Reasoning with Theoretical Models

Rui Francisco de Souza Só Maia

"A sentence uttered makes a world appear
Where all things happen as it says they do;
We doubt the speaker, not the tongue we hear:
Words have no words for words that are not true"
W. H. Auden (1991, p.624)

Erasmus Institute for Philosophy and Economics
Erasmus University Rotterdam

Research Master Thesis by Rui Francisco de Souza Só Maia
ECTS Credits: 30
Supervisor: Dr. Ir. Marcel J. Boumans
Advisor: Prof. Roger E. Backhouse
Date of Completion: 27/6/2016
Word Count: Circa 27700

Acknowledgements

I would like to thank both Marcel and Roger for their assistance in the writing of this thesis. I would also like to thank Philippe, Lukas and especially Attilia for the many fruitful discussions we had about the argument of the thesis. I would further like to express my gratitude to the members of the EIPE student community for the years we spent together. And finally, I would like to thank Aline, to whom I dedicate this thesis.

Contents

Introduction (Chapter 0)

It is always a challenge to understand the scholarly practices of academic economists. They seldom self-consciously and systematically reflect upon the academic activities they are engaged in, and it often falls within the responsibility of someone else, such as a philosopher of economics, to undertake thoughtful investigation of them. Some of these practices will be more pervasive and, consequently, more paradigmatic of what goes on in the discipline. Still, given their ubiquity, these paradigmatic practices might exhibit such a degree of intricacy as to give even the most analytically-minded philosophers headaches. Take for instance a topic which has occupied scholars of the discipline for a long while now, the (extensive) modelling practices of economists. Here one may ask a number of intriguing questions: how do economists evaluate economic models? What is it that models do for economists? Do they teach economists anything about the world? These are some non-trivial philosophical interrogations whose answers hold the potential to shed light on different aspects of this important set of economic practices. Here, of course, a good strategy is to narrow down the investigation by focusing on specific (and tractable) research questions. Naturally, the scope and difficulty of such questions will vary greatly depending on the nature of the topic the philosopher is interested in, but the hope is that in the end this strategy will assist in the development of a more comprehensive picture of what goes in the economic discipline.

Now, as it turns out, my broad concern in this thesis is exactly with investigating the modelling practices of economists. And indeed, my approach to the topic is quite specific. I develop the thesis around a single research question: how do economists reason with theoretical models? But before I outline the account I develop in response to this interrogation, it will be useful to provide a few qualifications since the project might seem at first sight to be a bit broad. Note first that the question is rather descriptive. The idea is not to determine what consists good or appropriate modelling practice but instead to put forward an account of model-based reasoning which is consistent with actual disciplinary practice. 'Reasoning' here is to be understood as the cognitive process by which inferences are made about the world, irrespective of whether these inferences are epistemically justified. This does not mean, however, that the project is one of examining the actual brain mechanisms responsible for these inferences. After all, this is a philosophical essay. The intention instead is to provide an account situated at the phenomenological level which is at least plausible given what we know about the nature of human cognitive processes. My approach is therefore quite naturalistic. Note second that I am not interested in any modelling practice. I am only interested in the practices of the theoretical type. The marking characteristic of these practices is that they are somewhat removed from empirical considerations. What I mean by this will be explicated in detail in the next chapter. Furthermore, as it will become clearer later on, my account is intended to be quite partial. Although I do expect to provide a proper answer to the said research question, this answer focuses primarily on (very) limited inferential aspects of

model-based reasoning. To give one example, I do not elaborate on potential ethical considerations affecting model-based reasoning. Finally, the account I put forward here is not intended to be wholly original. On the contrary, it could be understood as an attempt to appraise and bring together major existing accounts of (model-based) reasoning. These qualifications apart, I now turn to provide a brief outline of my account.

My starting point is that to understand how economists reason with theoretical models one must first understand the nature of these entities. I submit that theoretical models are (mostly non-empirical) reasoning devices which are constituted of two main components, a formal structure and a narrative. Loosely speaking, the former comprises formal manipulation constraints whereas the latter comprises a host of literary supplements which, among other things, imbue models with meaning. Given this composition, I maintain, reasoning may proceed along two lines. The first one (which has been significantly developed in the literature) proceeds in the interplay between formal structures and narratives. There a formal system is created, interpreted and manipulated in terms of real world entities. This exercise allows the economist to demonstrate answers to definite economic questions she finds of relevance. The second and more controversial one proceeds in the interplay between the narrative and the reasoner herself. Narratives, I contend, may contain a story which lists a sequence of integrated events describing a situation of interest. The story prompts the economist to create and manipulate a mental representation (thought experiment) of this situation which in turns yields a plausible answer to an economic question she finds of relevance. This exercise is different from the previous one in that no explicit formal manipulations are made: the economist simply makes use of her background knowledge to produce and evaluate a thought experimental outcome. The derivative nature of mental representations vouches for her confidence that such representations can stand for their target. Finally, I suggest, both lines (or types) of reasoning with theoretical models seem to be desired by economists. The reason is that they appear to complement and corroborate each other in a way that increases the economist's belief in the appropriateness of the inferences she draws from economic models.

The argument I put forward to support my account is structured in the following way. The first chapter motivates and contextualizes the research question which inspires this thesis. There I explain why model-based reasoning is a topic which deserves attention and defend the claim that theoretical models constitute reasoning devices. I also present one type of reasoning which is possible with theoretical models: I first uphold the distinction between formal structures and narratives and then explain the way in which the interplay between these two components enable this particular type of reasoning.

The second chapter is dedicated to supporting the claim that another (distinct) type of reasoning with theoretical models is also possible. There I show how narratives can prompt reasoning on their own. I argue that narratives may contain what I call stories. These are special units of text which prompt the reader to imagine certain (thought experimental) situations: they allow the reader to infer outcomes about the world on the basis of mental operations which are categorically different from operations performed on formal structures.

The third and the fourth chapters contain two case studies that I employ to support the account developed in the preceding chapters. Each chapter examines a singular and authoritative economic model. The idea is to perform a comprehensive and practice-informed investigation of the argumentative strategy in each case in order to show that my two-fold account of model-based reasoning can be an appropriate characterization of the theoretical modelling practices of economists.

The fifth chapter is an attempt to provide an epistemic justification of the two-fold account of model-based reasoning I put forward. There I argue that economists do not usually take inferences obtained from reasoning with any single type of reasoning to be warranted in themselves and that the apparent shortcomings of each single type of reasoning can presumably be curtailed when both types are employed simultaneously. Thus it is reasonable to expect that economists want to make use of both types of reasoning. A final chapter then concludes the argument.

Now, before I proceed to the substantive argument of this thesis, it will be interesting to explain the motives which led me to write on the topic. In my experience whenever an economist wants to be serious about a point she will generally resort to some formal model to support it. Although these formal models do seem to grant some sort of authority (or deductive entailment) to the point the economist is making, I have often had a feeling that these points could also be made without the formalism in them. Indeed, I figured that very often that is exactly what would happen. Namely, a point would be supported both by a formal and by a non-formal argument. The problem however is that I feel that economists (and philosophers) tend to neglect the importance of this second non-formal argument. This is a pity for, in my opinion, the latter also plays an important role in bolstering the inferences economists make from theoretical models. This thesis, in turn, constitutes an attempt to rationalize these general impressions of mine.

The Nature and Function of Theoretical Models (Chapter 1)

Here the research question is motivated and contextualized. Two main components of theoretical models are identified, the narrative and the formal structure. A first type of reasoning based on these components is presented and defended.

Academic economists, like the members of many other scholarly communities, spend much of their time making use of specific tools in the quest to produce and disseminate knowledge about their subject matter. This observation is important for it suggests that a possible starting point for someone interested in studying the practices of economists[1] is to pay attention to the way these academics work with their tools. Among the various tools employed by economists to advance knowledge, perhaps the most quintessential one is that of formal modelling[2]. Professional journals are filled with models and they occupy a central role in the everyday academic life of the economist (Lipsey, 2001). In fact, some would go as far as to say that any economic idea worth of general recognition within the profession must in some way or another make use of this particular tool: "If it isn't modelled, it isn't economics, no matter how insightful" (Colander et al., 2004, p.493). The upshot is that an account of the way economists work must devote appropriate attention to their modelling practices.

Furthermore, the modelling practices of economists are quite varied and encompass econometric, theoretical and simulation models. Nonetheless, the argument put forward in this thesis is not applicable to all varieties of modelling. I will confine myself to discussing the particular practice of theoretical modelling. Whenever specific models or modelling activities are mentioned, the reader must bear in mind that the variety being referred to is the theoretical one, unless indicated otherwise. This does not mean that theoretical models do not have any relation to econometric or simulation models, but simply that these relations will not be discussed in this thesis. But what sort of thing is a theoretical model?

Following Alexandrova (2008, p.383) and Morgan (2012, p.13), I conceive of theoretical models as objects that are diagrammatic, algebraic and arithmetic in form and which typically come with a causal interpretation such that some of their premises are treated as describing a putative cause and some of their deductive consequents a putative effect. No systematic empirical data goes into the construction of these models. However, they characteristically describe familiar entities such as agents with beliefs and desires, or populations and organisms, while also imposing various idealizing assumptions on their behaviour. Familiar

[1] Although it is true that many economists can be found outside the scholarly environment, this thesis will focus on the research practices of academic economists.

[2] Boumans (2005, p.1) makes precisely the same point. He then proceeds to explain that he is concerned with empirical as opposed to theoretical models. In contrast, my interests, as it will be made clear, lies instead with the latter type of models.

examples of theoretical models include Samuelson's (1958) consumption-loan model, Akerlof's (1970) quality uncertainty model and Lucas' (1972) neutrality of money model.

Now, whatever one may think of the significance of any particular model, it seems clear that the practice in general is highly esteemed by the economic profession. On the one end, understanding and manipulating models is one of the first things taught to new students of economics (for instance, see Mankiw, 2015, ch.2). On the other end, various Nobel Memorial Prizes in Economic Sciences, arguably the most prestigious evidence of academic achievement in the discipline, have been awarded to eminent theoretical modellers (such as Thomas Schelling and Gary Becker). The pervasiveness of this practice is thus a defining characteristic of modern economics.

Puzzles of Model-Based Inference (1.1)

Historians and philosophers of economics have attended to this pervasiveness and have devoted much of their time to making sense both of the current modelling practices of economists and of the development of these practices through time. Yet, despite the combined efforts of these scholars, some aspects of these practices remain puzzling. Perhaps the most widely discussed (and problematic) aspect is the inferential justification of economic models. That is the question of what epistemic basis justifies inferences made from models. Although inferential justification is not the main topic of this thesis, briefly discussing this aspect will prove useful later on in the sense that it serves as a basis for comparison when dealing with other ways of making sense of models. Now, the puzzling nature of this aspect stems from the fact that scholars have not been able to provide a satisfactory account of (whether and) what kind of inferences can legitimately be drawn from theoretical models. Various proposals have been made, but none appears to be capable of both explaining and (soundly) justifying the inferences made by economists on the basis of such models.

For instance, Daniel Hausman proposes that theoretical models be understood as definitions. For him, they offer "conceptual, logical and mathematical explorations of the consequences of" formally defined systems (Hausman, 1992, p.77). They make no claim to truth. Instead, their point lies "in providing the conceptual means for making claims that can be tested and can be said to be true or false" (ibid., p.78). Thus, only when the assumptions of a model are true of some portion of the world that this model's claims are to be taken as making truth assertions. If this is the case, the "model may be used to state a general theory, to explain or to predict, or merely to state a fact about an individual" (ibid.).

This proposal is often objected to on the grounds that it only succeeds in a very "limited range of cases" (Alexandrova, 2008, p.389). The reason is that the overwhelming majority of economic models are designed with a degree of idealization so high that it prevents their assumptions from being true in virtually any portion of the world. Worse, the falsehoods, generalizations and omissions found in these models contrast so much with the complexity of economic reality that "economists are not usually able to explicate [their] content" in terms of

real world counterparts (Morgan and Knuuttila, 2012, p.78). The problem here is that if one attends to the importance economists attribute to theoretical models, they appear to allow for much more inferential work than would be allowed by this account. What this means is that Hausman's proposal would render much of the inferences made by economists vacuous since they often take their models to make theoretical claims that apply way beyond the scope of the simple systems used to derive them (Hindriks, 2013, p.524).

Another proposal is that of Robert Sugden. For him, models describe simple credible worlds. They allow inferences to be made to the extent that they can be understood "as a description of how the world could be", given "what we know (or think we know) about the general laws governing events in the real world" (Sugden, 2000, pp.24-5). They provide "an explanation of the world by virtue of an inductive inference: roughly, from the similarity of effects we infer a similarity of causes" (Sugden, 2013, p.241). Thus, by being similar to the world in credible ways, models are able to tell something about its workings.

Again, this proposal has also not gone uncontested. Among the various objections mounted against it, two seem particularly damaging. The first is that "it is difficult to see how models as credible constructions can license inferences concerning the real world" (Morgan and Knuuttila, 2012, p.79). Although our knowledge of how the world works does seem to impose constraints on what is possible, the number of credible states of the world seems nevertheless immensely higher than the number of actual worlds. With that in mind, it is difficult to believe that credibility is a requirement strong enough to justify inferences from models. The second is that the type of inductive inference implied by this account requires conditions not usually met by economic models. In the words of Cartwright (2009, p.52), induction should not be done

> "when we have good reason to think there are differences that matter between the inductive base and the target. [...] In many models we know by inspection that assumptions about the structure matter. Not only do they play a necessary role in the derivations we do in the models, in many cases [...] we know that with different structures we get different results. Induction then is a very bad idea".

The problem is that the results of many theoretical models are just too sensitive to different specifications to provide the sort of stability required to justify inductions. The significance of this point will be explained in an argument I make later on in this thesis but for now it is enough to note that the appropriateness of the type of induction Sugden identifies is problematic. These observations imply that the credible worlds account renders many of the model inferences economists make unwarranted and it is thus unable to satisfactorily make sense of the modelling practice. Now, the shortcomings of these two proposals taken together indicate that the inferential justification of theoretical models remains very much disputed. Indeed, many other distinct accounts of inference justification have been proposed and this issue continues to be much debated in the literature[3]. This suggests that attempts to make

[3] Recent examples include Mäki (2009), Alexandrova (2008), Cartwright (1999, 2007), Grüne-Yanoff (2009), Ylikoski and Aydinonat (2014) and Gilboa et al. (2014).

sense of the modelling practices of economists run the risk of encountering much controversy, at least if they aim to approach the topic from a justificatory perspective.

Models as Composite Reasoning Devices (1.2)

Nonetheless, inference justification is not all that there is to the practice of theoretical modelling. Despite the evident importance of knowing when one is licensed in exporting results from models, there are still other questions about the practice which have an importance of their own. For instance, how does the practice of modelling in economics relate to those in other disciplines? What are theoretical models made of? How do economists work with these objects? What is it that they do for economists? Further, finding answers to these questions might help to elucidate the more controversial issues surrounding inference justification. In fact, some scholars in recent years have decided to take a different approach on the study of such practices and have focused instead on tackling these other questions.

A major scholar to have done so is Mary Morgan. While her work spans over a wide range of topics in the history and philosophy of economics, a particular proposal of hers will be especially relevant for this thesis. Morgan takes an interesting position with respect to the modelling practices of economists. In fact, she largely sidesteps the controversy around inference justification and focuses instead on the relationship between economists and their models. Now, her general claim is that modelling constitutes "a practical mode of reasoning to gain knowledge about the economic world" (Morgan, 2012, p.18). She does so by following the historian of science Alistair C. Crombie, who distinguishes "in the classical scientific movement [various] styles of scientific thinking, or methods of scientific inquiry and demonstration", one of which the "hypothetical construction of analogical models" (Hacking, 1992, p.4). In this respect, she hopes to establish a link between the academic practices of economists and that of scientists in other disciplines.

For the most part, it is not my intention to discuss in any detail the merit of this interdisciplinary characterization of the modelling practices of economists. I will, however, dwell on one relevant point. The interesting aspect of this proposal is that it attends to a feature of modelling practices that is relevant *prior* to the problem of whether and how models provide reliable knowledge about the economic world: the proposal holds that models are first and foremost reasoning devices. They are "ways of finding out about the world" which support the scientist in her enquiry (Morgan, 2012, p.15). On this perspective, models possess the cognitive significance of prompting inferences about the world which is relatively independent of whether the inferences they prompt are justified. But if that is the case, an important question to ask is: what sort of reasoning (that is, of process by which inferences are made) is that which is conducted by means of theoretical modelling in economics? This is the question which in a sense inspires the general project of this thesis.

But in order to elaborate on this point, it will be necessary to first discuss another important aspect of theoretical models in economics: their composition. As defined above, the observer

of the practices of economists will recall that an important property theoretical models have in common is the fact that they are diagrammatic, algebraic or arithmetic in form[4]. One might thus be tempted to conclude that the form of these models constitute the essence of the whole practice and that to understand a given modelling exercise all one has to do is understand the nature of forms which are being employed. Yet, this conclusion is misguided. It ignores the fact that these forms not the only component of these objects, and may not even be the most important. This argument can be traced back (at least) to a contribution by Gibbard and Varian (1978) and can be nicely introduced in their own words:

> "A model, we shall say, is a story with a specified structure: to explain this catch phrase is to explain what a model is. The structure is given by the logical and mathematical form of a set of postulates, the assumptions of the model. The structure forms an uninterpreted system [...] In economists' use of models, there is always an element of interpretation: the model always tells a story" (Gibbard and Varian, 1978, p.666)

The specific reasons these two authors give to support their claim are not so important for the purposes of this thesis. What is important, however, is to notice that what this claim boils down to is that in order to understand how models work we should observe that they cannot be reduced to the formal structure which is so characteristically associated with them, they also need an element of interpretation. This conclusion has been recently and prominently developed by Morgan, who provided a broader and more elaborate set of arguments to underpin it. These arguments will be subsequently introduced.

One Way of Reasoning with Theoretical Models: Components and Interactions (1.3)

In a relatively early essay, Morgan puts forward the claim that theoretical models invariably contain a literary component. In her own words: "Modelling involves a style of scientific thinking in which the argument is structured by the [formal] model, but in which the application is achieved via a narrative prompted by an external fact, an imagined event or question to be answered" (Morgan, 2001, p.361). For her, the typical formalism present in economic models can only be made sense of in the light of an interpretation provided by the model narrative. Thus, inspired by the earlier proposal of Gibbard and Varian, Morgan creates a distinction between what she calls 'formal structures' and 'narratives'[5]. This distinction is important for, according to her, it is in the interaction of these two components that a salient type of model-based reasoning is conducted in economics.

---

[4] While my definition of theoretical models exclude the (other) kinds of models without the required form from consideration, the full composition of theoretical models is still a contingent matter.

[5] To be fair, Morgan's terminology can be quite confusing at times. She sometimes speaks of 'narratives' and 'stories' as synonyms and sometimes she attributes different roles to them. The same is true of the terms 'model structure', 'structure', 'formalism' and 'formal structure'. I will prioritize clarity and thus use the terms 'narrative', 'formal structure' and 'story' in distinct senses. While the first two will be shortly defined, the last one will be left for the following chapter.

A formal structure is that which provides the formal rules with which the model is to be manipulated. It is a structure because it constrains the possible ways models can be manipulated: "when an economist reasons with any model, he or she must obey certain reasoning rules" (Morgan, 2012, p.26). Further, this structure is formal in that the constraints it imposes on the manipulation of the model are rigorous. That is, these constraints are not made up "each time the economist works with a particular model". Instead, "they are given and fixed by the substance of the model" (ibid.)[6]. Furthermore, results that are obtained through reasoning on a formal structure have the property of being deductively warranted, that is to the extent that they follow from applying formal manipulation rules to premises explicitly stated in the model. Thus, for instance, when the model involves geometric constructions, the economist must not violate principles of geometry when manipulating it. The formal structure guarantees that the economist will follow certain acceptable (e.g., logical, mathematical) ways of reasoning and therefore imposes a certain amount of consistency among the possible models that can be created.

Now, it should be clear from this characterization that formal rules by themselves say nothing about what they are to be applied to. As Morgan (echoing Hausman, Gibbard and Varian) argues, formal structures provide uninterpreted systems of rules, symbols, predicates, quantifiers, etc., which can be used to deductively demonstrate the consequences of certain premises (Morgan, 2001, pp.366-7). Yet, she continues, "when we use economic models, we do so to relate the general (theoretical or law like) claims back to the specifics of the world" (ibid., p.377). The problem is that a pure formal structure is empty in the sense that its manipulation produces results which have no economic meaning. What is needed therefore is an interpretative complement to give the formal structure an appropriate economic meaning. This, she argues, is done by a narrative[7].

While the general function of a narrative is to complement a formal structure, the actual ways in which this goal is achieved are varied. In fact, one can distinguish in Morgan's contribution at least three major roles played by narratives. First, narratives are "the place where [...] demonstrations are interpreted within the world of the model and in terms of the things in the world that the model describes" (Morgan, 2012, p.238). That is, they provide interpretations (definitions) of the elements of the structure in terms of real world entities. This in a sense is the most basic role a narrative can play. For instance, the fact that an arbitrary cost function represents the costs of inputs for a given firm is not obvious from the structure of the function alone. A numerical function *tout court* will only become a cost function once its elements (arguments, domain, operations, etc.) are given the proper interpretation. Once the elements of the formal structure are interpreted, one can proceed with its manipulation.

[6] I use this sense of the word 'formal' without prejudice to a second sense Morgan uses to describe models, namely "to give form to, to shape, or to provide an outline of something" (2012, p.20). After all, theoretical models overwhelmingly share the same general mathematical 'form'. Further, I also hold no prejudice to the various other senses of the word 'formal' which have been proposed in the literature (for instance, see Krugman, 1997; Backhouse, 1998).

[7] For a related account of how this particular separation of model components can be applied to game theory in particular, see Grüne-Yanoff and Schweinzer (2008).

Additionally, as mentioned above, the manipulation of formal structures must obey certain rules. This is where Morgan sees a second role for narratives: it provides rules of manipulation which come not from the formal structure, but from "the economics subject matter" (ibid., p.26). In this case, the narrative provides the economic content which determines how a model is to be manipulated. It does so by assigning properties and relations to the elements of the formal structure, which in turn inform the economist of how certain aspects of the model are supposed to function. For instance, that the prices of the output of firms are fixed in a simple model of perfect competition is not implied by the formal structure of that model, but by the fact that firms in perfect competition are supposed to take prices as given. This information is obtained externally: it is imported from the interpretation of the formal structure imputed by the narrative. Note that this role is different from the previous one in that defining a variable as 'price' does not in itself determine how this variable is to be manipulated. It is the economic context in which this variable is interpreted (in the example, perfect competition) that constrains the values it can assume.

Finally, the third role Morgan attributes to narratives is that of giving purpose to the whole exercise of manipulating interpreted formal structures. According to her, narratives provide the questions which set up and guide the enquiry. That is, economists "create a model to answer a set of questions they find of interest. They manipulate the model to demonstrate the answers to those questions" (Morgan, 2012, p.225). This role is more general than the other two in that 'questions' justify the development of an interpreted formal structure which can provide 'answers' when manipulated in certain ways. Thus, for example, a partial equilibrium model will only be useful if it has the resources to answer relevant questions posed by the investigator (e.g., what would happen to prices if the government would increase indirect taxes by a certain amount?).

Morgan then combines these major roles of narratives with that of formal structures to obtain a unified account of model-based reasoning. On this particular account, a model's formal structure and narrative relate in intimate ways in order to produce an outcome that is of value to the economist:

> "using the model involves answering questions, and the structure cannot demonstrate these answers by sheer deductive logic or unadulterated mathematics without the prompt given by the question that sets off the changes. The question gives us the starting point; and there may be many possible ones using the constraints of the structure. Each time one of these questions is asked, the economist gives an account which begins with the situation proposed in the question and is taken to be a question about the world defined by the economic model. The elements in the model have to be mentally or physically shifted around on the diagram, or the algebra has to be manipulated and solved through, to suggest an answer or demonstrate outcomes" (Morgan, 2001, p.368)

While there is significantly more detail to Morgan's account of economic models than is here suggested, this brief reconstruction is enough to provide a sufficiently clear picture of the

formal structure and the narrative as their two main components, and of the way economists reason with these components. Note that the concept of 'narrative' is less specific than that of 'formal structure': it encompasses the many literary aspects of the practice of modelling that stand in a direct relation to the latter. From this perspective, what 'questioning models', 'constraining manipulation' and 'defining elements' have in common is that they all relate to the attribution of meaning[8] to the formal structure. Once the interpretation of a suitable formal structure is complete, the model is capable of fulfilling its role as a reasoning device: a demonstrative exercise whose execution is bound by certain rules constraining "the kinds of right reasoning that are possible" (Morgan, 2012, p.27).

Now, although certainly rich and thoughtful, I believe that this account of model-based reasoning does not exhaust the disciplinary practices of economists[9]. It focuses too much on the direct interaction between formal structure and narrative, and leaves the possibility that the narrative component may also play more separate roles somewhat unexplored. This is not to say that economists cannot reason along the lines just presented. On the contrary, this way of reasoning is perhaps the one most visible to an observer of the modelling practices in the discipline[10]. Still, I reckon that economists can also reason in a much less formal manner. But the argument for this moderately quirky (or, depending on your convictions, absolutely obvious) claim will have to wait until the next chapter. For now, it is enough to reiterate that the account of reasoning just discussed (and especially the division of theoretical models into two components) seems quite applicable to the modelling practices of economists. In fact, my argument in the following chapters will be largely an attempt to expand on this account and I will therefore spend some time defending it from a major criticism found in the literature.

Objection: There Is No Such Thing as a Narrative (1.4)

In a recent review of Mary Morgan's latest book 'The Model in the World', Daniel Hausman fiercely objects to the idea that reasoning conducted through economic models is inextricably bound up with narratives (Hausman, 2015, p.244). When criticizing Morgan's account of modelling, he claims to be unable to see any role for narratives and attempts to use a brief example to illustrate his point:

> "Consider a simple model of consumer choice in an elementary text in which a perfectly informed consumer faces a choice between two commodities. The consumer seeks to maximize a symmetrical utility function such as $U(q_x, q_y) = (q_x q_y)^5$ subject to the budget constraint $q_x p_x + q_y p_y \leq 1$. Solving the constrained maximization problem, the student can derive the quantities of x and y that are consumed as a

[8] Both in the semantical and the functional sense of the word.

[9] Note that I am aware that Morgan does not believe that this account exhausts the modelling practices of economists. Indeed, she explicitly acknowledges that there are other ways of reasoning with theoretical models (see specifically Morgan, 2004). Still, I find it difficult to locate in her work (or in the literature in general) a *comprehensive* philosophical discussion of these other ways.

[10] Compare, for instance, Knuuttila (2011); Kuorikoski and Ylikoski (2015).

> function of their prices and the consumer's income. If the student is then asked what will happen if the consumer's income increases, the student can deduce that the consumption of both *x* and *y* will increase. Where is the narrative that, Morgan claims, is 'inextricably bound up with' this demonstration?" (ibid., pp.244-5)

Here I wonder whether Morgan and Hausman are talking about the same thing, for the answer to Hausman's question seems to be 'all over the example'. First, if one would ignore the interpretative services provided by the narrative, the example would become simply an exercise in constrained optimization. Indeed, it would become something like:

> Maximize the function $U(q_x, q_y) = (q_x q_y)^5$subject to $q_x p_x + q_y p_y \leq 1$. Now, maximize the same function subject to $q_x p_x + q_y p_y \leq n, n > 1$.

This latter is clearly a mathematical exercise on a formal structure. It has no economic content and arguably no use for the economist whatsoever. Thus, Hausman must agree that an interpretative text (provided in the narrative) is necessary in order for the exercise to become one of theoretical modelling. Second, the elements of Hausman's example are interpreted in terms of real world entities: U is taken to be a *consumer's* utility function; q are the *quantities of commodities demanded*; p are the *prices* of these commodities and 1 is the *income*. These are not implicit in the formal structure, they need to be brought in from the outside (from the narrative). Third, the functional constraint can only be made sense of in the light of being conceived as a budget constraint: a budget limits the amount of goods a consumer may buy in a very specific way. In the absence of this interpretation, the specific form of the constraint could be completely arbitrary (e.g., $p_x p_y + q_x^{q_y} > 0$). Thus, the narrative informs the economist about the possible ways she can manipulate the interpreted structure. Finally, it is the narrative that poses the question that guides the whole exercise: in this case, 'what will happen if the consumer's income increases'? This question frames the problem in a way that allows from manipulation to occur. It suggests to the economist that she must replace the 1 on the constraint for a higher number and thereafter optimize the function subject to the new constraint. These observations show that, at least according to interpretation of Morgan's contribution put forward in this thesis, Hausman's worries seem unfounded.

Moreover, Hausman's position becomes intriguing once it is recognized that, as mentioned above, he himself believes that formal systems must have their predicates interpreted before they can be applied. What Morgan does is simply observe that models also possess other complements to their formal structures. She then notes that these are virtually always found in a literary form, which she names narratives. Now, Hausman may perhaps be questioning the usefulness of this overarching concept of narratives. But if that is the case the arguments he provides are simply unconvincing. Morgan does make a careful attempt to show the importance of narratives and to disambiguate its various aspects. Whether she succeeds or not can be questioned, but at any rate Hausman's argument does not seem to touch on this issue: he simply attempts to somehow deny the existence of the narratives, instead of showing that the concept as such is not applicable (or not useful). The upshot is that narratives, at least in the sense they have been discussed in this chapter, do seem to play an important

complementary role in the practice of theoretical modelling. Hausman's example shows that the formal structure and the narrative are not only there, but that they do seem to prompt reasoning in the manner suggested above. Moreover, it clearly shows how an uninterpreted formal structure has no economic meaning and arguably no use for the economist.

Another Way of Reasoning with Theoretical Models (Chapter 2)

Here a subcomponent of the narrative is identified, the story. A second independent type of reasoning based on the story is presented and defended.

Now that a type of reasoning based on the distinction between narrative and formal structure has been upheld, I can turn my attention to a second type of reasoning which I think is left somewhat underexplored in the literature. To recapitulate, the former type of reasoning with economic models occurs in the manipulation of interpreted formal structures (call an 'interpreted formal structure' a formal structure which has been given meaning by a narrative) to demonstrate answers to certain economic questions. This is done through a process in which the economist goes back and forth between the formal structure and the narrative with which it is linked: "the question proposes an event, which changes something in the model, which suggests that something else happens (and maybe another round of changes result) and then we arrive at a final outcome" (Morgan, 2001, p.369). In this process, the narrative mediates between the formal structure and the knowledge economists have about the economic world: "whenever we ask and answer these questions, we tend to bring in the interpretative level and discuss the changes in the elements in terms of the things in the world we have represented in our model" (ibid.).

This account may give the impression that only those activities which are in one way or another bound by the formal structure should be counted as reasoning. On the one hand, this appears to be consistent with the observation that economists often hold formalism in high regard. On the other hand, however, it precludes narratives from playing any role that is not directly related to the formal structure. This is where I attempt to expand on the concept of narrative[11]. While demonstrating answers to certain questions might certainly be considered a way of reasoning, I believe it does not exhaust the types of reasoning that can be conducted through theoretical modelling in economics. The reason, I contend, is that narratives play an important additional role, one that is not directly dependent on the formal structure.

A Different Basis for Reasoning? (2.1)

Besides the complementary roles extensively discussed in the previous chapter, narratives can also play a role that is associated with what most people have in mind when they hear the

[11] To be honest, Morgan herself sometimes suggest that narratives may have separate import. For instance, she passingly remarks in her latest book that narratives by themselves may provide "a distinctive form and level of understanding" (2012, p.245). And even though she soon returns to the idea that narratives link formal structures "with the facts of the messy economic world we all live in" (ibid., p.246), I suspect she would not object to the point I put forward in this chapter. Indeed, she makes a related (though quite underdeveloped) point in Morgan (2004).

term: a description of a series of events[12]. In this respect, I agree with Hartmann (1999) that models contain what he calls stories[13]. For him, "a story is a narrative told *around* the formalism of the model" (ibid., p.344). Further, and this is what gives them a separate role to play, a story is not "a deductive consequence of the model" (ibid.). In other terms, stories are literarily cohesive portions of text which describe relevant unfolding events and which are contained in narratives. As such, they exist alongside formal structures. But they are nonetheless distinct and relatively independent entities. Their role – unlike the ones identified in the previous chapter – is not to instil meaning to these structures. Instead, they seem to play a somewhat independent role: they prompt a different type of reasoning. Finally, as Hartmann points out, recognizing the existence of stories is important. They seem to "help us gain understanding" in a way that manipulating interpreted formal structures does not (ibid.). Yet, simply asserting that these entities exist does not get one very far in making sense of how economists work with models. One also needs a more detailed portrayal of their nature and an account of the way one reasons with them. That is what I subsequently attempt to do.

Perhaps the best way to fully grasp the nature and significance of stories is to begin by investigating their most basic (and general) functions. As discourse units of text, stories serve to convey information about sequences of events to the reader (or listener). They do that by supplying sentences which are interpreted by the reader, who proceeds by integrating "prior knowledge from long-term memory and working memory [...] with the information extracted from the text" (León and Peñalba, 2002, p.158). Now, the particulars of the process by which readers comprehend a text are not so relevant for the purposes of this thesis. What is relevant, however, is that readers appear to construct a 'mental representation' of the situation described "in order to understand, to interpret, to explain why a consequence could be produced, and to predict a consequence from an antecedent described in the text" (ibid.) (see also Zwaan & Radvansky, 1998). The import of this observation will be explicated below. In the meantime, it is best to keep in mind that theoretical models in economics come with a causal interpretation: since stories may easily and effectively convey causal information, one wonders whether they can do that by themselves. This possibility will be shown to be crucial for establishing the sort of cognitive role I claim stories play.

But how is it even possible for one to reason on the basis of stories? Aren't they just static pieces of text? Well, not really. First of all, it must be recognized that although stories are presented as fixed units of text they can only function once they are read (and comprehended). The access to the information in a story depends actively "on the reader, and in particular, on the reader's knowledge" (Noordman and Vonk, 1998, p.192). It is in integrating existing knowledge with the information extracted from the story that a reader comprehends the text (ibid., pp.195-6). What this means is that text comprehension always

[12] Indeed, a careful examination of narratives (qua literary complements) will indicate that they can play a great many number of additional roles. These seem to include, among other things, "mathematical derivations that are left implicit" (Backhouse, 2002, p.206) and rhetorical devices such as "figures of speech" (McCloskey, 1998). I aim for modesty and thus I will only focus one particular role, to be spelled out subsequently.

[13] Although he is a making a point about models in Hadron physics, I will argue that it also applies to theoretical models in economics.

involves cognitive effort from the part of the reader. Second, 'mental representations' of the situations described in the story are hardly ever static: events must unfold. A story which describes some kind of causal process will inevitably require that the reader's 'mental representation' of the situation be manipulated to accord with the contingencies implied by the specificities of the causal process (Khemlani et al., 2014). The upshot is that stories provide appropriate cognitive inputs to allow for inferences to be made by means of mental operations. In this sense, stories do indeed form a basis for reasoning.

Stories and the 'World' in the Mind (2.2)

Still, attributing a cognitive role to stories in general discourse comprehension situations is not the same as arguing that they play an equivalent role in theoretical modelling in economics. The latter conceivably requires that one specifies (and defends) the cognitive relevance of stories in the context of academic enquiry. This is what I intend to do in the remainder of this chapter. My claim is that the stories within economic models constitute a distinct and important reasoning tool in the scientist's toolbox. They serve to prompt what is known in the philosophical literature as 'thought experiments' (Kuhn, 1977; Sorensen, 1992; Brown, 2004)[14]. Loosely speaking, these are imaginary situations that have been invented by scholars with the intent to support specific contingent claims. They are considered to have had great importance in the historical development of both natural science (Mach, 1976) and philosophy (Gendler, 2006). The reason is that they appear to display evidential import in the corroboration of hypotheses. But before elaborating on the relation between thought experiments and economic theorizing, it will be appropriate to first provide a more precise conceptualization of the former.

I submit that a thought experiment "is the construction of a dynamical [scenario] in the mind by the scientist who imagines a sequence of events and processes and infers outcomes" (Nersessian, 1992, p.292). This imaginary scenario functions by "evoking certain quasi-sensory intuitions" which base the formation of new beliefs about "contingent features" of the world (Gendler, 2004, p.1162).

This definition is quite dense and requires some unpacking. First, thought experiments are 'thought' because they exist in the mind. The reasoner creates and manipulates a particular scenario, but "the access to this scenario is via imagination rather than via observation" (ibid., p.1155). Second, this scenario is not static, for otherwise it would not make sense to call it an 'experiment'. The reasoner imagines a dynamic scene that unfolds over time according to the specific instructions in the story. Still, it is the reasoner who controls the development of the

[14] This is not the first time thought experiments are claimed to be utilized in economics. For instance, Schabas (2008), Reiss (2012) and Maas (2014) make similar claims. However, their conceptions of thought experimentation differ considerably from mine. Rather, this thesis is very much inspired by Thoma (2016). She also points to the thought experiments prompted by stories within economic models. Yet, she does not have much to say about the nature of these stories nor about the type of reasoning involved in thought experimentation. My claim in this chapter can thus be partly understood as an attempt to provide context to and elaborate on her contribution.

events in the scenario. She can stop or restart the scenario at will, though she must obey the instructions in the story if she is to conduct the same thought experiment. Third, the thought experimental scenario is a mental representation[15]. Thought experimenting is rooted in the human capacity for simulative thinking. That is, it is "rooted in the ability to imagine – to depict in the mind – both real-world and imaginary situations, and to make inferences about future states of these situations based on current understandings, with and in the absence of the things reasoned about" (Nersessian, 2008, p.91). In this sense, the activity of conducting thought experiments is quite mundane. It is a sophisticated outgrowth "of the kinds of cognitive strategies humans employ in coping with their environments and in problem solving of a more ordinary kind" (ibid., p.6).

Some people believe that this definition is too broad for it includes too many activities that should not be considered thought experimenting. They claim, for instance, that thought experiments must involve "a certain degree of visualization, complexity, or novelty" (Gendler, 2006, p.2) and that they must lead to large-scale conceptual change in scientific thinking. Now, I have absolutely no quarrels with denying the title of 'thought experiment' to the type of activity I am referring to in this thesis. My aim is simply to point out that this activity exists and that it possesses cognitive import, thus being valuable for academic economists. If one accepts this claim, then I do not care about whether one agrees with the name given to the activity or not. In any case, a considerable portion of the literature also refers to this activity as thought experimentation and hence I see no problem in also doing so. This disclaimer aside, the relevant question for the purpose of this thesis is, how do thought experiments work and how exactly are they related to stories?

To being with, thought experiments are prompted by stories. While the original inventor of a thought experiment could arguably create and manipulate it entirely in her own mind, making the thought experiment public requires that it somehow be made available to other minds. This is where a story comes into being:

> "The purpose of a thought experimental narrative is to get others to arrive at the same inferential outcome as the original experimenter, to understand or grasp the significance of the point (i.e. to interpret the data in the same way), and to accept the significance of the problem it unveils, and possibly, though not always, a proposed solution" (Nersessian, 2007, p.127)

By writing a thought experimental story, a scientist is able to communicate her reasoning process in a very special manner. To understand a story, the scientist is forced to call "upon a combination of pre-existing conceptual and real-world knowledge and [to employ] the tacit and recursive inferencing mechanisms of her cognitive apparatus to integrate this with the information contained in the narrative" (Nersessian, 1992, p.294). Note that this is very different from reasoning on the basis of a formal structure. While the latter requires manipulating a structure by observing certain explicit formal rules (with an interpretation

[15] In fact, thought experiments are referred to in the literature as 'mental models' (Miščević, 1992, Nersessian, 2007). Since the term model usually carries other connotations, I chose to use the term 'mental representation' instead.

superimposed), the former proceeds by manipulating an already interpreted mental representation of a target situation.

Imaginary Elephants in the Living Room (2.3)

In fact, and this might surprise some philosophers and economists, experimental research (Johnson-Laird, 2010) suggests that mental representations requires much less cognitive effort to be *mentally* manipulated than formal structures. The reason is that manipulation constraints are already built into the mental representation of a situation, thus "making many consequences implicit that would require considerable inferential work if carried out" on formal structures (Nersessian, 2007, p.146). One might point that this makes reasoning on mental representations subject to implicit constraints and thus much less transparent than reasoning on formal structures. Yet, this is precisely the reason why the former type of reasoning can be cognitively advantageous: it reduces the load on working memory. Not having to follow explicitly stated constraints greatly increases one's ability to reach conclusions about certain situations (Evans, 2002). In any case, what is important is that mental representations can be profitably dissociated from formal structures.

Now, the view that human reasoning proceeds on a formal basis might be so intuitive for some people that it may be difficult for them to assess the validity of my claim in the abstract. Perhaps attending to an example will make the point more compelling. The example is taken (and slightly modified) from Gendler (2004, p.1156):

> "Think about your next-door neighbour's living room, and ask yourself the following questions: [...] If you removed all its furniture, could four elephants fit comfortably inside? If you removed all but one of the elephants, would there be enough space to ride a bicycle without tipping as you turned?"

Assuming that the reader is able to follow the instructions on this example, what sort of mental operations can been carried to answer these questions? Take the first question. One could obtain a result by mentally drawing a three dimensional prism (as one would do with a ruler and a piece of paper) and thereafter calculating the number of brick-shaped elephants that would fit into it on the basis of explicit geometrical operations. Nonetheless, this does not seem to be the most sensible way of approaching this problem. Numerical measurements for imaginary rooms and elephants are arguably hard to conceive of, and the geometry of the room might make it difficult to mathematically calculate the simultaneous distribution of all elephants in it. This is not to say that this calculation cannot be done mentally, but simply to point that this operation requires substantial cognitive effort and might be more appropriate in other circumstances (for instance, if one had knowledge of the dimensions of the objects and could make use of pen and paper)[16].

[16] Incidentally, philosophers working in the area of 'extended cognition' greatly emphasize the importance of external aids (such as calculators, pen and paper) in relieving certain burdens of human cognitive processing (e.g.,

One could also obtain a result by creating a mental representation of the room and thereafter mentally manipulating it in accordance with the instructions. This is different from performing formal operations on a Euclidean space. For instance, what sort of explicit mathematical operations does one need to perform in order to mentally remove furniture from the room? Arguably, this can be done by simply (and literally) making the furniture disappear. And if you agree with this claim, then it is not difficult to see that mentally fitting elephants in a room can proceed in a similar fashion. Gendler shows this by making an analogy between mental representations and real world objects:

> "suppose I had given you a three-dimensional scale-model of the room, along with four similarly scaled plastic elephants (and suppose it wasn't immediately clear whether or not the elephants could be placed comfortably therein): wouldn't you proceed by putting the elephants into the room, and seeing whether they fit?" (2004, p.1158)

Thus, it is in mentally experiencing whether the elephants fit in the room that one obtains an answer to the original question. And the manipulation of the imaginary elephants makes the answer evident without requiring any prior explicit formal operation. Now, someone may argue that this is not quite true since a scale-model of the situation would plausibly require that the real objects be measured and thereafter shrunk to a plastic shape while maintaining the right proportions. Yet, this objection misses the point. Gendler is arguing that once we have an appropriate representation of the situation, the answer to the question is made evident by the manipulation of that representation. One simply (and literally) attempts to fit the (plastic) elephants in the (scaled) room. This would be the case even if the construction of a mental representation of the situation would require the use of formal operations (as it might be the case with an actual scale-model). Moreover, and this is important, creating a mental representation is not the same as creating a miniature scale model. As mentioned above, mentally representing something does not require that one goes about measuring this something's dimensions and proportions. It only requires that one has quasi-sensory (and instinctive) experiences of the scenario one is reasoning about. Of course, this not does mean that mental representations cannot capture spatial relations, but simply that the construction of these do not follow explicit formal procedures.

Perhaps the point is obscured by the nature of the first question in the example. The reason is that one may be misled by the phrasing of that particular question into believing that thought experiments require mental visualization of objects. Yet, this is not necessarily the case. Thought experimenting involves the processing of diverse "perceptual and motor mechanisms" which yield a mental representation and thus "should not be conflated with a 'picture in the mind'". Indeed, "the representations on which these mechanisms operate can be highly abstract and schematic" (Nersessian, 2007, p.133). The second question of the example is a case in point. Mentally riding bikes involves more than just visual images. It

---

Clark and Chalmers, 1998). Indeed, some would even attempt to define models as "components of distributed cognitive systems" (Giere, 2002, p.227) (see also Knuuttila, 2011; Kuorikoski and Ylikoski, 2015). I cannot examine this account here in detail, but it will become clear throughout this thesis that I do not believe that reasoning with theoretical models (at least in economics) *necessarily* requires external aids (unless of course stories are also considered external aids).

requires accessing overt and tacit information about how to properly ride bikes, and this includes varied non-visual sensorial inputs (such as balance and acceleration perception). Further, what one imagines is not really an object, but oneself performing an action. Indeed, this observation is consistent with findings in social cognition indicating that humans possess a highly developed multi-modal capacity to both appreciate their own actions and empathize with the actions of others on the basis of mentally experiencing the execution of these actions (Calvo-Merino et al., 2006; Decety and Grèzes, 2006; Barsalou, 2009)[17]. The conclusion is that the domain of things that can be targeted by thought experiments is not limited to physical objects and spatial relations.

Yet, this domain is not without limits. A thought experiment is supposed to say something about the world and therefore it must at some point be bound to conform to it. This is achieved by requiring that the reasoner follows through the "sequence of events or processes as one would in the real world as currently understood" (Nersessian, 2007, p.147). Consequently, the reasoner's understanding of the way the world works is fundamental to the success of the thought experiment. This understanding derives both from the reasoner's real-world experiences and from her (theoretical) conceptualizations of these experiences (e.g., one must be familiar with elephants in order to ascertain whether they fit a given living room). Indeed, it is the fact that the reasoner follows her understanding of the world that grants scientific thought experiments a compelling force. If one were to realize a thought experiment, one would expect the actualized sequence of events to be consistent with what one had in mind. This is exactly what happens when they are successful, e.g. in a mundane case where one is figuring out how to move a chair through a door. Nonetheless, the idea behind these experiments is that one need not perform them in the world and so the very expectation of success will often suffice to compel the reasoner to accept the results (Nersessian, 1992, p.296). In this sense, thought experiments perform an epistemic trick. They force the reasoner to deploy her most deeply held (tacit and overt) convictions about the world and then reach a result on the basis of these. This result therefore inherits the plausibility of prior convictions and this make it much easier for the reader to be convinced of it[18].

By now, the reader of this thesis must be wondering what on earth imaginary elephants have to do with economic models. My answer would be: a lot! The example shows how reasoning need not proceed on an explicitly formal manner. Again, this is important since my claim is that the stories inside the narratives in theoretical models prompt thought experiments. These stories constitute the basis for a type of reasoning which is of a different nature than that done on formal structures: they prompt the reader to create and manipulate a mental representation of a situation, on the basis of which she forms new beliefs about contingent features of the world.

So far, I have focused on introducing two different types of reasoning and on explaining how they work. Yet, I still have to show how this two-fold account fares with respect to the actual

[17] The import of this point becomes clear once one notices that individual actions permeate much of the subject matter of economics (e.g., selling, hiring and choosing).

[18] I will expand on the significance of this observation in chapter five.

modelling practices of economists. I will do so in the following chapters where I make use of a pair of case studies to support my claim. The idea is that these cases will furnish concrete evidence in a manner consistent with the requirements of a naturalistic philosophy of science. For now, however, I wish to consider certain strong objections which could be raised against the account of reasoning I put forward in this chapter.

Objection: Unknowns Knowns (2.4)

One might have noticed that so far I have spent a lot of time trying to deny that reasoning through thought experiments is equivalent to reasoning through formal structures. Indeed, this must be the case for otherwise I could not differentiate this account from the one introduced in the previous chapter. But one might be unconvinced that such difference exists. One might hold that, although explicit formal operations are ruled out in the construction and manipulation of a mental representation, the latter in fact involves a host of formal operations which are of an implicit nature. In the context of the elephant example, one might claim that the reasoner implicitly retrieves from memory appropriate measures of the dimensions of the room and the elephants, and thereafter proceeds to create a (mental) geometrical rendition of these objects, which are in turn utilized by the reasoner to make accurate calculations of the fit of the elephants. Of course, all of this would proceed unconsciously and so the reasoner would not be aware of doing any formal operations.

The allure of this way of thinking about mental representations may arguably be said to derive from the fact that it depicts these representations as an extension of the human ability to perform formal operations. This makes theorizing about reasoning easier since existing knowledge of formal structures has been greatly developed by logicians and mathematicians over the centuries. Yet, this objection becomes difficult to support once one begins to attend to what it entails. Take the elephant example again. The objection would imply that the reasoner (including those without any training in geometry) would have a profound knowledge of the dimensions of the objects, as well as an instinctive capacity to make simultaneous non-trivial geometrical calculations. And all of this while being completely unaware of doing so. How plausible is this? Worse, if one considers that mental representations may have as targets an indeterminate number of different types of situations, this objection would be equivalent to stating that every person secretly masters classical logic, set theory, calculus, topology, etc.[19]

The problem with this objection is that it relies on very implausible claims about human cognition. It attributes to the unconscious a power which looks very suspicious from a conscious viewpoint. And even psychologists seem to agree. Their research show, for instance, that mentally operating formal structures can be quite difficult (Wason, 1968). The reason again is that formal operations impose substantial load on working memory. In fact, these

[19] And that is without considering the fact that mental representation can simulate not only relations among objects, but also actions.

operations can be so difficult in some cases that it would be virtually impossible for any human being to perform them, though they might be easily done with the assistance of pen and paper (Johnson-Laird, 2008, ch.8). Of course, the evidence is only suggestive since what is being theorized about is the very nature of unconscious thought. In the end, it is one's metaphysical commitments that will determine which position one will take: if one doesn't see a problem in attributing a mysterious power to an unobservable portion of the human mind, then one might accept the objection. Here I think we reach the limits of a philosophical argument, for there is no longer a principled reason to support (or reject) the objection. Yet, as argued above, a theory that separates mental representations from formal structures seems much more cognitively plausible than one that supposes that every person is unknowingly (and mysteriously) a mathematical genius[20].

Objection: Mechanisms in Disguise (2.5)

One might also have noticed that at no point in this chapter have I referred to the philosophical literature on causal mechanisms (Craver, 2006; Hedström and Ylikoski, 2010). One might find this strange since stories can be seen as carrying causal information (or at least sequences of events) in a way that could be said to refer to mechanisms. One might then conclude that my discussion of the cognitive roles of thought experiments can be understood as an attempt to identify causal mechanisms in economic models. Indeed, one might even suggest that my argument amounts to nothing more than window dressing and that I would be better off by simply focusing on how stories convey mechanisms. However, these remarks overlook an important difference between mechanisms and mental representations.

Mechanisms are useful "to explain how a phenomenon comes about or how some significant process works" (Machamer et al., 2000, p.2). In this sense, the stories I talk about in this thesis do seem to relate to causal mechanisms. Yet, and this is quite relevant, mechanisms are sought for different reasons than thought experiments. That is, reference to them is "is motivated by ontic, descriptive, and epistemological concerns" (ibid., p.4). A mechanism may arguably have three functions: it can identify the nature of a real phenomenon, it can describe how a real phenomenon comes about and it can explain a real phenomenon by pointing to the more basic processes and entities that constitute it. Note that all of these functions focus on the relation between the phenomena in the world and the mechanisms which might be conveyed by a story. But that is not the relation I am concerned with in this thesis.

My interest lies instead on the relation between the story and the reasoner. That is, I am primarily concerned with how economists reason with models and not with how these models relate to the world. And while it remains the case that the target of a thought experiment is

[20] Here perhaps some mathematical Platonists will raise some objections. I will be honest and remark that I am not interested in discussing arguments for or against Platonism in this thesis (for that, see Horsten, 2015). And in any case, it is not clear to me how questions about the nature of mathematical objects relate to questions about the cognitive processes behind scientific enquiry. I will thus leave it for the Platonist to take the initiative to respond to the argument I make in this chapter.

the world itself, there is arguably a distinction between a mental representation of the world and the way it actually is. Now, this does not mean that the account put forward in this chapter is inconsistent with one which purports that stories convey mechanisms. In fact, a story may identify (or explain, or describe) a mechanism in the world and simultaneously prompt the reader to create (and manipulate) a mental representation of this mechanism. But as long as there is a distinction between what goes on in the world (the actual mechanism) and what goes in the mind (the mental representation of the mechanism), these two functions of stories can be kept apart. And in any case, mental representations need not necessarily be of a mechanistic nature[21]. For example, the elephant story (again) does not involve any causal mechanism: it is simply an instance of reasoning about spatial relations. Therefore I cannot see the need to base my argument on the philosophical concept of mechanisms.

Objection: Failure Will Not Be Tolerated (2.6)

The critic of thought experimentation might have one final objection. She might point that the just mentioned distinction between the mind and the world might cause problems for an account of reasoning through thought experimentation. The reason is that the link between a mental representation and its target may be too weak to warrant valid conclusions on the basis of the manipulation of the former. Thus, "a thought experiment could fail to yield the correct result even though it persuaded everyone to accept the result" (Sorensen, 1992, p.31). Indeed, this leads many philosophers of science exhibit a degree of suspicion towards the activity. Carl Hempel forcefully explains:

> "of course, intuitive experiments-in-imagination are no substitute for the collection of empirical data by actual experimental or observational procedures. This is well illustrated by the numerous, intuitively quite plausible, imaginary experiments which have been adduced in an effort to refute the special theory of relativity; and as for imaginary experimentation in the social sciences, its outcome is liable to be affected by preconceived ideas, stereotypes, and other disturbing factors" (1970, p.165)

Thus, the problem with reasoning through thought experiments is that it may lead one to form false beliefs. And it may do so by deceiving the reasoner into thinking that a mental representation can stand for its target, though in reality it cannot. For instance, one's mental representation of an elephant may be just too big and it may lead one into believing that only two elephants will fit into one neighbour's living room, whereas in fact elephants are somewhat smaller and thus the said living room will actually accommodate three of them. Fair enough. But the relevant question is, does the fallibility of thought experiments count against the possibility of reasoning through them?

[21] Note that I do not therefore deny that the type of models I talk about in this thesis come with a causal interpretation (and perhaps even a causal mechanism). The point is simply that there is no *a priori* reason to limit mental representations to causal (mechanistic) situations.

I seriously doubt so. Of course, thought experimentation is by its very nature plagued by implicit concepts and tacit knowledge, and such unexamined information may clearly affect the validity of the conclusions one obtains. Yet, saying that these problems render thought experimentation unwarranted is too much of a stretch. After all, all methods of reasoning *about the world* are, in one way or another, fallible. For instance, celebrated and widely employed deductive methods (derivations from theory, RCTs, etc.) may guarantee their conclusions but "the conditions under which they can give [sound] conclusions at all are very strict" (Cartwright, 2007, p. 26). These methods are extremely narrow in their range of application and there is simply no guarantee that the world will lend itself to be targeted by them. In this sense, they are also fallible. And the same is true, for instance, for induction. One cannot deny that induction is an important method of reasoning in the scientist's toolbox. Yet, the validity of conclusions obtained by inductive reasoning is questionable (to say the least). Indeed, this problem has been the subject of considerable philosophical debate at least since David Hume (Lipton, 2004, ch.1). Therefore, the objection that thought experimentation is not a type of reasoning is too strong. It overweighs fallibility in a way that renders much of actual scientific reasoning unwarranted. The upshot is that there is no ideal way of reasoning about the world.

Case Study One: Reasoning about Unemployment (Chapter 3)

Here the two-fold account of reasoning put forward in this thesis is examined in the light of a case study.

Finding evidence to support claims made of the economic discipline can be very challenging. On the one hand, the discipline is in a sense quite diverse and it is not easy to find commonalities among the various subfields and practices economists engage in. Still, one cannot simply examine the whole of the discipline in order to support a claim. A sensible way to proceed is then to look for representative examples that capture the most important aspects of the way economics is done. On the other hand, however, determining what constitutes a representative example involves a certain degree of discretion. It is not always obvious what the most important aspects of the discipline are and one must make choices regarding which of these aspects one will focus on. This is precisely the problem facing the argument of this thesis.

Although the practice of theoretical modelling is just one subset of the various practices economists engage in, the list of possible cases which could be used to support my claim is simply too long to be considered here. Therefore, I decided to focus on a pair of examples which were selected according to the following two restrictions. First, I want my claim to be applicable to the way economics is done today. I have thus excluded examples which, though could certainly be taken to be prime cases of theoretical modelling, are somewhat removed from current developments in the discipline (such as Lucas' islands model (1972) and Akerlof's lemon model (1970)). Second, I want my claim to be supported by examples of the highest academic standard. Thus, I will only consider contributions which are relatively well-known and which were written by authors whose credentials are presumably undisputed. Finally, even though concision and space demands that I restrict myself to just a pair of cases, each of which will receive a thorough chapter-long treatment. The idea is to identify in each case the existence of a formal structure and an accompanying narrative, and in the process examine whether these components (especially the narrative) prompt the two types of reasoning I mention.

The first case I discuss in order to support my claim is the article 'Loss of Skill During Unemployment and the Persistence of Employment Shocks' by Christopher Pissarides (1992). This is a somewhat recent contribution in the area of labour economics which tackles certain issues related to unemployment dynamics. The author of the article is a well-known scholar who has recently been awarded the Nobel Memorial Prize in Economic Sciences precisely for his works on labour economics. Further, the academic relevance of this particular article is attested by the fact that it was published in one of the top journals of the discipline (The Quarterly Journal of Economics) and by the relatively high number of citations it has received.

It is also relevant because of the fact that it has been the subject of previous analysis by an eminent philosopher of science (see Cartwright, 1999, ch.7), though for reasons which are not directly related to the main claim of this thesis (but which in due time will be shown to have some relevance for the argument I put forward).

Demonstrating Unemployment Persistence (3.1)

In 'Loss of Skill During Unemployment', Christopher Pissarides is interested in studying the aggregate economic phenomenon known as 'unemployment persistence'. Namely, the generally observed fact that deviations of the unemployed rate from trend are serially correlated. Pissarides proposes that a possible cause of this phenomenon is the loss of job skills that unemployed workers incur by not exercising their productive abilities. He then goes on to develop an argument to show how this particular cause brings about its effect. The argument is mostly centered on a series of related interpreted formal structures which are progressively manipulated in order to demonstrate certain results. Since I am interested in his argumentative strategy, it will be best to begin the analysis from the very first interpreted formal structure and then advance from there. This will provide evidence for the first type of reasoning identified in the account put forward in this thesis, namely through the creation and manipulation of interpreted formal structures.

After introducing his article with (what in principle looks like) an ingenuous discussion of (the literature on) the phenomenon he is interested in, Pissarides proceeds to present his first and basic interpreted formal structure. An interesting characteristic of the model is that the interpreted structure is presented in a very descriptive way. This is how it begins:

> "[…]The model is in discrete time, and the dynamics derive from two overlapping generations of workers, each of which is of fixed size L, and a variable number of jobs that last for one period only. In each period of their life workers are in one of two states, employed or unemployed. If employed, the output they produce depends on their previous status. Young workers and old, previously employed, workers produce two units of output each. Old and previously unemployed workers produce $2y$ units of output, where $0 < y < 1$" (Pissarides, 1992, p.1373)

Note that the distinction between narrative and formal structure can already be seen in this short initial presentation. The formal structure provides constraints (e.g., there exists a given set with a size L) while the narrative provides interpretation (e.g., the set which has a size L is a generation of workers). These two components are interwoven and this makes it so that the formal system is seamlessly defined and interpreted in terms of real world entities. Notice the language: L is the size of a *generation* of *workers*, whose *product* $2y$ depends on whether they were *previously employed*. Again, this is necessary if the model is to refer to the real world since formal structures are not self-interpreted. The meaning of the terms L and y must be given from the outside. Still, providing interpretation of theoretical terms is simply one of the

roles played by narratives. A second role is that of providing additional manipulation constraints. Consider then the subsequent passage:

> "Unemployed workers always produce zero output and enjoy no utility or income. If y is strictly less than one, workers lose some of their skill during unemployment. [...]Employment in each period is determined by a matching process. At the beginning of each period all workers and a number of jobs that come into the market take part in matching. [...]Two types of workers take part in matching in each period. The first type consists of the L new entrants and those of the previous period's entrants who became employed. Workers of this type can produce two units of output each. The second type consists of workers who spent the first period of their life unemployed and who can produce only $2y$ units of output each. I refer to the former group as the short-term unemployed and to the latter as the long-term unemployed" (ibid., pp.1373-74)

Here the second function of narrative becomes evident. Observe that defining $2y$ as the product of a worker does not fix the values this variable may take. This is true even if one further defines these workers as unemployed. The definition by itself gives us no information whatsoever about how much a worker may produce. But Pissarides doesn't stop after defining y, he also provides some constraints for y to be manipulated. Namely, $0 < y < 1$. The question then is, where does this inequality come from? Well, the answer becomes clear once one attends to the narrative. His argument is that previously unemployed workers produce less than previously employed (or new) workers because they lose job skills. Since he fixes the product of previously employed (or new) workers at 2, it becomes apparent that previously unemployed workers must produce less than 2. Thus $y < 1$. Further, it does not seem reasonable to believe that previously unemployed workers suddenly lose all of their productive skills. Also, only currently unemployed workers produce nothing. Thus $0 < y$. Clearly, then, these constraints are introduced and justified from outside the formal structure. And although it is possible to imagine that one could introduce certain arbitrary rules of manipulation (e.g., $2L > y^2$), these would not make sense in the light of Pissarides' account of unemployment dynamics. It is only in the interaction with the narrative that the formal structure can be meaningfully manipulated.

Indeed, this claim is true is more than one sense. Namely, the narrative also plays a third role of giving purpose to the creation and manipulation of the formal structure. To recapitulate, Pissarides' whole undertaking is one of trying to make sense of the observation that "deviations of real economic aggregates from trend are serially correlated" (ibid., p.1371). The creation and manipulation of a formal structure should be understood as an attempt in this direction. Yet his purpose is articulated right at the beginning of the article, prior to any formal structure being introduced. This is necessary for otherwise one would not know what to do with the interpreted formal structure. After all, structures do not manipulate themselves. This will become clearer once the presentation of Pissarides' first mathematical structure (for variety's sake, call an 'interpreted formal structure' a 'mathematical structure') is complete. This will now be my task.

Note that the two previously cited passages already contain important information about what Pissarides is planning to do. Remember that he hopes to "show that if unemployed workers lose some of their skills during unemployment, aggregate employment can exhibit persistence" (ibid.). In the interpreted structure, skill loss is manifested in the productivity of workers: "If y is strictly less than one, workers lose some of their skill during unemployment" (ibid., p.1373). Thus, what he has to show is that the value of y is somehow linked with unemployment dynamics. But to be able to do so, he must complement his mathematical structure with additional definitions and constraints. It is not wise to reproduce his mathematical structure in its entirety for that would consume too much space. Instead, I will present only the parts which seem to be necessary for comprehending how he derives the results. That is:

> "The total number of workers who take part in matching in each period is $2L$, so, with matching probability q_t, employment in period t is given by $2Lq_t$. To find out what determines q_t, consider the matching technology and the number of jobs. Let the number of jobs that come into the market be J_t. With $2L$ workers and J_t jobs, job matchings are given by the function,
>
> $X_t = \min \{x(J_t,2L), J_t,2L\}$
>
> where X_t is equal to the level of employment in period t, $2Lq_t$, so q_t is on average equal to the ratio $X_t/2L$. I assume that this ratio is also the actual transition probability for each worker" (ibid., p.1375)

Thus, X_t is a generic labour market matching function which determines the level of employment. He does impose some supplementary constraints on this function (and on the function $x(.\ ,\ .)$), but that is not so important. What is important is that he can now locate the effects of skill loss. The mathematical structure is constructed in such a way that the matching probability q_t is equivalent to the level of employment in period t. Persistence can be then be understood as a relation between the level of employment q in period t and the same variable in the previous period. What remains to be shown is how q_t and y are related. He does so by introducing firms in the mathematical structure:

> "A job moves from vacancy to employment with average probability X_t/J_t, or $2Lq_t/J_t$. Firms make profit one when the job is taken by a short-term unemployed and profit y when it is taken by a long-term unemployed [wages are determined by Nash bargaining][22]. In period t the number of short-term unemployed is the sum of new entrants, L, and those who got jobs when they entered in period $t - 1$, Lq_{t-1}. The long-term unemployed are those who were not matched to a job when they entered the market in period $t - 1$, $L(1 - q_{t-1})$. [...] Therefore, the firm's expected profit from coming into the market with a vacant job is

[22] Note that the narrative here is also playing a role which I have not discussed in this thesis. It is referring to certain demonstrations which are not explicitly performed by Pissarides. The author is implying that profit (and wage) determination under Nash Bargaining proceeds exactly as he claims, but he leaves the demonstration of this claim to the reader.

$\pi_t = [1 + q_{t-1} + (1 - q_{t-1})y](Lq_t/J_t)$.

In equilibrium the number of jobs is such that no firm can make a positive profit by opening one more job. Assume that the cost of opening a job for one period is a constant $1/k$. Then, the condition for the supply of jobs is

$k\pi_t \leq 1$" (ibid., pp.1375-76)

Pissarides then goes on to show that, given the constraints of his formal structure, this inequality becomes an equality in a non-trivial market equilibrium (that is, excluding equilibria with zero jobs being supplied). Then, by substituting the previous equations, he shows that the number of equilibrium jobs available at time t is:

$J_t = Lk[1 + y + (1 - y)q_{t-1}]q_t$

Observe that the supply of jobs is made responsive to various labour market circumstances. "It is higher when workers are more likely to find the jobs (q_t higher); when the previous period's employment level is higher (q_{t-1} higher), because then there are relatively more short-term unemployed workers; when the long-term unemployed are more productive (y higher); and when the cost of operating a job is less (k higher)" (ibid., p.1376). Still, Pissarides main interest lies not in the supply of jobs per se, but on the probability of successful job matchings (which is equivalent to the employment level). He has to show that, given the existence of skill loss arising from unemployment, the employment level on a given period is dependent on the employment level on the previous period. After making a final round of manipulation of his mathematical structure, Pissarides arrives at the following differential equation:

$\{1 - x_1(.)[k/2][1 + y + (1 - y)q_{t-1}]\}\partial q_t/\partial q_{t-1} = [k/2](1 - y)q_t x_1(.)$

From which he concludes: "By the homogeneity of $x(.)$, $\partial q_t/\partial q_{t-1} > 0$. Thus, the dynamics of q_t are characterized by persistence" (ibid., p.1377). This completes the demonstration of the result. Note that the dynamics of q_t are only characterized by dependence if there is skill loss during unemployment ($0 < y < 1$). If there is no skill loss ($y = 1$), the right hand side of the equation becomes zero and $\partial q_t/\partial q_{t-1} = 0$, which implies no unemployment persistence. Thus, by reasoning from y to q_t, Pissarides demonstrates the link between skill loss and unemployment persistence in his mathematical structure.

It should be clear by now that the narrative's function is essential to the creation and manipulation of the structure. To give one final example, note that the total equilibrium supply of jobs is dependent on the equilibrium condition ($k\pi_t \leq 1$). This inequality, in turn, only holds if one is sure that k is greater than zero. But k is the inverse of the cost of opening a job and since it is arguably always costly to open a job, it follows that $k > 0$. Again, this information comes from outside the formal structure. Further, the narrative gives purpose to the whole undertaking by indicating the direction Pissarides must take if the claim is to be demonstrated. His goal was to account for unemployment persistence and the interpreted formal structure was created and manipulated with that purpose. Without it, he would have no reason to operate a formal structure (unless perhaps to prove some sort of incidental

mathematical theorem). Indeed, it is hard to conceive of him being able to achieve anything valuable to the economist without that goal in mind.

Now, Pissarides does not end his argument after presenting this basic formal structure. He subsequently attempts to amend it by substituting the generic matching function with a Cobb-Douglas one to further study the properties of the differential equation in q_t (the basic equation is highly nonlinear and does not determine the nature and stability of the equilibrium). It will not be necessary to discuss this amendment in detail since it does not change the main results. That is, it remains the case that "the extent of persistence in q_t depends on the loss of skill during unemployment. If there is no loss of skill, $y = 1$, and q_t is independent of q_{t-1}" (ibid., p.1378). He also shows that "if $y < 1$, there is persistence that increases as y falls" (ibid.).

Thereafter, Pissarides proceeds to similarly amend his basic structure in in order to assess whether certain other factors can account for the occurrence of unemployment persistence. He finds that, for instance, falling probabilities of matching as unemployment duration increases cannot account for the high degree of persistence in the unemployment level as well as his basic model (ibid., p.1382); and that falling search intensities on the part of the unemployed cannot properly account for the said phenomenon either (ibid., pp.1386-87). Now, and this is what makes this model a theoretical one, Pissarides does not in any way attempt to systematically test his argument against empirical evidence. What he does instead is simply express the hope that his proposal be tested "in future work" (ibid., p.1390). More importantly, however, his whole argumentative strategy appears to fit well the first type of reasoning identified in this thesis. The way he reasons with the amended structures is exactly the same as with the basic structure: he manipulates a formal structure to demonstrate certain points, and relies on the narrative to give meaning to the exercise.

A Short Story about the Labour Market (3.2)

The reader might by now feel somewhat uneasy with my argument. What I have done was simply to use Pissarides' contribution as evidence for one particular type of model-based reasoning. I have yet to show how his contribution constitutes evidence for the second type of reasoning identified in this thesis, namely through the creation and manipulation of mental representations. The reader will have noticed that nothing in what I have written so far indicates that the contribution contains a story which could be seen as prompting a thought experiment. Indeed, I suspect that most people who read Pissarides' article with the intention of studying the formal structure supporting his argument will probably reach exactly the same conclusion. The problem, however, is that this inclination arguably leads the reader to downplay the importance of the parts of the text which are not (at first sight) related to the author's mathematical exercise. Key literary aspects of the contribution risk being overlooked, including the fact that Pissarides did include a story inside the narrative. But where exactly is this story?

As briefly mentioned above, the presentation of the interpreted formal structure is preceded by what in principle seems to be an ingenuous discussion of the phenomenon he is interested in. Yet an attentive reading will reveal that this discussion is not as ingenuous as it first seemed. It includes a description of the process which, according to Pissarides, 'drives the results' of his interpreted formal structure (that is, the postulated link between skill loss and unemployment persistence):

> "To see this, consider a one-period negative shock to employment that reduces hiring, and so lengthens unemployment durations. If unemployed workers lose some of their skills and become less attractive to firms, fewer jobs come into the market next period. The market becomes "thin," because job seekers as a whole have less human capital. But with the number of jobs below trend after the shock, the unemployment durations of the new cohort of unemployed also increase above trend. So the market remains thin, even if the old unemployed have all left unemployment. The thin market leads to more job shortage which in turn perpetuates the thinness. Thus, the effects of the shock persist" (ibid., pp.1371-1372)

Observe how peculiar this passage is. It invites the reader to see the author's conclusion via imagining a hypothetical situation (note the language: 'To *see* this, *consider*…'). This situation is described by means of a story which lists a sequence of linked events that suggest how skill loss can cause unemployment persistence. Further, this passage is located fairly early in the introduction. It precedes the discussion of the existing literature on the topic and the presentation of the mathematical structures he uses to demonstrate his results. In this sense, the passage stands alone. The sequence of events which it describes is supposed to underlie the results demonstrated later in the article, but no reference whatsoever is made to the way these results are obtained[23]. The precise and explicit instructions on how to use one's imagination are missing. The terms (e.g., employment and human capital) are not given a formal definition and many of the links between the events (e.g., that firms reduce their jobs offers when facing a job market with lower average human capital) are not made explicit. Still, the author somehow expects the reader to be able to reach the same conclusion.

Well, is this an unreasonable expectation? Certainly not. That is, not if one renounces the idea that reasoning is necessarily equivalent to the manipulation of formal structures. In fact, I contend, the reader is perfectly able to reason on the basis of the story. And she does so through thought experimentation. She uses the story to construct a mental representation of the events described in it and then reasons through manipulating this representation until she arrives at the same conclusion as the author. To give one possible account of what might happen: the reader begins by imagining a cohort of people who had just became unemployed. Perhaps she has an existing mental representation of a typical unemployed person and all she does is multiply this representation until she feels satisfied with the number. She then

[23] To be honest, the author does state that the process he describes is "a variant of the thin market externality that characterizes equilibrium models of search" (Pissarides, 1992, pp.1371). But he does not present any model of search to show how *his* specific externality is obtained. Indeed, the externality is simply a *variant* of existing ones. Anyhow, it remains the case that the process described in the passage should be intelligible on its own since Pissarides' very aim is to produce a model of search which captures the process he proposes.

considers the skill loss claim. She is probably convinced that most skills are such that if they are not exercised regularly, they will gradually be forgotten or otherwise diminished with the passage of time. She then imagines the unemployed losing their job skills. Perhaps she recalls her own experience of losing certain skills and reasons by analogy with job skills. Thereafter, she considers the side of firms. She is convinced that if people have lost some skill, they will probably have difficulties in performing tasks which demand the use of that skill. She then reasons that unemployed people who have lost productive skills will probably produce less. Further, she is convinced that firms will prefer to hire people who are more productive. Perhaps she thinks that hiring less productive people is unprofitable and that training them is prohibitively expensive. She then attempts to imagine the interaction between firms and unemployed. Perhaps she pictures herself as the manager of a firm who wants to hire people, but who notices that the job applicants are not so skilled and reduces the number of job offers. She will then extrapolate this observation to the whole market and conjecture that less job offers are being made compared to a situation in which workers were more productive (i.e., in which the initial unemployment shock had not occurred). She is convinced that many people will remain unemployed and that these will continue to lose job skills, thus further deteriorating their productive capacities and attractiveness to firms. Finally, she reasons that the cycle of unemployment and skill loss will be renewed and concludes that therefore the level of unemployment will exhibit persistence.

Now, although the specific details may vary from person to person, I submit that some mental representation of this sort must occur if the reader is to comprehend Pissarides' story. This mental representation is based both on the reader's knowledge of the world and on the information contained in the story, and it leads the reader to reach the intended conclusion about unemployment persistence. Here, of course, the sceptic may feel unconvinced. She may ask: is it actually necessary that that seemingly innocent story produces such elaborate mental gymnastics? My answer would be: What is the alternative? Suppose, for the sake of argument, that the reader partially entertains previously known mathematical structures when she encounters certain concepts in the story. For instance, suppose the term 'unemployment' induces the reader to retrieve from memory her preferred model of unemployment determination. All right, but now what? Presumably, this pre-existing mathematical structure relates some variables (say, the level of output) with the level or rate of unemployment. Yet, it is not self-evident how to link the latter with the rest of the story: levels and rates do not lose job skills, unemployed people do. And if one has to mentally simulate the actual phenomenon of unemployed people losing skills, then the pre-existing mathematical structure is of little help. In this sense, these pre-existing mathematical concepts are simply too partial to allow the reader to follow through the instructions of the story.

Suppose then that the reader mentally constructs a complete interpreted formal structure on the basis of the story and then proceeds obtain the desired result by manipulating it. Does this solve the problem? I don't think so. First, the story does not refer to any formal structure and it is not clear how readers would be able to create one anew. If anything, the sentences it contains must have a meaning in natural language. And that is exactly what they are supposed to. As mentioned above, Pissarides' own claim is that the events described in the story

somehow 'drives' the results he obtains later by manipulating his formal structures (and not the other way around). This can only be the case if the story can be independently comprehended. Second, as argued in the previous chapter, if the events described in the story are to be understood in terms of real world counterparts, then the story must prompt the reader to make use of her pre-existing concepts and knowledge about the world in order to mentally simulate these events. Otherwise the reader would not actually be able to apprehend the intended meaning of the sentences nor grasp the relevant connections among them. Third, if the reader can already form a mental mathematical structure of the situation which clearly demonstrates how skill loss may cause persistence, what is the point of Pissarides' demonstrations in the first place? He could have finished the article right after presenting the story: the demonstrations would have been redundant.

At this point the sceptic may perhaps voice one last complaint against the argument in this section. She may concede that the story in question prompts a thought experiment but still doubt that this is done independently of the formal structure later introduced in the article. She may claim that the reason why Pissarides introduced a story in the article was simply to make the formal structure intelligible. That is, to provide a basis for the choices he made with respect to which economic factors driving the phenomenon of unemployment persistence to include in the structure. For instance, the sceptic may claim, the introduction of firms in the mathematical structure only becomes understandable once one is capable of mentally visualizing the importance of their predisposition to hire highly skilled workers to the process leading to unemployment persistence. In this sense, reasoning with the story would not be truly independent from reasoning with the formal structure.

Now, while one can certainly argue that the story somehow grounds the specificities of the formal structure[24], I believe the objection is confused. The reason is that it mixes the dependent nature of the intelligibility of the formal structure with the independent nature of the intelligibility of the story. What the sceptic is suggesting is that the story is necessary for the comprehension of the formal structure. In this case, it is the formal structure that is directly dependent on the story, not the other way around. Moreover, by not objecting to the claim that the story prompts a thought experiment, the sceptic implicitly acknowledges that the formal structure has no bearing on reasoning with the latter. Thus, she is actually not objecting to the claim that stories function relatively independently of formal structures.

I conclude that the best way to make sense of the said passage is to recognize that it is indeed a story which prompts a thought experiment in the mind of the reader. The reader reasons from unemployment, through skill loss and back to unemployment by simulating the events described in the passage. This is done independently of the formal structure, which is introduced and manipulated only later in the article. Still, the reader presumably achieves the same result as by manipulating the interpreted formal structure, namely the realization that skill loss is a possible cause of unemployment persistence. Therefore, Pissarides' article constitutes evidence for my claim that model-based reasoning may proceed on two lines. The

[24] Indeed, I will have something to say about the supporting relation between the story and the interpreted formal structure in the fifth chapter (though the relation I propose is much less direct).

first through the interaction between narratives and formal structures and the second through thought experimentation prompted by stories. I now proceed to further defend this claim in the light of a second case study.

Case Study Two: Reasoning about International Trade (Chapter 4)

Here the two-fold account of reasoning put forward in this thesis is examined in the light of another case study.

Using only one case study to support a claim made of what is perhaps the most widespread practice among academic economists can be risky. As mentioned before, the practice is just too diverse to admit of bold generalizations. This is the reason why I now turn to the examination of a second case study. It is expected that the survey of a theoretical model in an altogether different disciplinary field lends additional support to the argument put forward in this thesis. Again, the idea here is also to identify the existence of a formal structure and an accompanying narrative, and in the process examine whether these prompt the two types of reasoning I discuss.

The second case which I investigate is the article 'The Impact of Trade on Intra-industry Reallocation and Aggregate Industry Productivity' by Marc Melitz (2003). The author is a relatively (as compared to Christopher Pissarides) young professor at Harvard University who specializes in international trade and investment. His contribution is more recent than the previously discussed one and it pertains the area of international economics. The prominence of the contribution is attested by the fact it was published in a top journal of the discipline (Econometrica) and by the very substantial number of citations it has received. Furthermore, the nature of the contribution poses a challenge to the argument of this thesis since, as will be shown below, the author's argumentative strategy is overwhelmingly based on a formal structure.

Demonstrating the Effects of Trade (4.1)

In 'The Impact of Trade', Marc Melitz is interested in studying certain intra-industry phenomena resulting from international trade. Namely, the "large and persistent productivity differences among establishments in the same narrowly defined industries" which "are strongly correlated with the establishment's export status: relatively more productive establishments are much more likely to export" (Melitz, 2003, p.1695). Melitz proposes that exposure of a country to international trade induces "the more productive firms to export while simultaneously forcing the least productive firms to exit" (ibid.), and then goes on to develop an argument to support this proposal. As with the previous case study, the argument is mainly focused around an interpreted formal structure which is progressively manipulated in order to show how trade and industry dynamics are linked. Since I am interested in the author's argumentative strategy, I will build my argument by following the structure of the text. The aim in this section is to provide evidence for the first type of reasoning identified in

this thesis, namely the one which proceeds through the creation and manipulation of interpreted formal structures.

The way in which Melitz presents his contribution is overall remarkably systematic. Unlike Pissarides, he is quite brief in discussing his own proposal in the introduction and instead spends much of the section discussing the existing literature on the topic and previewing certain aspects of the interpreted formal structure he presents later. In fact, the introductory section is mostly a comparison of the specificities of his structure with that of already existing ones. It is supposed to indicate to the reader the ways in which his own contribution constitutes an improvement. This is to some extent understandable since the journal in which the article is published is well-known for specializing in contributions to the development of economic theory in its relation to statistics and mathematics. Now, right after finishing this brief introduction, Melitz proceeds to present his elaborate formal structure. He begins by addressing what he calls the 'demand':

> "The preferences of a representative consumer are given by a C.E.S. [constant elasticity of substitution] utility function over a continuum of goods indexed by ω:
>
> $$U = \left[\int_{\omega \in \Omega} q(\omega)^{\rho}\, d\omega\right]^{\frac{1}{\rho}}$$
>
> where the measure of the set Ω represents the mass of available goods. These goods are substitutes, implying $0 < \rho < 1$ and an elasticity of substitution between any two goods of $\sigma = 1/(1 - \rho) > 1$" (ibid., p.1698)

Note how Melitz is much less descriptive than Pissarides in the presentation of his structure. For instance, he does not explain what C.E.S. utility functions are since he takes for granted that the reader is familiar with what he is talking about. Yet, the passage is still a nice illustration of how narratives function within economic models. Take the first role of narratives (namely providing interpretation of theoretical terms). The reader cannot interpret all of the constituents of this particular formal structure by herself. That is why Melitz has to explain what the terms Ω, ρ, and σ are supposed to mean: mass of available goods and two parameters indicating the degree to which these goods are substitutes, respectively.

Subsequently, Melitz turns to consumer behaviour and proceeds to formalize it by "considering the set of varieties consumed as an aggregate good $Q \equiv U$ associated with an aggregate price [function]" (ibid.)[25]. He uses these functions to derive the optimal consumption and expenditure decisions for individual varieties, respectively $q(\omega)$ and $r(\omega)$. He then addresses the 'production':

> "There is a continuum of firms, each choosing to produce a different variety ω. Production requires only one factor, labour, which is inelastically supplied at its aggregate level L, an index of the economy's size. Firm technology is represented by a cost function that exhibits constant marginal cost with a fixed overhead cost. Labour

[25] I will not reproduce the complete formal structure of Melitz' model since it is quite extensive. What I will do instead is reproduce only the parts which I consider necessary to understand how he obtains his main results.

used is thus a linear function of output q: $l = f + q/\varphi$. All firms share the same fixed cost $f > 0$ but have different productivity levels indexed by $\varphi > 0$. For expositional simplicity, higher productivity is modelled as producing a symmetric variety at lower marginal cost" (ibid., p.1699)

Observe the language used to present the model: there is a set of *firms*, each of which *chooses* to *produce* one different *variety of substitute products*. The structure is being interpreted in terms of real world entities. The narrative is again playing its first role. Moreover, these two passages also show that the narrative is playing its second role, namely providing additional manipulation constraints. Take a look at the first passage again. Notice that the inequality $0 < \rho < 1$ is not fixed by the form of the utility function per se. As the author himself recognizes, it is the fact that the products ω are substitutes that constrains the values that ρ may assume. This information, however, comes from the outside of the formal structure. It comes from the meaning 'substitute products' has in the actual world (e.g., products that could be roughly used for the same purpose). The same is true for the cost function $l = f + q/\varphi$ in the second passage. That it is the result of the sum of marginal costs and fixed costs is not fixed by the structure. Instead, it is based on information about the way costs are supposed to be incurred by firms in the actual world.

Furthermore, the third role of the narrative (namely giving purpose to the exercise) should already be somewhat clear. Melitz has some 'questions' which he hopes the mathematical structure will 'answer'. Remember that he is interested in studying the effects of exposure to international trade on industry productivity and organization (the 'question'): he wants to make the case that more productive firms are more likely to export and that the least productive firms will be forced to leave the market (the 'answer'). That is the whole reason why he creates and manipulates his mathematical system. And this reason is given at the beginning of the article, prior to any formal structure being introduced. Still, this role may not be completely clear since he has not yet made a case to support his claim. One must first understand how the results of the structure are demonstrated in order to see how it 'answers' the 'question'. I therefore proceed to (compactly) reproduce the main steps of Melitz' mathematical argument.

Note that the structure presented so far is very limited. Melitz must find a way to further expand this structure in order to account for the phenomenon he is interested in (firm and industry dynamics). To that end, he introduces certain additional constraints:

> "Regardless of its productivity, each firm faces a residual demand curve with constant elasticity σ and thus chooses the same profit maximizing markup equal to $\sigma/(\sigma - 1) = 1/\rho$. This yields a pricing rule
>
> $p(\varphi) = w/\rho\varphi$
>
> where w is the common wage rate hereafter normalized to one. Firm profit is then
>
> $\pi(\varphi) = r(\varphi) - l(\varphi) = [r(\varphi)/\sigma] - f$

where $r(\varphi)$ is firm revenue and $r(\varphi)/\sigma$ is variable profit" (ibid., p.1699)

Notice that by substituting the previously mentioned labour cost function into this profit condition Melitz obtains a link between firm profit (π) and firm productivity (φ). This is relevant since the phenomenon he is interested in is supposed to affect both of these firm attributes. By manipulating these constraints, Melitz is able to show that "a more productive firm (higher φ) will be bigger (larger output and revenues), charge a lower price, and earn higher profits than a less productive firm" (ibid., p.1700). Yet, he still has to link the competitive attributes of individual firms with that of whole industry (which he is also interested in). He does so by aggregating individual firms: "An equilibrium will be characterized by a mass M of firms (and hence M goods) and a distribution $\mu(\varphi)$ of productivity levels over a subset of $(0, \infty)$" (ibid., p.1700). This allows him to obtain an index of industry productivity:

"$\tilde{\varphi} = \left[\int_0^{\infty} \varphi^{\sigma-1}\mu(\varphi)\, d\varphi\right]^{\frac{1}{\sigma-1}}$

where $\tilde{\varphi}$ is a weighted average of the firm productivity levels φ and is independent of the number of firms M. These weights reflect the relative output shares of firms with different productivity levels. $\tilde{\varphi}$ also represents aggregate productivity because it completely summarizes the information in the distribution of productivity levels $\mu(\varphi)$ relevant for all aggregate variables" (ibid.)

He then performs a round of manipulations in the formal structure to obtain $\bar{r}$ and $\bar{\pi}$, which respectively "represent both the average revenue and profit per firm as well as the revenue and profit level of the firm with average productivity level $\varphi = \tilde{\varphi}$" (ibid., p.1701). Note that these definitions and manipulations supply him with the means to examine the firm and industry attributes (productivity and profitability) that are relevant to his claim. That is, he now knows where to locate the effects of international trade exposure in his model. Nonetheless, he must still show how φ, r and π are simultaneously determined in the model. He does so by introducing industry dynamics:

"There is a large (unbounded) pool of prospective entrants into the industry. Prior to entry, firms are identical. To enter, firms must first make an initial investment, modelled as a fixed entry cost $f_e > 0$ (measured in units of labour), which is thereafter sunk. Firms then draw their initial productivity parameter φ from a common distribution $g(\varphi)$. $g(\varphi)$ has positive support over $(0, \infty)$ and has a continuous cumulative distribution $G(\varphi)$. Upon entry with a low productivity draw, a firm may decide to immediately exit and not produce. If the firm does produce, it then faces a constant (across productivity levels) probability δ in every period of a bad shock that would force it to exit. [...] This paper only considers steady state equilibria in which the aggregate variables remain constant over time. Since each firm's productivity level does not change over time, its optimal per period profit level (excluding f_e) will also remain constant. An entering firm with productivity φ would then immediately exit if this profit level were negative (and hence never produce)" (ibid.)

What the passage indicates is that firms are not aware of their productivity level prior to entering the market, though they must be minimally productive in order to survive in it. Melitz defines this lowest productivity level (or cutoff level) as φ^*: "since $\pi(0) = -f$ is negative, $\pi(\varphi^*)$ must be equal to zero" (ibid., p.1702). Further, he shows that "the average productivity level $\tilde{\varphi}$ is completely determined by the cutoff productivity level φ^* [and so] the average profit and revenue levels are also tied to the cutoff level φ^*" (ibid., p.1703). Two equilibrium conditions are then derived. The first is a relation between average profit per firm and the cutoff productivity level ('Zero Cutoff Profit' condition):

$$\bar{\pi} = fk(\varphi^*), \text{ where } k(\varphi^*) = [\tilde{\varphi}(\varphi^*)/\varphi^*]^{\sigma-1} - 1$$

All incumbent firms (besides the cutoff firm) will have positive profits and the expectation of these leads prospective firms to contemplate entering the market: the higher the expected profit the more willing the prospective firms will be to attempt to enter the market. And since there exists a probability δ of leaving the industry, incumbent firms will be on average more profitable than prospective firms. This condition therefore establishes an inverse relation between productivity and profits. However, the sole expectation of profit does not by itself determine firm survivability. Only firms above a certain productivity level will be able to cover their costs. This capacity to cover costs increases with the productivity of the firm. This constitutes the second equilibrium condition ('Free Entry' condition):

$$\bar{\pi} = \frac{\delta f_e}{1 - G(\varphi^*)}$$

These two conditions determine an equilibrium which Melitz graphically represents as follows (see Figure 1):

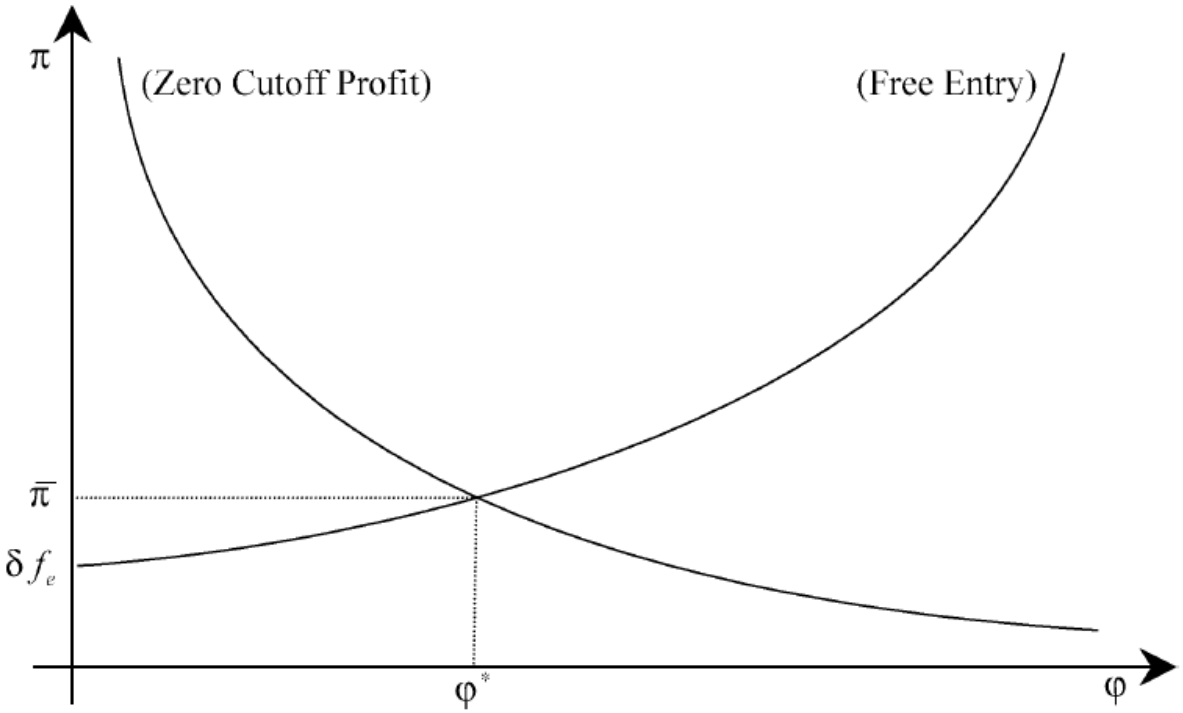

Figure 1. Reproduced from Melitz (2003, p.1704)

Where he shows the equilibrium $(\varphi^*, \bar{\pi})$ exists and is unique.

Now, all of these definitions and constraints combined supply a great variety of formal relations both at the firm and at the industry level. Yet, Melitz still has to address the central reason why such information is sought. That is, he must show how trade exposure is manifest in the model and then proceed to trace its effects. He does so by assuming that

> "the economy under study can trade with $n \geq 1$ other countries (the world is then comprised of $n + 1 \geq 2$ countries). Firms can export their products to any country, although entry into each of these export markets requires a fixed investment cost" (ibid., p.1707)

For notational simplicity, this cost is supposed to be incurred on every period and is defined as f_x. Furthermore, all firms which decide to export will also produce for their domestic market and thus "each firm's profit can be separated into portions earned from domestic sales, $\pi_d(\varphi)$, and export sales per country, $\pi_x(\varphi)$" (ibid., p.1709). This yields a new cutoff productivity level:

> "A firm who produces for its domestic market exports to all n countries if $\pi_x(\varphi) \geq 0$. Each firm's combined profit can then be written: $\pi(\varphi) = \pi_d(\varphi) + max\{0, n\pi_x(\varphi)\}$. [...] Additionally, $\varphi_x^* = inf\{\varphi : \varphi \geq \varphi^* \text{ and } \pi_x(\varphi) > 0\}$ now represents the cutoff productivity level for exporting firms. If $\varphi_x^* = \varphi^*$, then all firms in the industry export. [...] If $\varphi_x^* > \varphi^*$, then some firms (with productivity levels between φ^* and φ_x^*) produce exclusively for their domestic market" (ibid.)

Note that the decision problem faced by firms is now significantly more complex. They must first decide whether to enter the domestic market and then, given that they are successful, decide whether to export. As a result, Melitz must amend the Zero Cutoff Profit condition to account for the fact that these two decisions are made on the basis of two different productivity cutoff levels:

$\bar{\pi} = fk(\varphi^*) + p_x n f_x k(\varphi_x^*)$, where p_x is the fraction of firms that export

This new 'Zero Cutoff Profit' condition allows Melitz to complete the demonstration of the impacts of trade. Inspection of the new condition reveals that for every productivity level, the average profit will be higher: the Zero Cutoff Profit curve shifts up. Since Melitz shows that the Free Entry condition "is identical in both the closed and open economy", it follows that

> "exposure to trade induces an increase in the cutoff productivity level and in the average profit per firm. [The least productive firms] can no longer earn positive profits in the new trade equilibrium and therefore exit. Another selection process also occurs since only the firms with productivity levels above φ_x^* enter the export markets. This export market selection effect and the domestic market selection effect (of firms out of the industry) both reallocate market shares towards more efficient firms and contribute to an aggregate productivity gain" (ibid., p.1713)

Note that if a country is not exposed to trade, then $n = 0$ and the new Zero Profit Condition collapses into the previous one. Thus, by reasoning from n to π and φ, Melitz demonstrates his proposed link between exposure to trade, productivity and survivability.

Again, it should be clear by now that the narrative is essential in giving purpose to Melitz' exercise. Although his reasoning is quite formal, he makes sure to clearly motivate every single manipulation step in terms of his overall attempt to determine the impact of trade exposure. Indeed, that is undoubtedly one of the reasons why his formal structure is so elaborate: it has to account for various firm and industry attributes and show how they simultaneously respond to exposure to trade. Further, by demonstrating the results of his structure, Melitz is able to provide a principled 'answer' to the question of what the industry impacts of trade are. The narrative is therefore a fundamental component of his model.

Now, Melitz does not end his argument after demonstrating these results. He subsequently proceeds to amend his interpreted formal structure in order to make certain additional points. These points do not affect his main results and so it will not be necessary to discuss them in detail. In fact, they are simply further explorations of the impact of trade. To give one example, Melitz shows that an increase in number of trade partners will generate results that accentuate the impacts of trade exposure: "both market shares and profits are reallocated towards the more efficient firms. As was the case for the transition from autarky, this reallocation of market shares generates an aggregate productivity gain" (ibid., p.1717). Finally, just as in the previous case study, Melitz does not in any way attempt to empirically test his argument. What he does instead is make a very curious note: he concludes the article by claiming that he has "described and analysed" a "new transmission channel for the impact of trade on industry structure and performance" whose importance has already been highlighted by "recent empirical work" (ibid., p.1718). He speaks as if his mathematical structure was already inspired by empirical work and so did not need to be corroborated by evidence. In any case, the most important lesson to be drawn is that Melitz manipulates a formal structure to demonstrate his main points, and in doing so he relies on a narrative to give meaning to the exercise. Once again, this argumentative strategy appears to fit well the first type of reasoning identified in this thesis.

A Short Story about the Export Market (4.2)

Here the reader of this thesis might once more begin to experience some discomfort. I have given evidence for one particular type of model-based reasoning (based on interpreted formal structures) but I have not said anything about the second type (based on stories). This neglect however is intentional. The argumentative strategy pursued in the case examined in this chapter is overwhelmingly based on an interpreted formal structure and my (confessedly protracted) reconstruction of it is intended to make this absolutely clear. The zeal with which Melitz develops his formal argument appears to pose a strong challenge to my account: as shown above, he manages to demonstrate his main point (namely the industry effects of trade exposure) without resorting to any story whatsoever. Therefore, if I could still manage to find

a story in this case study then my claim that theoretical models may also prompt a type of reasoning which is not directly dependent on the formal structure would arguably have received strong evidential support. And indeed, I submit that this story can actually be found. As with the previous case study, a careful reading of the article will reveal that the argument contained in it is not exclusively formal. But again, where exactly is this story?

Well, in this case the story is not located in the introduction but instead at the end of the principal section of the article. Right after demonstrating his main point, Melitz introduces a sub-section entitled "Why Does Trade Force the Least Productive Firms to Exit?". It is a very short and discursive sub-section in which he expands on what he calls the "channels through which trade can affect the distribution of surviving firms" (Melitz, 2003, p.1715). Incidentally, it also includes a description of the 'channel' he claims to be operative in his just presented interpreted structure:

> "in the current model, all the effects of trade on the distribution of firms are channeled through a second mechanism operating through the domestic factor market where firms compete for a common source of labour: when entry into new export markets is costly, exposure to trade offers new profit opportunities only to the more productive firms who can "afford" to cover the entry cost. This also induces more entry as prospective firms respond to the higher potential returns associated with a good productivity draw. The increased labour demand by the more productive firms and new entrants bids up the real wage and forces the least productive firms to exit" (ibid., pp.1715-16)

Notice how curious this passage is. It claims to be related to the mathematical exercise but it has a quite literary character. It consists of an interconnected sequence of events that suggest how trade exposure can cause intra-industry reallocation. Further, the passage is supposed to have an elucidatory function: the idea is that the 'mechanism' described somehow explicates why the model yields the results that it does. Yet, this passage makes no explicit reference to the formal structure whatsoever. This is intriguing since, while the results of his theoretical model seemingly hold by virtue of the rigorous manipulation of the interpreted formal structure, the sequence of events described in the passage is not linked by explicit demonstrations. For instance, the term 'labour demand' is not given a formal definition and the precise process by which real wages induce the least productive firms to exit is not spelled out. Still, the reader is somehow supposed to understand the sequence of events and reach the same conclusion as the author, namely that exposure to trade causes intra-industry reallocations. How is it possible?

The answer, I claim, is that the passage is a story. The reader comprehends it not by manipulating the previously presented formal structure but by thought experimenting. She conducts a mental simulation of the situation described and reasons by following through the sequence of events until she reaches the desired conclusion. Again, to give one possible account of what might happen: the reader begins by imagining a cohort of firms in an industry. Perhaps she thinks of some good and then imagines herself as being the manager of

a firm which produces that good. She then multiplies the number of firms like hers and arrives at the simulation of the cohort. She subsequently attempts to imagine trade exposure. As the manager of a firm, she sees possible profit opportunities in the new export market. However, she believes that to export a firm must be very efficient: it must have a high productivity level in order to be able to face the costs of entering foreign market (here she might think of transport costs or of costs associated with establishing an outpost in the new country). Yet, the prospective profit opportunities are simply too good and seem to justify entry into the export market. Perhaps she imagines herself as a young entrepreneur who sees the higher profit opportunities in the export market and decides to take the chance. She multiplies the number of prospective entrants and reaches the conclusion that exposure to trade will increase the number of firms attempting to enter the industry. She subsequently proceeds to consider the input side. She knows that the number of workers domestically available is fixed (perhaps she thinks that the domestic population won't change much in the time span she is considering). She knows that firms entering the industry need employees to function and reasons that this will increase the demand for workers. She then imagines herself as a worker. She knows that the number of available jobs in the industry has increased and that she now has more leeway in choosing employer. She will reason that the best action is to choose the job which pays the most. She then considers the effect of this on whole the industry. She reasons that since the number of available jobs has increased, only firms which manage to offer jobs with higher wages will find employees. She knows that a firm must have high productivity in order to be able to afford higher wages. She then concludes that the increase in wages will induce the least productive firms to leave the market. Therefore, she reaches the same verdict as the author: exposure to trade will increase average firm productivity and profit by forcing the least productive firms to leave and enticing more productive ones to enter the industry.

Again, although the actual way in which the reader will reason might differ from this account, I claim that some sort of mental simulation like this must occur if the reader is to understand Melitz' passage. In this case, however, my claim might be significantly more controversial. The reason is that the passage is located right after the demonstration of the main results. And some of the language used in it does seem to refer to the interpreted formal structure (e.g., 'entry cost'). It might thus seem natural to take the passage as prompting the reader to entertain Melitz' interpreted formal structure. Nevertheless, it seems clear that if the function of the passage is simply to remind the reader of the formal structure, then it is both redundant and misleading. On the one hand, it is redundant since it adds nothing: the reader had already understood the effects of trade exposure from the previously demonstrated results. All the passage does is make the reader mentally rehearse the demonstration. And if the reader had not understood how the results were obtained, then the passage is not much of a help. The results in this case hold not in virtue of the passage, but in virtue of the formal structure. On the other hand, it is misleading since it does not do what Melitz claims it to be doing. Remember that Melitz' intent in the sub-section is to answer the question of 'why' trade forces the least productive firms to exit. He claims to be doing so by identifying the 'mechanism' which 'channels' the effects of trade on his interpreted formal structure. The problem,

however, is that if the passage simply reminds the reader of the formal structure, then there is no 'mechanism' separate from this structure to be had. The reason once again is that the results hold by virtue of demonstrations on the complete interpreted formal structure. Indeed, he needs an extensive structure to obtain his results. From this perspective, there would be no way to obtain a separate 'mechanism' without jeopardizing the results he obtains[26]. In this sense, the passage would not actually answer 'why' trade has the said effects since it would not identify the actual 'channel' through which the effects of trade are manifest. It would do nothing more than remind the reader of the just presented formal structure.

Here the sceptic may retort: well, if the passage isn't supposed to prompt the reader to somehow entertain the interpreted formal structure, why does it employ terms which were formally defined in the latter? Fair enough, but the passage also does not employ any explicit formal notation or demonstration. In any case, the important question is this: can the passage prompt the reader to create and manipulate a mental representation of the situation described? My answer is yes. The reason is that the terms which were formally defined also happen to have a pertinent meaning in natural language. Take, for instance, 'export markets', 'entry cost' and 'potential returns'. These terms are afflicted by a convenient ambiguity: they need not be formally defined in order to be properly understood. Furthermore, given the context in which these terms appear, they seem to perfectly warrant a natural-language-based comprehension of the story (perhaps along the lines I suggested above) which is done independently of the formal structure.

Still, the sceptic may remain unconvinced. She may insist that there appears to be one particular term which is not really ambiguous: 'productivity draw'. This term has arguably no applicable meaning in natural language and would thus pose a problem for my claim. True, but the point is: what is the term supposed to mean in that particular context? Although Melitz defines it in his formal structure as a productivity measure variable drawn from a probability distribution, he does take the time to explain the meaning of the term in terms of real world entities. In his own words, the term is supposed to capture "the fact that firms cannot know their own productivity with certainty until they start producing and selling their good" (ibid., p.1701). Now, if this is what the reader has in mind when trying to comprehend the passage, then the term does not refer to the formal structure but to real world entities. If this is case, then the meaning of the term does not prevent the reader from being able to reason through the events in the story without resorting to any manipulation of the formal structure.

I conclude that one can make sense of the said passage by recognizing that it is a story which prompts a thought experiment in the mind of the reader. The reader reasons from trade exposure, through the domestic labour market, to firm profitability and productivity by mentally simulating the events described in the passage. She then presumably achieves the

[26] Furthermore, these results follow from accepted formal rules of manipulation, not from some sort of underlying 'mechanism'. That is why they are 'demonstrated'. It is thus unclear how a 'mechanism' would directly underscore a formal result. Nonetheless, I will have more to say about this topic in the next chapter where I critically examine the nature of the results which formal structures yield and elaborate on the relation between these structures and stories.

same result as Marc Melitz, namely the realization that exposure to international trade will induce certain changes in industry organization and productivity. This can be done independently of the formal structure since the events described in the passage are ambiguous and can be perfectly understood in terms of real world entities. Yet, the results achieved by reasoning with both the story and the interpreted formal structure are consistent with each other. Thus, Melitz' contribution constitutes additional evidence for my two-fold account of model-based reasoning.

Justifying a Two-Fold Account of Reasoning (Chapter 5)

Here an epistemic grounding for the two-fold account of reasoning put forward in this thesis is attempted. A complementing and corroborating relationship between the two types of reasoning is identified as a justification for why economists may want to reason in this concocted manner.

So far, the main concern of this thesis has been to understand how economists reason with theoretical models. Two different types of reasoning have been identified. The first one revolves around the manipulation of interpreted formal structures. The second one revolves around the manipulation of mental representations prompted by stories. I have made a great effort to argue for the ubiquity of the latter type of reasoning and to differentiate it from the former type. And, indeed, if the reader has been convinced of this claim then my main goal in this thesis has been achieved. Still, at least as it now stands, this claim is remains somewhat wanting. The reason is that, as evidenced by the case studies, economists seem to employ both types of reasoning when working with theoretical models. Consequently, these models often have both an interpreted formal structure and a story. A natural follow-up question is then, why do economists employ both types of reasoning?

Now, I am afraid that a complete answer to this question would fall outside the limited scope of this thesis. The reason is that such answer would probably have to be based on a comprehensive enquiry on the historical development and philosophical grounding of the theoretical modelling practices of economists. Given this worry, what I will do instead is attempt to provide a sketch of an answer to a much more circumscribed question, namely: why may economists *want* to employ both types of reasoning? I expect that this tentative answer will also help to clarify an important aspect of model-based reasoning which has been so far neglected in this thesis, that is the relationship between stories and interpreted formal structures. I will proceed in a stepwise manner, first discussing the desirability of reasoning on the basis of interpreted formal structures before addressing the desirability of reasoning on the basis of stories (and the relationship between the two).

The Desirability of Interpreted Formal Structures (5.1)

The question of why economists may want to reason on the basis of the manipulation of interpreted formal structures is very likely the most uncontroversial one to address. In general, economists seem to hold that employing formal structures constitutes a necessary condition for conducting admissible academic enquiry. As Morgan (2012, p.27) puts it, formal structures "limit how each particular model can be used, and so, *constitute the kinds of right reasoning that are possible with that particular model*". But how exactly formal structures grant rightfulness to the reasoning of economists? Here it might be a good idea to examine a couple

of answers to this question provided by practitioners themselves, many of whom have incidentally taken the time to explain and justify the importance they attribute to formal structures. Take Nobel Memorial Prize Laureate Gérard Debreu:

> "Being denied a sufficiently secure experimental base, economic theory has to adhere to the rules of logical discourse and must renounce the facility of internal inconsistency. A deductive structure that tolerates a contradiction does so under the penalty of being useless, since any statement can be derived flawlessly and immediately from that contradiction. In its mathematical form, economic theory is open to an efficient scrutiny for logical errors" (1991, pp.2-3)

From this perspective, the fact that economic theorizing is so removed from secure empirical verification helps to justify the appeal to formal structures. On the one hand, these structures force the economist to be logically consistent when reasoning with a model. In doing so they limit the scope for certain unwanted 'errors'. On the other hand, these structures make it so that the theoretical results obtained by the economist are deductively warranted. Curiously, this deductive rigor also seems to increase the economist's belief that the results of a theoretical model are somehow useful in making inferences about the world (Grüne-Yanoff, 2013; Gilboa et al., 2014), though elaborating on this effect is not so relevant for the purposes of this thesis. What is relevant is the fact that formal structures are taken to provide some sort of rigor to economic theorizing. Parenthetically, the conception of rigor here referred to is the one closely linked to axiomatization. Although different conceptions of rigor have can be identified in the historical development of mathematical economics (Weintraub, 2001), it seems fair to say that this axiomatic conception is the one contemporary economists have in mind when they extol the virtues of reasoning with formal structures (see Boumans, 2005, ch.6). Still, that is only one justification of why economists want to reason on the basis of formal structures.

A second justification for wanting to reason on the basis of formal structures is that these force the economist to be conceptually precise. What this means is that economic concepts must be given an exact formal characterization before they can be appropriately explored by means of rigorous theoretical modelling. This is often taken to be an advantageous aspect of reasoning on the basis of formal structure. For instance, Paul Romer recently explained:

> "Words that are evocative and ambiguous better serve factional interests than words that are analytical and precise. [...] Tight links between words from natural language and symbols from the formal language of mathematics encourage the use of words that are analytical and precise" (2015, p.89)

For him, the precision obtained by making use of formal structures helps to increase 'scientific consensus' "around theoretical and empirical statements that are true" (ibid.). In turn, he claims, the deleterious effects of disagreements induced by loose use of concepts are reduced. Again, whether the sort of conceptual precision imposed by formal structures does indeed have the effect Romer claims is not so relevant. What is relevant is the fact that economists do in fact seem to value this sort of conceptual precision.

Taken together, these two justifications (rigor and precision) should suffice to support the claim that economists may indeed want to reason on the basis of formal structures. At any rate, it is expected that this claim is already relatively uncontroversial. Moreover, arguing for this claim is not the primary aim of this thesis[27]. Instead, the important aim is to argue for the more controversial (and challenging) claim that economists also want to reason on the basis of stories. This is what I subsequently proceed to attempt.

The Desirability of Stories – Part One: Curtailing Fragile Inferences (5.2)

What is it about stories that makes them desirable for economists? The answer, I suspect, has to do with a topic which I have been so far reluctant to address in this thesis: the justification of model-based inferences. Since, as I have been arguing, models are composite entities it seems reasonable to reckon that all of their components might somehow be implicated in inference justification. Indeed, my suggestion is that the desirability of stories stems from the nature of the inferences they prompt (as compared to those of interpreted formal structures). This is for two interrelated reasons. First, inferences from interpreted formal structures can be fragile in a way that inferences from stories are not. And second, stories contain substantially richer causal information than interpreted formal structures and this expands the range of model inferences the economist feels justified to make. Each of these reasons will be in turn developed.

First, in what way can inferences from formal structures be fragile? One interesting answer to this question has been alluded to by D. McCloskey in what she called 'The A-Prime/C-Prime Theorem':

> "For each and every set of assumptions A implying a conclusion C, there exists a set of alternative assumptions, A', arbitrarily close to A, such that A' implies an alternative conclusion, C", arbitrarily far from C" (1993, p.235)

Although no formal axiomatic proof to this 'theorem' is provided, it still points to an important problematic feature of theoretical modelling in economics, namely the fact that the results obtained by manipulating interpreted formal structures can be extremely sensitive to their specific constraints. This problem was subsequently expanded upon in the philosophical literature and has been given a particularly thoughtful treatment by Nancy Cartwright, who has named it as the problem of 'overconstraint'.

Now, it is not my intention to elaborate on the context in which Cartwright's treatment of the problem appears (namely her metaphysical account of causality) nor to engage in the debate regarding the epistemic nature of the constraints (or assumptions) of economic models. Instead, my aim is simply to point that the problem of 'overconstraint' can be a serious one

[27] Of course, this does not mean that there are no interesting remarks to be made on the topic. For instance, see Weintraub (2002) for a historical account of how mathematical argumentation came to play a prominent role in the academic practices of economists. Nonetheless, this issue is not directly related to the point I want to make in this thesis and thus will be left unexplored.

and that this has consequences for how economists work with their models. But how exactly is this problem manifest?

Take the first case study again. Remember that Christopher Pissarides' aim in that article was to show that unemployment persistence could be caused by the loss of job skills incurred by unemployed workers. Yet, a careful study of the argument will reveal that this does not quite exhaust what goes on in the model's interpreted formal structure. The reason is that the results Pissarides obtains depend not only on the precise formal characterization of both skill loss and unemployment persistence, but also on a great many number of additional constraints which allow for the main result to be deductively obtained. These include (see Figure 2):

1. Discrete time.
2. Two overlapping generations.
   a. Each of fixed size, L.
   b. Each generation is in the job market exactly two periods.
3. Each job lasts one period only and must be refilled at the beginning of every period.
4. The number of jobs, J_t, available at beginning of period t is endogenous.
5. Workers in each of their two life periods are either employed or unemployed.
6. a. Output for young workers and old, previously employed workers = 2.
   b. Output for old, previously unemployed workers = $2y$, $0 < y < 1$.
   ($y < 1$ represents skill loss during unemployment.)
7. Unemployed workers have 0 output, no utility, no income.
   (This is relevant to calculating wages and profits.)
8. In each period all workers and some jobs are available for matching.
9. Each job must be matched at the beginning of a period to be filled in that period.
10. In each period workers and jobs meet at most one partner.
11. The number of matches between a job and a worker is designated by χ, where
    a. χ is at least twice differentiable.
    b. First derivatives in χ are positive; second, negative.
    c. χ is homogeneous of degree 1.
    d. $\chi(0, 2L) = \chi(J_t, 0) = 0$.
    e. $\chi(J_t, 2L) = \max(J_t, 2L)$.
12. There is a probability that a worker meets a job at the beginning of t, designated by f_t.
    a. f_t does not depend on what a worker does nor on whether the worker is employed or unemployed.
    b. f_t is a function only of J_t and L.
13. There is a probability that a job meets a worker at the beginning of period t.
    a. This probability is independent of what jobs do.
    b. This probability is a function only of J_t and L.
14. The cost of opening a job and securing the output as described in 6, is equal to $1/k$
    (whether the job is filled or not).
15. Wages are determined by a Nash bargain.
16. Workers and employers maximise expected utility.

Figure 2. Reproduced from Cartwright (1999, p.171)

These constraints are necessary for otherwise Pissarides might not have been able to construct the link between skill loss and unemployment. Nonetheless, these constraints are distressingly arbitrary. Although many of them are given an interpretation in terms of real world entities, the interpretations themselves do not fix the form of the constraints. Instead, the form of the constraints are fixed on the discretion of the author. But this decision is not necessarily innocuous, it may affect the results the author obtains. For instance, why do the generations of workers have the same size L and stay in the market for exactly two periods? Pissarides does not even seem to bother with this question and yet this particular formal characterization of the labour market is in principle essential to derive his main results. Were the generations of workers to be specified differently, there would be simply no guarantee that the results would hold.

Note that the problem is not simply that the constraints are false or abstract, but that the result may be just too sensitive to them. The same is true *mutatis mutandis* for the second case study. Now, given the potential sensitivity of the results to changes in the constraints, it becomes difficult to believe that the formal structure licenses inferences that can be exported to the actual world. This is because the economist can no longer be sure that the putative causes do in fact produce the putative effects in the formal structure: perhaps the effects are much more dependent on an arbitrary constraint (such as the exact form of a utility or a job matching function) than on the cause itself. And since no systematic empirical work goes into the development of the formal structure (remember that the models discussed in this thesis are theoretical), determining whether one is licensed in making inferences on the basis of over sensitive formal structures becomes very difficult. This is the sense in which inferences from interpreted formal structures can be fragile.

Here, however, one may point that this fragility can be at least partially overcome by means of a practice known in the literature as 'robustness analysis' (Kuorikoski and Lehtinen, 2009; Ylikoski and Aydinonat, 2014). That is, the "practice of deriving the same result using different modelling assumptions" (Kuorikoski and Lehtinen, 2009, p.125). The general idea behind it is to investigate how the results of the interpreted formal structures change with respect to variations in the form of arbitrary constraints. The hope is that this provides valuable "information about the relative importance of the assumptions with respect to the result of interest" (ibid., p.126). Now, although it is not clear whether every theoretical economist engages in this practice, it is fair to say that many do. For instance, Melitz in the second case study does investigate (especially in the footnotes) whether changing specific constraints in his formal structure would change his results in a substantial way.

But can the practice of robustness analysis solve the fragility problem? I don't think so. On the one hand, the aim of the practice is to see how certain constraints affect the results of interest. It can't by itself guarantee that the results will be robust to such changes. Robustness, by definition, will always depend on the particular nature of the formal constraints. In this sense, the practice is simply an attempt to further understand the nature of the fragility problem. On the other hand, robustness analysis too often reveals that results from formal structures are not at all robust (Cartwright, 2009; Alexandrova, 2008, p.388; Odenbaugh, 2011; Odenbaugh

and Alexandrova, 2011). This is to be expected since economic models usually have a complex formal structure whose results hold by virtue of the intricacies of this structure. The second case study is again a case in point. While firm productivity (φ) would admit a couple of different specifications without changing the main result (Melitz, 2003, p.1699), the form of the (constant elasticity of substitution) utility function does affect the result (ibid., p.1715). In this case, as Melitz himself recognizes, the significance of the results he obtains should "be interpreted with caution" (ibid., p.1716). Thus, robustness analysis does not effectively solve the problem, it only serves to diagnose its extension and severity[28].

The upshot is that the fragility of results obtained on the basis of formal structures alone may seriously curtail the plausibility of the economist's inference from these results to the world. The epistemic grounding of these inferences may be simply too thin. In the words of Cartwright:

> "The [formal] assumptions can provide a way to secure deductively validated results [...]. But these create their own problems. For the validity of the conclusions appears now to depend on a large number of very special interconnected assumptions. If so, the validation of the results will depend then on the detailed arrangement of the [formal] structure of the model and is not, prima facie at least, available otherwise. We opt for deductive verification of our claims in order to achieve clarity, rigour and certainty. But to get it we have tied the results to very special circumstances; the problem is how to validate them outside" (2007, p.229)

Now, one advantage of reasoning on the basis of stories is that it allows one to sidestep the fragility problem. As argued in chapter two, stories work by inviting the reader to combine her current knowledge of the world with the information provided in the story in order to create and manipulate a quasi-sensorial representation of the events described. In doing so, the reader imagines and evaluates the sequence of events described in the story and presumably reaches the conclusion intended by the author, namely that the putative causes produce the putative effects. Here is where stories play a crucial role. Remember that formal structures within economic models are interpreted in terms of real world entities (and that includes the causal interpretation superimposed on the structure). But the putative causes and effects are purportedly the same in both the story and the formal structure. This, I suggest, induces the economist to transpose her belief in the appropriateness of the results obtained from the story to the results obtained from the formal structure: the story corroborates the results of the formal structure[29]. And again, since the reader's knowledge of the way the world works is invoked in reasoning with the story, it may become easier for her to believe that the results in

[28] Of course, changing the specificities of the formal structure can also give valuable information about the domain over which its results obtain. This constitutes a practice is similar is to what Hausman (1992) named 'conceptual exploration'. However, as argued in chapter one, this account does not really seem to capture what goes on in the practice of economic modelling.

[29] The corroboration relationship between the story and the interpreted formal structure is similar to that which obtains in the research strategy known in the social sciences as 'triangulation'. That is, "the combination of methodologies in the study of the same phenomenon" (Denzin, 1978, p.291). Yet, economists are usually unclear about the research strategy they follow and thus I choose to remain silent on whether the relationship in question constitutes triangulation.

question can be exported to the world. From this perspective, reasoning on the basis of stories can help to attenuate the fragility problem which afflicts reasoning on the basis of formal structures.

Of course, this does not mean that the results obtained by manipulating mental representations can always be reliably exported to the world. As argued in chapter two, the appropriateness of these results will depend on the correctness of both the reader's knowledge of the world and the information contained in the story. If either of them is incorrect, then the mental representation might lead one to form an inappropriate belief about the world. Furthermore, a story (and its results) can appear to be quite plausible for a reader while at the same time being externally invalid. But in any case, the relevant point here is not that inferences from stories are more warranted than from formal structures. Rather, the point is simply that stories can make the results of formal structures more *plausible*.

The two case studies nicely illustrate this claim. In the first case study, Pissarides goes as far as to remark that the story he presents "drives the results" of his interpreted formal structures (Pissarides, 1992, p.1371). Whatever connotation the verb 'drive' may have in this context, it seems clear that Pissarides is claiming that the results of the formal structure are in some unspecified manner supported by the story. In the second case study, Melitz suggests that the story he presents explains 'why' the causal link (or channel "through which trade can affect the distribution of surviving firms") which is 'operative' in his interpreted formal structure obtains (ibid., p.1715). Again, whatever the terms 'why' and 'operative' may connote in this context, it is clear that Melitz believes that the said results are somehow supported not only by the formal structure as such, but also by the story. Hence, in both cases the stories can be said to at least help to convince the economist of why the respective formal results are justified. They underpin (or drive, or explain) these inferences and in doing so they increase their plausibility[30]. In this sense, stories can indeed make formal results less fragile. This is one reason why economists may also want to make use of stories.

The Desirability of Stories – Part Two: Enriching Inferences (5.3)

Another reason why economists might want to make use of stories is that they provide a considerably broader (and more useful) array of (causal) information than interpreted formal structures. This has to do with the fact that stories are built with concepts from natural language. These concepts differ from those defined in formal structures in two relevant senses. First, as argued in chapter two, they can be successfully employed without being formally defined: they have a pre-existing shared meaning. Second, and more importantly, these natural language concepts can have a thick nature. But what does that mean? In a nutshell, thick concepts are "complex and vague terms [which] lack the conceptual precision that

[30] Although it would be certainly interesting to investigate the precise nature of the 'supporting' relationship going on here, I fear it would take me too far away from the point I am trying to make. In principle, it seems that such investigation would touch on foundational aspects of the nature of the human ability for mental representation (for an introductory discussion of the problem, see Cummins, 1988).

scientific concepts get within axiomatized and mathematized theories" (Cat, 1995, p.222). They were initially referred to as *Ballungen* by Otto Neurath in his criticism of Karl Popper's (and the logical positivists') attempt to logically reconstruct the scientific method. Neurath intended to combat the idea that science could be described as a collection of "well-defined theories built up of clean [logical] statements" (Neurath, 1983, p.121). He instead stressed the importance of "complex (messy) statements of little cleanliness – '*Ballungen*' – [which he claimed to be] the basic material of the sciences" (ibid., p.128). Now, it is not my intention to discuss the historical context in which the notion of thick concepts was developed. Instead, my aim is simply to suggest that these concepts can be useful when making inferences about the world.

The type of thick concepts which are relevant for the claim put forward in this thesis are those which deal with causal relations, that is thick causal concepts. The thickness of these particular concepts stem from the fact that they are rich in causal content. They contain a multitude of details regarding how and under what conditions the causal relations they capture hold[31]. As such, they convey a "vast amount of [...] important, useful information that can help us with crucial questions of design and control" (Cartwright, 2007, p.20). This information can be vital for the economist's attempt to make inferences about the world. However, the richness of these concepts is lost when one attempts to define them in formal structures. This is because the use of these concepts in natural language is very general and this imposes great difficulties to (non-trivial) attempts to capture their meaning in formal language[32].

To see this, take the story in the first case study. Observe the following extract: "If unemployed workers lose some of their skills and become less attractive to firms, fewer jobs come into the market next period..." (Pissarides, 1992, p.1371). This sentence may seem very innocent at first but once one reflects upon it one sees that there are concepts in it which provide an incredible amount of detailed information about the causal link between unemployment and the supply of jobs in the world. For instance, take the natural language concept of 'losing skills'. It is there to complete the link between job supply and unemployment. But the concept itself is thick. It describes a process in the world whereby individuals effectively become unable to perform certain tasks as well as they used to simply because they are no longer performing the task. The meaning of the concept in that context is thus very rich.

In fact, it is so rich that Pissarides does not really completely capture it with his formal structure. What he does instead is link the unemployment level (q) with the reduced productivity (y) resulting from losing skills, and then with job supply (J). He cannot do more without risking the tractability of his demonstrations. Yet, this is very uninformative. The result in the formal structure holds by virtue of the partial derivative of q on J. But the

[31] This point is inspired by Anscombe's (1993) notorious argument that the causation only instantiates itself in very concrete circumstances and that knowledge of these circumstances is what matters in the proper understanding of causal concepts.

[32] For a seminal philosophical discussion of the difficulties involved in formalizing certain concepts of natural language, see Wittgenstein (2009).

derivative contains very little information: it only shows how *q*, *y* and *J* are formally related[33]. The problem is that the formal nature of this result seriously constrain the inferences the economist can make regarding how to alleviate unemployment persistence. Here is where the thickness of the concept of 'losing skills' becomes convenient. It conveys to the economist detailed knowledge of how and where the causal link between unemployment and the supply of jobs holds (and can thus be severed or interfered with). For instance, it allows the economist to infer that the link can be undone should unemployed workers be offered job (re)training, thereby stopping their productivity loss and their attractiveness to firms. Again, this inference cannot be made on the basis of the interpreted formal structure alone. It is made on the economist's knowledge of the thick concept of 'losing skills': the loss of skills may occur whenever one does not exercise particular skills and this can be prevented by regularly exercising these skills, irrespective of whether one is employed or not. In this sense, the richness of the thick causal concepts contained in stories may indeed assist the economist in making inferences about the world. This is another reason why economists may want to make use of stories.

In summary, economists work with theoretical models in order to make inferences about the world. They make use of interpreted formal structures because they want to be precise and rigorous. They want to avoid logical errors and want to employ concepts with a well-defined meaning. However, formal structures may yield results that are quite fragile. These results may be too dependent on the specific constraints of the formal structures and this reduces the belief of economists that they can be exported to the world. This is one reason why stories may be desirable. Results from stories are not fragile in the same way as those from formal structures and as such they can be used to corroborate the latter. Furthermore, well-defined formal concepts can be very uninformative. This constrains the range of inferences that can be made from interpreted formal structures. This is another reason why stories may be desirable. They contain thick concepts which supply substantial additional information about the causal process which is being captured in the formal structure. This information can expand the range of inferences the economist may make about the world. Therefore, it is reasonable that economists may also want to reason on the basis of stories.

[33] Now, I'm not saying that information about formal relations among economic variables have no use (which they obviously have). Instead, I am arguing that, for questions of control and design, this information only supports a limited range of inferences.

Conclusion (Chapter 6)

This thesis has been an attempt to comprehend certain aspects of the modelling practices of economists. It has been argued that theoretical models are better understood as reasoning devices: they are used by economists to make inferences about the world. This realization has been shown to raise one interesting question, namely 'how do economists reason with theoretical models?', which has in turn instigated the argument put forward in this thesis. In order to address this question, the composition of theoretical models has been analysed. It has been argued that these models contain two main elements, a formal structure and a narrative. The former provides certain formal constraints of manipulation whereas the latter furnishes a literary complement which (either directly or indirectly) supplements the former. Based on this distinction, two possible types of model-based reasoning have been identified.

The first one proceeds mostly in the interplay between narratives and formal structures. In this case, a formal structure is constructed, interpreted and manipulated in terms of real world entities. The economists asks a 'question' of this interpreted structure and then 'answers' it by following acceptable rules of manipulation. The 'answer' is therefore seen as being deductively warranted. This constitutes the most visible way in which economists reason with theoretical models.

The second one proceeds mainly in the interplay between the economist and the narrative. In this case, the narrative is seen as containing stories. These are units of text which describe a sequence of events leading to an outcome of interest. The economist reasons with these stories by thought experimenting. She uses her knowledge of the way the world works to create and manipulate a mental representation of the situations described, and then reaches certain outcomes of interest. This is a different type of reasoning in that it requires no explicit operations on formal structures.

This two-fold account of model-based reasoning has been supported by examining two case studies: one theoretical model investigating a possible labour market cause of unemployment persistence and another theoretical model investigating possible industry effects of trade exposure. In both cases, a narrative and a formal structure have been distinctly identified. Moreover, the first type of model-based reasoning was shown to play a prominent role in the argumentative strategy of both contributions. Still, a careful examination of the two cases revealed that a relevant story could also be found in each of them. Reasoning on the basis of these stories generated outcomes that were consistent with the ones obtained by reasoning on the basis of the corresponding interpreted formal structures. This has been shown to constitute evidence for the claim that reasoning with theoretical models can proceed along the two proposed lines.

Finally, an attempt to epistemically justify the desirability of these two types of reasoning was made. It has been argued that the first type of reasoning is valued by economists for being precise and rigorous. Yet, reasoning on the basis of interpreted formal structures alone was

shown to run the risk of producing fragile or limited inferences about the world. These shortcomings were shown to be reduced by simultaneously making use of the second type of reasoning. This is because the latter may increase the plausibility and range of inferences made on the basis of the former (though not necessarily their soundness). From this perspective, it makes sense to expect that economists may want to employ both types of reasoning.

Now, I want to conclude by making a brief remark about the repercussion of the argument developed in this thesis. Julian Reiss has recently suggested that economic models appear to perform "some sort of epistemic hat-trick": they seem to give us knowledge about the economic world "without involving new observations or experiments" (Reiss, 2013, pp.280-81). For him, this looks puzzling because models are seen to confirm hypotheses without providing new data: an animal is mysteriously pulled out of the hat. Now, irrespective of whether there is indeed an inferential puzzle to be had, some of the magic surrounding the hat-trick has hopefully been dispelled by this thesis. As we have seen, reasoning with economic models need not be based on purely formal operations on interpreted formal structures. It can also involve active cognitive effort from the part of the economist, who tries to further her knowledge about the world by using this very knowledge as a basis with which to work. In this sense, there is hardly anything magical about theoretical models, they can be seen as an attempt to make use of our most basic cognitive strategies in order to aid economic enquiry. So the creature was not conjured out of thin air, it was concealed in the hat all along. Of course, this does not mean that the trick is trivial. On the contrary, as the controversy around inference justification has shown, the precise nature of the animal remains very much disputed: it is not at all clear whether the creature is a majestic white dove, an inconspicuous grey mouse or a specious red herring.

References

Akerlof, G. A. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, *84*(3), 488.

Alexandrova, A. (2008). Making Models Count. *Philosophy of Science*, *75*(3), 383–404.

Anscombe, G. E. M. (1993). Causality and Determination. In E. Sosa & M. Tooley (Eds.), *Causation*. Oxford ; New York: Oxford University Press.

Auden, W. H. (1991). *Collected poems*. (E. Mendelson, Ed.) (1st Vintage International ed). New York: Vintage International, Vintage Books.

Backhouse, R. E. (1998). If Mathematics is Informal, Then Perhaps we Should Accept That Economics Must be Informal Too. *The Economic Journal*, *108*(451), 1848–1858.

Backhouse, R. E. (2002). Economic models and reality: the role of informal scientific methods. In U. Mäki (Ed.), *Fact and Fiction in Economics* (pp. 202–213). Cambridge: Cambridge University Press.

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1521), 1281–1289.

Boumans, M. (2005). *How economists model the world into numbers*. London ; New York: Routledge, Taylor & Francis Group.

Brown, J. R. (2004). Why Thought Experiments Transcend Empiricism. In C. Hitchcock (Ed.), *Contemporary debates in philosophy of science*. Malden, MA: Blackwell Pub.

Calvo-Merino, B., Grèzes, J., Glaser, D. E., Passingham, R. E., & Haggard, P. (2006). Seeing or Doing? Influence of Visual and Motor Familiarity in Action Observation. *Current Biology*, *16*(19), 1905–1910.

Cartwright, N. (1999). *The dappled world: a study of the boundaries of science*. Cambridge, UK : New York, NY: Cambridge University Press.

Cartwright, N. (2007). *Hunting causes and using them: approaches in philosophy and economics*. Cambridge ; New York: Cambridge University Press.

Cartwright, N. (2009). If No Capacities Then No Credible Worlds. But Can Models Reveal Capacities? *Erkenntnis*, *70*(1), 45–58.

Cat, J. (1995). The Popper-Neurath debate and Neurath's attack on scientific method. *Studies in History and Philosophy of Science Part A*, *26*(2), 219–250.

Clark, A., & Chalmers, D. (1998). The Extended Mind. *Analysis*, *58*(1), 7–19.

Colander, D., Holt, R., & Rosser, B. (2004). The changing face of mainstream economics. *Review of Political Economy*, *16*(4), 485–499.

Craver, C. F. (2006). When mechanistic models explain. *Synthese*, *153*(3), 355–376.

Cummins, R. (1989). *Meaning and mental representation*. Cambridge, Mass: MIT Press.

Debreu, G. (1991). The Mathematization of Economic Theory. *The American Economic Review*, *81*(1), 1–7.

Decety, J., & Grèzes, J. (2006). The power of simulation: Imagining one's own and other's behaviour. *Brain Research*, *1079*(1), 4–14.

Denzin, N. K. (1978). *The research act: a theoretical introduction to sociological methods* (2d ed). New York: McGraw-Hill.

Evans, J. S. B. T. (2002). Logic and human reasoning: An assessment of the deduction paradigm. *Psychological Bulletin*, *128*(6), 978–996.

Gendler, T. S. (2000). *Thought experiment: on the powers and limits of imaginary cases*. New York: Routledge.

Gendler, T. S. (2004). Thought Experiments Rethought--and Reperceived. *Philosophy of Science*, *71*(5), 1152–1163.

Gendler, T. S. (2006). Thought Experiments. In L. Nadel (Ed.), *Encyclopedia of Cognitive Science*. Chichester: John Wiley & Sons, Ltd.

Gibbard, A., & Varian, H. R. (1978). Economic Models. *The Journal of Philosophy*, *75*(11), 664–677.

Giere, R. N. (2002). Models as Parts of Distributed Cognitive Systems. In L. Magnani & N. J. Nersessian (Eds.), *Model-Based Reasoning* (pp. 227–241). Boston, MA: Springer US.

Gilboa, I., Postlewaite, A., Samuelson, L., & Schmeidler, D. (2014). Economic Models as Analogies. *The Economic Journal*, *124*(578), F513–F533.

Grüne-Yanoff, T. (2009). Learning from Minimal Economic Models. *Erkenntnis*, *70*(1), 81–99.

Grüne-Yanoff, T. (2013). Appraising Models Nonrepresentationally. *Philosophy of Science*, *80*(5), 850–861.

Grüne-Yanoff, T., & Schweinzer, P. (2008). The roles of stories in applying game theory. *Journal of Economic Methodology*, *15*(2), 131–146.

Hacking, I. (1992). "Style" for historians and philosophers. *Studies in History and Philosophy of Science Part A*, *23*(1), 1–20.

Hartmann, S. (1999). Models and stories in hadron physics. In M. S. Morgan & M. Morrison (Eds.), *Models as Mediators* (pp. 326–346). Cambridge: Cambridge University Press.

Hausman, D. M. (1992). *The inexact and separate science of economics*. Cambridge ; New York: Cambridge University Press.

Hausman, D. M. (2015). Much ado about models. *Journal of Economic Methodology*, *22*(2), 241–246.

Hedström, P., & Ylikoski, P. (2010). Causal Mechanisms in the Social Sciences. *Annual Review of Sociology*, *36*(1), 49–67.

Hempel, C. G. (1970). *Aspects of scientific explanation, and other essays in the philosophy of science*. New York: Free Press.

Hindriks, F. (2013). Explanation, understanding, and unrealistic models. *Studies in History and Philosophy of Science Part A*, *44*(3), 523–531.

Horsten, L. (2015). Philosophy of Mathematics. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2015). Retrieved from http://plato.stanford.edu/archives/spr2015/entries/philosophy-mathematics/

Johnson-Laird, P. N. (2008). *How we reason*. Oxford; New York: Oxford University Press.

Johnson-Laird, P. N. (2010). Mental models and human reasoning. *Proceedings of the National Academy of Sciences*, *107*(43), 18243–18250.

Khemlani, S. S., Barbey, A. K., & Johnson-Laird, P. N. (2014). Causal reasoning with mental models. *Frontiers in Human Neuroscience*, *8*.

Knuuttila, T. (2011). Modelling and representing: An artefactual approach to model-based representation. *Studies in History and Philosophy of Science Part A*, *42*(2), 262–271.

Krugman, P. (1998). Two Cheers for Formalism. *The Economic Journal*, *108*(451), 1829–1836.

Kuhn, T. (1977). A Function for Thought Experiments. In *The Essential Tension: Selected Studies in Scientific Tradition and Change* (pp. 240–265). Chicago: University of Chicago Press.

Kuorikoski, J., & Lehtinen, A. (2009). Incredible Worlds, Credible Results. *Erkenntnis*, *70*(1), 119–131.

Kuorikoski, J., & Ylikoski, P. (2015). External representations and scientific understanding. *Synthese*, *192*(12), 3817–3837.

León, J., & Peñalba, G. (2002). Understanding causality and temporal sequence in scientific discourse. In J. Otero, J. A. León, & A. C. Graesser (Eds.), *The psychology of science text comprehension* (pp. 155–178). Mahwah, N.J: L. Erlbaum Associates.

Lipsey, R. G. (2001). Successes and failures in the transformation of economics. *Journal of Economic Methodology*, *8*(2), 169–201.

Lipton, P. (2004). *Inference to the best explanation* (2nd ed). London ; New York: Routledge/Taylor and Francis Group.

Lucas, R. E. (1972). Expectations and the neutrality of money. *Journal of Economic Theory*, *4*(2), 103–124.

Maas, H. (2014). *Economic methodology: an historical introduction*. London ; New York: Routledge.

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about Mechanisms. *Philosophy of Science*, *67*(1), 1–25.

Mach, E. (1976). On Thought Experiments. In E. Mach & E. N. Hiebert, *Knowledge and Error* (pp. 134–147). Dordrecht: Springer Netherlands.

Mäki, U. (2009). Realistic Realism about Unrealistic Models. In H. Kincaid & D. Ross (Eds.), *The Oxford handbook of philosophy of economics*. Oxford ; New York: Oxford University Press.

Mankiw, N. G. (2015). *Principles of economics* (Seventh edition). Stamford, CT: Cengage Learning.

McCloskey, D. N. (1993). The A-Prime/C-Prime Theorem. *Eastern Economic Journal*, *19*(2), 235–238.

McCloskey, D. N. (1998). *The rhetoric of economics* (2nd ed). Madison, Wis: University of Wisconsin Press.

Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, *71*(6), 1695–1725.

Miščević, N. (1992). Mental models and thought experiments. *International Studies in the Philosophy of Science*, *6*(3), 215–226.

Morgan, M. S. (2001). Models, stories and the economic world. *Journal of Economic Methodology*, *8*(3), 361–384.

Morgan, M. S. (2004). Imagination and Imaging in Model Building. *Philosophy of Science*, *71*(5), 753–766.

Morgan, M. S. (2012). *The world in the model: how economists work and think*. Cambridge ; New York: Cambridge University Press.

Morgan, M. S., & Knuuttila, T. (2012). Models and Modelling in Economics. In *Philosophy of Economics* (pp. 49–87).

Nersessian, N. J. (1992). In the Theoretician's Laboratory: Thought Experimenting as Mental Modeling. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, *1992*, 291–301.

Nersessian, N. J. (2007). Thought Experimenting as Mental Modeling: Empiricism without Logic. *Croatian Journal of Philosophy*, *7*(2), 125–161.

Nersessian, N. J. (2008). *Creating scientific concepts*. Cambridge, Mass.: MIT Press.

Neurath, O. (1983). Pseudorationalism of Falsification. In O. Neurath, R. S. Cohen & M. Neurath (Eds.), *Philosophical Papers 1913–1946* (pp. 121–131). Dordrecht: Springer Netherlands.

Noordman, L. G. M., & Vonk, W. (1998). Memory-based processing in understanding causal information. *Discourse Processes*, *26*(2-3), 191–212.

Odenbaugh, J., & Alexandrova, A. (2011). Buyer beware: robustness analyses in economics and biology. *Biology & Philosophy*, *26*(5), 757–771.

Odenbaugh, J. (2011). True Lies: Realism, Robustness, and Models. *Philosophy of Science*, *78*(5), 1177–1188.

Pissarides, C. A. (1992). Loss of Skill During Unemployment and the Persistence of Employment Shocks. *The Quarterly Journal of Economics*, *107*(4), 1371–1391.

Reiss, J. (2012). Genealogical Thought Experiments in Economics. In M. Frappier, L. Meynell, & J. R. Brown (Eds.), *Thought experiments in philosophy, science, and the arts*. New York: Routledge.

Reiss, J. (2013). The explanation paradox redux. *Journal of Economic Methodology*, *20*(3), 280–292.

Romer, P. M. (2015). Mathiness in the Theory of Economic Growth. *American Economic Review*, *105*(5), 89–93.

Samuelson, P. A. (1958). An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money. *Journal of Political Economy*, *66*(6), 467–482.

Schabas, M. (2008). Hume's monetary thought experiments. *Studies in History and Philosophy of Science Part A*, *39*(2), 161–169.

Sorensen, R. A. (1992). *Thought experiments*. New York: Oxford University Press.

Sugden, R. (2000). Credible worlds: the status of theoretical models in economics. *Journal of Economic Methodology*, *7*(1), 1–31.

Sugden, R. (2013). How fictional accounts can explain. *Journal of Economic Methodology*, *20*(3), 237–243.

Thoma, J. (2016). On the Hidden Thought Experiments of Economic Theory. *Philosophy of the Social Sciences*, *46*(2), 129–146.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, *20*(3), 273–281.

Weintraub, E. R. (2001). Measurement, and Changing Images of Mathematical Knowledge. *History of Political Economy*, *33*(Suppl 1), 303–312.

Weintraub, E. R. (2002). *How economics became a mathematical science*. Durham ; London: Duke University Press.

Wittgenstein, L. (2009). *Philosophische Untersuchungen - Philosophical investigations*. (G. E. M. Anscombe, P. M. S. Hacker, & J. Schulte, Trans.) (Rev. 4th ed). Chichester, West Sussex, U.K. ; Malden, MA: Wiley-Blackwell.

Ylikoski, P., & Aydinonat, N. E. (2014). Understanding with theoretical models. *Journal of Economic Methodology*, *21*(1), 19–36.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162–185.