact that de-idealization is always bound to the conceptual framework of the model. In order to make necessary transformations possible, de-idealization has to be understood in a less restrictive way. One obvious way to do so would be to understand de-idealization as a general method of approximating the real world. Any modification that serves this aim could be called de-idealization, including a transformation of the essential features of the model. In principle, this conception is very broad and would require some further refinement

BIBLIOGRAPHY

Acheampong, Alex O., and Eric E. O. Opoku. 2023. "Environmental Degradation and Economic Growth: Investigating Linkages and Potential Pathways." *Energy Economics* 123:106734.

Bird, Alexander. 2022. "Thomas Kuhn." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2022. Metaphysics Research Lab, Stanford University. Accessed May 26, 2025https://plato.stanford.edu/archives/spr2022/entrieshomas-kuhn/

Brock, William A., and M. Scott Taylor. 2010. "The Green Solow Model." *Journal of Economic Growth* 15 (2): 127-153.

Earman, John, and John Roberts. 1999. "Ceteris Paribus, There is No Problem of Provisos." *Synthese* 118 (3): 439–78.

Faber et al. 2002. "Homo Oeconomicus and Homo Politicus in Ecological Economics." *Ecological Economics* 40 (3): 323–333.

European Commission. "Green Transition - European Commission." Accessed May 26, 2025. https://reform-support.ec.europa.eu/what-we-do/green-transition_en.

Hirshleifer, Jack. 1985. "The Expanding Domain of Economics." *The American Economic Review* 75 (6): 53–68.

Jebeile, Julie, and Ashley G. Kennedy. 2015. "Explaining with Models: The Role of Idealizations." *International Studies in the Philosophy of Science* 29 (4): 383–92.

Knight, Frank H. *Risk, Uncertainty and Profit*. University of Chicago Press, 1921.

Knuuttila, Tarja, and Mary S. Morgan. 2012. "Models and Modelling in Economics." *Handbook of the Philosophy of Economics, U. Mäki, Ed., Forthcoming*.

———. 2019. "Deidealization: No Easy Reversals." *Philosophy of Science* 86 (4): 641–61.

Kuhn, Thomas S. *The Structure of Scientific Revolutions*. University of Chicago Press, 1962.

Lawson, Tony. *Economics and Reality*. Routledge, 1997.

MacKenzie, Donald. *An Engine, Not a Camera: How Financial Models Shape Markets*. The MIT Press, 2006.

McMullin, Ernan. 1985. "Galilean Idealization." *Studies in History and Philosophy of Science Part A* 16 (3): 247–73.

Morgan, Mary S. 2006. "Economic Man as Model Man: Ideal Types, Idealization and Caricatures." *Journal of the History of Economic Thought* 28 (1): 1–27.

Morrison, Margaret. 2005. "Approximating the Real: The Role of Idealizations in Physical Theory." *Poznan Studies in the Philosophy of the Sciences and the Humanities* 86 (1): 145–72.

Nguyen, James, and Roman Frigg. *Scientific Representation*. Cambridge University Press, 2022.

Nowak, Leszek. 1992. "The Idealizational Approach to Science: A Survey." *Poznan Studies in the Philosophy of the Sciences and the Humanities* 25 (1): 9–63.

Portides, Demetris. 2021. "Idealization and Abstraction in Scientific Modeling." *Synthese* 198 (24): 5873–95.

Prescott, Edward C. 1988. "Robert M. Solow's Neoclassical Growth Model: An Influential Contribution to Economics." *The Scandinavian Journal of Economics* 90 (1): 7–12.

Reiss, Julian. 2012. "Idealization and the Aims of Economics: Three Cheers for Instrumentalism." *Economics & Philosophy* 28 (3): 363–83.

Reutlinger et al. 2024. "Ceteris Paribus Laws." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta and Uri Nodelman. Metaphysics Research Lab, Stanford University. Accessed May 26, 2025. https://plato.stanford.edu/archives/spr2024/entries/ceteris-paribus/.

Rol, Menno. 2008. “Idealization, Abstraction, and the Policy Relevance of Economic Theories.” *Journal of Economic Methodology* 15 (1): 69–97.
Rorty. Richard. *Philosophy and the Mirror of Nature*. Princeton University Press, 1979.
Slattery et al. 2013. “Neoclassical Economics: Science or Neoliberal Ideology?,” *European Journal of Economics, Politics and Economic Law* 10 (3): 313–26.
Solow, Robert M. 1956. “A Contribution to the Theory of Economic Growth.” *The Quarterly Journal of Economics* 70 (1): 65–94.
Urbina, Dante A., and Alberto Ruiz-Villaverde. 2019. “A Critical Review of ‘Homo Economicus’ from Five Approaches.” *The American Journal of Economics and Sociology* 78 (1): 63–93.
Weisberg, Micheal. 2007. “Three Kinds of Idealization.” 2007. *The Journal of Philosophy* 104 (12): 639–59.